Dissecting the Function of the *Leishmania* Flagellum in a CRISPR-Cas9 Knockout Screen



A thesis submitted for the degree of
Doctor of Philosophy at the University of Oxford

Trinity Term 2019

Tom Beneke
St Edmunds Hall
Sir William Dunn School of Pathology

“The only true wisdom is in knowing you know nothing.”

— Socrates

Abstract

Eukaryotic cilia and flagella are conserved organelles with specialized structures and functions. Cilia and flagella can be well studied in parasites of the genus *Leishmania*, which possess at different stages in their life cycle either a motile 9 + 2 or immotile 9 + 0 flagellum. Much remains unknown about *Leishmania*-specific flagellar biology, for example whether flagellar motility is required throughout the *Leishmania* life cycle. More fundamentally, it is still not fully known how flagellar bending is controlled at a molecular level in motile cilia and flagella.

To dissect flagellar function in *Leishmania* with regard to these open questions, the Cas9 gene editing toolbox LeishGEdit was developed, enabling a knockout screen of ~450 proteins identified in the promastigote flagellum of *L. mexicana*. This identified 14 mutants with a flagellar curling phenotype. A systematic analysis of the curling phenotype showed that flagellar bending in *Leishmania* is supported by the extra-axonemal paraflagellar rod structure and regulated through central pair projections and inner dynein arm (IDA) f/I1 components. Further analysis of the IDA f/I1 tether/tether head (T/TH) complex identified LAX28 as a novel IDA f/I1 assembly factor. However, using the CRISPR-Cas9 toolbox not only allowed a dissection of the regulation of flagellar bending *in vitro*, but also enabled assessing the requirement for motility *in vivo*. By passaging a range of barcoded *Leishmania* motility mutants in pooled populations through sand flies and mice, flagellar-driven directional motility was shown to be essential for colonialization of the sand fly vector, but dispensable for survival in the mammalian host. Thus, this thesis shows the implementation of a powerful Cas9 toolbox that provided new insights into the regulation of flagellar bending and the importance of *Leishmania* motility *in vivo*.

Acknowledgement

I would like to start by thanking my supervisor Eva Gluenz for so many things: scientific advice, helpful discussions, patients with my temper at lab meetings and lots of extremely useful comments on manuscripts, abstracts and especially this thesis, but most importantly for giving me the freedom during my PhD to research on what I found most interesting and not dictating me what I should do instead.

I also would like to thank all the members of the Gluenz, Gull and now Sunter and Wheeler lab, for helpful discussions over all these years. A special thanks goes to my fellow PhD students and friends, Laura Makin and Jessica Valli who were always there and turned the music louder when it was needed late at night in the lab. A special thanks goes also to Edward Hookway and Ross Madden, who both contributed each an essential piece of technology to methods used and developed in this PhD. Many thanks also to everyone in the Dunn School, in research facilities and elsewhere in Oxford who contributed in any way to the completion of this thesis. I also really enjoyed many collaborations during my PhD and I am especially grateful to Jeremy Mottram's and Petr Volf's lab for help with experiments.

I would also like to thank my family and especially my parents for all her support. I would not be where I am today without their help. Many thanks also to my wonderful daughter, Malea, who always makes me smile, no matter what. Finally, I cannot express how much I would like to thank my wife Maria. Thank you for going through this with me, for being a loving wife, for supporting me, for understanding when it took longer in the lab, for putting up with me when I moan about not working experiments or annoying colleagues, for being the best mother and for looking after Malea while I was writing this thesis. I owe you everything.

Table of Contents

Abstract ii

Acknowledgement iii

Table of Contents iv

List of Figures xii

List of Tables xvi

Abbreviations xvii

Declarations xix

1 Introduction 2

1.1 Eukaryotic cilia and flagella 2

1.1.1 Motile cilia and flagella 3

1.1.2 Non-motile cilia and flagella 7

1.1.3 Flagellar and ciliary assembly 10

1.1.4 Ciliopathies 12

1.2 *Leishmania* parasites 13

1.2.1 Leishmaniasis 13

1.2.2 The *Leishmania* life cycle and interactions with host and vector 15

1.2.3 The *Leishmania* flagellum – a multifunctional organelle 20

1.2.4 The *Leishmania* flagellar proteome 21

1.3 The challenges and key aims 28

1.3.1 Project questions 28

1.3.2 Project aims 29

1.3.3 Project strategy 29

2 LeishGEdit - a CRISPR-Cas9 toolbox for kinetoplastids 32

2.1 Introduction 32

2.1.1 Chapter summary and aim 32

2.1.2 Loss-of-function strategies in *Leishmania* before CRISPR 33

2.1.3 CRISPR-Cas9 gene editing 35

2.1.4 CRISPR-Cas9 gene editing in *Leishmania* and other kinetoplastids. 37

2.2 Results 42

2.2.1 Constitutive expression of Cas9 and T7 RNAP in *Leishmania* 42

2.2.2 sgRNA templates with a T7 RNAP promotor and a modular plasmid system allow for rapid PCR generation of transfection constructs 44

2.2.3 LeishGEdit: An online resource for gene editing in kinetoplastids and a streamlined high-throughput gene editing pipeline 49

2.2.4 Efficient gene editing of PF16 in *Leishmania mexicana* 52

2.2.5 Whole genome sequencing of *Leishmania* parasites confirms precise mutations 61

2.2.6 Constitutive expression of Cas9 and T7 RNAP in *Leishmania* and *Trypanosoma* spp. 74

2.3 Discussion 77

2.3.1 Limitations of CRISPR-Cas9 gene editing in *Leishmania* 77

2.3.2 Possible applications for LeishGEdit 81

3 A systematic analysis of the flagellar curling phenotype provides new insights into the regulation of flagellar bending 84

3.1 Introduction 84

3.1.1 Chapter summary and aim 84

3.1.2 Asymmetric distribution of dynein activity results in flagellar bending 85

3.1.3 Regulation of dynein activity and flagellar bending 88

3.1.4 Extra-axonemal structures and their function in flagellar bending . 92

3.1.5 Functional analysis of flagellar proteins identifies the flagellar curling phenotype 94

3.2 Results 97

3.2.1 Discovery of mutants with curled flagella 97

3.2.2 A subset of 49 flagellar mutants for the flagellar curling analysis .. 99

3.2.3 Flagellar curling restricts motility 100

3.2.4 Loss of IDA and CPC proteins causes flagellar curling 106

3.2.5 Failed recovery of flagellar curling by ATP reactivation suggest a curled flagellum is locked at one extreme of bending 110

3.2.6 A subset of secondary deletions resolves flagellar curling 113

3.2.7 CPPs influence flagellar curling 119

3.2.8 The paraflagellar rod may determine the flagellar curl direction .. 124

3.3 Discussion 129

3.3.1 A model for flagellar curling gives new insights into the regulation of flagellar bending 129

3.3.2 IDA f/I1 components and CPPs are vital for progression of the flagellar beat in *Leishmania* 132

3.3.3 A possible model in which CPPs control the periodic switching of dynein activity 136

3.3.4 The paraflagellar rod amplifies the principal bend 139

4 LAX28 is required for assembly of the IDA f/I1 and T/TH within the *Leishmania flagellum* 143

4.1 Introduction 143

4.1.1 Chapter summary and aim 143

4.1.2 Regulation and function of IDAs 144

4.1.3 The IDA f/I1 tether and tether head complex (T/TH) 146

4.2 Results 148

4.2.1 Identification of proteins missing from *ΔCFAP44* mutant salt extracted axonemes 148

4.2.2 Axonemal localisation of LAX28 depends on CFAP44 and IC140 154

4.2.3 Axonemal localisation of CFAP44 and IC140 depends on LAX28 159

4.2.4 LAX28 is an inner arm dynein component required for flagellar function 166

4.2.5 LAX28 shows similarity to human testis expressed protein TEX47... 167

4.3 Discussion 173

4.3.1 LAX28 is an IDA component required for assembly of IDA f/I1 ... 173

4.3.2 IDA f/I1 may control activity of remaining IDA subspecies 176

4.3.3 The IDA f/I1 and T/TH complex shows diversity across flagellated eukaryotes 177

5 Dissecting the requirement of flagellar motility for survival *in vivo* by using a CRISPR-Cas9 bar-seq strategy 181

5.1 Introduction 181

5.1.1 Chapter summary and aim 181

5.1.2 The *Leishmania* flagellum and its role in parasite-host and parasite-vector interactions 182

5.1.3 The principles of bar-seq experiments 185

5.2 Results 192

5.2.1 A bar-seq strategy allows pooled mutant screens 192

5.2.2 Setting up a bar-seq *in vivo* and *in vitro* fitness screen 201

5.2.3 A bar-seq fitness screen in sand flies identifies proteins required for colonialization of sand flies 209

5.2.4 Paralysed and uncoordinated motility phenotypes are cleared from sand flies 212

5.2.5 A bar-seq fitness screen in mice identifies proteins required for infection 216

5.3 Discussion 221

5.3.1 Flagellar-driven directional motility in *Leishmania* is essential for colonialization of the vector 221

5.3.2 The *Leishmania* amastigote flagellum has a role in parasite survival in the mammalian host 225

5.3.3 The *Leishmania* flagellum is a multifunctional organelle 229

6 Conclusions and Outlook 231

6.1 Reflection of research strategies for project aims 231

6.2 The future of gene editing in *Leishmania* 232

6.3 The *Leishmania* flagellum as a model for motile cilia and flagella gives new insight into flagellar bending 236

6.4 General conclusion 240

7 Materials and Methods 242

7.1 Cell culture 242

7.2 LeishGEdit system and CRISPR mutants 242

7.2.1 Construction of pT knockout, pPLOT tagging and Cas9 T7 RNAP expression plasmids 242

7.2.2 CRISPR-Cas9 gene knockouts and tags 243

7.2.3 Diagnostic PCR for knockout verification 243

7.2.4 Addback construction 245

7.3 Microscopy 245
7.3.1 Light microscopy 245
7.3.2 Fixation and cytoskeleton preparation for TEM and SEM 246
7.3.3 TEM 246
7.3.4 SEM 247
7.3.5 Image processing 247
7.3.6 Motility analysis 248
7.4 ATP reactivation assay 248
7.5 Bar-seq pooling assays and infections 249
7.5.1 Pooling of mutants 249
7.5.2 *Lutzomyia longipalpis* infections, dissection and DNA extraction 250
7.5.3 BALB/c infections and DNA extraction 250
7.6 DNA sequencing 251
7.6.1 Library preparation for bar-seq samples 251
7.6.2 Library preparation for genome samples and Illumina sequencing 251
7.6.3 Sanger sequencing 254
7.7 Computing 254
7.7.1 Barcode counting 254
7.7.2 Bam file assembly 256
7.7.3 Automated design of sgRNA and primers using LeishGEdit v4.2 256

7.7.4 BLASTp searches 256
7.8 Proteomics 257
7.8.1 Western blot 257
7.8.2 Salt extracted axonemes of *ΔCFAP44* mutants and parentals ... 257
7.8.3 Proteomics 260
7.8.4 Proteomic analysis 261
7.8.5 Solubilisation of CFAP44 261
8 References 262
9 Appendix 288
9.1 List of supplementary material 288

List of Figures

Figure 1. The *Leishmania* promastigote flagellum. 5

Figure 2. The *Leishmania* amastigote flagellum. 9

Figure 3. The *Leishmania* life cycle..... 18

Figure 4. Isolation of promastigote flagella and deflagellated cell bodies. 24

Figure 5. Proteomic analysis of flagella and deflagellated cell bodies. 25

Figure 6. Enrichment plot of flagella and cell body identified proteins. 26

Figure 7. Summary of tagged proteins in Beneke et al. (2019b)..... 27

Figure 8. Gene replacement efficiency in *Leishmania mexicana*. 41

Figure 9. Effect on growth by Cas9 and T7 RNAP expression in *L. mexicana*. .. 43

Figure 10. Efficient sgRNA production for CRISPR gene editing..... 45

Figure 11. A modular plasmid system for gene editing in *Leishmania*. 46

Figure 12. PF16 pPLOT tagging. 48

Figure 13. sgRNA mismatch and perfect match count for TritrypDB41. 54

Figure 14. Current LeishGEdit online resources. 55

Figure 15. Streamlined LeishGEdit pipeline for mutant generation..... 58

Figure 16. Gene editing of PF16 by using the LeishGEdit toolbox. 60

Figure 17. Sequencing of *ΔPF16* mutants and *L. mexicana* Cas9 T7 parental cells. 67

Figure 18. PCR verification and Illumina sequencing of *ΔPKAC1* and *ΔLmxM.08_29.1000*. 69

Figure 19. DNA alignment of PKAC1 and 2 ORF. 70

Figure 20. Analysis of chromosome ploidy in mutants and parental cells..... 71

Figure 21. *ΔCD047* sgRNA analysis. 73

Figure 22. Assessing versions of Cas9 and T7 RNAP expressing cell lines. 76

Figure 23. Dynein power stroke conformations in the switch inhibition model. ... 87

Figure 24. Current models for dynein regulation. 89

Figure 25. Classes of mutant phenotypes identified in a functional knockout screen. 96

Figure 26. Criteria for counting flagellar curling. 98

Figure 27. A subset of 49 flagellar mutants for the flagellar curling analysis. ... 104

Figure 28. Motility analysis of mutants with curled flagella. 105

Figure 29. Loss of IDA and CPC proteins causes flagellar curling. 109

Figure 30. ATP reactivation and live cell video microscopy of *Δhydin* mutants. 112

Figure 31. Illumina sequencing of *Δhydin* mutants. 114

Figure 32. PCR verification of deletions in *Δhydin* mutants. 116

Figure 33. A subset of secondary deletions resolves flagellar curling. 118

Figure 34. Mutant verification of CPC single and co-deletions. 121

Figure 35. Analysis of CPC co-deletions..... 123

Figure 36. Position of the paraflagellar rod within bent flagella..... 127

Figure 37. Position of the PFR in *ΔPF16* and *Δhydin* flagellar curling mutants. 128

Figure 38. Proposed model for flagellar curling and flagellar bend regulation. . 131

Figure 39. Solubilisation of CFAP44. 149

Figure 40. Salt extracted flagellar axonemes of *L. mexicana.* 151

Figure 41. Null mutant verification..... 156

Figure 42. Testing localization dependencies for LAX28. 158

Figure 43. Western blot analysis. 160

Figure 44. TEM of *ΔLAX28* mutants. 161

Figure 45. Tracking IC140::eYFP and CFAP44::mNG signal over the *Leishmania* cell cycle in *ΔLAX28* null mutants. 164
Figure 46. Tracking LAX28::mNG signal over the *Leishmania* cell cycle in *ΔIC140* and *ΔCFAP44* null mutants. 165
Figure 47. Illumina sequencing of *ΔLAX28* null mutants. 169
Figure 48. Flagellar curling and motility analysis of LAX28. 171
Figure 49. Protein alignments of LmxM.24.1310 (LAX28). 172
Figure 50. Model: LAX28 is required for assembly of the IDA f/I1 complex within the flagellum. 175
Figure 51. Motility analysis of mutants. 191
Figure 52. Overview of bar-seq strategy for mutant fitness phenotyping in pools. 195
Figure 53. Sanger sequencing of barcoded mutants and control cell lines. 198
Figure 54. Testing the *in vitro* bar-seq screen. 199
Figure 55. Pilot bar-seq *in vivo* screen in sand flies. 200
Figure 56. Null mutant PCR validation. 204
Figure 57. Overview for *in vitro* and *in vivo* bar-seq screen. 205
Figure 58. Promastigote growth mutant analysis. 207
Figure 59. Sanger sequencing the integrated barcode cassette in *ΔDRC2* mutants. 208
Figure 60. Bar-seq *in vivo* screen in sand flies. 211
Figure 61. Development of *ΔPF16*, *ΔMBO2* and *ΔBBS2* mutants in *L. longipalpis* infections. 213
Figure 62. Mapping mutants with reduced passage in sand flies with their motility phenotype. 215

Figure 63. Bar-seq *in vivo* screen in mice. 219

List of Tables

Table 1. Identification of *L. mexicana* orthologues of well-characterised axonemal proteins. 102

Table 2. A subset of 49 flagellar mutants for the flagellar curling analysis. 103

Table 3. Quantitative proteomics of *ΔCFAP44* mutant derived flagellar axonemes. 153

Table 4. *Leishmania* genes identified to be required for colonialization of the sand fly and infection of mouse. 220

Table 5. Primer sequences used for diagnostic PCR to show presence of DNA. 244

Table 6. Custom designed p5 and p7 primers for Illumina amplicon sequencing. 253

Table 7. Primer sequences for Sanger sequencing. 255

Table 8. Anitbodies and buffers used in Western blots. 258

Abbreviations

9v	9 – variable
AB	Addback
AMG	Abdominal midgut
Bar-seq	Barcode analysis by sequencing
BBS	Bardet-Biedl Syndrome
cAMP	Cyclic adenosine monophosphate
CAR	Cardia
Cas9	CRISPR-associated gene 9
Ci	Cell body insoluble
CPC	Central pair complex
CPP	Central pair projections
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
Cs	Cell body soluble
CSC	CaM and Spoke associated Complex
DB	Database
DiCre	Dimerizable Cre recombinase
DSB	Double strand break
ESP	Endoperitrophic space
Fi	Flagellar insoluble
Fs	Flagellar soluble
GDP-MP	GDP-mannose pyrophosphorylase
GPI	Glycosylphosphatidylinositol
HF	Homology flank
hSpCas9	Humanized *Streptococcus pyogenes* Cas9
ICLC	Intermediate chain-light chain complex
IDA	Inner dynein arm
IFT	Intraflagellar transport
LPG	Lipophosphoglycan
MIA	Modifier of inner arms

MM	Mismatches
MMEJ	Microhomology-mediated end joining
N-DRC	Nexin-dynein regulatory complex
NHEJ	Non-homologous end joining
ODA	Outer dynein arm
PAM	protospacer adjacent motif
PCD	Primary Ciliary Dyskinesia
pf	Paralyzed flaglla
PFR	Paraflagellar rod
PG	Phosphoglycan
PMI	Phosphomannose isomerase
PMM	Phosphomannomutase
RNAi	RNA interference
RSP	Radial Spoke
SaCas9	Cas9 derived from Staphylococcus aureus
sgRNA	Single-guide RNA
SNV	Single nucleotide variance
SSA	Single-strand annealing
SpCas9	Cas9 derived from Streptococcus pyogenes
SV	Stomodeal valve
T/TH	Tether/tether head complex
tacrRNA	Trans-activating crRNA
TALEN	Transcription activator-like effector nucleases
TMG	Thoracic midgut
ZNF	Zinc finger nucleases
Δ	Null mutant (knockout)

Declarations

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification at this or any other university or institute of learning.

Data shown in this thesis has been published in Beneke et al. (2017), Beneke et al. (2019b), Beneke et al. (2019c) and Beneke and Gluenz (2019) or has been deposited as a manuscript in bioRxiv (Beneke et al. (2019a), bioRxiv 782888; doi:https://doi.org/10.1101/782888).

Excluding bibliography, appendices, diagrams and tables, the text of this thesis consists of 49,999 words.

CHAPTER 1

INTRODUCTION

1 Introduction

1.1 Eukaryotic cilia and flagella

Eukaryotic cilia and flagella are highly conserved structures. They are found in single cells and multicellular organisms. These include the green algae *Chlamydomonas*, protozoan parasites of the genera *Leishmania* and *Trypanosoma*, specialized cells in humans and other vertebrates, as well as cells in non-vertebrates such as *Drosophila*. The term flagella and cilia describe essentially the same structure and are often used interchangeably. However, historically the different names were used according to the number of cilia and flagella a cell possess. Typically, a cell has many cilia or few flagella (Lodish et al., 2000). This is also reflected by the translation of their latin name, as cilium translates into “eyelash” and flagellum into “a whip”. Flagella and cilia are typically classified into motile and non-motile organelles and their ultrastructure is described by the microtubule arrangement found within the axoneme, meaning the number of outer microtubule doublets and inner microtubule singlets. The axoneme, defined as the cilia and flagellar skeleton, starts as a bundle of triplet microtubules in the basal body. Then the A and B tubule are extended to form the outer doublets while the C tubule terminates at the transition zone (Ishikawa, 2017).

Even though cilia and flagella are composed of structures and proteins that are highly conserved, microtubule arrangements can be incredibly diverse, ranging from motile flagella with 3 + 0 arrangements e.g. in male gametes of the parasitic protozoan *Diplauxis hatti* (Prensier et al., 1980) to more common motile 9 + 2 cilia and flagella. Meanwhile, non-motile cilia have typically a 9 + 0 microtubule arrangement. However, there are exceptions known, as for example nodal 9 + 0

cilia in embryos, which are believed to be involved in left-right symmetry breaking, are found to be motile (Wagner and Yost, 2000). Also 9 + 0 sensory cilia in Johnston's organ neurons of *Drosophila melanogaster*, which are required for sound-evoked electrical responses (Sarpal et al., 2003), actively generate motion (Gopfert and Robert, 2003). Diversity among the outer microtubule doublet arrangement and absence or presence of a central pair, allow cilia and flagella to fulfil different roles in context of their biological function, such as signaling and motility.

1.1.1 Motile cilia and flagella

Motile cilia are among the earliest known eukaryotic cell organelles, first described by the pioneering work of Antony van Leeuwenhoek around 1647 (Fred, 1933). They generate motion and fulfil various roles in different biological aspects, including transport of fluids, movement of individual cells and also sensing (Mitchell, 2017).

The flagellar beat is generated and propagated by the axoneme itself, as demembranated cilia and flagella can be reactivated by supplementation of ATP (Brokaw, 1975; Gibbons and Gibbons, 1972). However, it remains largely unknown how cilia and flagella regulate movement and control flagellar bending. There are various hypotheses, including purely mechanistic or signaling proposals, as well as mechanosignaling models that describe how motor proteins in cilia and flagella are regulated to slide microtubules against each other and generate the ciliary or flagellar beat, as further outlined in the introduction of chapter 3.

Motile cilia and flagella typically have nine outer microtubule doublets and usually a central pair (Figure 1 A and B). This includes the promastigote (insect form)

flagellum of *L. mexicana*, but not that of the amastigote form in the mammal. While doublets consist of an A and B microtubule, which are each formed by α/β – tubulin heterodimers, the central pair consists of a C1 and C2 microtubule. This ultrastructure is referred to as canonical 9 + 2 microtubule axoneme (Ishikawa, 2017) (Figure 1 A and B).

Attached to the A microtubule are dyneins, a family of ATPase motor proteins that are arranged in units along the axoneme (King, 2016). In the cilia model organism *Chlamydomonas* axenomal dynein heavy chains assemble with other small and intermediate chains into complexes that are divided into outer dynein and inner dynein arms (ODAs and IDAs). ODAs are homogeneously arranged in 24 nm repeat units, while IDAs are organized in seven subspecies (dyneins a, b, c, d, e, f/I1 and g), which are each uniquely targeted within a 96 nm repeat unit (Bui et al., 2012; Heuser et al., 2012) (Figure 1 A and B). Dyneins undergo ATP - powered conformational changes, which cause microtubules to slide against each other and consequently result in flagellar bending (Lin and Nicastro, 2018), as further outlined in chapter 3.

In *Leishmania* and other 9 + 2 motile cilia and flagella, the nine outer doublets are connected through nexin links and linked with the central pair through flexible connections composed of the radial spoke (RSP) complex and central pair projections (CPP). In *Chlamydomonas* at least 23 RSP proteins build a T-shaped complex which is anchored at the A microtubule and connects the central pair through a thin stalk with the outer microtubule doublets (Goodenough and Heuser, 1985; Warner and Satir, 1974; Witman et al., 1978; Yang et al., 2006). The RSP complex interacts through its bulbous head domain with the central pair surrounding CPPs.

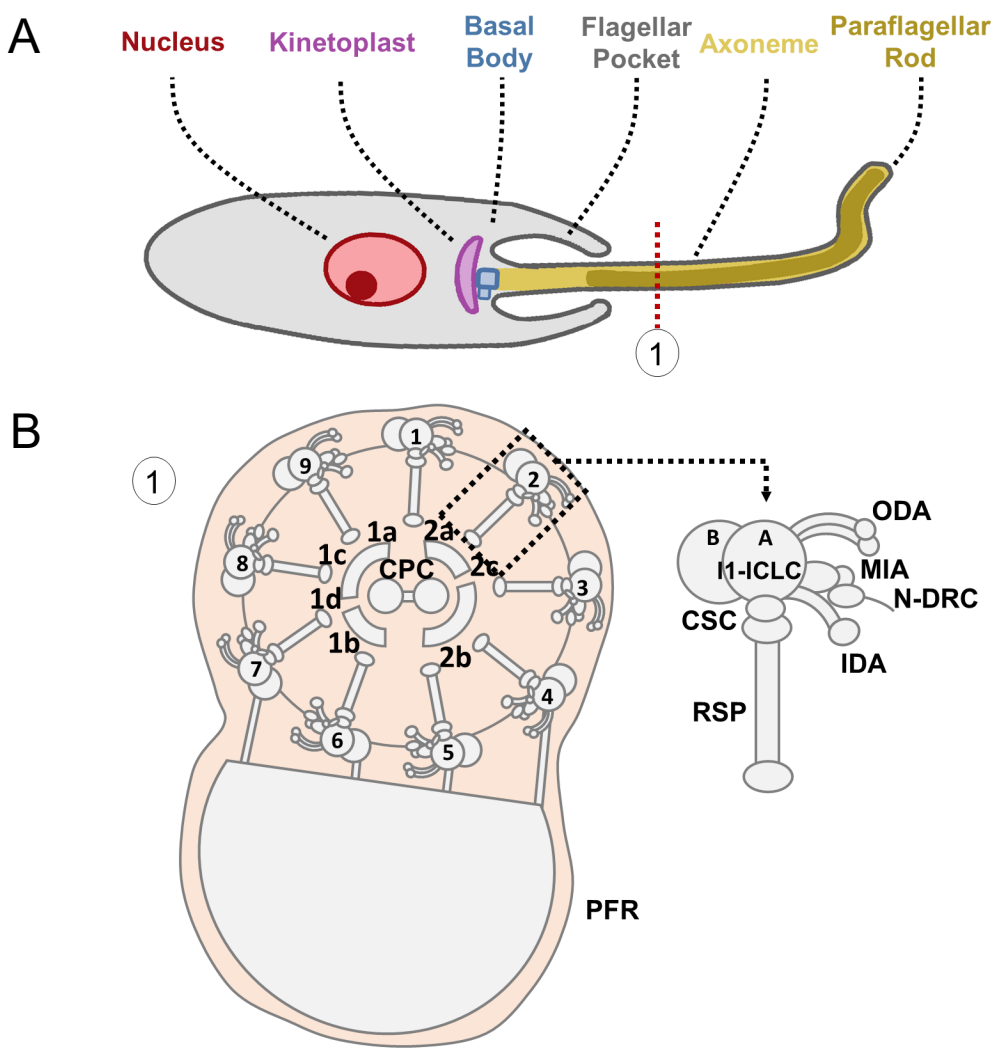


Figure 1. The *Leishmania* promastigote flagellum.

(A) Overview of selected promastigote (insect stage) *Leishmania* organelles and **(B)** schematic transverse axonemal section of the promastigote flagellum. Numbers in **(A)** correspond to numbers in **(B)**, highlighting the position of the transverse axonemal section. **(B)** Section shows 9 + 2 microtubule arrangement of the motile promastigote flagellum, with schematic drawing of CPPs numbered on the left according to Carbajal-Gonzalez et al. (2013). Highlighted on the right is a microtubule doublet with axonemal components. CPP: Central pair projections; CPC: Central pair complex; PFR: Paraflagellar rod; RSP: Radial Spoke; CSC: CaM and Spoke associated Complex; IDA: Inner dynein arm; ODA: Outer dynein arm; MIA: modifier of inner arms; N-DRC: nexin-dynein regulatory complex.

These projections allow to distinguish between C1 and C2 central pair microtubules. Previous work in *Chlamydomonas* and *Tetrahymena* has shown that the C1 microtubule has longer CPPs, namely C1a, C1e C1c, C1d, C1f and C1b, than the C2 microtubule, on-which CPP C2a, C2e, C2c, C2d and C2b attach (Figure 1 A and B) (Carbajal-Gonzalez et al., 2013). They repeat either every 16 or 32 nm (Mitchell, 2003; Mitchell and Sale, 1999) and together with the central pair microtubules and central pair bridge, which connects the C1 and C2 microtubule, they build the central pair complex (CPC) (Carbajal-Gonzalez et al., 2013; Loreng and Smith, 2017; Zhao et al., 2019).

Dyneins are believed to be regulated by various complexes within the axoneme. These are known to be either additional structures that connect the RSPs with IDAs, such as the CaM and Spoke associated Complex (CSC) (Dymek et al., 2011), or structures that are in close proximity to IDAs and are shown to modulate their activity, such as the modifier of inner arms (MIA) complex (Yamamoto et al., 2013) and the nexin-dynein regulatory complex (N-DRC) (Heuser et al., 2009; Huang et al., 1982; Oda et al., 2015; Ralston et al., 2006) (Figure 1 A and B). The regulation of dyneins is essential for the coordination of flagellar bending with the direction of movement. This direction is typically the result of either a ciliary or a flagellar beat, which are also referred to as asymmetric and symmetric waveforms (Lindemann and Lesich, 2010). Both beat types are usually found in motile cilia of *Chlamydomonas* and organisms of the *Trypanosomatidae* family, but also in vertebrates (Mitchell, 2017). It has been early shown that cilia and flagella in various organism, such as *Crithidia* (Sugrue et al., 1988) and *Drosophila* (Yang and Lu, 2011), can also switch their bending pattern, for example mediated through

calcium signaling (Bessen et al., 1980; Eckert and Lux, 1976; Fukuyama and Epstein, 1973).

Additional to axonemal structures, several organism in major eukaryotic groups, also assemble extra-axonemal structures. These are structures that can be arranged asymmetrically or symmetrically around the axoneme. For example, euglenozoan protists, such as *Leishmania*, have an asymmetrically arranged extra-axonemal paraflagellar rod (PFR), which is a lattice-like structure (Figure 1 A and B) (see section 3.1.4 for a detailed description of extra-axonemal structures). While the function of extra-axonemal structures remains largely unknown, motile cilia and flagella have been shown to not only have a mechanical function of moving cells or fluids over surfaces, they also can have a sensory function (Baldari and Rosenbaum, 2010). Sensory proteins and pathways in motile cilia have been identified in *Paramecium* (Preston and Saimi, 1990) and *Chlamydomonas* (Huang et al., 2004; Pan and Snell, 2000) *species* and also in human cells. For example motile cilia of the human airway epithelia contain bitter taste receptors that react to bitter compounds by increasing calcium influx and thus stimulate ciliary beat frequency (Kinnamon and Reynolds, 2009; Shah et al., 2009).

1.1.2 Non-motile cilia and flagella

Non-motile cilia, also often referred to as primary cilium in mammalian cells (Sorokin, 1968), are distinct from motile cilia. This type of cilium is typically described as having a 9 + 0 ultrastructure, which lacks the CPC, RSP, N-DRC and dynein arms (Sun et al., 2019). However, there are non-motile cilia that deviate from this definition as described above (section 1.1). Additionally, primary cilia in neuroepithelial bipolar cells that are part of the embryonic nervous system,

collapse gradually towards their distal end, going from a 9 + 0 to 7 + 0, then to 5 + 0 and finally reaching 2 + 0 at their tip (Cohen and Meininger, 1987). The amastigote flagellum of *Leishmania* is also a non-motile flagellum and has striking similarities to the primary cilium of vertebrates (Figure 2 A and B). For example, the nine-fold microtubule symmetry of the primary cilium in kidney cells collapses towards the distal end in a similar manner to the amastigote flagellum (Figure 2 A and B). Sometimes it even follows into a reduction of doublets or a reduction of doublets to singlets. This variation is referred to as 9 – variable (9v) flagellum (Gluenz et al., 2010b). However, the amastigote flagellum deviates from primary cilia as ODAs and RSPs remain visible at the proximal, but not distal end of the flagellum (Wheeler et al., 2015a) (Figure 2 A and B).

Primary cilia have been shown to be involved in regulating signaling pathways and in mediating mechano- and chemosensation in many cell types (Hoey et al., 2012). For example fluid flows within kidneys bend primary cilia, which in-turn stimulates the influx of intracellular calcium (Praetorius and Spring, 2001; Praetorius and Spring, 2003a; Praetorius and Spring, 2003b). This is believed to be controlled by a mechanosensory complex at the base of the cilium (Nauli et al., 2003). However, mechanosensation by primary cilia through fluid flow remains controversial as a complete lack of mechanically induced calcium increases in primary cilia has also been reported (Delling et al., 2016).

A similar mechanism has been observed in liver cholangiocytes (Masyuk et al., 2007) and in bones from humans and mice, although at least in bones this pathway differs to cilia mediated mechanosignaling in liver and kidney.

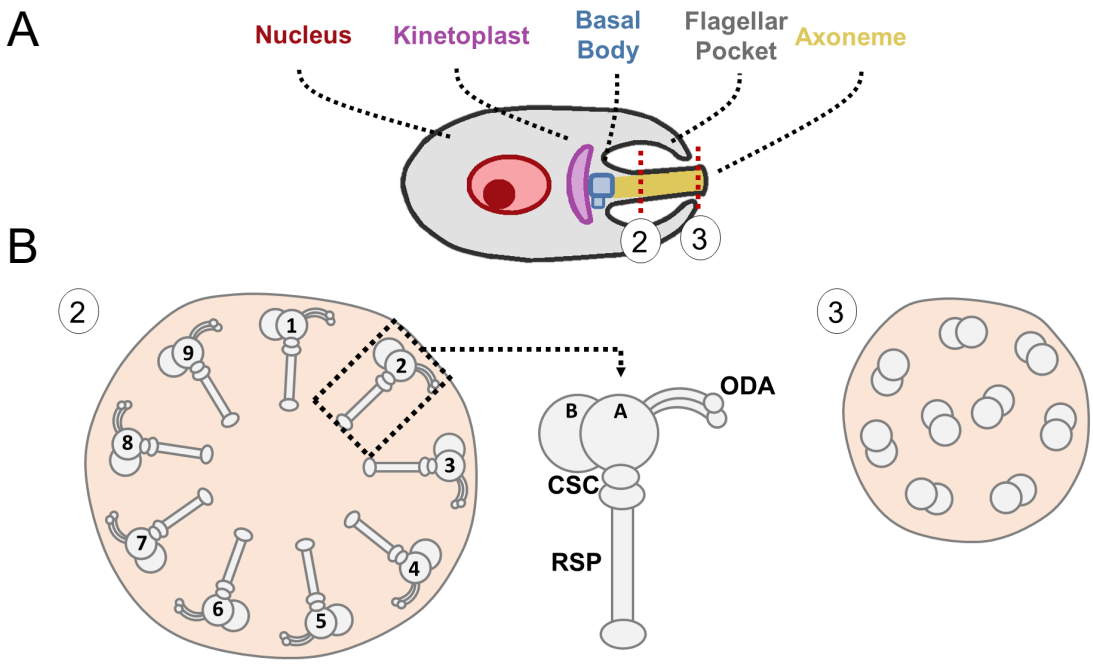


Figure 2. The *Leishmania* amastigote flagellum.

(A) Overview of selected amastigote (mammalian stage) *Leishmania* organelles and **(B)** schematic transverse axonemal section of the amastigote flagellum. Numbers in **(A)** correspond to numbers in **(B)**, highlighting the position of the transverse axonemal section.

(B) Sections on the left show the 9 + 0 microtubule arrangement of the immotile amastigote flagellum and on the right the collapsed 9v arrangement according to Gluenz et al. (2010b) and Wheeler et al. (2015a). RSP: Radial Spoke; CSC: CaM and Spoke associated Complex; ODA: Outer dynein arm.

Bone cells do not mediate flow-induced intracellular calcium, but signal a decrease in the second messenger cyclic adenosine monophosphate (cAMP) (Kwon et al., 2010; Malone et al., 2007a; Malone et al., 2007b; Xiao and Quarles, 2010).

However, mechanosensing has not only been shown in the context of bending, as non-motile cilia in *Drosophilia* (Ernstrom and Chalfie, 2002) or primary cilia of cholangiocytes and chondrocytes (Gradilone et al., 2007) can also react to pressure, touch and vibration (reviewed in Hoey et al. (2012)). Consequently, non-motile cilia are considered in many organisms to act as an antenna that reacts to environmental stimuli.

1.1.3 Flagellar and ciliary assembly

Ciliary and flagellar complexes are incorporated into the axoneme of motile and non-motile cilia as repeated units and are mostly targeted independently from each other when the axoneme is assembled. The intra-flagellar transport (IFT) machinery thereby moves non-membrane-bound and membrane-associated particles from the cell body out to the flagellar tip to incorporate components into the growing cilium or flagellum and then returns to cell body. This way IFT proteins maintain and assemble cilia or flagella (Rosenbaum and Witman, 2002). In *C. reinhardtii* IFT particles have been resolved into two sub-complexes: IFT-A and IFT-B (Cole et al., 1998). It was shown that the IFT-A complex consists of six components (IFT43, IFT122A, IFT122B, IFT139, IFT140 and IFT144), while the IFT-B complex consists of 16 components, separated into IFT-B1 (IFT22, IFT25, IFT27, IFT46, IFT52, IFT56, IFT70, IFT74, IFT81, IFT88) and IFT-B2 (IFT20, IFT38, IFT54, IFT57, IFT80, IFT172) (reviewed in Taschner and Lorentzen (2016)). IFT particles are moved either anterograde (from the base to the tip) or retrograde

(from the tip to the base) by associated IFT motor proteins (kinesin-2 or IFT-dynein) (reviewed in Scholey (2008)). IFT-B components are believed to be required for the anterograde transport, while the IFT-A complex seems to be involved in returning IFT-B particles. However, this remains controversial (Prevo et al., 2017). IFT transport of particles is also thought to require a third functional component, the BBSome complex (namely derived from the associated genetic disorder Bardet-Biedl Syndrome) (reviewed in Wingfield et al. (2018)). The complex consists of 7 components (BBS1, BBS2, BBS4, BBS5, BBS7, BBS8, BBS9, (Nachury et al., 2007)) and other associated proteins, such as ARL6/BBS3 (Liew et al., 2014). The BBSome complex is believed to be an IFT transport adapter for cilia and flagellar membrane proteins (Berbari et al., 2008; Liu and Lechtreck, 2018) and to be required for export of flagellar signaling proteins (Lechtreck et al., 2009). However, the BBSome seems to have more a role in assisting IFT protein transport, as it is expendable for flagella and cilia assembly in most systems (Liu and Lechtreck, 2018; Williams et al., 2014).

In *Leishmania* at least two other proteins are also involved in flagellar protein targeting. For example, KHARON1 has been shown in loss-of-function experiments to be required for targeting the flagellar localized glucose transporter 1 in *L. mexicana* promastigotes, which is not expressed in the amastigote form (Tran et al., 2013). Also, flagellar length is shortened when the small G protein ARL-3A is overexpressed in the promastigote stage of *L. donovani*, *L. major* or *L. amazonensis*, indicating its possible role in flagellar integrity (Cuvillier et al., 2000; Sahin et al., 2008; Sahin et al., 2004).

1.1.4 Ciliopathies

Dysfunctions in motile and/or non-motile cilia in humans are linked to dozens of diseases. A link between defects in cilia function was first drawn in 1976, when male infertility and chronic respiratory tract conditions were identified to result from immotile cilia (Afzelius, 1976a; Afzelius, 1976b), a disease termed Primary Ciliary Dyskinesia (PCD). Cornillie et al. (1984) later established the term ciliopathy to describe generally diseases associated with defects in motile and/or non-motile cilia. To-date at least 187 ciliopathy-associated genes have been identified and defects have been shown to be associated with retinal dystrophy, anosmia, hearing loss, organ laterality defects, central obesity, skeletal anomalies, infertility, hypogonadism, genital anomalies, ataxia, epilepsy, mental disability, brain malformations, hydrocephalus, facial anomalies, congenital heart defects, chronic respiratory problems, renal anomalies and hepatic disease. These defects are associated with various syndromes, such as Bardet-Biedl Syndrome, Joubert Syndrome and Meckel-Gruber Syndrome. While for some defects the genetic cause is known, for many it remains unknown (Mitchison and Valente, 2017; Reiter and Leroux, 2017). For example male infertility affects about 7% of men, but the genetic cause for this disease has been identified for only 25% of these 7%. While more than 3000 genes have been associated with spermatogenesis, less than 0.01% of these have been linked to infertility (reviewed in Neto et al. (2016)).

Compared to other organism, the *Leishmania* flagellum offers the unique opportunity to study both, motile and non-motile cilia. Much remains unknown about the function of the amastigote flagellum and about many features of the promastigote flagellum. This makes the *Leishmania* flagellum of great interest to study.

1.2 *Leishmania* parasites

Leishmania spp. belong to the major eukaryotic group of Excavata, specifically to the class of *Kinetoplastida*. They possess a large single mitochondrion with a network of circular DNA, which is called the kinetoplast (Shapiro and Englund, 1995). These parasites are closely related to *Trypanosoma* spp., of which the causative agent of sleeping sickness *T. brucei* is one of the most widely used organism to study fundamental biological mechanisms of protozoan parasites. *Leishmania* parasites have mostly a diploid genome organization (Figure 20), but some species such as *L. donovani*, are known to vary their chromosomal copy number significantly (Mannaert et al., 2012). Protein-coding genes are often arranged in tandem repeats and are constitutively transcribed into long polycistronic transcripts. *Leishmania* parasites and other kinetoplastids lack transcriptional control and rely mostly on post-transcriptional regulation (reviewed in Clayton (2019)).

1.2.1 Leishmaniasis

Leishmania parasites represent an infection risk for approximately 350 million people worldwide and cause a broad spectrum of diseases in 97 endemic countries (World Health Organization, 2019). The parasite, *L. donovani*, was first discovered in India by Leishman and Donovan in 1903 (Donovan, 1903; Leishman, 1903), although Cunningham was the first one to identify the parasite in ulcers of patient samples. He delivered the first drawings of a intracellular *Leishmania* amastigote cell when studying leishmaniasis in India, referred to as Dehli Boil at the time by locals (Cunningham, 1885).

There are various *Leishmania* species that cause, depending on the geographical area and strain, several diverse clinical manifestations of the disease leishmaniasis. Leishmaniasis may have been described as early as 2000 BCE as “Nile Pimple” and even the Old Testament may have referred to it by describing it as boils: “It will spread like fine dust overall the land of Egypt and cause boils to break out upon people and animals alike, throughout the land.” (Oumeish, 1999).

Visceral leishmaniasis (kala azar; e.g. *L. donovani*, *L. infantum*) is fatal in 95% of cases if left untreated. It is characterized by an aggressive visceral infection of the reticuloendothelial system. Cutaneous leishmaniasis (e.g. *L. mexicana*, *L. major*, *L. tropica*) is the most common form and causes cutaneous lesions or ulcers that heal spontaneously. There are between 0.6 – 1 million new cases annually mainly in Europe, Africa and Central Asia, but most commonly in the Middle East. Some *species* (notably in the *Viannia* subgenera, e.g. *L. braziliensis*) are also able to infect the nasopharyngeal tissues, which can lead to destruction of cutaneous and subcutaneous tissues. Therefore, this type of leishmaniasis is referred to as mucocutaneous leishmaniasis and is most common in Latin America. Additionally, there are about 26,000 to 65,000 deaths reported annually, mostly due to visceral leishmaniasis (World Health Organization, 2019).

Currently, there are a limited number of drugs available for treatment. Pentavalent antimonials, amphotericin B, liposomal amphotericin B and paromomycin are administered through injections, whereas miltefosine is the only orally-available leishmanicide. But these drugs are associated with problems of cost, toxic side effects, duration of treatment and require often hospitalization (Kumar and Engwerda, 2014). Furthermore, they all show different efficiency to different strains (el-On and Hamburger, 1987) and drug resistance is a major concern (Chakravarty

and Sundar, 2010). For pentavalent antimonials wide-spread resistances are known (Croft, 2001; Sundar et al., 2000). In regions of India resistance to amphotericin B was shown (Srivastava et al., 2011) and in Nepal failure of miltefosine in the treatment of visceral leishmaniasis was reported (Rijal et al., 2013). During the last decades, there have been efforts to develop effective vaccines against leishmaniasis. These include vaccines of the first (killed or live attenuated parasites), second (antigens or recombinant proteins) and third generation (DNA vaccines). While several vaccines have been shown to confer long-term protection against leishmaniasis in animals, several remain to be tested on humans in clinical trials (reviewed in Moafi et al. (2019) and Ghorbani and Farhoudi (2018)).

To facilitate further the development of drug therapies and vaccination approaches against leishmaniasis, it remains to be important to improve the understanding of pathogenesis. This will help to identify novel drug targets and/or vaccine candidates.

1.2.2 The *Leishmania* life cycle and interactions with host and vector

Leishmania parasites have a digenetic life cycle, proliferating in the digestive tract of its insect vector as promastigote form and primarily intracellularly in its vertebrate host as amastigote form (Figure 3).

The vector of *Leishmania* species are dipterans of the genus *Lutzomyia* in the Americas and *Phlebotomus* in Europe and Asia. Sand flies reside in tropical and sub-tropical areas and the duration of the life cycle within the fly varies from 4 to 18 days, but can be even longer or shorter depending on temperatures (Torres-Guerrero et al., 2017). Female sand flies take up parasites during a bloodmeal

from an infected host and *Leishmania* amastigotes quickly differentiate into procyclic promastigotes (Figure 3: 1 and 2). They further develop into nectomonad promastigotes in the abdominal midgut (Figure 3: 3) before migrating to the anterior parts of the fly where they proliferate as leptomonads in the alimentary tract (Figure 3: 4) (Killick-Kendrick et al., 1974). They finally reach the stomodeal valve where they mature to highly infective metacyclic parasites (Figure 3: 5) (Bates, 2007). During a second blood meal parasites get re-injected into a mammalian host (Figure 3: 6), prompting promastigotes to rapidly change their morphology and switch their gene expression patterns. This affects many flagellum-associated, but also other proteins (Fiebig et al., 2015a). *Leishmania* parasites are predominantly phagocytosed by macrophages, but also by neutrophils and dendritic cells. Once phagocytosed they complete their morphological differentiation into their amastigote form within 10 – 24 hours post injection and further maturate for another 96 hours within the phagolysome-derived parasitophorous vacuoles (Barak et al., 2005). Low pH and increased temperatures within the macrophage host stimulates differentiation (Zilberstein and Shapira, 1994), along with other stimuli, such as oxidative stress and iron uptake (Mittra et al., 2013; Mittra et al., 2017). Once fully matured, parasites proliferate inside a phagolysosome (Chang and Dwyer, 1976) (Figure 3: 7 - 9). However, not all *Leishmania* spp. seem to be obligate intracellularly parasites when in their host, as shown for the subgenera *Sauroleishmania* which are transmitted into reptiles (e.g. *L. tarentolae*) (Raymond et al., 2012).

During this complex life cycle *Leishmania* parasites are exposed to rapidly changing harsh environments, forcing them to develop clever strategies to complete their life cycle anyway.

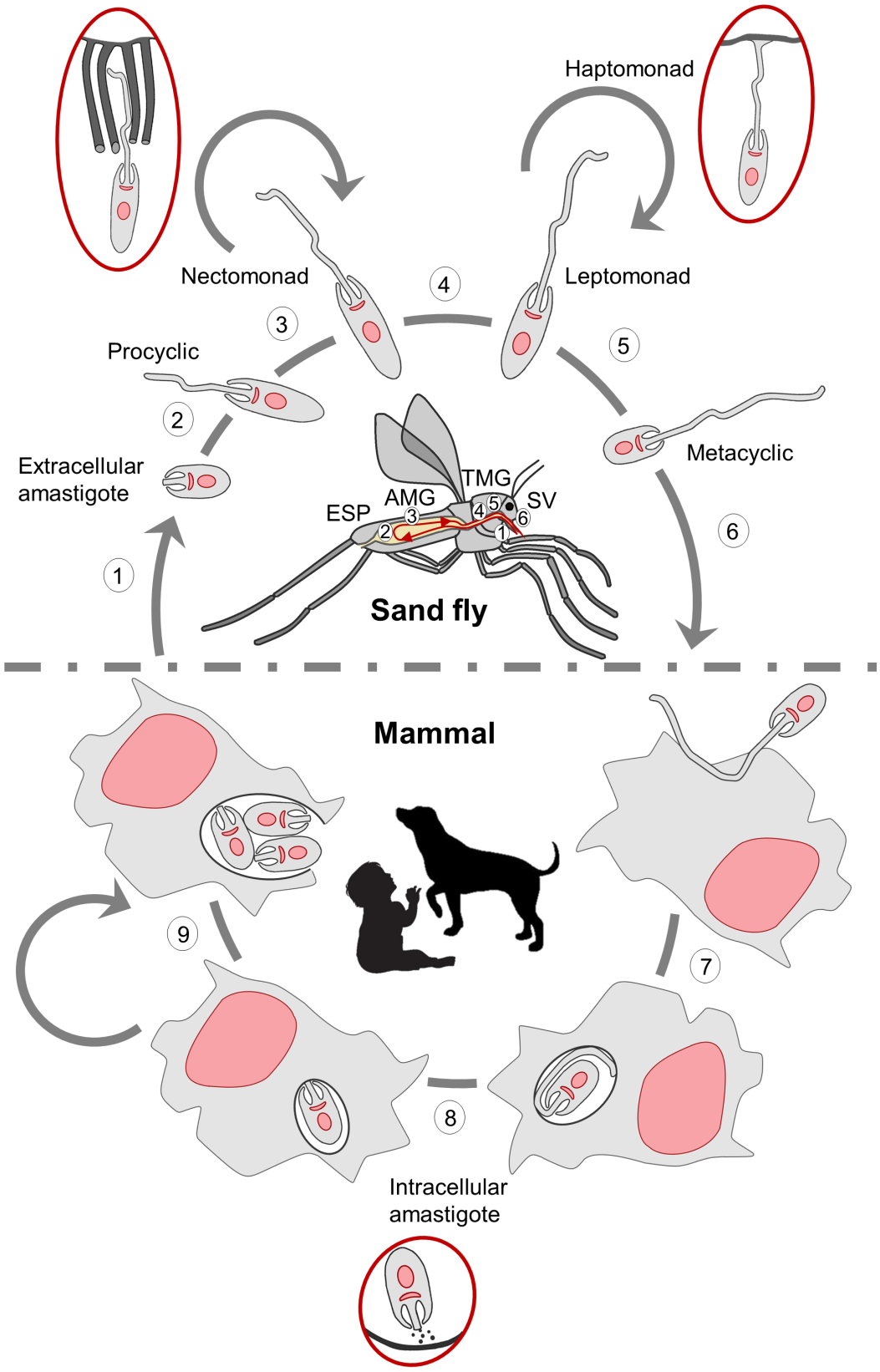


Figure legend on following page.

Figure 3. The *Leishmania* life cycle.

Schematic overview of the digenetic life cycle of *Leishmania* parasites. Numbers shown on sand fly life cycle stage correspond to sand fly organelles. ESP: endoperitrophic space; AMG: abdominal midgut; TMG: thoracic midgut; SV: stomodeal valve. **(1)** Intracellular amastigotes are taken up during a blood meal of sand flies. **(2)** In the ESP parasites differentiate into procyclics and further in the **(3)** AMG to nectomonads. Nectomonads attach via their flagellum to the microvilli and proliferate in the sand fly gut. **(4)** Leptomonads proliferate in the TMG and attach as haptomonads with their flagellum to the chitin lining of the stomodeal valve. **(5, 6)** Metacyclics are then injected into the mammalian host upon a second blood meal, **(7-9)** where they then differentiate into amastigotes and proliferate in the phagolysosome.

For example, *Leishmania* survival within the sand fly is associated with the interaction of parasites with the vector's peritrophic matrix, an extracellular envelope, which separates the gut lumen from the midgut epithelium, consisting of chitin, proteins, and glycoproteins (Lehane, 1997). The peritrophic matrix protects parasites during differentiation from amastigotes to procyclics from proteolytic stress (Pimenta et al., 1997), but then once parasites are fully differentiated, *Leishmania* migrate through posterior openings that exist in the peritrophic matrix and migrate through the fly (Sadlova and Volf, 2009). Other mechanisms to progress through the sand fly vector include strategies to escape proteolytic stresses secreted by flies to digest the blood meal. Serine proteases, such as trypsin- and chymotrypsin-like enzymes are detected in significant levels at about 6 hours post blood meal. However, *L. major*, *L mexicana* and *L. infantum* fly infections have shown to suppress or delay trypsin activity in the midgut, allowing more time for parasites to differentiate into procyclics (reviewed in Dostalova and Volf (2012)). Procyclics cells are thought to be fully resistant to proteolytic stress, as their coat of glycoconjugates, most prominently the lipophosphoglycan (LPG), has been shown to be essential for parasite attachment and resistance against proteolytic enzymes in the sand fly midgut (Jecna et al., 2013; Sacks et al., 2000). *Leishmania* LPG consists of a GPI anchor (glycosylphosphatidylinositol), a phosphosaccharide core, a linear phosphoglycan (PG) chain and a oligosaccharide cap (Turco and Descoteaux, 1992).

Interestingly, glycoconjugate surface molecules are not only believed to be essential in the fly. PG-containing surface glycoconjugates, but not LPG, are expressed at high levels in amastigotes and play also an essential role in parasite-host interactions. They are thought to shield *Leishmania* parasites against

complement-mediated lysis and from reactive oxygen and nitrogen species produced by the host to kill the invading parasite (reviewed in Franco et al. (2012)). However, the role of some surface molecules, such as LPG, remains controversial in the field, as discussed further in chapter 5.

1.2.3 The *Leishmania* flagellum – a multifunctional organelle

While these examples above show the diversity of parasite–host and parasite–vector interactions, there are also several lines of evidence that in particular the flagellum is an essential organelle for both main life cycle stages. For instance, the flagellum is involved in attachment to the sand fly gut, as nectomonad promastigotes attach via their flagellum to the microvilli in the abdominal midgut and as haptomonads attach to the alimentary tract with their flagellar tip (Killick-Kendrick et al., 1974) (Figure 3). But it remains so far unknown how the precise molecular interactions between flagellum and the microvilli of the gut lining functions. While the LPG coat is believed to play a major role in parasite attachment to the sand fly midgut (Jecna et al., 2013; Sacks et al., 2000), a purely geometric nature of flagellar attachment, meaning the flagellum just perfectly fits into the elaborations of the lining of the gut, remains also possible.

The significance of the *Leishmania* flagellum has also been proposed for host cell invasion, as the flagellum is believed to injure host cells to promote phagocytosis (Forestier et al., 2011). Furthermore, parasites with defects in flagellar membrane protein targeting complexes, such as BBS1 (Price et al., 2013) and KHARON1 mutants (Tran et al., 2013; Tran et al., 2015) have been shown to be avirulent. Thus, structural similarities between primary cilia and the *Leishmania* amastigote flagellum have led to the hypothesis that the amastigote flagellum could function

as an antenna in parasite–host interactions, when parasites reside intracellularly (Gluenz et al., 2010b).

Overall, much remains unclear about the role of the amastigote and promastigote flagellum in parasite–host and parasite–vector interactions. The ultrastructural differences between long motile flagella in promastigotes (Figure 1 A and B) and short non-motile flagella in amastigotes (Figure 2 A and B) may suggest that the function and role of both flagella is different in their life cycle stages.

1.2.4 The *Leishmania* flagellar proteome

The lack of knowledge about the role of the flagellum *in vivo* is despite the wealth of data on ultra-structural features and flagellar proteomic analyses from *Leishmania* (Dupe et al., 2015), other kinetoplastids (Broadhead et al., 2006; Oberholzer et al., 2009; Subota et al., 2014), human sperm cells (Wang et al., 2013) and many other organisms. These data sets deliver a comprehensive list of conserved flagellar proteins that when subjected to null mutant analysis allows dissection of the mechanism and function of the flagellum.

However, the majority of these existing proteomes have used salt extraction methods to depolymerize subpellicular microtubules and isolate flagellar axonemes. Consequently, they lack flagellar membrane and matrix proteins in their analysis (Broadhead et al., 2006; Dupe et al., 2015; Oberholzer et al., 2009; Robinson et al., 1991; Schneider et al., 1987). Salt extraction of flagellar axonemes has been a reliable way for isolating flagella since the early 60s (Child, 1959) and has been adapted for a variety of organisms, including *T. brucei* and *C. reinhardtii*. Isolated flagella from *T. brucei* can be expected to contain more than 500 proteins, including especially axonemal, PFR and basal body proteins (Broadhead et al.,

2006). As salt extraction is known to destabilize ODAs and the CPC (Dutcher et al., 1984), other approaches have been developed to isolate intact flagella. For example by knockdown of Fla1 in *T. brucei*, flagella can be detached by mechanical shearing. These isolated flagella appear to have a normal ultrastructure by electron microscopy. However, they are still derived from Fla1 RNA interference (RNAi) mutants and thus isolated flagella may lack other Fla1 associated proteins (Subota et al., 2014).

In contrast, the recent proteomic study of the promastigote flagellum of *L. mexicana* in the Gluenz lab used a “deflagellation” method in which the flagellum is sheared off by a combination of a calcium shock and mechanical shearing, allowing the subsequent analysis of the proteomic composition of isolated intact flagella. This flagellar proteome was generated in my undergraduate research project (Beneke, 2014) and recently published in Beneke et al. (2019b) and a method chapter describing the flagellar isolation protocol is currently in press (Beneke et al., 2019c). In this proteome purified flagella and deflagellated cell bodies were also additionally separated into detergent soluble and insoluble fractions for analysis of all four fractions by mass spectrometry. Figure 4 and Figure 5 summarize the deflagellation method and show in particular how the flagella purification process is being tracked by an *in situ* tagged flagellar membrane localizing protein SMP-1 (small myristoylated protein) (Tull et al., 2010; Tull et al., 2004). The fractions obtained contained the external part of the flagellum (average break point 2.7 µm distal to the basal body) comprising the axoneme, flagellar membrane and matrix. Using a label-free quantification method SINQ (Trudgian et al., 2011) this proteomic analysis yielded a total of 2711 proteins across two biological replicates, with 701 unique proteins enriched in flagellar vs cell body

fractions (Figure 5 B). A key advantage of this proteome is also that it allows a multiway comparison of identified proteins in the cell body soluble and insoluble fraction against proteins in the flagellar fractions, thus filtering out cell body associated contaminants and yielding many possible yet undiscovered new flagellar proteins. These unknown proteins were identified among already known flagellar components, such as flagellar tip, flagellar membrane, IFT, CPC, ODA, IDA, PFR, N-DRC and RSP proteins (Figure 6). Additionally, Figure 7 summarizes how this proteomic analysis was validated by *in situ* tagging 85 flagellar enriched proteins at their N- and C-terminus, showing that ~80% of these proteins actually localized to the flagellum.

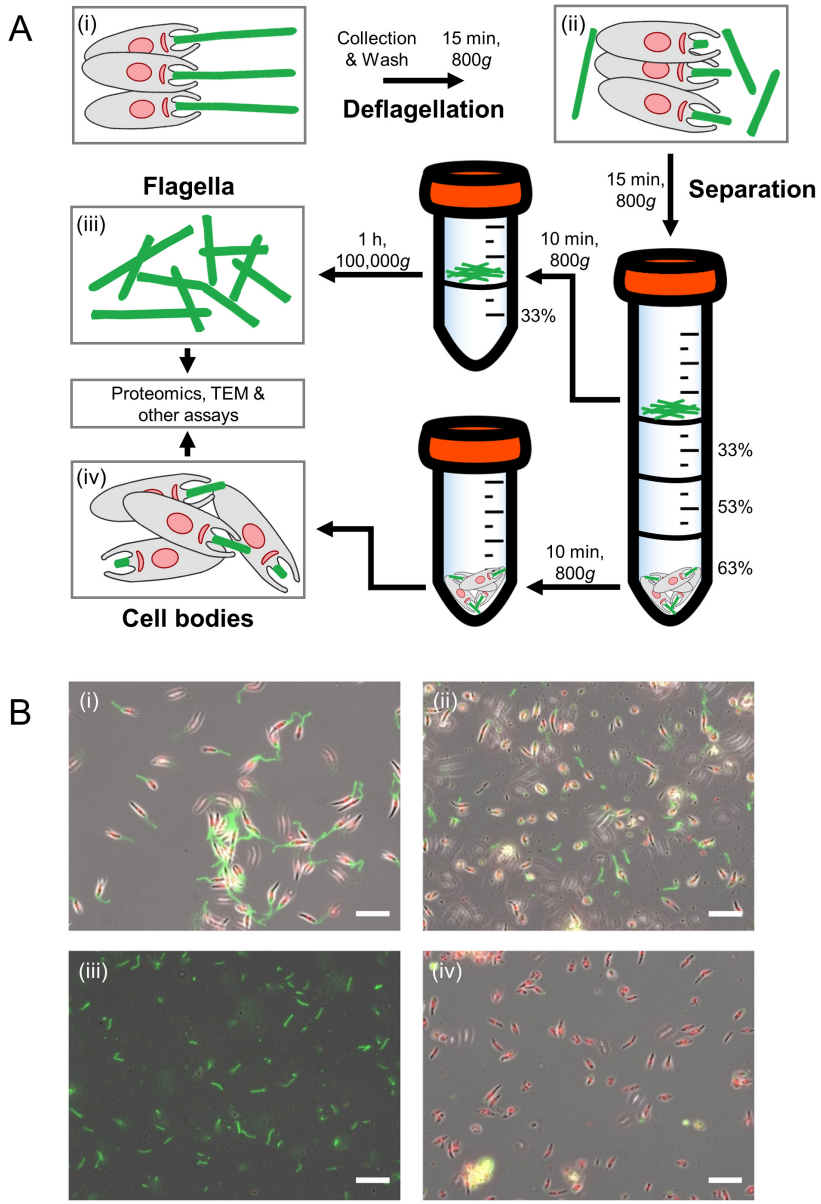


Figure 4. Isolation of promastigote flagella and deflagellated cell bodies.

Figure made by Tom Beneke, taken from Beneke et al. (2019c). **(A)** Overview of the deflagellation and flagella isolation protocol. Percentage sucrose concentration (w/v) and parameters for centrifugation steps are indicated. Used cell line expresses fusion protein SMP-1::GFP. **(B)** Micrographs show merged fluorescent and phase channel (SMP-1::GFP in green, Hoechst DNA stain in red, phase in grey) for each isolation stage (i-iv) depicted in **(A)**. **(i)** *L. mexicana* cells before deflagellation, **(ii)** cells after deflagellation, **(iii)** isolated flagella and **(iv)** isolated deflagellated cell bodies. Scale bar is 10 µm.

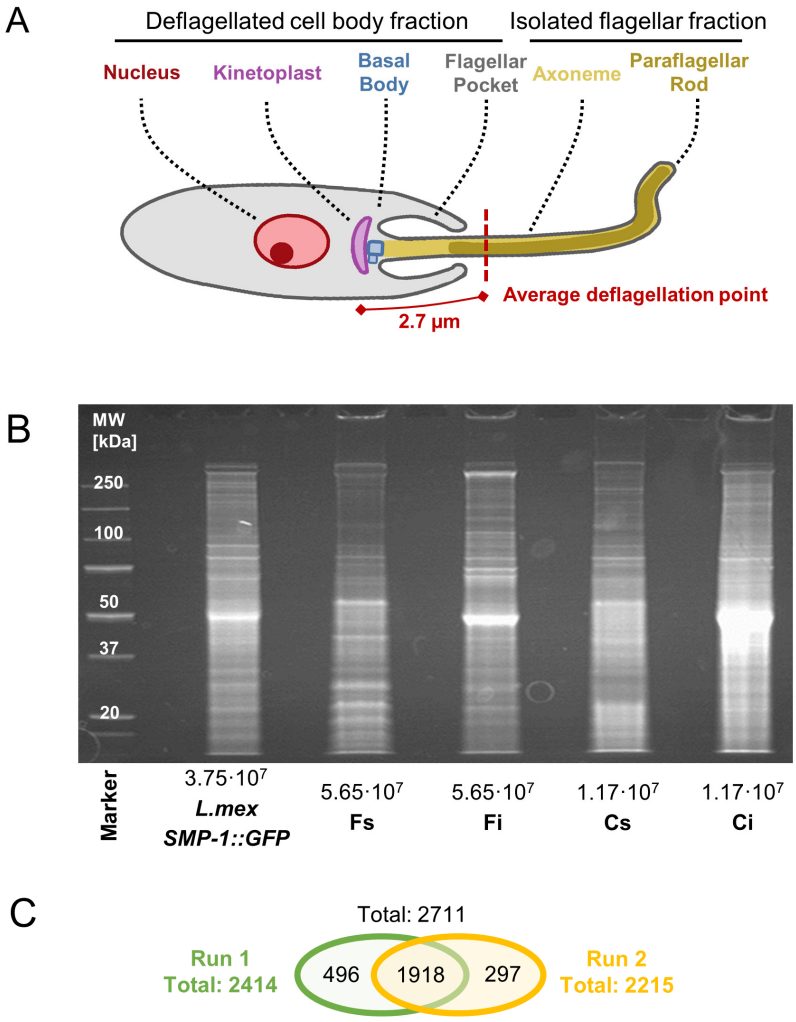


Figure 5. Proteomic analysis of flagella and deflagellated cell bodies.

Figure made by Tom Beneke, adapted from Beneke et al. (2019c). Outcome of deflagellation and flagella isolation protocol. **(A)** Cartoon of promastigote cell showing selected organelles associated with deflagellated cell body fraction and isolated flagellar fraction. The measured average deflagellation point in Beneke et al. (2019b) is indicated. **(B)** Protein samples of flagellar soluble (Fs), flagellar insoluble (Fi), cell body soluble (Cs) and cell body insoluble (Ci) fractions, as well as *L. mexicana SMP-1::GFP* cells are run on a polyacrylamide gel and stained with SYPRO Ruby. Molecular weights of marker lane are indicated. Cell equivalents used for loading of each lane are indicated. **(C)** Two biological replicates of all four protein fractions have been analysed by mass spectrometry. Van diagram showing overlap between both replicates and proteins uniquely identified in each replicate.

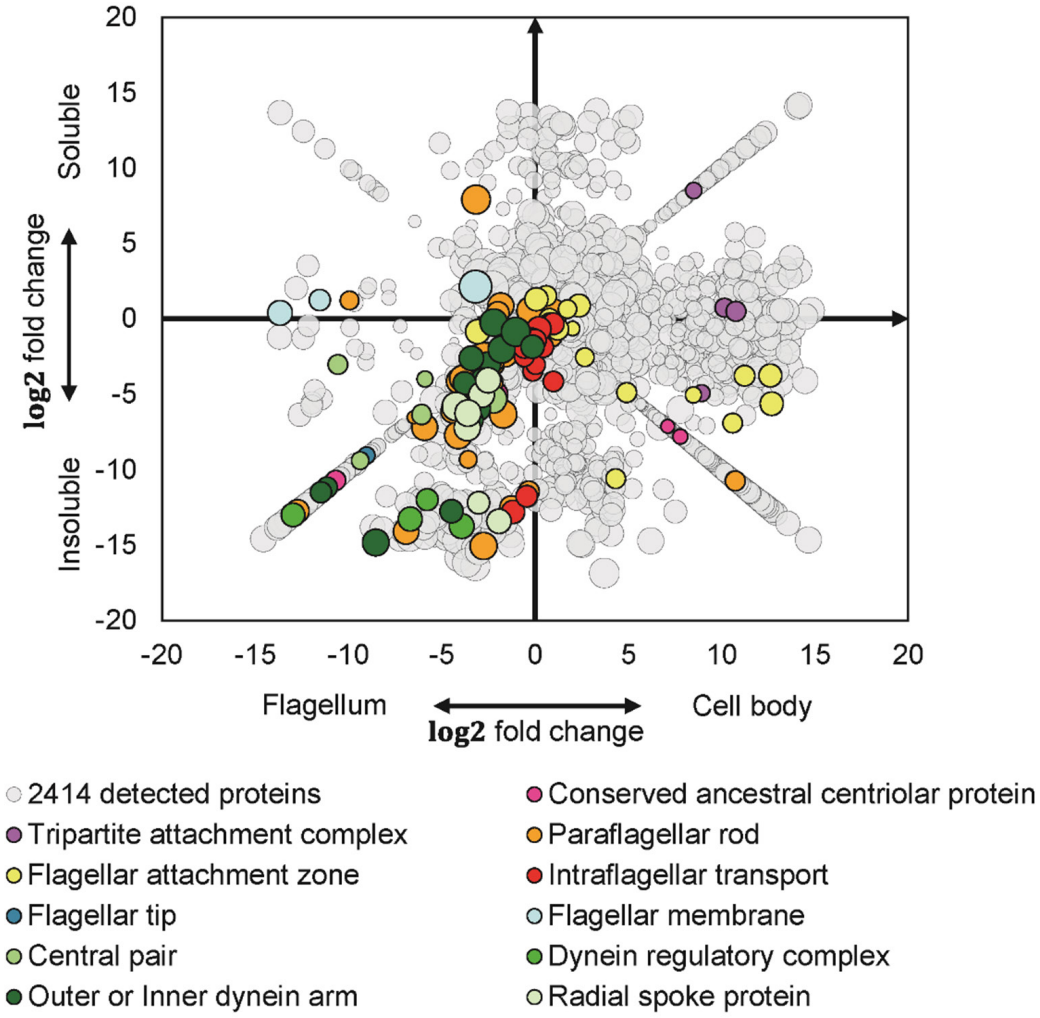


Figure 6. Enrichment plot of flagella and cell body identified proteins.

Figure made by Tom Beneke, taken from Beneke et al. (2019b). Detected proteins in all four fractions of replicate 1 were quantified with SINQ. Plotted on the X-axis is the log2 fold-change of the spectral index (C_S+C_I / F_S+F_I) and on the Y-axis the log2 fold change of the spectral index (F_S+C_S / F_I+C_I). C: cell body; F: flagella; I: insoluble and S: soluble. A pseudo spectral index of 10^{-10} was inserted for proteins not detected in all four fractions. Bubble size reflects protein abundance. Shown are all detected proteins in run 1 (grey) and known protein annotations.

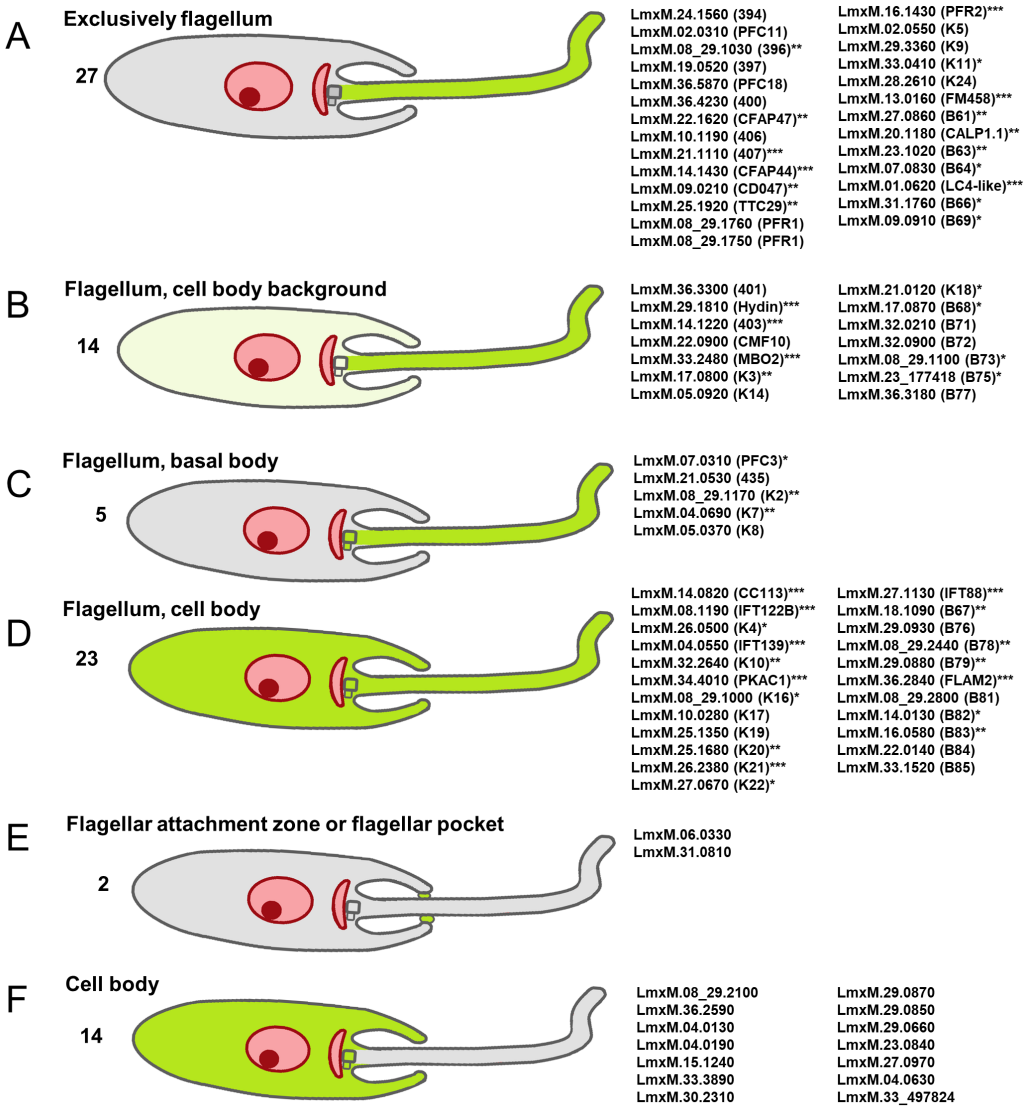


Figure 7. Summary of tagged proteins in Beneke et al. (2019b).

This figure was produced for this thesis and summarizes N- and C-terminal protein tagging localizations from Beneke et al. (2019b). Genes have been assigned to a category according to their protein localization. Localization categories include: **(A)** exclusively flagellum, **(B)** flagellum with little cell body background, **(C)** flagellum with strong basal body signal, **(D)** flagellum and cell body, **(E)** flagellar attachment zone or flagellar pocket and **(F)** cell body only. Asterisk on gene ID indicates whether the null mutant of the respective gene was significantly reduced in the motility analysis performed in Beneke et al. (2019b) (student's t-test: * $p \leq 0.05$, ** $p \leq 0.01$; *** $p \leq 0.001$).

1.3 The challenges and key aims

1.3.1 Project questions

The *Leishmania* flagellar proteome delivers a comprehensive list of more than 500 flagellar proteins. This data set offers the opportunity to dissect mechanism and function of the *Leishmania* flagellum because it not only consist of components of the motile axoneme and PFR, but also of proteins associated with the flagellar membrane and matrix. Studying the function of these proteins may answer questions on the *in vivo* role of the *Leishmania* flagellum. For example: (1) Is there a signalling cascade between flagellar-associated surface molecules and the macrophage host? (2) Is the interaction between flagellum and host purely mechanistically in which e.g. phagocytosis is promoted by flagellar mediated host cell injury? (3) Is flagellar motility required for invading the mammalian host and (4) is motility generally required for survival *in vivo*, during the insect and mammal stage?

But it also remains unknown how motility in itself is regulated, especially how dyneins are controlled to produce flagellar beating. Although several competing models for flagellar bend and dynein regulation have been proposed, it is not completely understood how the flagellar beat is coordinated. Also, it remains unclear how unknown flagellar proteins contribute to the regulation of flagellar bending. Unknown flagellar proteins have been identified by proteomic analyses in *C. reinhardtii*, *Drosophila*, *T. brucei* and other model organisms, but also in our flagellar proteome.

1.3.2 Project aims

The two overarching aims were to dissect flagellar function in *Leishmania* (1) *in vitro* and (2) *in vivo* to gain insight into general and *Leishmania*-specific flagellar biology, specifically:

1. Dissecting mechanism and function of conserved proteins to understand how motile flagella beat and how their beat is regulated. This includes many highly conserved proteins with orthologues in *C. reinhardtii*, as it is the major flagellar model organism, and humans, as of obvious medical interest.
2. Understanding the lineage-specific adaptations of the *Leishmania* flagellum and dissecting how the flagellum contributes to survival and life cycle progression. In particular, in how far motility contributes to the completion of the life cycle.

1.3.3 Project strategy

To dissect flagellar function (1) *in vitro* and (2) *in vivo*, genetic mutants were produced to study loss-of-function phenotypes. Gene candidates for knockout experiments were chosen to address project questions and aims above. The data source for these knockout candidates was our flagellar proteome and many flagellar proteins previously characterized in *C. reinhardtii*, *Drosophila*, *T. brucei* and other model organisms. Knockout candidates amounted to several hundreds of proteins and therefore demanded the development of a high-throughput gene editing toolkit, capable of producing large numbers of mutants and phenotyping them in a scalable format.

Recent advances in gene editing by using clustered regularly interspaced short palindromic repeats (CRISPR) and especially the CRISPR-associated gene 9 (Cas9) have enabled precise introduction of double strand breaks (DSB). Nucleases, such as Cas9, are guided by sequence-specific single-guide RNAs (sgRNA), a RNA chimera molecule consisting of originally the Cas9 binding backbone trans-activating crRNA (tracrRNA) and the sequence recognizing CRISPR RNA (crRNA). The use of CRISPR-Cas9 systems allows us to dramatically increase the editing efficiency in many different organisms, including *Leishmania* and *Trypanosoma* (Fernandez-Prada et al., 2018; Lander et al., 2016; Lander et al., 2015; Peng et al., 2014; Rico et al., 2018; Soares Medeiros et al., 2017; Sollelis et al., 2015; Vasquez et al., 2018; Zhang et al., 2017; Zhang and Matlashewski, 2015). However, previously developed CRISPR gene editing systems available for kinetoplastids (see section 2.1.4) either rely on labour intensive cloning procedures for sgRNA production or require *in vitro* sgRNA transcription.

Therefore, it was first of interest to design a CRISPR-Cas9 gene editing system that allows a scalable production of null mutants (chapter 2). This toolkit was then used for a systematic dissection of a curling phenotype (chapter 3) and detailed characterization of a novel protein that interacts with IDA f/I1 (chapter 4). Finally, the Cas9 toolbox and a bar-seq strategy was used to dissect the requirement of *L. mexicana* flagellar motility for survival *in vivo* (chaper 5).

CHAPTER 2

LEISHGEDIT - A CRISPR-CAS9 TOOLBOX FOR KINETOPLASTIDS

2 LeishGEdit - a CRISPR-Cas9 toolbox for kinetoplastids

Data presented in this chapter has been published in Beneke et al. (2017) and Beneke and Gluenz (2019). The *L. mexicana Cas9* (pRM006) cell line was generated by Ross Madden. All data and figures presented in this chapter are wholly my own work unless stated otherwise.

2.1 Introduction

2.1.1 Chapter summary and aim

The aim of the work shown in this chapter, was to develop a CRISPR-Cas9 toolbox that allows high-throughput generation and phenotyping of *Leishmania* mutants in order to dissect flagellar function. The toolbox developed here was termed LeishGEdit (Beneke and Gluenz, 2019; Beneke et al., 2017).

A *Leishmania mexicana* cell line was engineered to express constitutively Cas9 nuclease and T7 RNA polymerase. Using this cell line, a strategy was designed, in which linear sgRNA and donor DNA constructs were transfected in parallel, allowing for sgRNA *in vivo* transcription and integration of donor DNA constructs with 30 nt homology flanks (HF) identical to the target locus. The method does not involve any cloning procedures, PCR purifications or *in vitro* transcription prior to transfections. To streamline donor DNA production a modular plasmid system (pPLOT plasmids for tagging, pT plasmids for knockouts) was designed, allowing to re-use the same primer pairs for insertion of a variety of tags or knockout constructs. Additionally, the online platform www.leishgedit.net was developed for automated primer design for a variety of kinetoplastids.

Overall, data presented here shows that LeishGEdit is a high-throughput gene editing system that can be easily scaled up to mutant production on a 96 well plate format and used for screening hundreds of gene candidates from the *L. mexicana* flagellar proteome presented in section 1.2.4.

2.1.2 Loss-of-function strategies in *Leishmania* before CRISPR

Next-generation gene editing approaches, including CRISPR-Cas9 but also other programmable editing systems, such as transcription activator-like effector nucleases (TALEN) and zinc finger nucleases (ZNF), have become powerful tools for genome engineering (reviewed in Jaganathan et al. (2018)). CRISPR has become the genome editing technology of choice in animals, plants and many other organisms. CRISPR is much easier to design than TALENs and ZNFs editing approaches, more cost effective (Jaganathan et al., 2018) and recent improvements, such as the development of high-fidelity CRISPR-Cas9, have reduced significantly off-target cleavage (Kleinstiver et al., 2016). The introduction of CRISPR-Cas9 gene editing in *Leishmania* and *Trypanosoma* parasites has involved various approaches, rapidly cutting the time to obtain transgenic parasites and enabling loss-of-function experiments for a larger number of gene candidates than previously. However, before the golden era of next-generation gene editing started, genetic manipulations were much more labor intensive in many organisms, including *Leishmania* and in some organisms even impossible (reviewed e.g. in Housden et al. (2017)).

Transcript knockdown by RNAi is probably the most widely used approach for loss-of-function analysis. However, most *Leishmania* parasites, apart from some *Leishmania* spp. of the *Viannia* subgenus (Lye et al., 2010), have evolutionary lost

their RNAi machinery (Ullu et al., 2004). This is a major disadvantage when working with *Leishmania* compared to other kinetoplastids, such as *T. brucei* in which RNAi systems are widely utilized (Alsford et al., 2011; Broadhead et al., 2006; Ngo et al., 1998; Ralston et al., 2006). Hence, loss-of-function and also gain-of-function studies in *Leishmania* (tools for *Leishmania* reviewed in Jones et al. (2018) and Duncan et al. (2017)) relied in the past on mostly gene replacement approaches by using homologous recombination, a genetic recombination pathway in which a DNA sequence is exchanged or replaced with an identical or similar (homologous) DNA segment (Ceccaldi et al., 2016). Since the first reports of genetic manipulations using gene replacement strategies in *Leishmania* parasites nearly three decades ago (Cruz and Beverley, 1990) about 200 genes have been subjected to gene deletion in numerous labs (Jones et al., 2018). However, constructs for gene deletion by gene replacement required either a two-step cloning procedure to generate cassettes with HF and drug resistance markers (Roberts, 2011) or more recently fusion PCR approaches, in which linear PCR products that contained HFs were fused by PCR to drug resistance marker cassettes (Dean et al., 2015; Merritt and Stuart, 2013). HFs usually contain sequences that are identical to the target locus and thus mediated precise integration of donor DNA. Cassettes were then transfected and mutants isolated in the presence of selective drugs. Depending on the chromosome ploidy several rounds of transfection (usually two rounds for diploid *Leishmania* spp.) were required to acquire null mutants, making not only the generation of constructs labor intensive, but also the process of obtaining mutants. Even more time-consuming is to address essential genes with this strategy, meaning genes that are required for survival under the coditions used for selection. It required one additional

transfection, in which the target gene is complemented with an episomal addback construct prior to the actual gene deletion. Additionally, inducible gene deletion strategies can be also used to show that a gene is essential. For example in the “Forced Plasmid Shuffle” system, null mutants can be forced to lose their episomal addback copy after the locus deletion is complete. Cells are negatively selected by using ganciclovir, which enacts a fitness cost on parasites that maintain the addback plasmid (Murta et al., 2009). Alternatively, inducible gene deletion can be achieved by integrating loxP sites upstream and downstream of the target gene. By inducing dimerizable Cre recombinase (DiCre) the loxP flanked gene will be deleted (Duncan et al., 2016).

While manipulation of parasites via the homoglous recombination pathway is an effective approach for obtaining transgenic parasites, their use remains labor intensive. New approaches, such as CRISPR-Cas9 gene editing, are required to streamline loss-of-function and gain-of-function approaches in *Leishmania*. Most importantly, new genetic tools have to (1) reduce production time of tagging or knockout constructs and (2) increase the transfection efficiency so that multiple alleles can be edited at once.

2.1.3 CRISPR-Cas9 gene editing

CIRSPR repeats were first described in archaea and bacteria in the 1980s and early 1990s (Ishino et al., 1987; Mojica et al., 1993) and were later found to serve as an adaptive immunity-like memory in prokaryotes, cleaving nucleic acids, transmitted by bacteriophages and other viruses (Barrangou et al., 2007; Mojica et al., 2005; Pourcel et al., 2005). CRISPR-Cas9 was soon applied as a universal gene editing method by first discovering the requirement of a PAM (protospacer

adjacent motif) sequence, which has to be immediately adjacent to the crRNA target sequence, and then later starting to program the Cas9-crRNA complex *in vitro* and *in vivo* (Gasiunas et al., 2012; Jinek et al., 2012). With the introduction of CRISPR-Cas9 gene editing in mammalian cells, this technology finally introduced a new era of gene editing for nearly all organisms (Cong et al., 2013).

Cas9 is probably the most prominent used CRISPR system to-date. The reason for its popularity is due to its simplicity for its DNA cleavage activity. Cas9 is a class 2 type II CRISPR associated protein, meaning it does not need any other effector protein for binding the crRNA and tracrRNA. Binding the target sequence and cleaving DNA are both mediated by Cas9 itself. Cas9 can be modified for a large variety of applications, such as transcription blocking, pull-downs and DNA localization (reviewed in Zhang (2019)). Additionally, Cas9 creates blunt ends after introduction of a DSB, which is different to sticky end creating enzymes, such as Cpf1 (also known as Cas12a) (Makarova et al., 2015). Depending on the organism and substrate, blunt end DSBs enable many different DNA repair pathways, most prominently homologous recombination, non-homologous end joining (NHEJ) and microhomology-mediated end joining (MMEJ) (further described below in section 2.1.4).

While Cas9 and Cpf1 both target DNA, it has been shown that Cas9 can be also re-programmed to target RNA when using a complementary antisense oligo *in vitro* (O'Connell et al., 2014). Meanwhile, other proteins, such as Cas13a (formerly also known as C2C2), target and cleave as primary substrate RNA. However, once Cas13a binds to its target, the RNA cleavage domain develops a “collateral damage” activity, cleaving non-specifically any RNA molecule (Abudayyeh et al., 2017; Cox et al., 2017). Further protein engineering of several Cas13 orthologous

has led to the development of CasRx, which seems much more promising for programmable RNAi–like knockdown approaches, as it has been used successfully to knockdown many endogenous transcripts in human cells without detectable “collateral damage” activity (Konermann et al., 2018).

2.1.4 CRISPR-Cas9 gene editing in *Leishmania* and other kinetoplastids

In the past few years, introduction of CRISPR in the field of kinetoplastids has helped to eliminate time-consuming procedures, such as long cloning protocols, for generation of mutants. As highlighted in several reviews (Bryant et al., 2019; Duncan et al., 2017; Lander et al., 2016) CRISPR-Cas9 gene editing in protozoan parasites requires at least two, but mostly three components to function and effectively mutate a genomic locus of interest. These include (1) Cas9, (2) a sgRNA construct (either *in vivo* or *in vitro* transcribed) and (3) an optional corresponding donor DNA construct. In the following, this section will give an overview of how these three components are used in *Leishmania* and *Trypanosoma* parasites for CRISPR-Cas9 gene editing.

(1) There are three main strategies for delivery of Cas9 into the cell, *in vivo* expression of Cas9 from an episome, from an integrated cassette or use of recombinant Cas9. In *T. cruzi* (Lander et al., 2016; Lander et al., 2015; Peng et al., 2014), *L. major* (Sollelis et al., 2015) and *L. donovani* (Zhang and Matlashewski, 2015) CRISPR-Cas9 was first used by *in vivo* expressing SpCas9 (Cas9 derived from *Streptococcus pyogenes*) from episomal plasmids, which were then complemented with their target-specific sgRNA. However, Cas9 has been also successfully used and *in vivo* expressed when integrated into the SSU locus (Ishemgulova et al., 2018). An alternative approach to *in vivo* expression also

exists. For example, it was shown that transfection of a sgRNA-SaCas9 ribonucleotide complex (Cas9 derived from *Staphylococcus aureus*) enables marker free manipulation of wild type *Trypanosoma* strains (Soares Medeiros et al., 2017).

(2) Meanwhile, sgRNA molecules can be synthesized (Fernandez-Prada et al., 2018) or transcribed from template DNA (Beneke et al., 2017; Peng et al., 2014; Zhang et al., 2017), using a variety of promoters, including ribosomal RNA promoters (Lander et al., 2016; Zhang et al., 2017; Zhang et al., 2013), U6 promotors (Ishemgulova et al., 2018; Sollelis et al., 2015; Zhang et al., 2017) or T7 promotors (Beneke et al., 2017; Peng et al., 2014; Soares Medeiros et al., 2017). Furthermore, for T7 promotors sgRNAs can be *in vitro* (Peng et al., 2014; Soares Medeiros et al., 2017) or *in vivo* transcribed from linear (Beneke et al., 2017) or also circular constructs, as for example shown in *Plasmodium* (Wagner et al., 2014).

(3) Most systems co-transfect sgRNAs also along with donor DNA, as transfections without donor DNA result in low mutation frequency. This is typically less than 1% within 14 days after a CRISPR-Cas9 specific transfection (Zhang et al., 2017; Zhang and Matlashewski, 2015). Thus, additional transfection of donor DNA was shown to significantly increase the transfection efficiency and produce precise modifications of the target locus (Zhang et al., 2017; Zhang and Matlashewski, 2015; Zhang and Matlashewski, 2019). In *Leishmania*, integration of donor DNA used to rely on homologous recombination, as outlined above (section 2.1.2). A previous study (Dean et al., 2015) showed that HFs need to be at least 100 nt long for integration of donor DNA in *Leishmania*. However, it has been shown that there is a correlation between the number of correct integration events per transfected

cell and HF length, meaning HFs need to be longer than 300 nt for efficient integration of at least one correct transfectant per million transfected cells in *L. mexicana* (Dean et al., 2015) (Figure 8 A). Therefore, producing tagging or knockout constructs originally required multiple steps of intensive cloning procedures or fusion PCR approaches. All of this is in contrast to *T. brucei*, in which DNA constructs can be integrated with much shorter HFs (Barnes and McCulloch, 2007; Dean et al., 2015). Using Cas9 mediated DSBs allows significant reduction of HFs. Integration of donor DNA in *Leishmania spp.* with as short as 25 nt HF has been reported (Zhang et al., 2017; Zhang and Matlashewski, 2015). Furthermore, research of a previous lab member, Ross Madden, showed that the frequency of correct integration is independent of HF length (Beneke et al., 2017). He tagged the central pair protein PF16 in *L. mexicana* Cas9 (pRM006) cells with *in vitro* transcribed sgRNAs and a series of donor DNA cassettes with HFs ranging from 24 to 360 nt (Figure 8 B). This showed that HFs as short as 24 nt result in correct and efficient tagging of PF16 with the same transfection efficiency as 360 nt HFs, while wild type cells transfected with identical constructs did not yield transfectants when using less than 360 nt HFs. Additionally, tagging with this Cas9 approach was found to be 10 times more efficient than the traditional homologous recombination approach in wild type cells (Figure 8 B).

While it is known that the NHEJ pathway has been evolutionary lost in *Leishmania* (El-Sayed et al., 2005; Ivens et al., 2005; Nenarokova et al., 2019; Passos-Silva et al., 2010), it remained until very recent unclear whether integration of donor DNA with HFs as short as 24 nt is still mediated via the homologous recombination pathway, MMEJ or other DNA repair pathways. Previously it was suggested that MMEJ plays a major role in integration of donor constructs when using short HFs

in *Leishmania* (Zhang et al., 2017). However, new findings show that more than 90% of DSBs in *Leishmania* are predominantly repaired by single-strand annealing (SSA), another pathway that uses typically substrates of 200 to 2,500 bp repeat sequences for DNA repair (Ubeda et al., 2014). Since *Leishmania* has a highly repetitive genome (El-Sayed et al., 2005; Ivens et al., 2005), substrates for SSA are widely available for DNA repair. Meanwhile, MMEJ was found to account for less than 10% of DNA repairs. In *Leishmania* repeat sequences as small as 29 bp were shown to be used by SSA (Zhang and Matlashewski, 2019). Additionally, the DSB repair machinery in *Leishmania* was also found to have an overall low efficiency, resulting in cell death for more than half of Cas9 created DSBs (Zhang and Matlashewski, 2019). This explains why, compared to for example mammalian cells, the *Leishmania* field has not yet seen genome-wide CRISPR-Cas9 libraries that rely only on sgRNAs and do not require additional donor DNA with a drug selectable maker for gene editing (see discussion in 6.2).

While use of additional donor DNA is not strictly required to introduce mutations with CRISPR-Cas9 in *Leishmania* and *Trypanosoma*, these recent findings (Zhang and Matlashewski, 2019) explain why a combination of Cas9, sgRNA and donor DNA has been the chosen strategy of choice for deleting, tagging or gene editing in other terms (e.g. introduction of single nucleotide variances (SNV)) a gene of interest (Fernandez-Prada et al., 2018; Lander et al., 2016; Lander et al., 2015; Peng et al., 2014; Zhang et al., 2017; Zhang and Matlashewski, 2015).

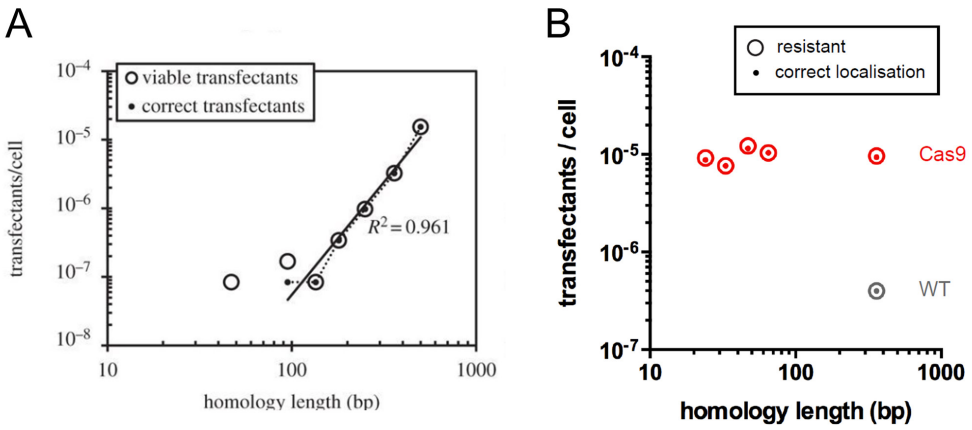


Figure 8. Gene replacement efficiency in *Leishmania mexicana*.

Figures taken from **(A)** Dean et al. (2015) and **(B)** Beneke et al. (2017). Counts of correct and viable transfectants per transfected cells using various lengths of HFs for C-terminal tagging of PF16 in *L. mexicana*. Y: axis transfectants/cell; X axis: used HF length for each transfection. The mean of biological triplicates is shown for: **(A)** Counts made in Dean et al. (2015) using homologous recombination for tagging of PF16. **(B)** Counts made in Beneke et al. (2017) by Ross Madden for transfecting PF16 donor tagging construct and *in vitro* transcribed sgRNA in Cas9 expressing cells (in red) and wild type cells (grey).

2.2 Results

2.2.1 Constitutive expression of Cas9 and T7 RNAP in *Leishmania*

To design a scalable, economic and efficient CRISPR-Cas9 gene editing toolbox that could be used for generation of a large number of mutants, a previously generated *L. mexicana* Cas9 cell line (pRM006, generated by Ross Madden), expressing hSpCas9 (humanized *Streptococcus pyogenes* Cas9; Cong et al. (2013)) from a transfected episome conferring resistance to hygromycin, was transfected with a T7 RNA polymerase (T7 RNAP) expression cassette (pVY087, Kraeva et al. (2014)). This cassette was inserted into the small subunit ribosomal locus (SSU) allowing for *in vivo* transcription of T7 promotor driven sgRNA templates.

Previously studies showed a minor growth defect for expression of T7 RNAP in *T. brucei* (Poon et al., 2012) and a major growth defect for constitutively expression of Cas9 in *T. cruzi* (Peng et al., 2014) was suggested. Therefore, the growth rate of *L. mexicana* Cas9 T7 cells was determined first. Each transfection of an additional expression cassette, first Cas9 (pRM006) and then T7 RNAP (pVY087), increased the doubling time of *L. mexicana* parasites in log phase culture (Figure 9 A). However, the doubling time increase was low and found to be of a similar size as shown for expression of T7 RNAP in *T. brucei* (Poon et al., 2012). Cells still reached stationary phase and the same density as the wild type cells after 96 hours in culture (Figure 9 B). Cell morphology of *L. mexicana* Cas9 T7 cells appeared normal, as assessed by microscopy. Consequently, constitutive expression of both enzymes appeared to be well tolerated by *L. mexicana* promastigotes.

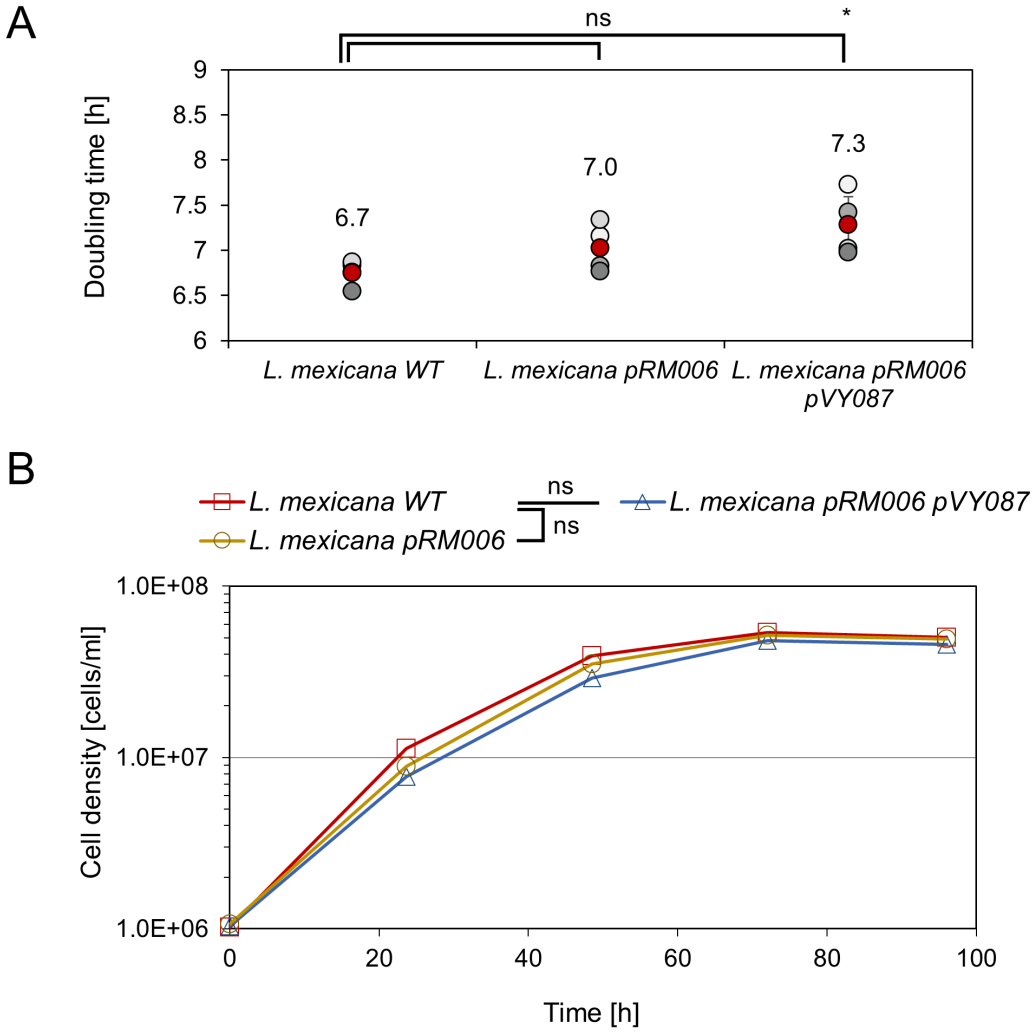


Figure 9. Effect on growth by Cas9 and T7 RNAP expression in *L. mexicana*. Measuring promastigote growth of three different cell lines: *L. mexicana* wildtype cells, Cas9 expressing cells (*pRM006*) and Cas9 T7 RNAP expressing cells (*pRM006* and *pVY087*). Student's t-test results are indicated. **(A)** Doubling times were measured four times (grey) by repeatedly diluting them to 10^6 cells/ml every 24 hours over four days and their average has been determined (red, value indicated above). Error bars show standard deviation. **(B)** Stationary growth curves were performed without replicates. Cells were cultured over 96 hours and measured every 24 hours.

2.2.2 sgRNA templates with a T7 RNAP promotor and a modular plasmid system allow for rapid PCR generation of transfection constructs

Constitutive expression of Cas9 and T7 RNAP in *L. mexicana* allows *in vivo* transcription of sgRNAs from transfected linear 124 nt PCR fragments made by using two overlapping single oligonucleotides (Bassett and Liu, 2014b). While one oligo contains a T7 RNAP promotor sequence and the 20 nt target sequence adjacent to the PAM site, the other oligo contains the entire sgRNA Cas9 backbone sequence (Figure 10 A). Both oligos overlap to enable an end-filling PCR reaction. Endogenous promotors, such as U6 or rRNA promotors, are much longer and do not fit onto a single oligonucleotide. Using short linear double-stranded DNA fragments for sgRNA production allows streamlining the gene editing process into a robust pipeline and makes cloning and *in vitro* transcription procedures redundant.

Additionally, linear sgRNA templates are rapidly degraded, which reduces chances for non-specifc CRISPR-Cas9 targeting. To test how quickly the sgRNA templates were lost from a transfected population, DNA isolated at different time points post transfection was analysed by PCR. This showed that sgRNA templates remained present up to 8 h post transfection but were undetectable after 48 h (Figure 10 B and C), indicating loss of the sgRNA template DNA.

To streamline donor DNA production, a modular plasmids system (Figure 11 A and B), which serves as template for amplification of donor DNA, was designed next. The development of this system was based on pPOT tagging plasmids for *T. brucei* (Dean et al., 2015). pPLOT and pT plasmids allow for loss- or gain-of-function experiments and can be integrated at the target locus after a DSB has been introduced by using a compatible sgRNA.

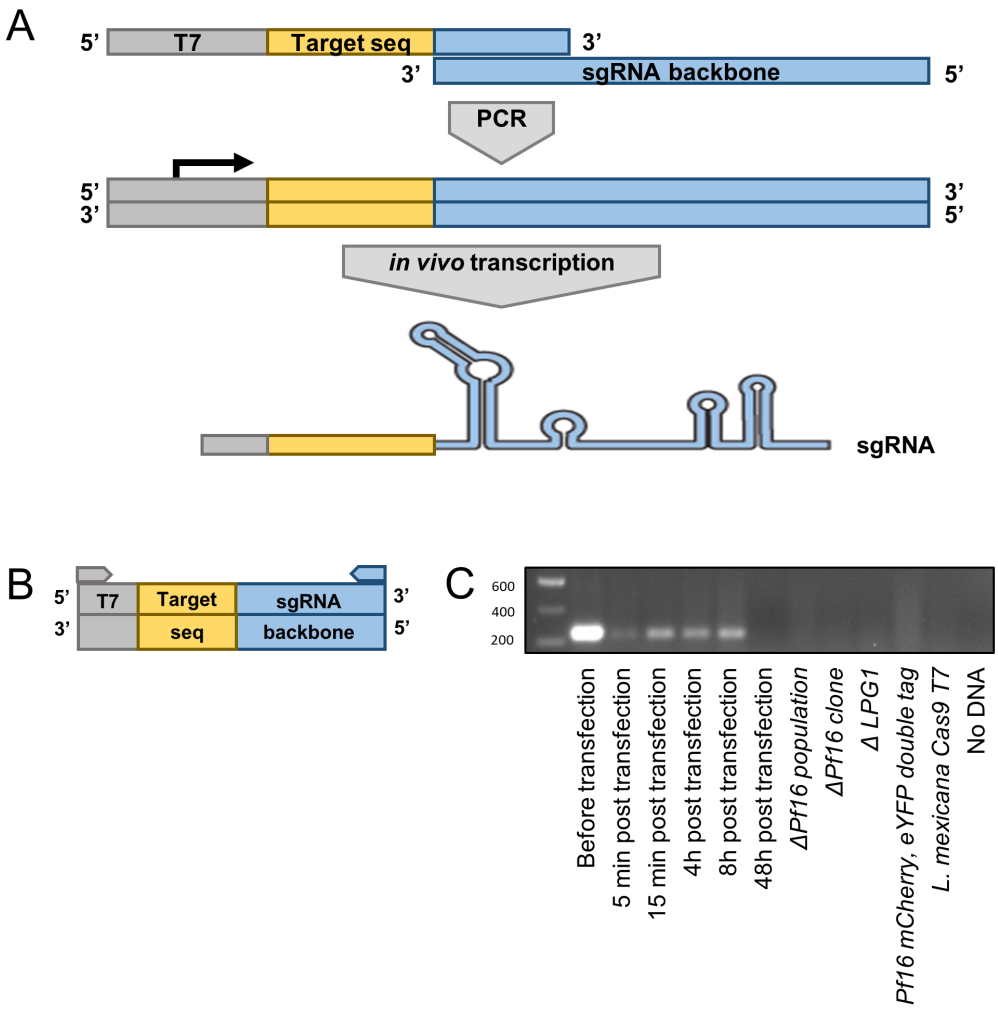


Figure 10. Efficient sgRNA production for CRISPR gene editing.

Figure made by Tom Beneke, adapted from Beneke et al. (2017). **(A)** Strategy for sgRNA production. Top: Oligo encoding the T7 promoter, 20nt target sequence and complementary sequence to 3' end of the sgRNA backbone sequence. Oligo overlaps to second oligo, which contains the sgRNA backbone sequence (Bassett and Liu, 2014a). Middle: Both oligos have been used for an end-filling PCR reaction, which is the template for *in-vivo* transcription shown below. **(B)** sgRNA template DNA is amplified using two primers (arrows) aligning to start and end of construct. **(C)** PCR reaction in **(B)** is performed to test presence of sgRNA template DNA. Various time points and knockout cell lines are tested as indicated. Marker lanes indicates amplicon length.

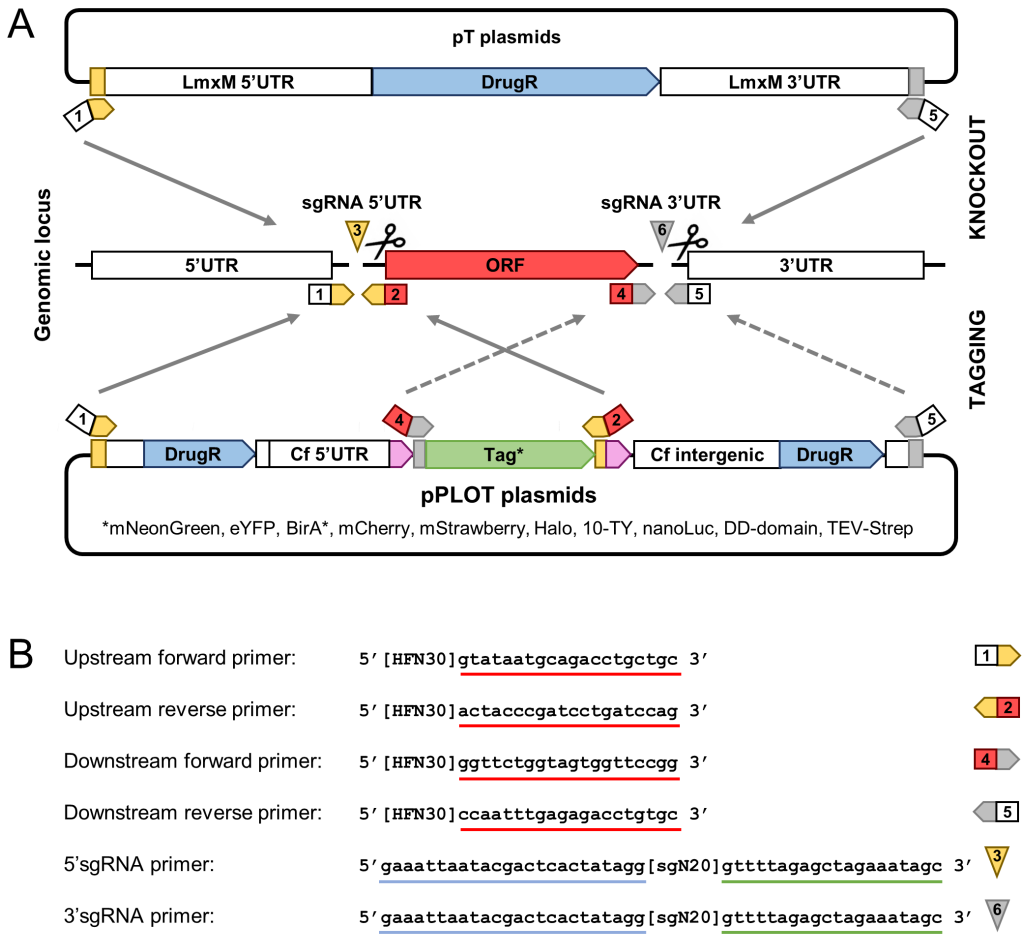


Figure 11. A modular plasmid system for gene editing in *Leishmania*.

Figure made by Tom Beneke, taken from Beneke and Gluenz (2019). Shown is the PCR strategy for donor DNA amplification from pT and pPLOT plasmids. **(A)** pT plasmids consist of a *L. mexicana* derived 5' and 3'UTR and a drug resistance marker gene to allow gene replacement by drug selection. pPLOT plasmids contain drug resistance markers and *Crithidia* and *T. brucei* UTRs, as well as myc epitope tags in-frame with various protein tags. pPLOTs can be used for amplification of tagging cassettes, allowing generation of fusion proteins with epitope tags fused at the N- or C-terminus. Various tags are available and are indicated. pT and pPLOT plasmids are amplified using four different primers, which contain **(B)** 30nt HFs immediately adjacent to sgRNA target sites (underlined in red). sgRNAs guide Cas9 to introduce a DSB at the genomic locus of interest and contain the T7 promotor sequence (underlined in blue) and an overhang sequence to the sgRNA backbone sequence (Bassett and Liu, 2014a) (underlined in green).

As previously shown, using CRISPR-Cas9 gene editing allows integration of donor DNA by using only 30nt HF (section 2.1.4), which compared to previous gene editing strategies in *Leishmania*, represents a decrease of more than 90% in HF length for efficient gene editing (Dean et al., 2015) and thus eliminates the need for cloning procedures.

pPLOT plasmids serve as templates for amplification of various N- or C-terminal tagging cassettes (available tags include BirA*, Halo tag, mCherry, mNeonGreen, eYFP, 10TY, nanoLuc, destabilization domain, TEV-Strep) (Figure 11 A and B). Additonally, all pPLOT tags are fused to three myc epitope tags to simplify detection of fusion proteins on Western blots, as shown for PF16 tagging with different pPLOT vectors (Figure 12). pT plasmids serve as template for PCR-amplification of knockout cassettes with different drug-selectable markers (Figure 11 A and B). UTRs used for expression of drug-resistance genes in pT plasmids are derived from *L. mexicana* and have been selected based on length (maximal 850 nt) and transcript abundance (100–300 fragments per kilobase of transcript per million mapped reads, with similar expression levels in promastigotes and amastigotes (Fiebig et al., 2015a)) to ensure sufficient expression of drug selectable markers independent of target locus and life cycle stage. The UTRs in each pT plasmid are different to minimize the risk of recombination between KO constructs.

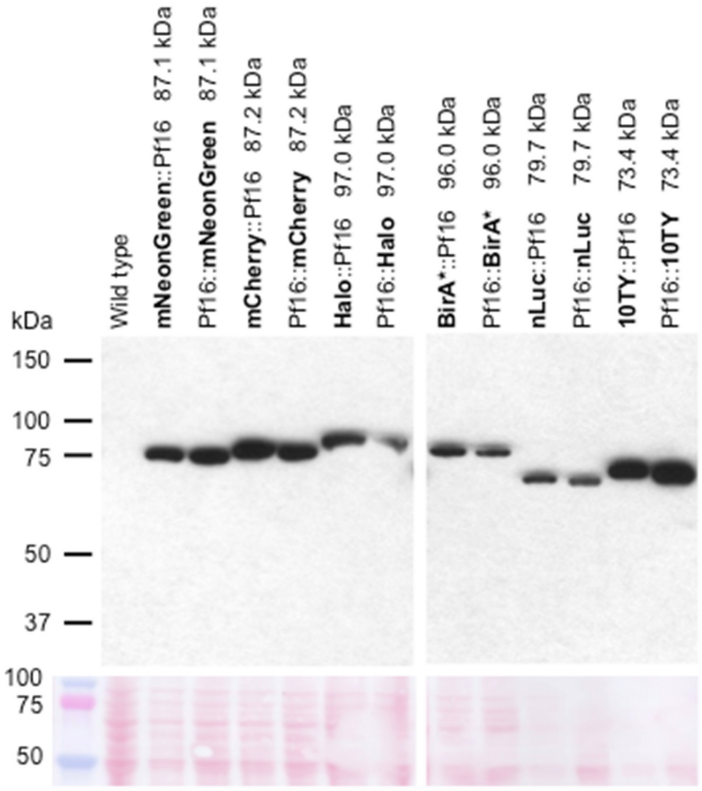


Figure 12. PF16 pPLOT tagging.

Figure made by Tom Beneke, taken from Beneke et al. (2017).PF16 has been tagged with different epitope tags at its N- and C-terminus using a pPLOT plasmids with puromycin N-acetyltransferase resistance gene. Molecular weights are indicated on the left. Top: Western blot using antibody against myc epitope tag. Expected molecular weight of fusion proteins are indicated above. Bottom: Ponceau stains of membranes before blocking membrane.

2.2.3 LeishGEdit: An online resource for gene editing in kinetoplastids and a streamlined high-throughput gene editing pipeline

To streamline this gene editing toolbox further into a high-throughput pipeline, a primer design tool, plasmid resource and mutant database was designed next and termed LeishGEdit. For primer design a single and batch mode bash script (current version LeishGEdit v4.2) was written that designs sgRNA primers by using a CCTop (Stemmer et al., 2015) and CRISPRater (Labuhn et al., 2018) standalone pipeline and their adjacent donor DNA primers (Supplementary: Script 1, Supplementary: Script 2). Previous script versions 1, 2 and 3 used a EuPaGDT pipeline for sgRNA design (Peng and Tarleton, 2015). Six primers are designed for each given ORF in the genome to enable CRISPR-Cas9 gene editing. Four of these primers are designed for pPLOT and pT plasmid amplification and can be used in different combinations to produce donor DNA. These primers include the: upstream forward primer (#1), upstream reverse primer (#2), downstream forward primer (#4) and downstream reverse primer (#5). Donor DNA primers contain the 30 nt HF sequence immediately adjacent to the sgRNA target sequence and its PAM site, as well as primer binding sites compatible with pT, pPLOT and pPOT plasmids (pPOT plasmids can be used for CRISPR-Cas9 mediated gene editing in *T. brucei*) (Figure 11 A and B). Two sgRNA primers are designed, one targeting the 5'UTR of the target gene (#3) and one the 3'UTR (#6). sgRNA primers consist of a T7 RNAP promotor sequence, a 20 nt sgRNA target sequence to introduce the double strand break at a locus of interest and a 20 nt overlap to the CRISPR-Cas9 backbone sequence allowing generation of sgRNA templates by PCR. An additional universal primer, containing the entire sgRNA backbone sequence (Bassett and Liu, 2014b) is needed to amplify both sgRNAs (Figure 11 A and B).

Consequently, only six primers are needed to tag a gene of interest at its N- or C-terminus with numerous available tags and to delete both alleles of the ORF.

Additionally, the upstream forward primer (#1) can be barcoded to allow for barcode analysis by sequencing (bar-seq) of mutants in pools, thus streamlining the phenotyping process. Bar-seq mutant analysis is a technique originally developed to analyze libraries of thousands of *Saccharomyces cerevisiae* (Smith et al., 2009) or *Schizzosaccharomyces pombe* (Han et al., 2010) gene deletion mutants, but has now been applied for other organisms, including the malaria parasite *Plasmodium berghei* (Bushell et al., 2017; Gomes et al., 2015) (further outlined with principles of bar-seq technique in section 5.1.2). The upstream forward primer is modified by adding a 17 nt barcode and a 20 nt constant region to the existing 30 nt HF and 20 nt pT-pPLOT-pPOT primer binding site in primers. Using the additional 20 nt constant region and 20 nt pT-pPLOT-pPOT primer binding site allows reading out barcode abundance by using Illumina amplicon sequencing strategies (further discussed and illustrated in chapter 5, Figure 52). Barcodes were initially generated by using barcode generator 2.8 (Comai and Howell, 2009) and then customized to select barcodes with 40-60% GC content (for unbiased Illumina sequencing), at least 3 mismatches (MM) between barcodes and allowing no blast hit in *Lutzomyia longipalpis*, in *Mus musculus* and in *L. mexicana*, yielding a total of 15,489 barcodes (Supplementary: File 1). This allows fitness screens in commonly used laboratory models of *in vivo* infections of sand flies and mammals. Thus, under these parameters a minimum barcode length of 17 nt is essential for generating enough barcodes for the number of ORFs in most *Leishmania* spp. (e.g. *L. mexicana* 9168 annotated ORFs (Fiebig et al., 2015a)). Each gene in the *L. mexicana* genome has been assigned a barcode, but

generated barcodes can be also used for any bar-seq experiments in other organisms.

The sgRNA target sequence required for integration of barcodes and other donor DNA constructs is selected from a 130 nt search window upstream of the start codon or downstream of the stop codon. The highest scoring sgRNA within this window is chosen based on the CCTop scoring pipeline. CCTop takes thereby the amount of sgRNA alignments to the genome with MM as main scoring criteria. Meaning, CCTop finds potential sgRNA target sites that have up to 2 MM in the first 12 nt upstream of the PAM site or up to 4 MM in the entire sgRNA target sequence and sorts these by least MM for the highest scoring sgRNA (Stemmer et al., 2015). There is experimental evidence that Cas9 sgRNA complexes are functional to introduce DSBs when their target site has up to 4 MM (Cong et al., 2013; Hsu et al., 2013; Sternberg et al., 2014; Tsai et al., 2015).

Additionally, selected sgRNAs are verified by counting the number of perfect matches within the genome. Since this is computed independently from the CCTop pipeline, this step also verifies the initial sgRNA design. While the upstream forward primer (#1) and downstream reverse primer (#5) position is always variable depending on the chosen sgRNA, the upstream reverse primer (#2) and downstream forward primer (#4) are designed at the same positions for each gene. To show that the LeishGEdit v4.2 script can be used as an efficient primer design tool, the batch mode script has been used to design primers for all currently available TritrypDB genomes (for DBv41) (Supplementary: File 2). sgRNA MM counts, as well as perfect match counts show that the large majority of designed sgRNAs have only one perfect match within the genome and no additional non-perfect matches with 1-4 MM (Figure 13 A and B).

To design a simple scalable method to check if the targeted ORF has been successfully deleted, a strategy for amplification of a small fragment of the ORF was designed next. To design primers for ORF amplification Primer3 was used. Primers were designed for all TritrypDB genomes (DBv41) and set to amplicon lengths of 100 to 350 bp (Koressaar and Remm, 2007; Untergasser et al., 2012) (Supplementary: File 3).

Next, outcomes from the LeishGEdit primer design were uploaded on LeishGEdit (www.leishgedit.net) to make primers publicly available. Meanwhile, primers for bar-seq assays, the primer3 search outcome and a primer design tool for arranging primers on 96-well plates were uploaded to LeishGEditDB, which is not publicly available yet (Figure 14).

2.2.4 Efficient gene editing of PF16 in *Leishmania mexicana*

For the first tests of functionality of the LeishGEdit toolbox, PF16 was subjected to gene deletion and tagging, yielding illustrative examples of generated mutants. Previously performed PF16 RNAi experiments in *T. brucei* (Branche et al., 2006; Ralston et al., 2006) and PF16 tagging experiments in *L. mexicana* (Wheeler et al., 2015a), as well as the editing experiments of the PF16 gene in *L. mexicana* shown here (Beneke and Gluenz, 2019; Beneke et al., 2017), demonstrate that PF16 is the gene of choice for testing LeishGEdit tools in any new kinetoplastid species. Its product has a defined localization in the flagellar axoneme and a null mutant is expected to produce a distinctive phenotype of reduced motility.

To test the functionality of LeishGEdit and *L. mexicana* Cas9 T7 cells, PF16 was first tagged separately at its N- and C-terminus using a pPLOT eYFP plasmid and the streamlined LeishGEdit pipeline (Figure 15) (Beneke and Gluenz, 2019).

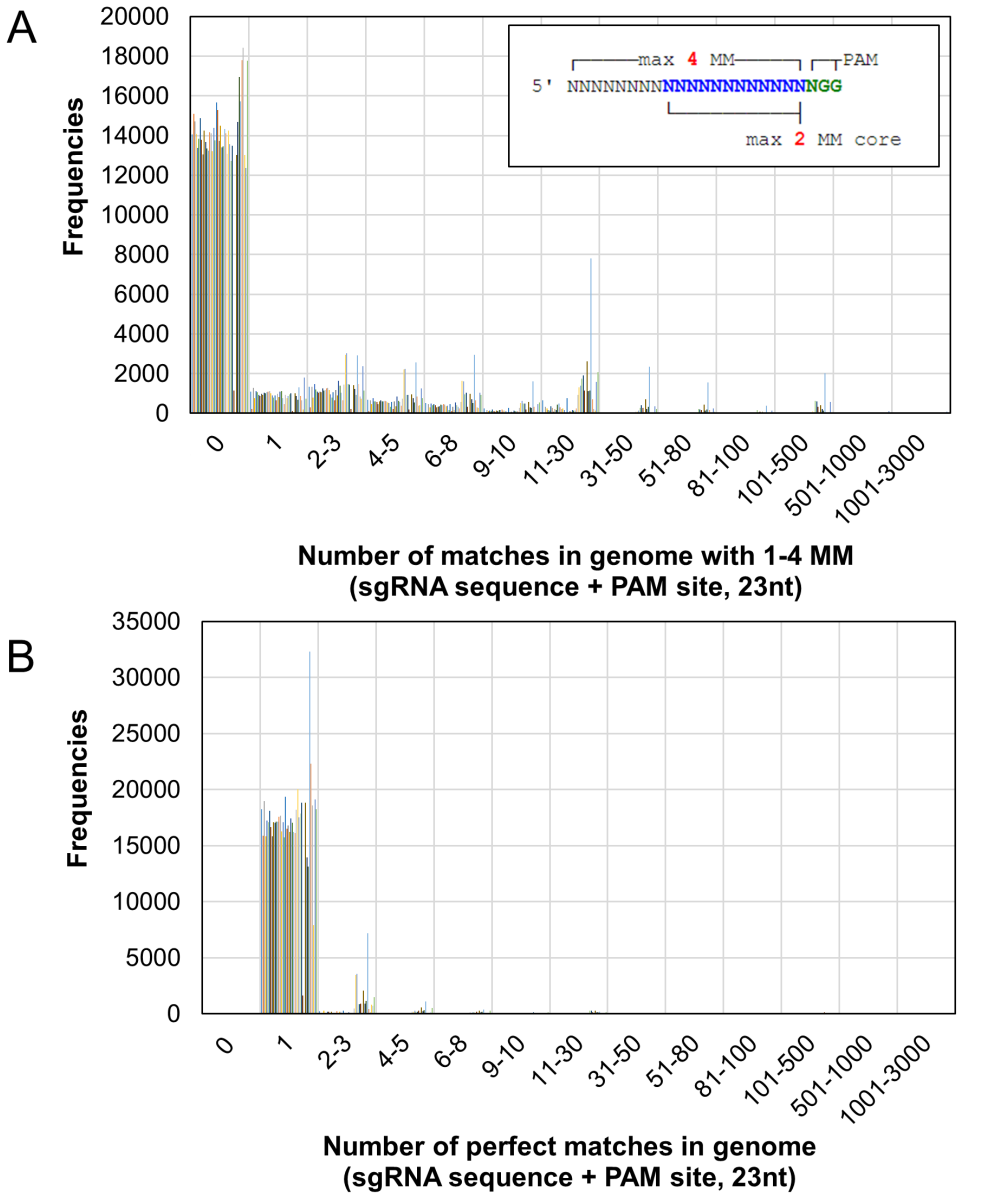


Figure legend on following page.

Figure 13. sgRNA mismatch and perfect match count for TritrypDB41.
LeishGEdit primer design script v4.2 has been run over all TritrypDB genomes (DB release 41). Y axis: Frequencies of perfect match and mismatch counts. X axis: Number of perfect match and mismatch count per gene. Designed sgRNAs have been verified by **(A)** using a mismatch counter, which counts all sites that match the sgRNA target and its PAM sequence with 1-4 MM as illustrated on the top right and **(B)** using a perfect match counter, which counts all sites that match the sgRNA target and its PAM sequence with 0 MM.

A



Public resource

- Primer design for CRISPR Cas9 gene editing (Beneke, Madden et al., Roy. Soc. Open Sci. 2017)

B



Currently lab resource

- Primer design for CRISPR Cas9 gene editing (Beneke, Madden et al., Roy. Soc. Open Sci. 2017)
- Barcoding of mutants (Beneke et al., PLoS Pathogens 2019)
- Primer design on 96 well plates (Beneke and Gluenz al., Methods in Biology 2019)
- Phenotype database

Figure 14. Current LeishGEdit online resources.

Overview of LeishGEdit website resources, including **(A)** the public resource www.LeishGEdit.net and **(B)** the currently only lab internal resource LeishGEditDB. The purpose of each resource is indicated.

As shown in Figure 15 primer sequences were retrieved from LeishGEdit and *L. mexicana* Cas9 T7 promastigote cells were co-transfected with two PCR-generated DNA fragments, the sgRNA template and the pPLOT donor DNA containing 30nt HF, a drug resistance cassette and an eYFP tag for in-frame tagging. Cells co-transfected with donor DNA and sgRNA template yielded drug resistant cells expressing fusion proteins that localized to the flagellar axoneme, which is consistent with the known association of PF16 with the C1 tubule of the central pair microtubule (Dutcher et al., 1984) (Figure 16 A, B and C). Cells transfected with donor DNA alone or without any DNA, did not become drug resistant. Next, pT plasmids and *L. mexicana* Cas9 T7 cells were used to generate null mutants in a single transfection round. PF16 knockdown by RNAi results in loss of motility (Branche et al., 2006; Ralston et al., 2006), but PF16 deletion had never been attempted in *Leishmania* before. *L. mexicana* Cas9 T7 promastigotes were co-transfected with four PCR-generated DNA fragments, two sgRNA templates, guiding CRISPR-Cas9 upstream and downstream of each target gene and two pT donor DNA constructs with 30nt HF, containing each a different drug resistance cassette (Figure 16 D).

All possible combinations pTBlast/pTNeo, pTBlast/pTPuro and pTNeo/pTPuro were used and yielded similar results, when employing dual drug selection to eliminate any heterozygous cells that still harbored a copy of PF16. Double-drug resistant cells emerged after 6-7 days post transfection and mutants were found to be immotile in culture and with aberrant orientations of the central pair microtubules or with missing parts or entirely missing central pair microtubules (Beneke et al., 2019b; Beneke and Gluenz, 2019; Beneke et al., 2017).

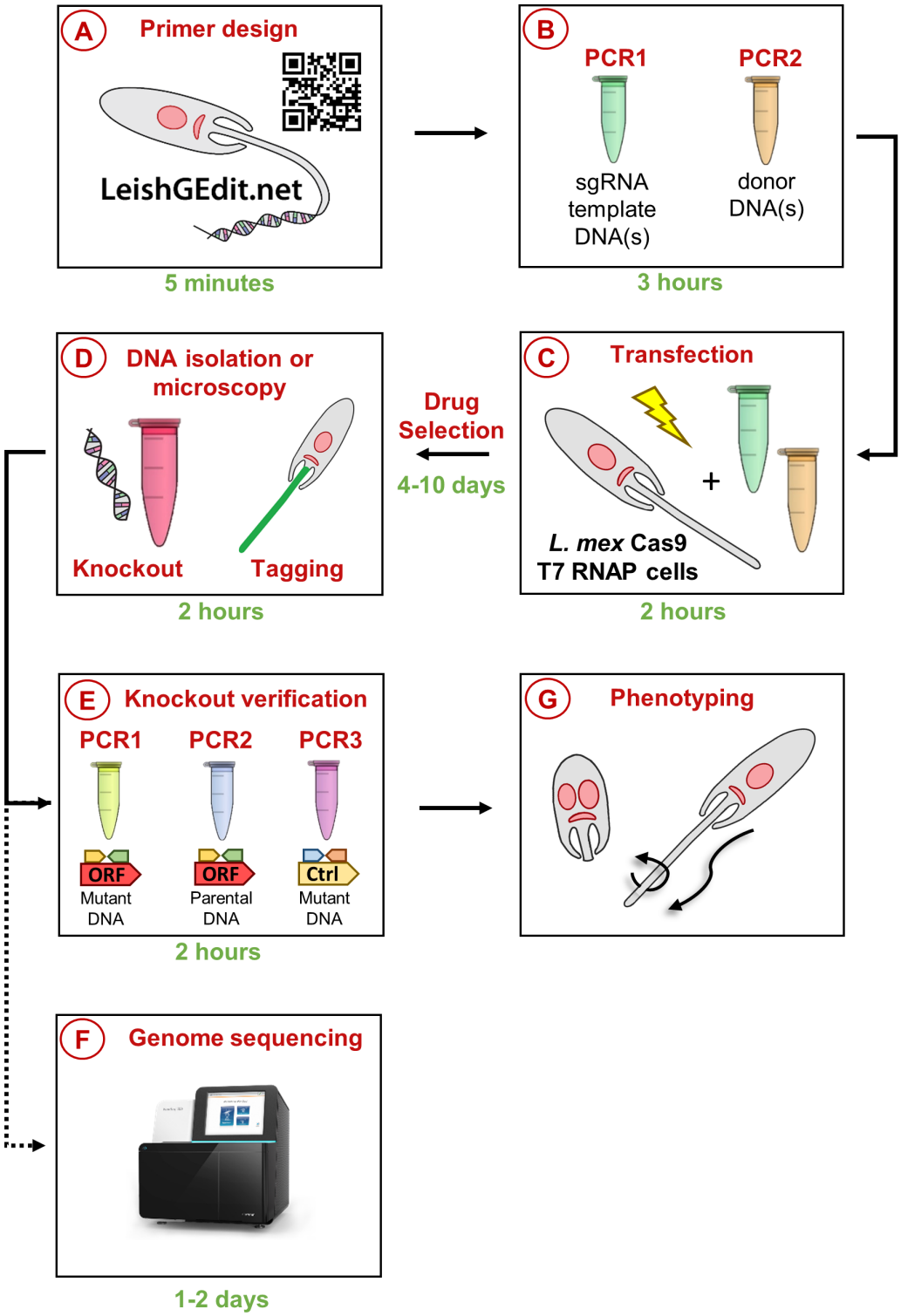


Figure legend on following page.

Figure 15. Streamlined LeishGEdit pipeline for mutant generation.

Figure made by Tom Beneke, adapted from Beneke and Gluenz (2019). Pipeline for generation of Cas9 mutants using the LeishGEdit protocol. Time for each protocol step is indicated in green below each protocol step. The pipeline applies for 96 well plates and single cuvettes. **(A)** Primers are design using the public resource www.LeishGEdit.net and used **(B)** to amplify donor DNA and sgRNA template DNA constructs. **(C)** PCR products are transfected in *L. mexicana* cells that are constitutively expressing Cas9 and T7 RNAP. **(D)** After 4 – 10 days drug selection mutants can be assessed for example by imaging tagged cell lines or isolating DNA from null mutants to perform knockout verification PCRs. **(E)** Three PCR reactions are performed for mutant verification, including (PCR1) amplification of the ORF using mutant DNA and (PCR2) parental DNA, as well as (PCR3) amplification of a control gene using mutant DNA to show presence of DNA. **(F)** Optional, mutants can be assessed by whole genome sequencing. **(G)** Confirmed null mutants are phenotyped by any follow-up assay.

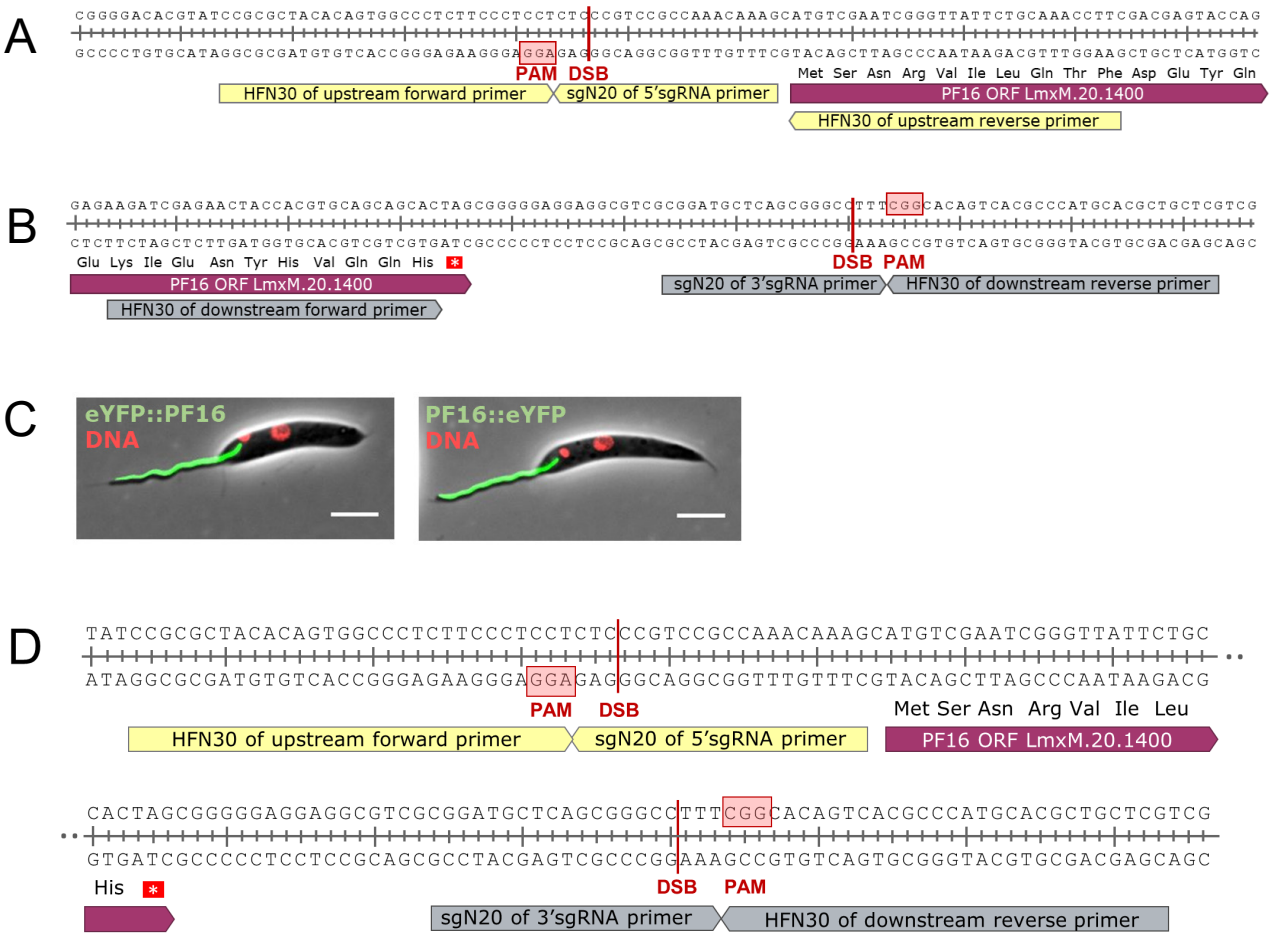


Figure legend on following page.

Figure 16. Gene editing of PF16 by using the LeishGEdit toolbox.

Figure made by Tom Beneke, adapted from Beneke and Gluenz (2019). Sequences show PF16 locus in *L. mexicana*. PAM site and location of DSB within the locus are indicated in red. K: knockout; P: parental. A DSB is introduced at the **(A)** upstream and **(B)** downstream region of the PF16 ORF to insert an N- or C-terminal eYFP tag in the PF16 locus. HFs used for pPLOT primers and sgRNA target sites are shown **(A)** in yellow and **(B)** in grey. **(C)** Outcome of **(A)** N- and **(B)** C-terminal tagging, showing merged images of PF16 fused to eYFP in green, DNA Hoechst stain in red and phase channel in grey. Scale bar is 5 µm. **(D)** Upstream forward primer, downstream reverse primer and sgRNAs from **(A)** and **(B)** are used to target the PF16 ORF for gene replacement. Agarose gels of PCR verification reactions illustrated in cartoons show **(E)** presence of the PF16 ORF in parental cells and absence in knockout cells, as well as **(F)** presence of DNA in both cell lines by amplifying a control gene.

Deletion of the PF16 ORF was confirmed by PCR amplifying small segments of the ORF in parental and knockout cell lines as shown for the pTBlast/pTPuro combination (Figure 16 E). The blasticidin-S deaminase ORF was PCR amplified as a control to show presence of DNA (Figure 16 F).

2.2.5 Whole genome sequencing of *Leishmania* parasites confirms precise mutations

While PCR amplification for verification of null mutants is scalable and is a highly sensitive approach with the capability of detecting a single molecule, it ultimately does not prove whether a deletion attempt was actually successful. Only presence or absence of the amplified ORF fragment can be assessed. Additionally, PCR amplification of a fragment of the deleted ORF may result in non-specifc amplicon bands. But it is unclear whether this is due to non-specifc primer binding or whether the targeted ORF re-integrates into a different locus. Meanwhile, whole genome sequencing provides much better insight into the genome of a mutant. In the following, DNA sequencing of *ΔPF16*, *ΔPKAC1* and *ΔLmxM.08_29.1000* mutants will demonstrate (1) that integration of the donor DNA construct is precise and (2) illustrate how paralogous genes can give misleading results when PCR amplification is used for assessing null mutants.

Illumina sequencing of genomic DNA from the *L. mexicana* Cas9 T7 parental cell line and of a pTBlast/pTPuro *ΔPF16* mutant population showed that the PF16 ORF was completely deleted. Only two reads, which were both paired, aligned to the PF16 ORF in *ΔPF16* mutants. However, this read pair had low quality scores and many SNV (single nucleotide variance) (Figure 17 A). Blast analysis showed that these two reads were the only ones still present in the *ΔPF16* mutant genome, while more than 750 reads aligned to the PF16 ORF in parental cells (Figure 17

B). It is possible that both reads were either wrongly aligned to this locus or falsely de-multiplexed. Also possible is a contamination during library preparation. Both sequenced cell lines had similar read coverage (~100x coverage) and the blasticidin-S deaminase and puromycin N-acetyltransferase resistance ORF was only found in the *ΔPF16* mutant genome (Figure 17 B). Sequencing of PCR amplicons spanning the resistance cassette integration sites in a pTNeo/pTPuro *ΔPF16* mutant population and a *ΔPF16* clone confirmed Illumina sequencing results of pTBlast/pTPuro *ΔPF16* mutants, showing that integration of the cassettes is precise, with no detectable insertions or deletions (Figure 17 C, D and E). Overall, these sequencing results strongly suggest that the PF16 ORF is indeed absent from the altered mutant genome.

While deletion of PF16 was shown to be precise, there are of course limits of CRISPR-Cas9 gene editing in *Leishmania*. Recently, the first 100 null mutants of a large scale knockout screen of approximately 500 flagellar protein coding genes (Gluenz lab unpublished), have been published, showing the potential of this gene editing toolbox for larger screens (Beneke et al., 2019b) (see section 3.1.5 for more information). However, out of 100 null mutants four were found to be incomplete deletions. One of these four mutants, *ΔPKAC1*, showed significantly reduced swimming speed even though the tested ORF fragment could be still detected by PCR (Figure 18 A). One likely explanation for incomplete deletions are randomly amplified chromosomes, also referred to as chromosomal copy number variation (Lachaud et al., 2014; Mannaert et al., 2012; Sterkers et al., 2012). It is unclear whether CRISPR-Cas9 gene editing specifically triggers chromosome amplification, but it has been suggested numerous times that *Leishmania* duplicates essential genes if knockouts are attempted (see further discussion in

section 2.3.1). To assess unclear PCR verification results of *ΔPKAC1*, mutant DNA was also subjected to whole genome Illumina sequencing. Sequencing results showed that reads aligned to HFs used for deletion of PKAC1 and then abruptly discontinued exactly were drug resistance cassettes are integrated (Figure 18 C). However, 660 reads still aligned to the ORF of PKAC1 (Figure 18 E) and the pTBlast/pTPuro resistance cassette was clearly integrated into the PKAC1 locus. Further analysis showed that all 660 reads that still aligned to the PKAC1 ORF had a mapping score of zero (read highlighted in white in Figure 18 C), meaning that these reads do align to another locus within the sequenced *ΔPKAC1* mutant genome. Blast analysis showed that these reads aligned to PKAC2 and that primers used for amplification of PKAC1 to verify mutants do also amplify a fragment of the PKAC2 ORF, as they are also matching the PKAC2 locus (Figure 19 A and B). Aligment of both ORFs showed that the PKAC1 ORF is highly similar to the PKAC2 ORF (Figure 19 B).

Meanwhile, Illumina sequencing of *ΔLmxM.08_29.1000*, a mutant that was also published in Beneke et al. (2019b) and showed non-specifc amplification bands in mutant PCR verifications, demonstrated that the targeted ORF is not detectable in mutants (Figure 18 D and E), suggesting that non-specifc bands seen in Figure 18 B are likely to be a result of non-specifc primer binding.

Sequencing *ΔPKAC1*, *ΔLmxM.08_29.1000* and *ΔPF16* mutants shows that integration of the cassettes is precise, with no detectable insertions or deletions.

Illumina sequencing showed also that there were about 33,000 SNVs between the *L. mexicana* Cas9 T7 parental cell line (WHO strain MNYC/BZ/62/M379) and the current TritrypDB reference genome (MHOM/GT/2001/U1103), which is derived from a different strain. All three sequenced mutants shared most of these SNVs

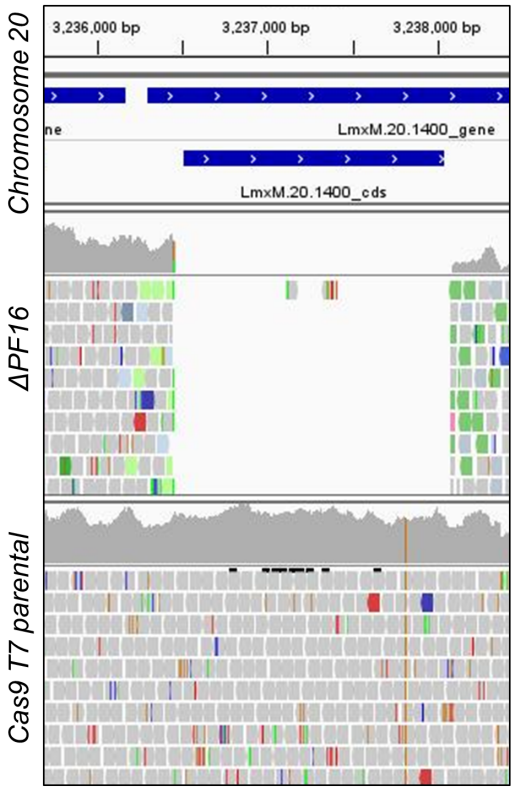
with a small number of different SNVs. Similar SNV rates were also observed after passaging *L. mexicana* Cas9 T7 parental cells through BALB/c mice. These numbers are comparable with a recent analysis of Tunisian and Algerian *L. major* samples, in which 36,726 SNVs were identified compared to the reference genome. This analysis also showed that only a small number of different SNVs (3.5%) between both analyzed samples was found (Bussotti et al., 2018).

To analyze chromosome ploidy of mutants and parental cells the read coverage per chromosome was calculated next. Read coverage for all chromosomes in *ΔPKAC1*, *ΔLmxM.08_29.1000, ΔPF16, Δhydin* (discussed in chapter 3) and *ΔLAX28* (discussed in chapter 4) mutants and parental cell lines was found to be mostly represented at equal levels. However, chromosomes 3 and 16 showed a consistent increase in read coverage of 20 – 50%. Additionally, read coverage for chromosome 30 was increased up to 100% compared to other chromosomes (Figure 20 A). Given that the majority of chromosomes in the *L. mexicana* MNYC/BZ/62/M379 strain are found to be diploid (Mannaert et al., 2012) (Figure 20 B), this suggests that chromosomes 3 and 16 are likely to be triploid or in an intermediate stage between diploid and tetraploid, while chromosome 30 may be tetraploid. Only results for *ΔLmxM.08_29.1000* deviated from these observations, as read coverage for chromosome 3 was not found to be elevated. While, the cause of this remains elusive for now, it actually is consistent with previous reports, stating that chromosome 3 is diploid in *L. mexicana* MNYC/BZ/62/M379. Meanwhile, results for chromosome 16 (intermediate or triploid) and 30 (tetraploid) are consistent with previous reports (Mannaert et al., 2012).

Nucleotide MM between the TritrypDB reference genome and *L. mexicana* Cas9 T7 parental cells also gave substantial evidence in *Leishmania* that sgRNAs are

still fully functional to guide Cas9 to target sites even though sgRNAs do not perfectly match. Out of the published 100 null mutants (Beneke et al., 2019b) both sgRNAs used to generate *ΔCD047* mutants were later found to not actually perfectly match their target locus (Figure 21 A and B). However, a null mutant was still successfully generated (Figure 21 C). This was the only null mutant where both sgRNAs were later found to have imperfect matches. While specificity for integration of drug resistance markers in null mutants is also achieved by using 30 nt HFs matching the target locus, this result shows that MM are tolerated. Sequence matching a sgRNA partially elsewhere in the genome could also be cleaved and result in off-target mutations which could affect the phenotype. Therefore, sgRNA MM must be carefully considered when using CRISPR-Cas9 gene editing in *Leishmania*. This motivated the new design of v4.2 of the LeishGEdit primer design tool (v4.2) with its improved pipeline to design sgRNAs with fewer non-perfect matches (see section 2.2.3 above).

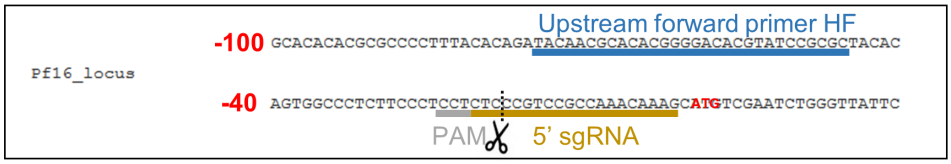
A



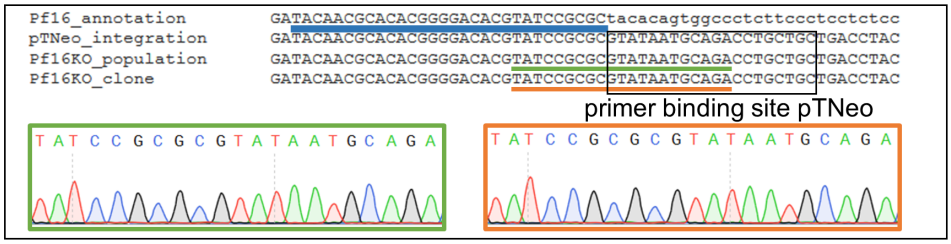
B

ORF blast query	ΔPF16	Cas9 T7 parental
	Aligned number of reads to ORF query	
PF16	2	>750
Blasticidin	83	0
Puromycin	148	0
Neomycin	0	0

C



D



E

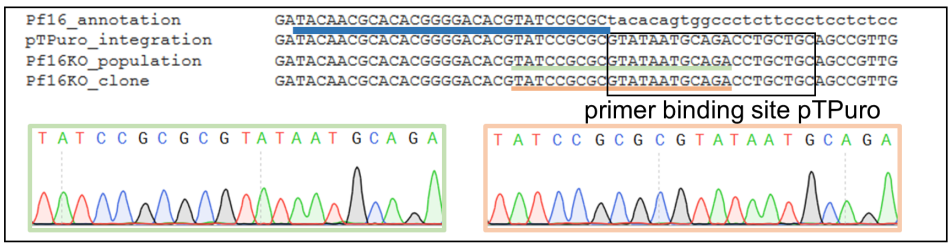
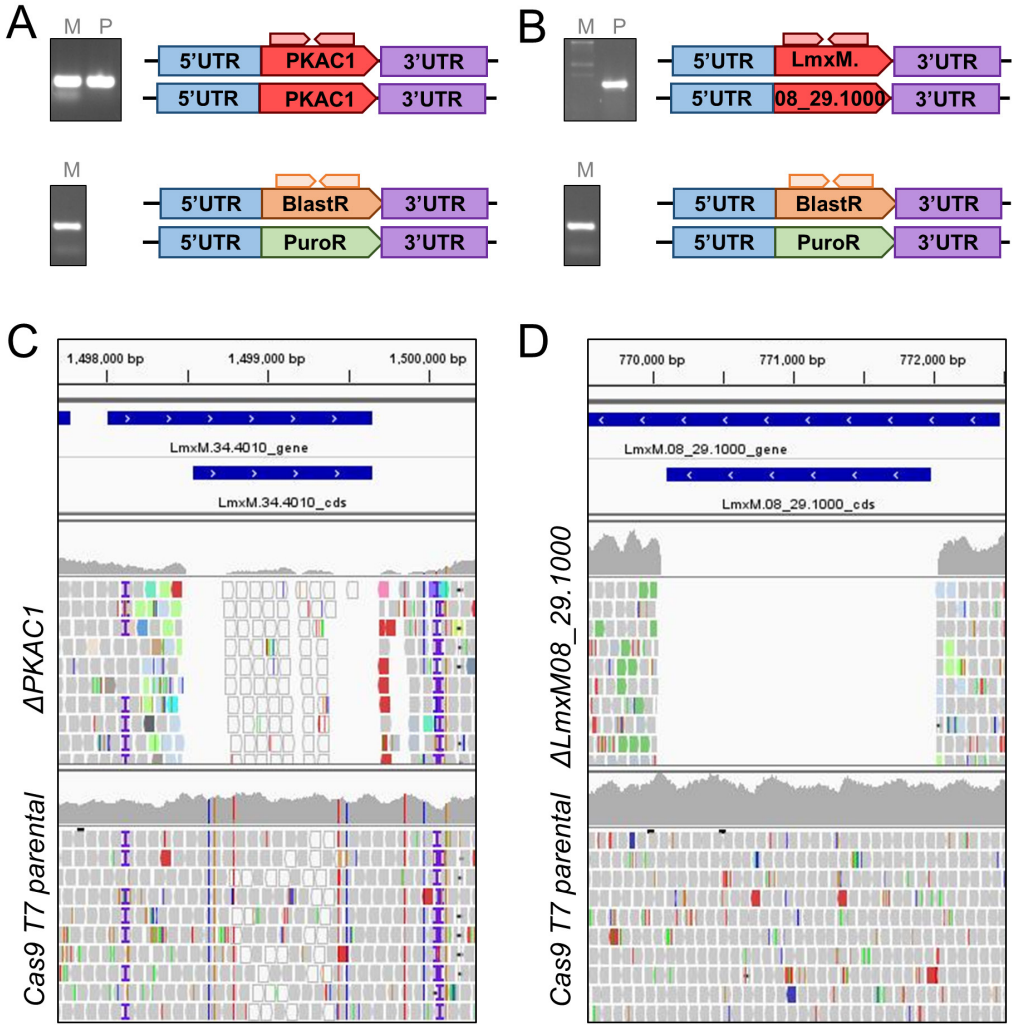


Figure legend on following page.

Figure 17. Sequencing of *ΔPF16* mutants and *L. mexicana* Cas9 T7 parental cells.

Figure **(C)**, **(D)** and **(E)** made by Tom Beneke, taken from Beneke et al. (2017). **(A)** Bam files of *ΔPF16* pTPuro pTBlast mutants and *L. mexicana* Cas9 T7 parental cells were viewed in IGV genome browser (Robinson et al., 2011), showing in each panel to the locus matching reads. Mapped reads are highlighted in grey as histogram (read coverage), in grey as arrows (high unique mapping score, perfect alignment to the locus), in white as arrows (zero unique mapping score, perfect alignment to the locus) and in various colours as arrows (high mapping score, alignment with SNVs to the locus as defined in IGV manual). **(B)** Blast queries of the PF16 ORF and blasticidin-S deaminase, puromycin N-acetyltransferase and neomycin-kanamycin phosphotransferase resistance gene against *ΔPF16* mutants and *L. mexicana* Cas9 T7 parental cells were performed. Two reads of PF16 with low quality scores and several SNVs were identified in *ΔPF16* mutants. The upstream HFs and their adjacent sequences in the *ΔPF16* pTNeo pTPuro mutant was subjected to Sanger sequencing. **(C)** Highlighted in the PF16 locus are the upstream forward primer HF sequence (underlined in blue), the sgRNA sequence (underlined in yellow) and the PAM site (underlined in grey), as well the location of the DSB. Sequencing results are shown for the HF integration site in the **(D)** pTNeo and **(E)** pTPuro allele. Highlighted sequences in green and orange refer to Sanger sequencing reads shown.



E

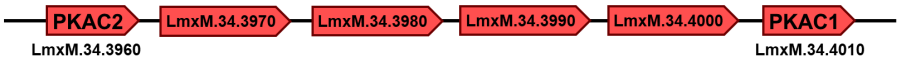
ORF blast query	*ΔLmxM08_29.1000*	*ΔPKAC1 (LmxM.34.4010)*	*Cas9 T7 parental*
	Aligned number of reads to ORF query		
LmxM.08_29.1000	0	>750	>750
PKAC1 (LmxM.34.4010)	>750	660	>750
Blasticidin	184	107	0
Puromycin	251	209	0
Neomycin	0	0	0

Figure legend on following page.

Figure 18. PCR verification and Illumina sequencing of *ΔPKAC1* and *ΔLmxM.08_29.1000*.

M: mutant; P: parental. PCR verification reactions show **(A)** presence of the PKAC1 ORF in parental cells and mutant cells (top agarose gel), as well as presence of DNA in mutant cells (bottom agarose gel) and **(B)** presence of the LmxM.08_29.1000 ORF in parental cells and non-specifc bands resulting from the PCR amplification in mutant cells (top agarose gel), as well as presence of DNA in mutant cells (bottom agarose gel). Arrows indicate location of primers used in reactions. **(C)** and **(D)** Bam files of both mutants and *L. mexicana* Cas9 T7 parental cells were viewed in IGV genome browser (Robinson et al., 2011) (as described in Figure 17) and **(E)** blast queries of both ORFs and blasticidin-S deaminase, puromycin N-acetyltransferase and neomycin-kanamycin phosphotransferase resistance genes were run against mutant cell lines, showing aligned number of reads to each blast query.

A



B

```
PKAC1       -----ATGTCGTCTGCTGCCAA--GGACAGCTATCCGGCGGAG---------TCATTGGG     44
PKAC2       ATGCTAGTGAGTGTGGAGCCGAAACCTGAGGTGGCTGGCACAGTCCACGCACCCCTTCGG     60
                 *    ** **  *** *      ** *  * ***  **          * ** **

PKAC1       CCAACTCAGCAGCGAGAAGGCGCA-----------------GACGTCGAAGGCTCCCTCA     87
PKAC2       CCCAGGAGGACGAGTGTACGCGGTGCGTCTCCCATTATCCCCACTCCGTCCAGATTGGTG     120
            ** *    *  * * * * ***                     **  **

PKAC1       AGTCATCCCGACTCCCCGAAGGGGGCTGTCACGCCAAATGCCAAGGCCACCTATTCCTTT     147
PKAC2       CATGCTGCGTCCCCTAGGACGTGGTCCTCCTCTTTGAAACCCAAGGTCGGGTGCACCTTT     180
              *  * *   * *   ** * ** *   * *    **  ****** *   *   *****

PKAC1       GAGAGGCCGGACACCTCCAGCTGGAAATTCTCGGACTTCGAACTGAAAAACACGCTCGGC     207
PKAC2       GAGAGGCCGGACACCTCCAGCTGGAAACTCTCGGACTTCGAATTGAAAAACACGCTCGGC     240
            ************************* ************** *****************

PKAC1       ACCGGCTCCTTCGGCCGCGTGCGCATCGCTCACCGCAAGGGCACGGAGGAGTACTACGCG     267
PKAC2       ACCGGCTCCTTCGGCCGCGTGCGCATCGCTCACCGCAAGGGCACGGAGGAGTACTACGCG     300
            ************************************************************

PKAC1       ATCAAGTGCCTGAGAAAGCGTGAGATCATCAAGATGAAGCAGCAGCAGCACGTTGCGCAG     327
PKAC2       ATCAAGTGCCTGAGAAAGCGTGAGATCATCAAGATGAAGCAGCAGCAGCACGTTGCGCAG     360
            ************************************************************

PKAC1       GAAAAGGGGATCCTAATGGAGGTGTGTCACCCGTTCATCGTGAACATGATGTGCTCCTTC     387
PKAC2       GAAAAGGGGATCCTAATGGAGGTGTGTCACCCGTTCATCGTGAACATGATGTGCTCCTTC     420
            ************************************************************

PKAC1       CAGGACGAGAAGAAGGTGTACTTTGTGCTGGAGTTCGTGATGGGCGGCGAGATGTTCACG     447
PKAC2       CAGGACGAGAAGAAGGTGTACTTTGTGCTGGAGTTCGTGATGGGCGGCGAGATGTTCACG     480
            ************************************************************

PKAC1       CACCTGCGCACTGCTGGGCGGTTCCCGAATGACGTTGCGAAGTTCTACCACGCGGAGCTG     507
PKAC2       CACCTGCGCACTGCTGGGCGGTTCCCGAATGACGTTGCGAAGTTCTACCACGCGGAGCTG     540
            ************************************************************

PKAC1       GTGCTTGCGTTCGAGTACCTGCACTCGCTGGACGTGATCTACCGCGATCTGAAGCCGGAG     567
PKAC2       GTGCTTGCGTTCGAGTACCTGCACTCGCTGGACGTGATCTACCGCGATCTGAAGCCGGAG     600
            ************************************************************

PKAC1       AATCTGCTGCTGGACAACAAGGGGCACGTGAAGGTGACGGACTTTGGGTTTGCGAAGAAG     627
PKAC2       AATCTGCTGCTGGACAACAAGGGGCACGTGAAGGTGACGGACTTTGGGTTTGCGAAGAAG     660
            ************************************************************

PKAC1       GTGCCGGACCGGACGTTCACGCTGTGCGGGACACCGGAGTATCTTGCGCCGGAGGTGATC     687
PKAC2       GTGCCGGACCGGACGTTCACGCTGTGCGGGACACCGGTGTATCTTGCGCCGGAGGTGATC     720
            ********************************** ***********************

PKAC1       CAGAGCAAGGGCCACGGGAAGGCGGTGGACTGGTGGACGATGGGCGTGCTGCTGTACGAA     747
PKAC2       CAGAGCAAGGGCCACGGGAAGGCGGTGGACTGGTGGACGATGGGCGTGCTGCTGTACGAA     780
            ************************************************************

PKAC1       TTCATTGCTGGATACCCGCCGTTCTACGATGACACGCCGTTCCGGATTTACGAGAAGATC     807
PKAC2       TTCATTGCTGGATACCCGCCGTTCTACGATGACACGCCGTTCCGGATTTACGAGAAGATC     840
            ************************************************************

PKAC1       CTTGCTGGGCGGCTGAAGTTCCCGAACTGGTTTGACGGCCGCGCTCGCGACCTTGTGAAG     867
PKAC2       CTTGCTGGGCGGCTGAAGTTCCCGAACTGGTTTGACGGCCGCGCTCGCGACCTTGTGAAG     900
            ************************************************************

PKAC1       GCGCTGCTGCAGACAGACCATACGAAGCGCCTGGGGACGCTGAAAGGCGGGCCGGCGGAC     927
PKAC2       GCGCTGCTGCAGACAGACCATACGAAGCGCCTGGGGTCGCTGAAAGGCGGGCTAGCGGAT     960
            ********************************** ***************  *****

PKAC1       GTGAAGAACCACCCGTACTTCCACGGTGCGAACTGGAGCAAGCTGTACGCCCGCTACTAC     987
PKAC2       GTGAAGAACCACCCGTACTTCCACGGCGTGAACTGGAGCAAGCTGTACGCCCGCTACTAC     1020
            ************************* * ******************************

PKAC1       CCCGCGCCGATCCCTGTGCGGGTGAAGAGCCCTGGCGACACGAGCAACTTCGAAAAGTAC     1047
PKAC2       CCCGCGCCGATCCCTGTGCGGGTGAAGAGCCCTGGCGACACGAGCAACTTCGAAAAGCAC     1080
            ************************************************************ **

PKAC1       CCGGACAGCCCGGTCGACCGCACGCCTGCGCTGACGTCGGCGCAGCAGGCCGAGTTCAAT     1107
PKAC2       CCGGACAGCCCGGTCGACCGCACGCCTGCGCTGACGTCGGCGCAGCAGGCCGAATTCAAG     1140
            ************************************************** *****

PKAC1       GGCTTTTAG      1116
PKAC2       GGGTTCTAA      1149
            ** ** **
```

Figure 19. DNA alignment of PKAC1 and 2 ORF.

(A) Cartoon of PKAC locus. Red arrows indicate ORFs. **(B)** Alignment of PKAC1 and 2 ORF. Primer sequences used for PCR verification are highlighted in yellow.

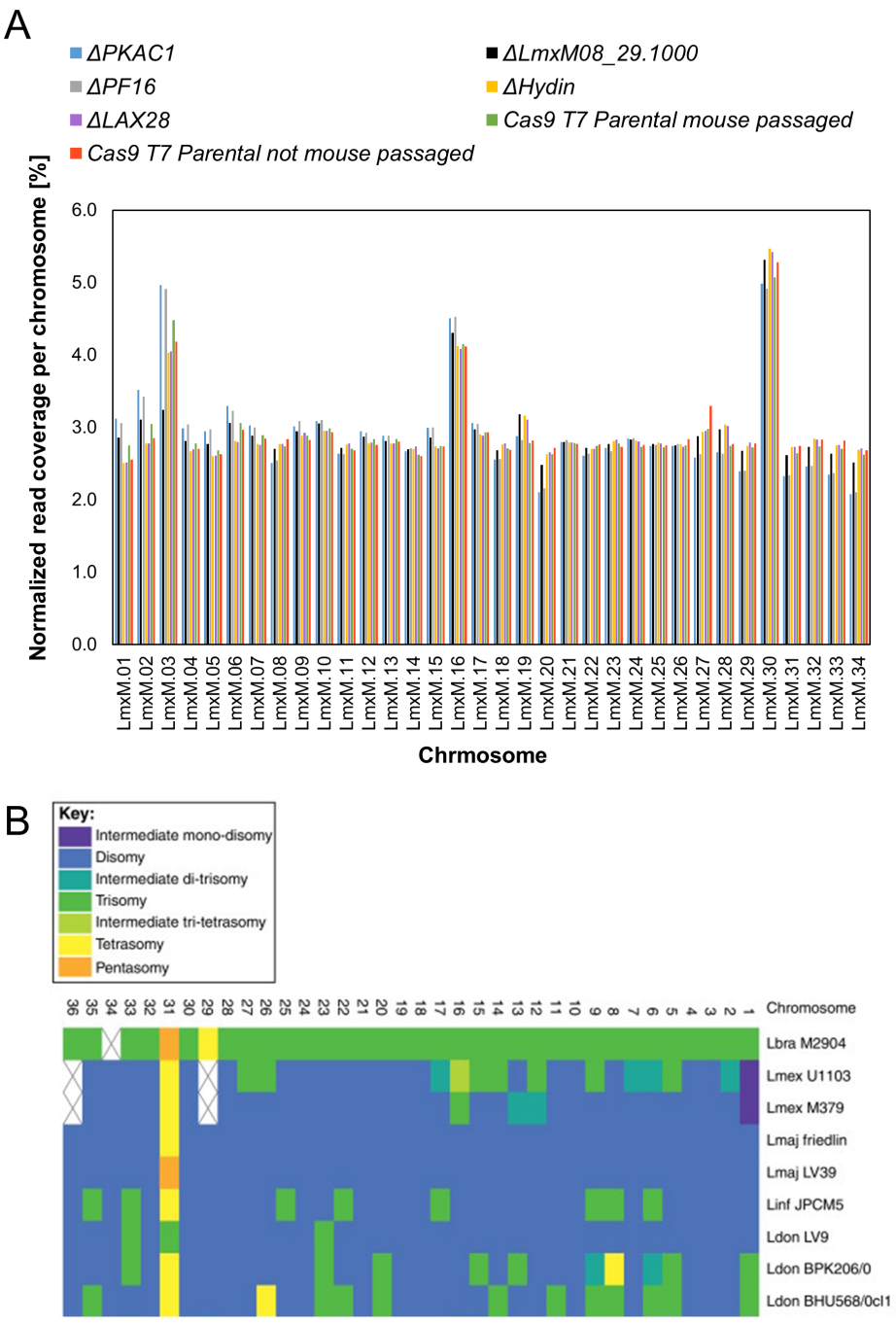


Figure 20. Analysis of chromosome ploidy in mutants and parental cells.

(A) Total reads per chromosome have been normalized with total reads per sequenced sample and plotted for all chromosomes of the *L. mexicana* MNYC/BZ/62/M379 genome annotation. **(B)** Figure reprinted from Mannaert et al. (2012) with permission from Elsevier, showing chromosome ploidy for several *Leishmania* species and strains.

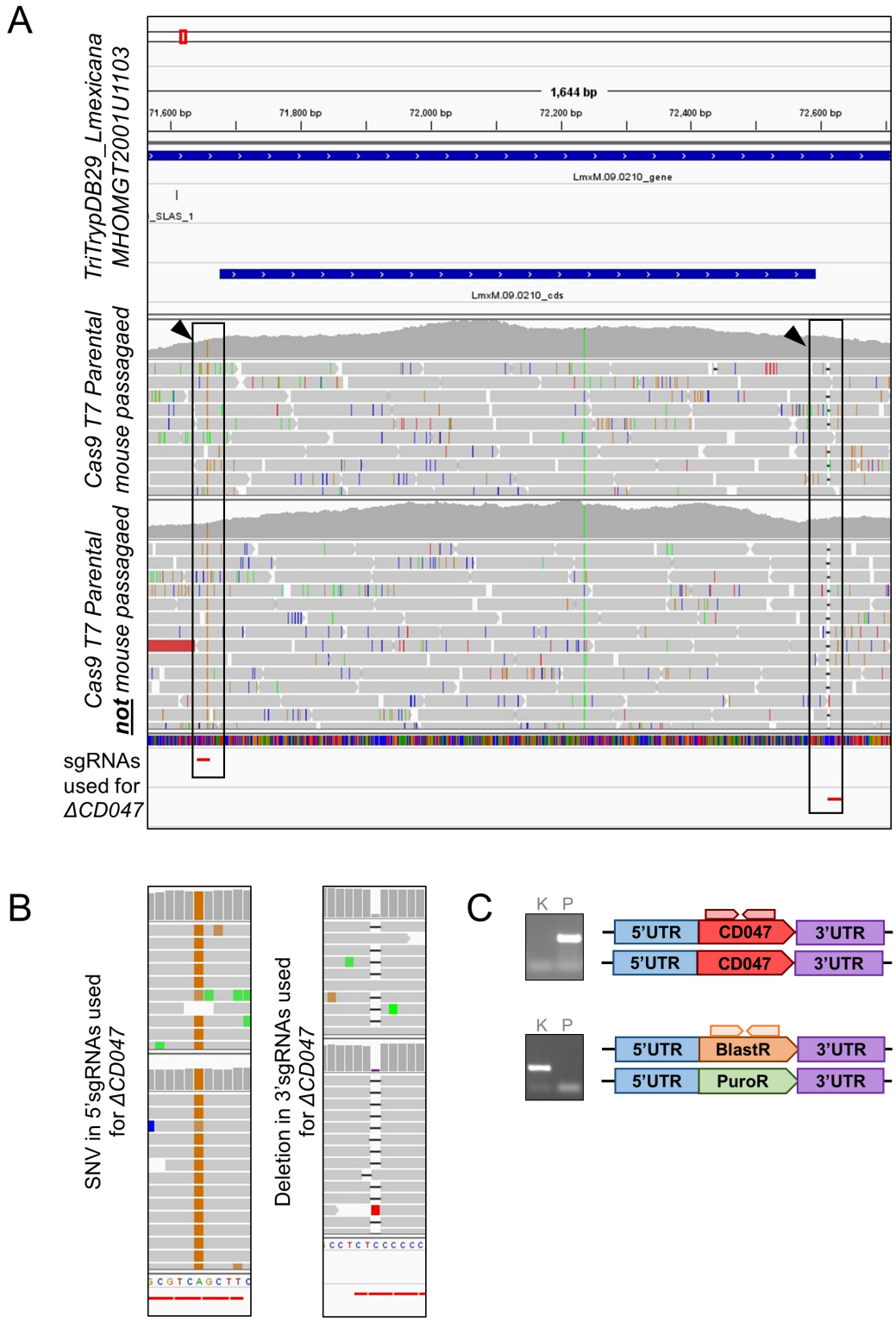


Figure legend on following page.

Figure 21. *ΔCD047* sgRNA analysis.

(A) *L. mexicana* Cas9 T7 parental cells, sequenced before and after mouse passage, were viewed in IGV genome browser (Robinson et al., 2011) (as described in Figure 17) using *L. mexicana* TriTrypDB29 reference genome annotation and viewed at the CD047 locus. **(B)** sgRNAs used for the deletion have a SNV or deletion and do not perfectly match their locus in sequencing data from parental cells. **(C)** Mutants are confirmed as knockouts as shown in agarose gels of PCR verification reactions (reactions illustrated in cartoons), showing presence of the CD047 ORF in parental cells and absence in knockout cells (top gel), as well as presence of the blasticidin-S deaminase ORF in knockout cells and absence in parental cells (bottom gel).

2.2.6 Constitutive expression of Cas9 and T7 RNAP in *Leishmania* and *Trypanosoma* spp.

To make the toolbox more versatile and to facilitate generation of Cas9 T7 cell lines in different kinetoplastid species and strains, three additional Cas9 and T7 RNAP plasmids were designed: (1) pTB007, a single marker CRISPR-Cas9 and T7 RNAP plasmid that can be transfected as an episome or integrated into the β-tubulin locus (HFs derived from *L. major*), (2) pTB008, a variation of the pVY087 plasmid in which the T7 RNAP is expressed under control of a CYS2 3'UTR rather than a calmodulin 3'UTR and (3) pTB011, a CRISPR-Cas9 expression plasmid for *T. brucei*, targeting the α-β-tubulin locus.

L. major and *L. mexicana* promastigotes that have been transfected with pTB007 express Cas9 and T7 RNAP (Figure 22 A) and show no difference in their growth rate compared to wild type cells. They also have slightly lower doubling times than cells that express Cas9 and T7 RNAP from two different loci, as e.g. pRM006 pTB008 cell lines (Figure 22 B), indicating that expression of fewer drug resistance markers might be beneficial.

Transfection of pTB011 into the T7RNAP-expressing cell lines SmOx B4 (bloodstream form) and SmOx P9 (procyclic form) introduced SpCas9 expression into *T. brucei* (Figure 22 A). A minor, but not significant increase in doubling time was noted upon transfection of pTB011 (Figure 22 B).

As shown in Beneke et al. (2017), *L. major* and *T. brucei* Cas9 T7 RNAP expressing cell lines can be used for gene editing using the LeishGEdit toolbox. This has been demonstrated by knocking out and tagging PF16 in *L. major pTB007* cells and GPI-PLC in *T. brucei pTB011 SmOx P9* and *SmOx B4* cells (Beneke et al., 2017).

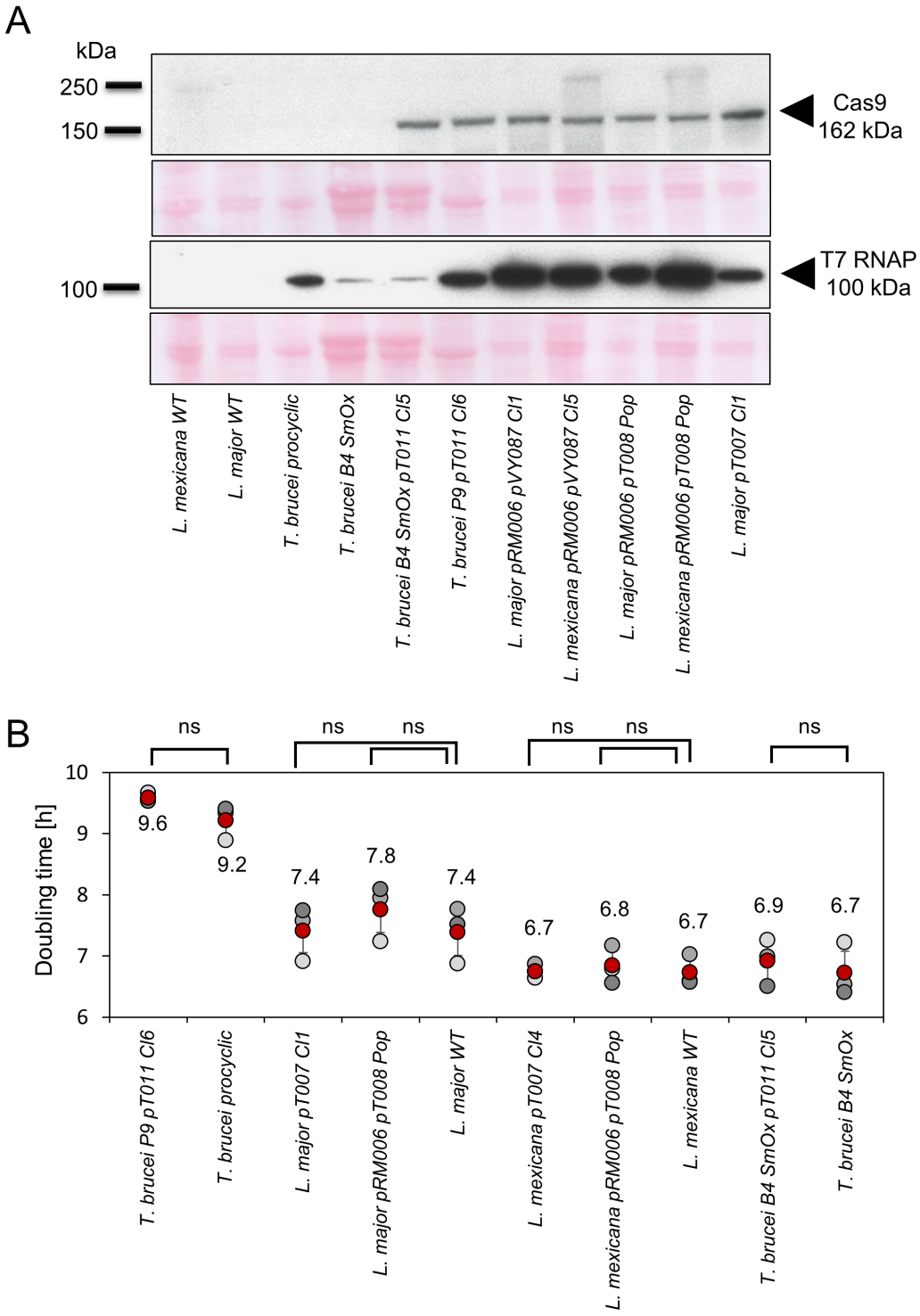


Figure legend on following page.

Figure 22. Assessing versions of Cas9 and T7 RNAP expressing cell lines.
L. mexicana and *L. major* parental cell lines have been transfected with Cas9 and T7 RNAP expression plasmids (Cas9: pRM006, T7 RNAP: pVY087 or pTB008, Cas9 T7 RNAP: pTB007). *T. brucei* procyclic and bloodstream form *SmOx* cell lines have been transfected with Cas9 expression plasmid (Cas9: pTB011). POP: parasites selected as population. All other cell lines are derived from single clones. **(A)** A western blot of several cell line versions using a flag epitope tag for Cas9 detection and a T7 RNAP antibody for detection of T7 RNAP has been performed. Molecular weights of Cas9 and T7 RNAP are indicated. Ponceau stains of membranes are shown in pink. **(B)** Doubling times were measured repeatedly three times (grey) and their average has been determined (red, value indicated above). Student's t-tests are indicated and error bars show standard deviation.

2.3 Discussion

2.3.1 Limitations of CRISPR-Cas9 gene editing in *Leishmania*

The aim of the work shown in this chapter was to develop a CRISPR-Cas9 toolbox, allowing high-throughput generation of mutants. Results in this chapter show how the LeishGEdit CRISPR-Cas9 toolbox can be effectively used for gene editing. Additionally, this toolbox has now been used successfully by myself (as shown in chapter 3, 4 and 5) (Beneke et al., 2019b; Beneke and Gluenz, 2019; Beneke et al., 2017) and by many other laboratories in the field (Basmaciyan et al., 2019a; Basmaciyan et al., 2019b; Cabello-Donayre et al., 2019; Chikne et al., 2019; Costa et al., 2018; de Lima et al., 2019; Dean et al., 2019; Edwards et al., 2018; Grewal et al., 2019; Halliday et al., 2019; Martel et al., 2017; Rojas et al., 2019; Schadeli et al., 2019; Shaw et al., 2019; Shrivastava et al., 2019; Szoor et al., 2019; Walker and Wheeler, 2019). This has contributed to major discoveries, such as showing the role of oligopeptide signalling through TbGPR89 in trypanosome quorum sensing (Rojas et al., 2019) or showing how flagellar cAMP signaling controls trypanosome progression through host tissues (Shaw et al., 2019). To our knowledge, in all organisms tested so far genes can be deleted in a single transfection round using LeishGEdit. However, limitations for *Leishmania* and other organisms remain, including especially null mutant validation issues and potential toxicity of Cas9.

PCR amplification of fragments of the target ORF is a scalable approach for null mutant verification and has been used in the first large scale knockout screen in *Leishmania* (Beneke et al., 2019b). For 96 null mutants no remaining fragment of the depleted ORF was detectable by PCR after deletion. This gives strong

evidence that the amplified ORF fragment is absent and therefore successfully deleted. However, since a single diagnostic PCR could fail because of technical reasons, such as SNVs in the primer binding site, this ultimately does not prove whether a transfectant is actually a null mutant. Illumina sequencing of *ΔPKAC1* and *ΔLmxM.08_29.1000* demonstrated how unclear PCR verification results can be further investigated. This includes incomplete deletions and deletions that result in non-specifc PCR products when the deleted target ORF is amplified.

PCR verification of *ΔPKAC1* suggested that mutants are incomplete in their deletion. But sequencing shows that the PKAC1 ORF is clearly lost and that the amplified ORF fragment may actually be derived from the PKAC2 ORF, which is not altered in PKAC1 null mutants. Although it seems unlikely, it cannot be excluded that these PKAC1 and PKAC2 matching reads might also result from excised PKAC1 fragments that re-integrated into another locus. This can be followed up in the future by using already available data and searching for reads that only partially match the PKAC1 ORF. Thus, incomplete mutants, as classified by PCR, can have multiple causes, such as: (1) a remaining ORF or ORF fragment, which re-integrated into another locus (highly unlikely as HFs are needed for integration), (2) primers used for verification PCR do anneal to other regions within the genome (as for PKAC1 primers) and (3) an additional chromosomal copy is present.

Next-generation sequencing can be used to determine whether subsets of chromosomes are amplified. However, this does have the flaw that copy numbers are averaged out over the sequenced population. Fluorescence *In Situ* Hybridization (FISH) assays are much more precise and have shown that some *Leishmania* species, especially *L. donovani*, vary significantly their chromosomal

copy number within one cell population (Lachaud et al., 2014; Mannaert et al., 2012; Sterkers et al., 2012). In the absence of transcriptional control, the parasite adapts itself with this mechanism to stressful conditions within its two hosts (all reviewed in Leprohon et al. (2014)). Additionally, genes can be amplified extra chromosomally, as circular (Beverley et al., 1984) or linear (Beverley and Coburn, 1990) elements. It is currently unknown whether CRISPR-Cas9 gene editing in *Leishmania* parasites specifically triggers chromosome or gene amplification, as it has been observed when attempting to delete essential genes in the past. For example knockout attempts of dihydrofolate reductase-thymidylate synthase in *L. major* showed that entire chromosomes can be amplified (Cruz et al., 1993). However, this has been observed when using homologous recombination in several transfection rounds as a replacement strategy, meaning each allele was deleted at a time giving cells the opportunity to adapt in-between transfections. This is different from the LeishGEdit CRISPR transfections which delete both alleles in a single transfection round, consequently reducing the time available for chromosome amplification. Alternatively, it is also possible that the gene editing process itself, whether through CRISPR or homologous recombination, does not cause the chromosome amplification, but selects for cells within the population that had varying numbers of chromosomes to start with.

However, read coverage for chromosomes in *ΔPF16*, *ΔPKAC1* and two other mutants did not deviate from coverage found for parental cell lines, suggesting that there aren't any abnormalities in these mutants. However, since read coverage for chromosome 3 in *ΔLmxM.08_29.1000* mutants is different from other mutants, a detailed systematic dissection of CRISPR mutants and wild type cells should be performed to investigate whether CRISPR-Cas9 may actually induce

chromosomal copy number variation. In a possible experiment, mutants should be produced in several replicates by CRISPR-Cas9 and original gene replacement approaches. At different time points after transfection sequencing samples from mutants should be isolated and analyzed along with parental and wild type cells that have been in culture for the same period of time.

What may also remain controversial is whether constitutive expression of Cas9 is toxic for parasites, as suggested by Peng et al. (2014) for *T. cruzi* and also others in different fields. For example, constitutive expression of Cas9 has been found to be toxic in *C. reinhardtii* (Jiang et al., 2014) and in human pluripotent stem cells (hPSC) (Ihry et al., 2018). To avoid toxicity of Cas9 gene editing, alternative more labor intensive systems have been developed, allowing for example transient transfections of ribonucleoprotein complexes, composed of recombinant Cas9 and *in vitro* transcribed sgRNAs (Soares Medeiros et al., 2017). However, data shown here shows only a minor growth defect for *Leishmania* Cas9 and T7 RNAP expressing cells. This defect might be negligible giving the opportunities that result from the constitutive expression of both proteins. A similar compromise was made previously for T7 RNAP expression in *T. brucei* (Poon et al., 2012) and has enabled large functional RNAi screens that allowed studying many fundamental cellular processes (Alsford et al., 2011). Additionally, using the LeishGEdit system in *T. cruzi* showed that expressing Cas9 from the tubulin array did not result in growth arrest compared to wild type parasites (Costa et al., 2018). Other systems have also not encountered growth defects in Cas9 expressing *T. cruzi* parasites (Lander et al., 2015). Consequently, the observed fitness phenotype in Peng et al. (2014) for constitutive expression of Cas9 may be well caused by the chosen type of

expression, in this case from a plasmid that was under the control of the strong rRNA promoter.

2.3.2 Possible applications for LeishGEdit

Expressing Cas9 and T7 RNAP constitutively in the LeishGEdit system enables key advantages, such as no requirement for any cloning procedures, PCR purifications or *in vitro* transcription prior to transfections. As shown for *ΔPF16* deletions, gene editing with the LeishGEdit toolbox is precise and cost effective, as only six unique primers and one universal primer are needed to generate N- and C-terminal tagged protein versions, as well as deleting the ORF.

Consequently, the main application of LeishGEdit include, introduction of epitope or protein tags, gene deletion on a small (single transfections) or large (96 well plate transfections) scale and barcoding of mutants for bar-seq assays. Since LeishGEdit can produce mutants on a large scale, this toolbox might have the potential to enable genome-wide knockout screens in *Leishmania* (discussed in section 6.2).

LeishGEdit can be also used for marker free gene editing, including marker free protein tagging and introductions of SNVs at a locus of interest. Additionally, this toolbox can be utilized for various kinetoplastids. In collaboration with Gerald Spaeth it has been used to study kinase CK1.1 in *L. donovani* (Martel et al., 2017), with Peter Bütikofer it has been used for studying cardiolipin depletion in *T. brucei* (Schadeli et al., 2019) and with John Kelly and Martin Taylor LeishGEdit was adapted for use in *T. cruzi* (Costa et al., 2018).

However, the current state of the system does not include an inducible component, making it impossible to study the loss-of-function phenotype of essential genes

and also making it a requirement to complement mutants with addback copies to fully link a phenotype with its mutation. But using the destabilization domain as a inducible protein tag (Banaszynski et al., 2006) or combining LeishGEdit with the available DiCre system (Duncan et al., 2016) or Forced Plasmid Shuffle system (Murta et al., 2009) might be actually an attractive approach to study the loss-of-function phenotype of essential genes (further outline in section 6.2).

Overall, this toolbox is versatile and can be utilized for various kinetoplastids. It also was essential for many experiments in the following chapters and has enabled high-throughput manipulation of parasites, allowing to dissect flagellar function in *Leishmania*.

CHAPTER 3

A SYSTEMATIC ANALYSIS OF THE FLAGELLAR CURLING PHENOTYPE PROVIDES NEW INSIGHTS INTO THE REGULATION OF FLAGELLAR BENDING

3 A systematic analysis of the flagellar curling phenotype provides new insights into the regulation of flagellar bending

Mutants shown in this chapter were created by me, unless stated otherwise. Scanning electron microscopy was carried out by obtaining help from Errin Johnson and Raman Dhaliwal (Dunn School Bioimaging Facility). All data and figures presented in this chapter are wholly my own work unless stated otherwise.

3.1 Introduction

3.1.1 Chapter summary and aim

There are competing proposals and models regarding how flagellar bending is regulated and the periodic switch of dynein activity is controlled. The previously proposed central pair-spoke axis model predicts that the CPPs distribute signals via adjacent RSPs to control dynein activity (Omoto et al., 1999). The aim of the work shown in this chapter was to use the flagellar curling phenotype, which I discovered in a specific cohort of *L. mexicana* mutants in the current on-going CRISPR-Cas9 knockout screen in the Gluenz lab, to dissect the regulation of flagellar bending in the *Leishmania* flagellum.

Therefore, a set of 49 axonemal and extra-axonemal proteins was chosen to perform an in-depth analysis of the curling phenotype. This showed that especially IDA f/I1 and CPPs component null mutants exhibit increased curling rates, which is likely to reflect a defect in progression of the normal flagellar beat. When subjecting the null mutant of the CPP gene hydin, which displays the strongest

curling phenotype of all mutants, for co-deletions with all 49 chosen candidates, curls can be either resolved (e.g. when hydin is co-deleted with IDA f/I1, ODA and RSP) or remain persistent (when hydin is co-deleted with MIA, N-DRC, most of CSC and most of PFR components). Interestingly, deletion of PFR2 also resolves curling and flagellar curls can remain persistent or be resolved when hydin is co-deleted with other CPP genes.

Overall, these observations agree with the central pair-spoke axis model, pointing towards a role of CPP proteins in the regulation of flagellar bending. Additionally, these results propose a model in which the PFR supports flagellar bending by amplifying the principal bend.

3.1.2 Asymmetric distribution of dynein activity results in flagellar bending

Ciliary and flagellar bending is a result of microtubule doublet sliding caused by inner and outer dynein arm motor activities. Dyneins undergo a mechano-chemical cycle of pre- and post-power stroke conformational changes powered through the hydrolysis of ATP (King, 2018). Since doublet microtubules are connected by tektin protofilaments and are anchored to the basal body, conformational changes of dyneins cause the microtubules to slide against each other, resulting in flagellar bending (reviewed in Lindemann and Lesich (2010)). In *Chlamydomonas* and other ciliated organisms the flagellar bending direction is restricted to be perpendicular to the planar arrangement of the central pair microtubules (Carbajal-Gonzalez et al., 2013) and of the two-part 5-6 bridge between microtubule doublet 5 and 6 (Afzelius, 1959; Lin et al., 2012). In flagella of kinetoplastids, the PFR, which is fixed relative to the axoneme (Gull, 1999), might additionally stabilize and restrain the beating plane through connections with doublets 4, 5, 6 and 7 (Dawe

et al., 2005; Portman and Gull, 2010). The type of beat, either flagellar-type (symmetric) or ciliary-type (asymmetric), and waveform directionality defines the direction of movement. *Leishmania* and *Trypanosoma* can switch the waveform direction, meaning they switch between tip-to-base and base-to-tip waveforms, but tip-to-base is the predominant beat type in promastigote culture (Baron et al., 2007; Edwards et al., 2018; Gadelha et al., 2007; Wheeler, 2017). Consequently, the swimming directions of e.g. *Leishmania* (tip-to-base waveform) and human sperm (base-to-tip waveform) are typically reversed.

There are various hypotheses regarding how flagellar bending is regulated. What is generally believed is that dynein activity is periodically switched between both sides of the axoneme. This asymmetric distribution of dynein activity between microtubule doublets 1 to 4 and 6 to 9 results in unidirectional bending and was previously described as the “switch-point hypothesis” (Satir and Matsuoka, 1989; Wais-Steider and Satir, 1979). Recent results from Lin and Nicastro (2018) indicate that this oscillating conformational change of dyneins is regulated by inhibition. Lin and Nicastro (2018) proposed a “switch-inhibition model” in which dyneins are active by default, meaning they are in pre-power stroke conformations with a large angle between the dynein linker and stalk domain. In contrast, inhibited dyneins are in the post-power stroke conformation with a small angle between the dynein linker and stalk domain. Using cryo-ET reconstruction of sea urchin sperm and *C. reinhardtii* flagella all of the outer dynein arms and IDAs f/I1, a, d and g either on microtubules 2 to 4 or 7 to 9 were shown to be in inactive post-power stroke or intermediate conformations when flagella were in a bend position. However, dyneins on all other microtubules remained in their active pre-power stroke conformations (Figure 23 A, B and C).

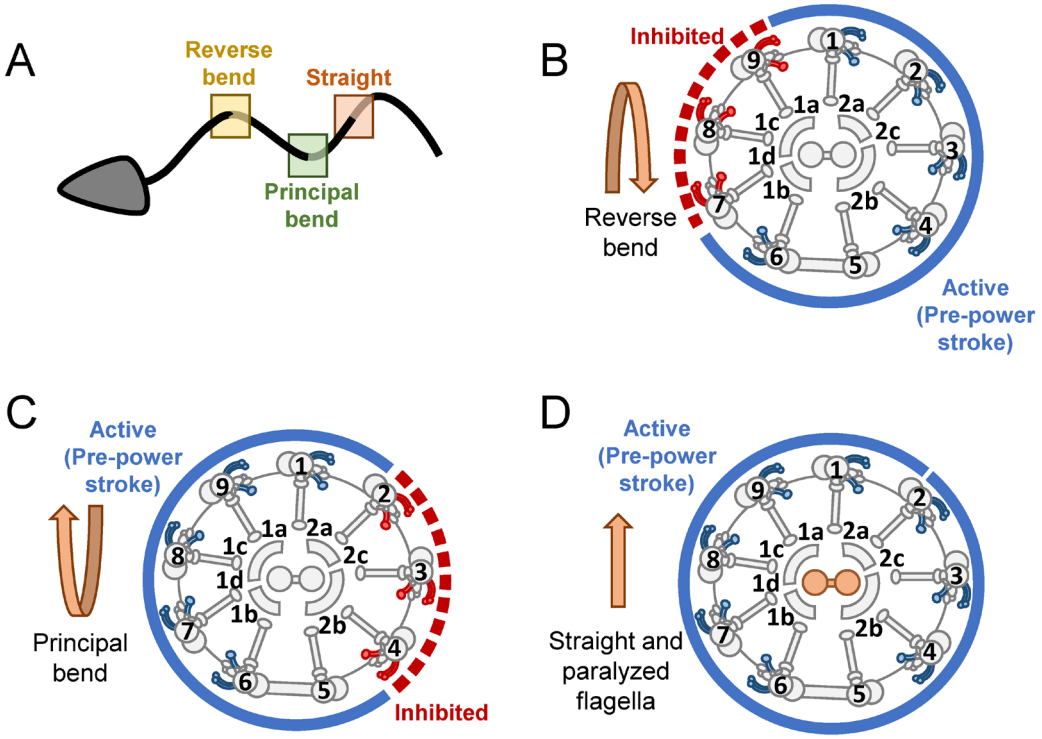


Figure 23. Dynein power stroke conformations in the switch inhibition model. Overview of dynein power stroke conformations in 9 + 2 motile flagella during flagellar bending with CPPs highlighted. Figure inspired by illustrations of King (2018) and Lin and Nicastro (2018). **(A)** Positions of reverse and principal bend, as well as of straight flagella, are indicated on the example of a sperm flagellum and are shown in transverse axoneme sections with dyneins being inhibited (red) or in pre-power active stroke conformations (blue). This conformation applies to ODAs and IDAs f/I1, a, d and g. **(B)** Reverse bend: dyneins on doublets 7 – 9 are inhibited and on doublets 1 – 6 active. **(C)** Principal bend: dyneins on doublets 2 – 4 are inhibited and on doublets 5 – 1 active. **(D)** Straight flagella: all dyneins are active. This applies also to paralyzed central pair (highlighted in orange) mutant flagella.

Post- and pre-power conformations switched between principal and reverse bend, resulting in dynein activity switching and consequently in a flagellar beat (Figure 23 A, B and C). Interestingly, dyneins in straight flagella and dyneins in paralyzed *pf19* mutant flagella, in which the CPC is not assembled and absent, are generally found to be in the pre-power stroke conformation, meaning they are active and the flagellum is straight because of balanced dynein forces (Lin and Nicastro, 2018) (Figure 23 A and D). Consequently, regulators of dynein activity and flagellar bending are predicted to act as an “off-switch” rather than “on-switch”.

3.1.3 Regulation of dynein activity and flagellar bending

It remains unknown how this proposed side-selective inhibition of dynein motors is regulated. There are several hypotheses that suggest how flagellar bending may be regulated (Figure 24). While some models seem purely mechanistic, others rely on signal transduction pathways to be functional.

Flagellar bend regulation by the curvature control model is a classic example for mechanistic modelling of dynein activity regulation (Brokaw, 1971; Brokaw, 1972a; Brokaw, 1972b). In this model dyneins are regulated by the degree of bending, meaning a sufficient curvature inactivates dyneins on one side and activates them on the opposite side. Although this model doesn’t fully account for some flagellar behaviors, such as flagellar beat initiation (Brokaw, 1985), recent evidence for this model was found when analyzing wild type asymmetric beating flagella and *mbo2* (move backwards only) symmetric beating mutant flagella in *Chlamydomonas* (Sartori et al., 2016). Both types of waveform fit this model better than other mechanistic curvature models, such as the geometric clutch model.

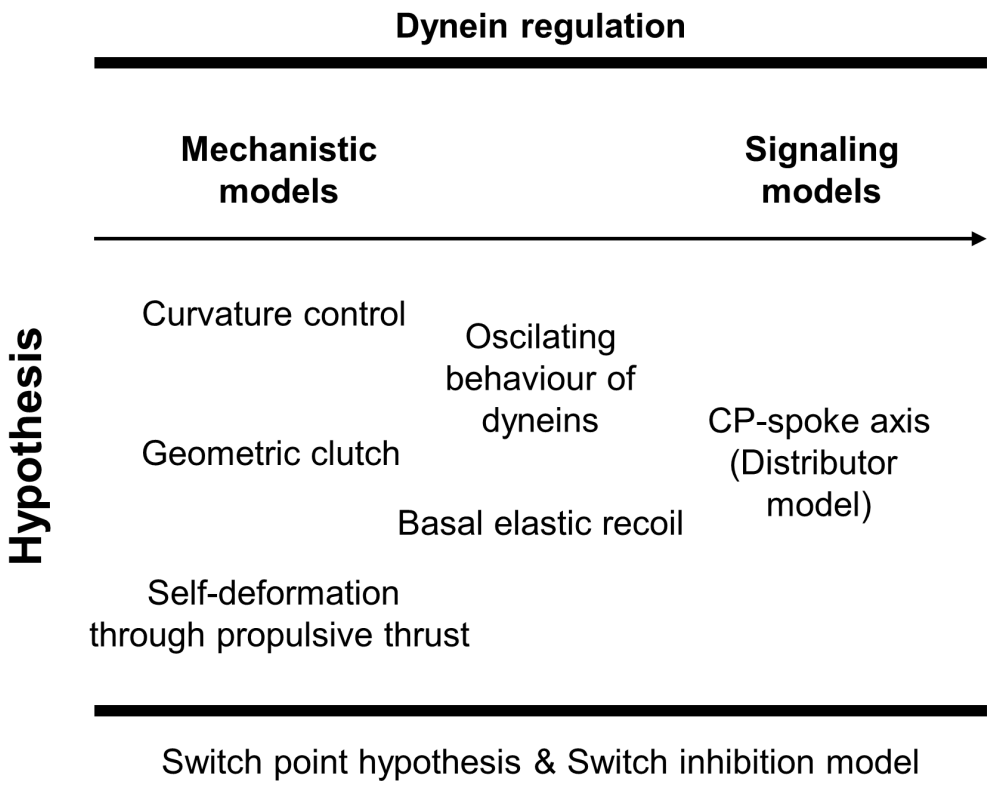


Figure 24. Current models for dynein regulation.

Overview of current dynein regulation models separated into mechanistic and signaling models. If a mechanistic model shares properties of signaling models it is shifted more towards signaling models. Meanwhile, what is generally believed is that dynein regulation follows the switch point hypothesis and switch inhibition model. See main text for reference and description of each model.

The geometric clutch model assumes that sliding of microtubule doublets generates a longitudinal force, which results in tension and thus pulls neighboring doublets apart. The resulting space between doublets detaches dyneins from their adjacent microtubules, while at the same time dyneins on the opposite site attach and initiate a new bend (Lindemann, 1994). Thus, it relies on static curvatures (Bayly and Wilson, 2015), which vary greatly between the *mbo2* mutant and wild type axonemes (Sartori et al., 2016).

Although there are several lines of evidence that flagellar bending follows mechanistic curvature models (Lindemann and Lesich, 2010), both models, curvature control and geometric clutch, do not completely explain meander waves (a series of regular sinuous curves or bends) that can be observed in the ~200 µm long 9 + 2 flagellum of quail spermatozoon. Thus, it has been proposed that the initiated bend deforms the flagellum by a propulsive thrust, initiated at the base of the flagellum. This force drives the flagellum into up to 12 meander waves (Woolley, 2007; Woolley, 2010) (propulsive thrust model).

The curvature control, the geometric clutch and the propulsive thrust model require all constitutively active dyneins, which would fit with the proposed model from Lin and Nicastro (2018). This is in contrast to models that suggest an oscillating behavior of dyneins, meanining each dynein undergoes by default and without the need of further regulation a specific cycle of attachment, power stroke and detachment. This model is also referred to as dynein cross-bridge cycle (Sugino and Naitoh, 1982). However, an oscillating behavior of dyneins would not fit with the observations that dyneins in straight and paralyzed flagella are by default in the active pre-power stroke conformation. Other models for flagellar bending regulation, such as the basal elastic recoil model, which proposes that the flagellar

beat is controlled at the base of the flagellum (reviewed in Wolley 2010), would also not follow the observed conformational changes of dyneins during principal and reverse bend of flagella.

Another model that aligns with current findings on power stroke conformational changes is the central pair-spoke axis or also named distributor model. It postulates that dynein arms are activated by their adjacent radial spokes which in-turn receive signals from CPPs (Omoto et al., 1999). There are several lines of evidence that RSPs (Barber et al., 2012; Curry et al., 1992; Diener et al., 1993; Ralston et al., 2006; Williams et al., 1989; Yang et al., 2006; Yang et al., 2004) and the CPC (Adams et al., 1981; Dawe et al., 2007; Dutcher et al., 1984; Lechtreck and Witman, 2007; Oda et al., 2014) influence activity of dyneins through a series of enzymatic and mechanical interactions. An early mechanistic proposal of this model, assumed that the CPC controls these interactions through rotation of the central pair microtubules (Omoto and Kung, 1980). However, in *Leishmania* and *Trypanosoma* the central pair microtubules do not rotate (Gadelha et al., 2006) and in *Chlamydomonas* the rotation of the CPC follows the flagellar beat rather than determines its direction (Mitchell and Nakatsugawa, 2004). It is much more likely that RSPs and CPPs coordinate sliding through spoke tilting, meaning the angle in which both complexes are connected is changed (Smith and Yang, 2004). Alternatively, dyneins might be inhibited when RSP heads and CPPs are pressed together through bending stress (Lindemann and Lesich, 2010).

Thus, these central pair-spoke axis regulation model versions are of a mechanistic nature. There is, however, also significant evidence for signaling via the CPPs, RSPs and other components that link central pair microtubules with inner and outer dynein arms. Calcium signaling through several RSP proteins (Smith, 2002; Yang

et al., 2006) and phosphorylation through IC138 (Hendrickson et al., 2004; VanderWaal et al., 2011) are only two examples of signal cascades that have been shown to control dynein activity (further discussed below and in chapter 4).

However, central pair-spoke axis models do not work for motile flagella with a missing central pair, such as 9 + 0 sperm flagella of eel (Woolley, 1997) or 3 + 0 flagella of male gametes of *Diplauxis hatti* (Prensier et al., 1980). Additionally, signaling and mechanistic central pair-spoke axis models are not supported by data from some suppressor mutations. For example, absence of N-DRC 3, 4, 5, 6 and 7 in *C. reinhardtii pf2* (paralyzed flagellar) mutants were shown to suppress the phenotype of paralyzed RSP and CPC mutants, indicating that another control mechanism can take over dynein regulation in the absence of central pair-spoke axis regulation (Heuser et al., 2009; Huang et al., 1982).

3.1.4 Extra-axonemal structures and their function in flagellar bending

It is also not understood to what extent extra-axonemal structures contribute to flagellar bending because the function of extra-axonemal structures is largely unknown. Several loss-of-function and biochemical studies have indicated they may have a physical, metabolic, homeostatic, regulatory and sensory function (reviewed in Oberholzer et al. (2007a) and Portman and Gull (2010)). Consequently, both, a signaling or a mechanical contribution to flagellar bending, is possible.

Organisms in major eukaryotic groups, including Metazoa and Excavata, have been shown to assemble extra-axonemal structures. However, organisms of the kingdom Fungi (Dick, 1997) and Plantae (Hodges et al., 2012) seem to miss these if they possess a cilium or flagellum at all. Extra-axonemal structures can be

arranged symmetrically or asymmetrically relative to the axoneme. For example mammalian sperm tails have a symmetrically arranged mitochondrial sheath, extra-axonemal outer dense fibers and a fibrous sheat (reviewed in Lehti and Sironen (2017)). However, organisms in the order Diptera, such as *Hermetia illucens* (black soldier fly), have asymmetrically arranged mitochondrial derivatives within their sperm tails (Kotze et al., 2019). Furthermore, some insect sperm tails show also additional outer microtubule doublets, which is referred to as 9 + 9 + 2 axoneme (Dallai et al., 2006). Also organisms in the class of *Kinetoplastida* and *Euglenida*, including *Leishmania*, possess an asymmetrical extra-axonemal structure, the PFR.

The PFR is a lattice-like structure, which is variable in size between species and in some organism significantly reduced in its dimensions. At least 40 proteins have been associated with the PFR (reviewed in Portman and Gull (2010)). PFR1 and PFR2 are the two main PFR components and are required for PFR assembly (Maga et al., 1999). A null mutant of PFR2 in *L. mexicana* (Santrich et al., 1997) and a PFR2 RNAi knockdown in *T. brucei* (Bastin et al., 1998) showed a decrease in flagellar wave frequency and beat amplitude, showing that the PFR is required for normal motility. The PFR has been proposed to act as a thickening or stiffening agent during flagellar bending (Fuge, 1969), but also to fulfil a regulatory function via signaling pathways, as proteins with a role in calcium signaling and adenine nucleotide homeostasis have been linked to the PFR (reviewed in Portman and Gull (2010)).

Overall, a model for flagellar bend regulation remains to be defined that can account for flagella with: (1) extra-axonemal structures that run either along the full length of the flagella or only along a portion, (2) flagella with extra-axonemal

structures that are either symmetrical or asymmetrical arranged, (3) flagella that miss the central pair and are still motile and (4) “normal” motile flagella that have a typical 9 + 2 microtubule arrangement without extra-axonemal structures.

3.1.5 Functional analysis of flagellar proteins identifies the flagellar curling phenotype

Loss-of-function approaches have been essential to dissect flagellar function and in particular to analyse the regulation of flagellar bending in many model organisms, including most prominently the green alga *Chlamydomonas*, but also the nematode *Caenorhabditis elegans*, ciliated organisms such as *Paramecium* and *Tetrahymena*, kinetoplastids such as *Trypanosoma* parasites and vertebrate models such as mouse and zebrafish, as well as non-vertebrates such as *Drosophila* (reviewed in (Vincensini et al., 2011)). *Leishmania* parasites are among these model organisms, but have been used far less to study broader flagellar biology than other organisms. However, since scalable gene editing approaches, such as LeishGEdit (see chapter 2), are now available for a functional dissection of flagellar protein coding genes in *Leishmania*, this is likely to change, as further discussed in section 6.3.

To-date, the LeishGEdit toolbox has been used in the Gluenz lab to generate more than 450 flagellar mutants (as of 10th of June 2019: 451 mutants; unpublished) and analyse their phenotype by various approaches, including motility analysis, *in vitro* bar-seq fitness screens and flagellar beat analysis. This has been especially useful to achieve my thesis aims, in particular to dissect broader aspects of flagellar function, such as the regulation of flagellar bending (see section 1.2.4). The first 100 mutants of this knockout screen were recently published (Beneke et al., 2019b) and showed several examples of various flagellar phenotypes. Mutants

were created in a combined lab effort by many members in the Gluenz lab and appeared either with no obvious phenotype or were found to be aflagellated, paralyzed, uncoordinated, slow or fast when measuring their swimming behaviour (Figure 25). Overall, 56 mutants of all 100 analysed mutants showed a motility phenotype, meaning their swimming speed was altered.

Among all 451 generated flagellar mutants, I found that the mutant of the CPC protein coding gene hydin, and subsequently also other mutants, exhibited a flagellar curling phenotype upon deletion (Figure 25). Since we proposed that failed flagellar beat propagation was seen as a possible explanation for flagellar curling, the curling phenotype was explored as an interesting opportunity to study the regulation of flagellar bending in *Leishmania*.

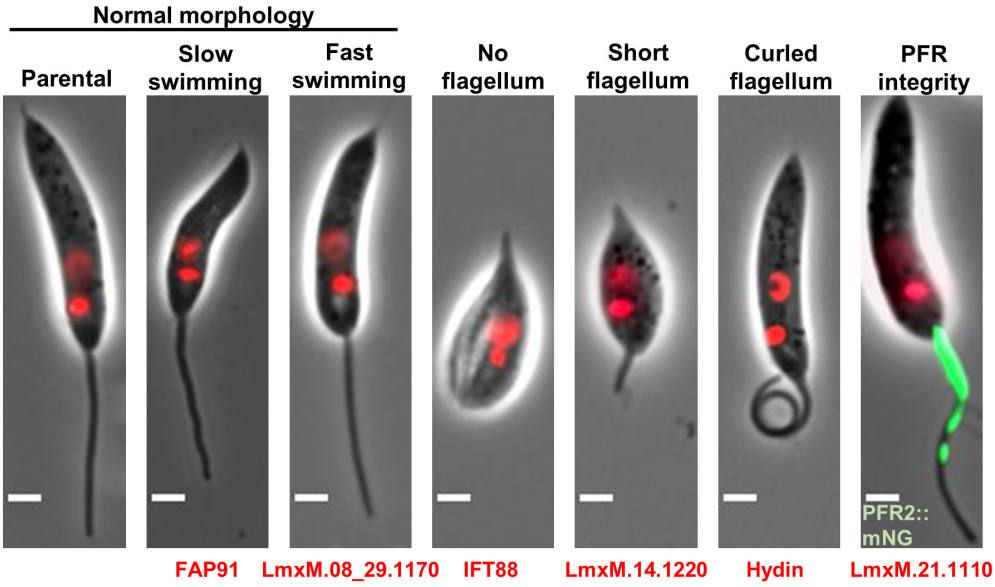


Figure 25. Classes of mutant phenotypes identified in a functional knockout screen.

Figure made by Tom Beneke, taken from Beneke et al. (2019b). Micrographs show merged fluorescent and phase channels (Hoechst DNA stain in red, phase in grey and for LmxM.21.1110 PFR2::mNG in green). Each cell is representative for an identified mutant category, including: slow swimming or fast swimming mutants with normal morphology and mutants with abnormal morphology as indicated. Cell line shown in micrograph indicated in red below. Scale bar is 2 μm.

3.2 Results

3.2.1 Discovery of mutants with curled flagella

The flagellar curling phenotype was first observed when imaging *L. mexicana Δhydin* mutants (Beneke et al., 2019b) alongside their Cas9 T7 parental cell line. By counting the flagellar curling rate across three micrographs and 187 cells, the rate of *Δhydin* mutants was found to be 60.9% (±4.8%) compared to 1.5% (±0.4%) in the Cas9 T7 parental cell line. Flagellar curls were counted using conservative criteria. A flagellar curl was defined as a flagellum that makes at least one fully closed circle within its curl (Figure 26 B). Other flagellar conformations were not counted as flagellar curl, these included e.g. hook-like conformations (Figure 26 B (i)) or flagella that were in extreme bend positions causing the flagellum to be in parallel to its cell body (Figure 26 B (ii)). Using scanning electron microscopy (SEM) to image also glutaraldehyde fixed *Δhydin* mutants (Figure 26 C and D) showed that curls were not introduced by microscopy sample preparation methods used for imaging live cells.

Next, all flagellar mutants produced to-date (as of 10th of June 2019: 451 mutants; unpublished) were examined under the same conditions to search for additional flagellar curling phenotypes. Interestingly, only 14 null mutants showed more than 15% flagellar curling within their population. Three of these mutants were deficient for expression of a central pair complex protein (hydin, PF16 or C1a-34), six for IDA components (IC138, IC140, CFAP43, CFAP44, DHC 1α and DHC 1β) and five had a yet unknown localisation within the *Leishmania* axoneme (MBO2, LmxM.33.0560, PACRGA, CMF15 and LAX28), although data presented in chapter 4 suggests that LAX28 is part of the IDA f/I1 complex.

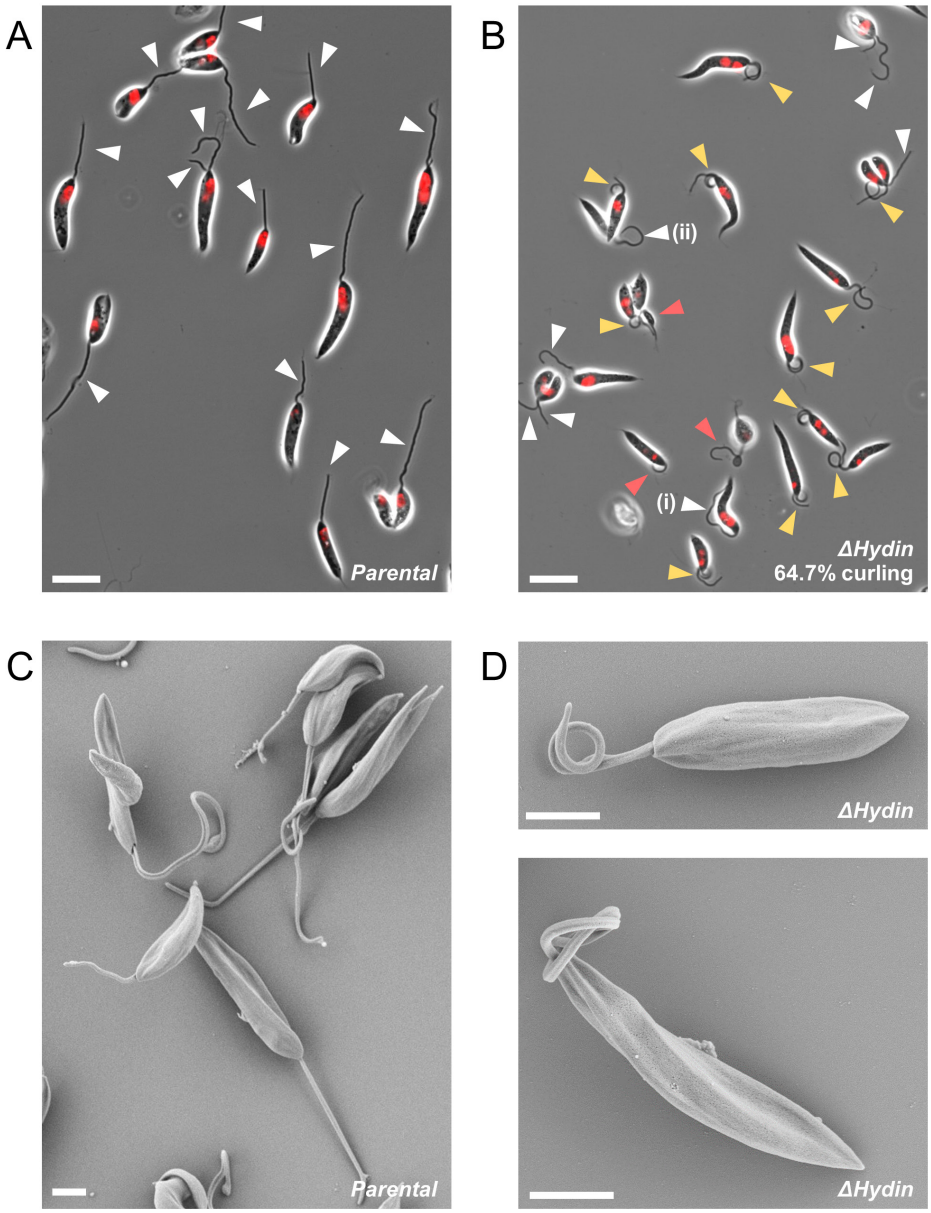


Figure 26. Criteria for counting flagellar curling.

Flagellar curling has been assessed in parental cells (**(A)** and **(C)**) and *Δhydin* deletions (**(B)** and **(D)**). **(A)** and **(B)** Micrographs show live cells with merged phase (grey) and fluorescent channel (Hoechst DNA stain in red). The rate of flagella curling is determined by counting: non-curling flagella (white arrows), curling flagella (yellow arrows) and flagella not counted (red arrows). Only curls are counted that make at least one fully closed circle within its curl, thus excluding **(i)** hook-like flagellar conformations and **(ii)** flagella in extreme bend positions. Scale bar is 10 µm. **(C)** and **(D)** SEM images of fixed cells. Scale bar is 2 µm.

3.2.2 A subset of 49 flagellar mutants for the flagellar curling analysis

It was initially unclear what causes flagellar curling, whether there are any particular axonemal structures that are involved in curling or not involved in curling and whether there are any secondary mutations that could release existing curls. The initial hypothesis was that flagellar curling might reflect a regulatory defect where dyneins are stuck in a certain part of the beat cycle and stop the flagellar beat propagation, causing an extreme form of flagellar bending. It was also unclear why in previous studies for *Δhydin* mutants in other organisms, such as *C. reinhardtii* (Lechtreck and Witman, 2007), a curling phenotype was not reported. This allowed us to speculate that the PFR could play a role in curling. Consequently, flagellar curling was seen as an opportunity to study the regulation of flagellar bending in *Leishmania* and also more broadly.

To start a systematic dissection of the curling phenotype the analysis was focused on a subset of proteins. To compile this protein list 34 well characterized axonemal proteins, consisting of CPC, CSC, N-DRC, MIA, IDA, ODA and RSP components (Wheeler et al., 2015a; Wickstead and Gull, 2007), were chosen. Since these proteins are best characterized in the cilia model organism *Chlamydomonas*, sequences from *C. reinhardtii* were used to identify the *Leishmania* orthologues by reciprocal best protein blast search (States and Gish, 1994) (Table 1). This showed that seven genes among all 34 searched genes, did not yield a reciprocal best protein blast search (C1a-86, LC4-like, RSP5, RSP7 and RSP11-LmxM.09) or did not identify any high-confidence *L. mexicana* orthologues (C1a-18 and LC7) (Table 1). Nevertheless, these genes were still included in the analysis. To also cover extra-axonemal structures in the analysis, 13 potential PFR components, identified in *T. brucei* and derived from Portman et al. (2009), were also chosen. Additionally,

two proteins, MBO2 and LmxM.21.1110, both shown to influence the integrity of the PFR upon deletion (Beneke et al., 2019b) but unknown in their flagellar localization, were also added to the list.

Thus, the final list consisted of a total of 49 flagellar protein coding genes (Table 2 and Figure 27). Null mutants of these genes were taken from the current CRISPR-Cas9 knockout screen in the Gluenz lab and used to do an in-depth analysis of the flagellar curling phenotype.

3.2.3 Flagellar curling restricts motility

To characterize the flagellar curling phenotype with these mutants in detail, at first the motility phenotype was analyzed to see whether there is any correlation between curling rate, swimming speed and directionality. All 49 mutants (Table 2 and Figure 27) were subjected to particle tracking analysis (Wheeler, 2017). Mutant speed and directionality were found both to decrease with increasing flagellar curling (Figure 28 A and B), indicating that a higher proportion of curling might restrict motility. Additionally, motility analysis showed that the majority of curling mutants exhibited rather slow swimming motility phenotypes (Figure 28 C). However, three mutants with curling defect (LmxM.33.0560, PACRGA and CMF15) showed slow to normal swimming behaviours.

Gene name	NCBI sequence ID *Chlamydomonas reinhardtii*	GeneID *Leishmania mexicana*	Structure	BLASTp e-value against *C. reinhardtii*	BLASTp e-value against *L. mexicana*	Reciprocal best blast (RBB)?
Hydin	XP_001689997.1; PNW88344.1	LmxM.29.1810	CPC	3.00E-108	5e-117; 5e-106	Yes
PF16	AAC49169.1; XP_001694680.1; PNW78912.1	LmxM.20.1400		0	0; 0; 0	Yes
PF20	P93107.1	LmxM.18.0470		4.00E-115	3.00E-103	Yes
IP4/C1a-18	AAZ31184.1	LmxM.10.0195		1E-26 (7 others >1E-10)	3.00E-25	Yes
IP2/C1a-34	XP_001696622.1	LmxM.36.1040		2.00E-26	6.00E-26	Yes
IP1/FAP101/ C1a-86	AAZ31187.1	LmxM.21.0180		1e-08 (6 others >1E-04)	not found	No
CaM-IP4/FAP215	PNW85421.1; XP_001691834.1	LmxM.25.1600	CSC	1E-117; 2e-67	3E-117; 1e-69	Yes
CaM-IP3 /FAP61/RSP19	XP_001703513.1	LmxM.26.0650		2.00E-18	9.00E-18	Yes
CaM-IP2 /FAP91/RSP18	XP_001690436.1; PNW80814.1	LmxM.32.2220		8.00E-22	7e-15; 7e-15	Yes
DRC7/FAP50	PNW72944.1; XP_001699034.1	LmxM.05.1020	N-DRC	5E-36; 1e-31	6e-30; 9e-30	Yes
DRC5/FAP155	PNW88806.1; XP_001689724.1	LmxM.08_29.2810		2.00E-20	9e-22; 1e-21	Yes
DRC2/FAP250/ CCDC65	XP_001699103.1; A8JB22.2	LmxM.28.0050		2e-47; 5e-47	1e-47; 8e-47	Yes
DRC3/FAP134	XP_001693032.1	LmxM.29.3640		2.00E-34	1.00E-31	Yes
DRC4/pf2/GAS8/ GAS11/trypanin	Q7XJ96.1; XP_001697295.1	LmxM.34.1810		1E-92; 4e-81	1e-90; 1e-79	Yes
FAP73/MIA2p	M1V4Y8.1	LmxM.13.0240	MIA	5.00E-29	7.00E-30	Yes
FAP100/MIA1p	XP_001699777.1	LmxM.31.1980		6.00E-21	7.00E-21	Yes
IC138/BOP5	XP_001696921.1	LmxM.32.2630	IDA	4.00E-53	1.00E-53	Yes
IC140/ida7	XP_001695786.1	LmxM.27.1630		2.00E-65	4.00E-61	Yes
FAP44 (IDA-T/TH)	PNW78215.1; XP_001695270.1	LmxM.14.1430		2e-79; 8e-79	3e-77; 4e-77	Yes
DHC 1α /PF9/DNAH10	XP_001703170.1	LmxM.33.3880		0.0 (Score 3703; Ident 44%)	0.0 (Score 3681; query cover 95%; Ident 44.45%)	Yes
DHC 1β	Q9MBF8.1	LmxM.23.1310		0.0 (Score 3132; Ident 40%)	0.0 (Score 3126; query cover 95%; Ident 40.27%)	Yes
β/βDHC/ODA4	XP_001695126.1	LmxM.13.1650	ODA	0.0 (Score 2897; Ident 39%)	0.0 (query cover 97%; Ident 39.12%)	Yes
alpha ODA	dynein gamma chain Q39575.1	LmxM.25.0980		0.0 (Score 3526; Ident 43%)	0.0 (Score 3496; query cover 96%; Ident 42.48%)	Yes
LC4-like	XP_001690249.1	LmxM.01.0620		4e-10;	4.00E-09	No
DC2/ODA1	AAK72125.1; XP_001701436.1	LmxM.31.2900		5E-34; 6e-34	4e-32; 5e-32	Yes
LC7/ODA15	XP_001694381.1; XP_001694098.1	LmxM.18.1010		3E-23; 4e-19 (LmxM.34.1750 3e-22 and 1e-29)	1e-23; 2e-19	Yes
RSP2/pf24	XP_001702718.1	LmxM.08.0960	RSP	2.00E-06	1.00E-05	Yes
RSP3/AKAP97	XP_001695406.1	LmxM.27.0520		2.00E-68	5.00E-65	Yes
RSP4/6	XP_001700728.1; XP_001700729.1	LmxM.13.0430		5E-23; 2e-21	1e-21; 9e-19	Yes
RSP5	XP_001694084.1	LmxM.21.0170		0.99	not found	No
RSP7	ABC02026.1	LmxM.32.1710		0.007	not found	No
RSP9	XP_001690441.1	LmxM.32.2500		3.00E-13	2.00E-10	Yes
RSP11-LmxM.34	PNW81221.1; XP_001698630.1	LmxM.34.4060		2e-10; 5e-07	4e-10; 2e-06	Yes
RSP11-LmxM.09	XP_001698630.1	LmxM.09.1530		2e-07	7.00E-07	No

Table legend on following page.

Table 1. Identification of *L. mexicana* orthologues of well-characterised axonemal proteins.

Overview of reciprocal best protein blast results between well characterized axonemal *C. reinhardtii* genes and their putative *L. mexicana* orthologue. Gene names are derived from *C. reinhardtii*, but their *T. brucei* name may also be given if applicable. E-values of BLASTp were recorded as described in 7.7.4. For some genes two *C. reinhardtii* sequence IDs were found. For these two e-values are given.

Gene name	GeneID *Leishmania mexicana*	Structure	Mutant produced by	Mutant published in Beneke et al., Plos Pathogens 2019?
Hydin	LmxM.29.1810	CPC	Tom Beneke	YES
PF16	LmxM.20.1400		Tom Beneke	YES
PF20	LmxM.18.0470		James Smith	NO
IP4/C1a-18	LmxM.10.0195		James Smith	NO
IP2/C1a-34	LmxM.36.1040		James Smith	NO
IP1/FAP101/C1a-86	LmxM.21.0180		James Smith	NO
CaM-IP4/FAP215	LmxM.25.1600	CSC	James Smith	YES
CaM-IP3/FAP61/RSP19	LmxM.26.0650		James Smith	YES
CaM-IP2/FAP91/RSP18	LmxM.32.2220		James Smith	YES
DRC7/FAP50	LmxM.05.1020	N-DRC	James Smith	YES
DRC5/FAP155	LmxM.08_29.2810		James Smith	NO
DRC2/FAP250/CCDC65	LmxM.28.0050		James Smith	YES
DRC3/FAP134	LmxM.29.3640		James Smith	YES
DRC4/pf2/GAS8/ GAS11/trypanin	LmxM.34.1810		James Smith	YES
FAP73/MIA2p	LmxM.13.0240	MIA	James Smith	YES
FAP100/MIA1p	LmxM.31.1980		James Smith	YES
IC138/BOP5	LmxM.32.2630	IDA	James Smith	NO
IC140/ida7	LmxM.27.1630		James Smith	YES
FAP44 (IDA-T/TH)	LmxM.14.1430		Tom Beneke	YES
DHC 1α/PF9/DNAH10	LmxM.33.3880		Tom Beneke	NO
DHC 1β	LmxM.23.1310		Tom Beneke	NO
β/βDHC/ODA4	LmxM.13.1650	ODA	Tom Beneke	NO
α/αDHC	LmxM.25.0980		Tom Beneke	NO
LC4-like	LmxM.01.0620		Heather Jeffery	YES
DC2/ODA1	LmxM.31.2900		James Smith	YES
LC7/ODA15	LmxM.18.1010		James Smith	NO
PFC1	LmxM.24.1560	PFR	James Smith	NO
PFC2	LmxM.10.0180		James Smith	NO
PFC4	LmxM.29.2850		James Smith	NO
PFC5	LmxM.27.1850		James Smith	NO
PFC8	LmxM.29.2390		James Smith	NO
PFC10	LmxM.32.2780		James Smith	NO
PFC15	LmxM.19.0520		James Smith	YES
PFC16	LmxM.32.0610		James Smith	NO
Tb5.20	LmxM.31.1910		James Smith	NO
ADKB	LmxM.21.1250		James Smith	NO
PI3K	LmxM.31.1460		James Smith	NO
Myosin XXI	LmxM.31.3870		James Smith	NO
PFR2	LmxM.16.1430 (array of 3 genes)		Tom Beneke	YES
407 (LmxM.21.1110)	LmxM.21.1110	Unknown	Tom Beneke	YES
410 (MBO2, CMF8)	LmxM.33.2480		Tom Beneke	YES
RSP2/pf24	LmxM.08.0960	RSP	James Smith	NO
RSP3/AKAP97/pf14	LmxM.27.0520		James Smith	NO
RSP4/6	LmxM.13.0430		James Smith	NO
RSP5	LmxM.21.0170		James Smith	NO
RSP7	LmxM.32.1710		James Smith	NO
RSP9	LmxM.32.2500		James Smith	NO
RSP11-LmxM.34	LmxM.34.4060		James Smith	NO
RSP11-LmxM.09	LmxM.09.1530		James Smith	NO

Table 2. A subset of 49 flagellar mutants for the flagellar curling analysis. Overview of all chosen proteins for flagellar curling analysis. The researcher who produced the mutant is indicated.

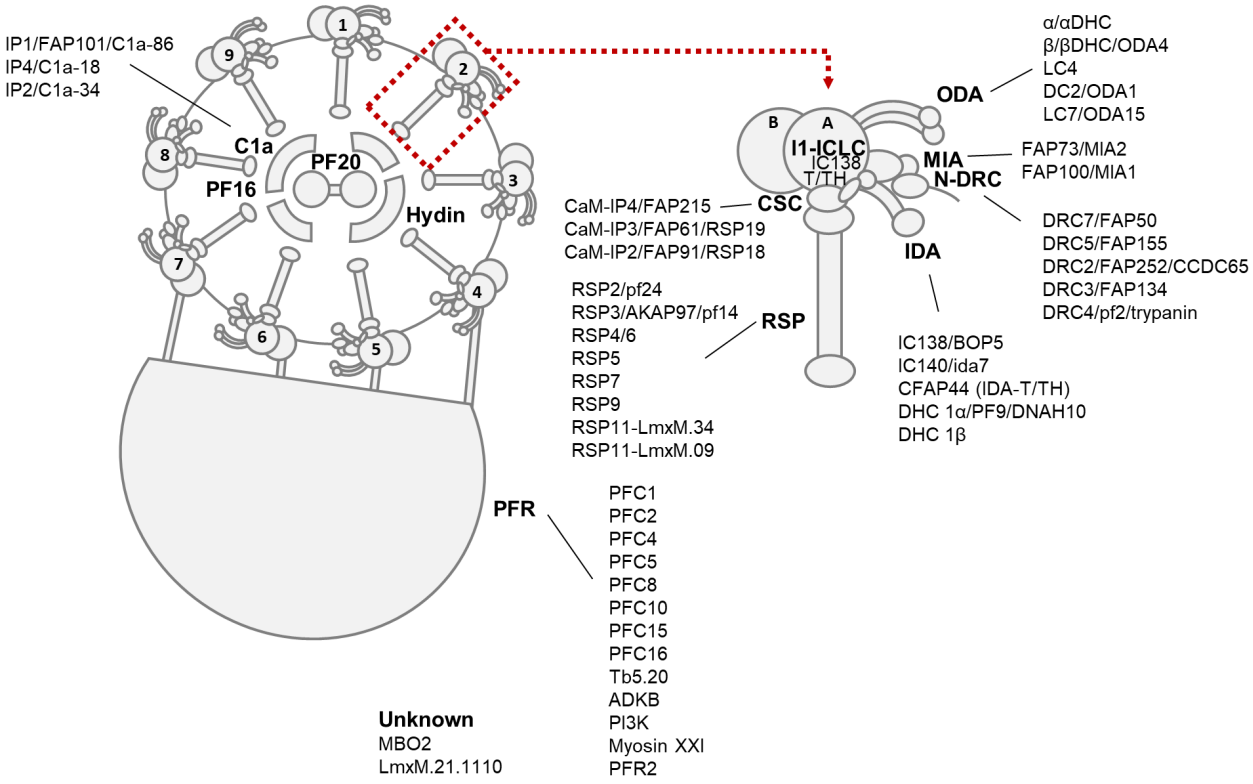


Figure 27. A subset of 49 flagellar mutants for the flagellar curling analysis.

Schematic cross-sectional view on *Leishmania* promastigote 9 + 2 axoneme with all chosen proteins for flagellar curling analysis.

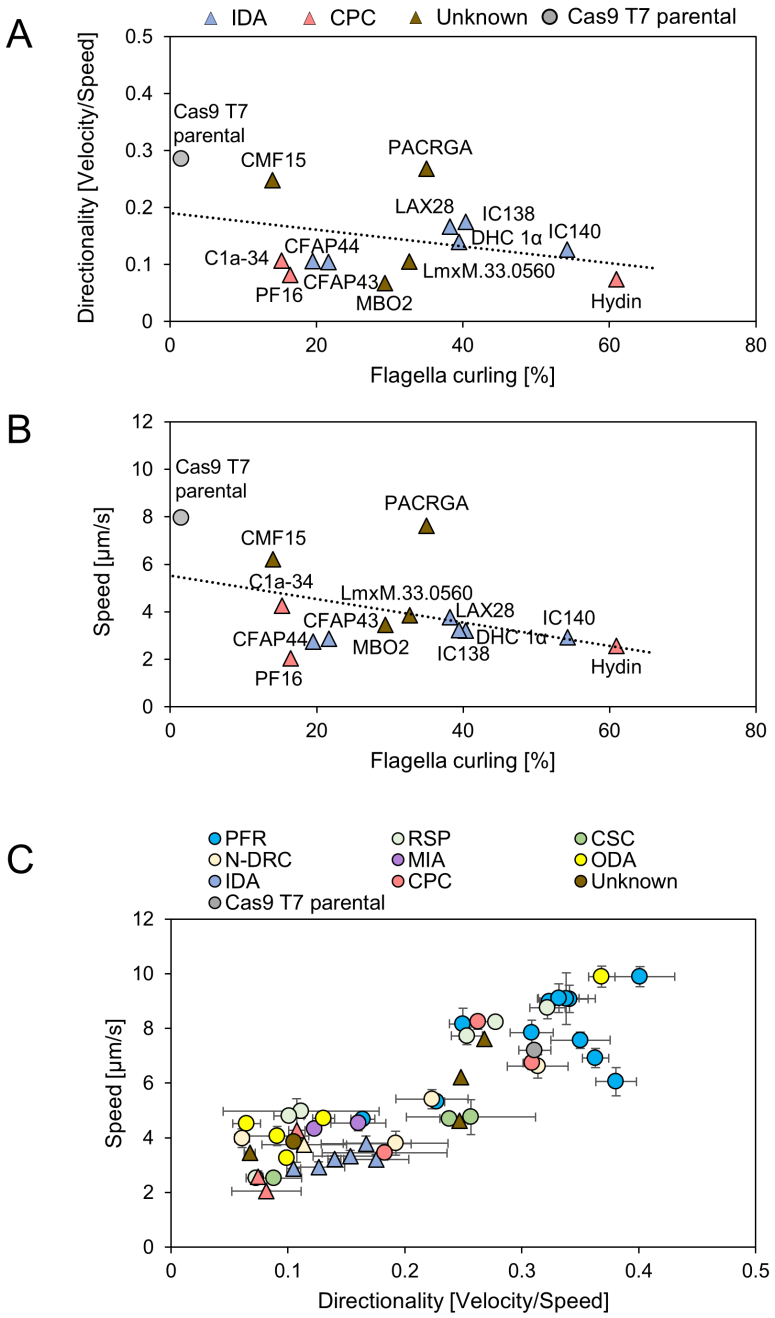


Figure 28. Motility analysis of mutants with curled flagella.

Triangles indicate mutants with curled flagella, circles are mutants with no curled flagella. Flagellar curling rates of mutants with >15% curling were plotted on the X axis against measured **(A)** average swimming speed [µm/s] and **(B)** directionality [velocity/speed]. The trendline for each data series is shown. **(C)** Speed and directionality was also plotted for the chosen subset of all 49 null mutants. Mutants were colour coded according to their axonemal structure.

Interestingly, although the majority of paralyzed mutants that still have a flagellum do show the curling phenotype not all paralyzed mutants have a curled flagellum, as e.g. paralyzed *ΔFAP91* mutants (FAP91 is a CSC component; Table 2, Figure 27) with normal flagellar length do not exhibit the curling phenotype (Figure 28 C). Motility analysis showed also that some mutants with curled flagella, such as *ΔMBO2* and *ΔC1a-34* mutants, can be also uncoordinated in their movement, meaning they have very little directionality, while still showing slow swimming speed. This was also found to be the case for ODA and MIA mutants. However, for N-DRC, RSP, CPC, CSC and PFR mutants the motility phenotype seemed to be more variable. IDA mutants tend also to form tight clusters when plotting swimming speed against directionality and exhibited slow swimming behaviours (Figure 28 C). Interestingly, CPC deletions *ΔPF16* and *Δhydin* were both paralyzed and clustered tightly together in the motility plot, even though *ΔPF16* mutants showed four times less curling than *Δhydin* mutants (Figure 28 A, B and C).

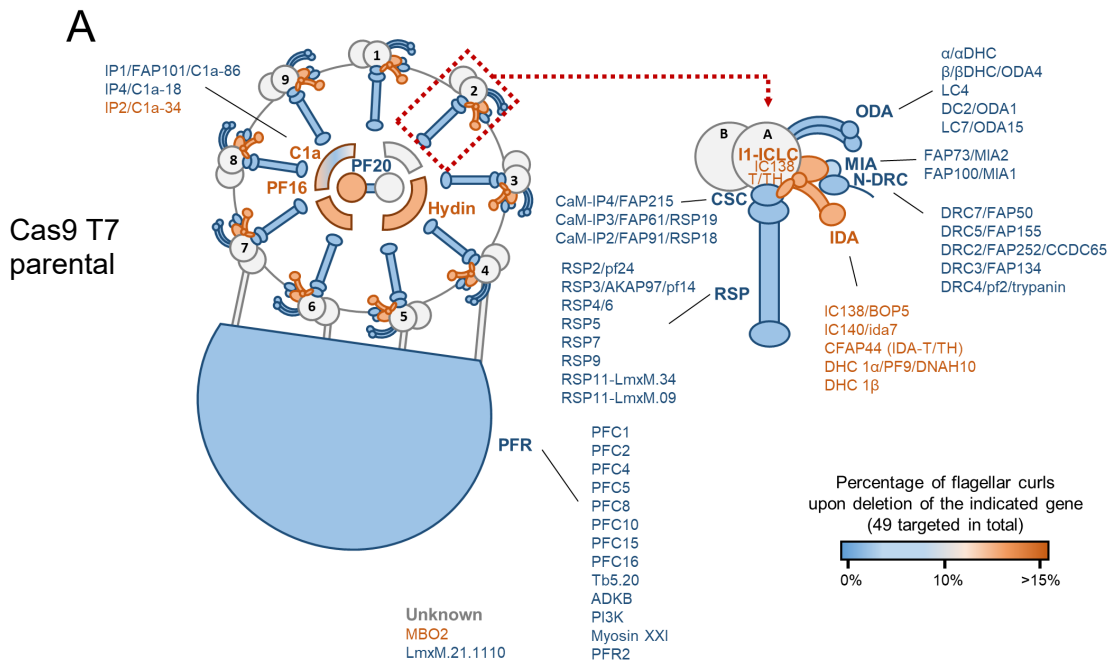
3.2.4 Loss of IDA and CPC proteins causes flagellar curling

Counting curls in all 49 knockout mutants (Table 2 and Figure 27) in a blinded assay showed that the loss of CPC and IDA components was most clearly associated with flagellar curling (Figure 28 C, Figure 29 A and B). This is in line with all 14 mutants known to-date that cause the curling phenotype. Three are deficient for central pair complex proteins and seven are deficient for IDA proteins (see section 3.2.1). However, not all CPC and IDA mutants that have been created to-date in the Gluenz lab show the curling phenotype.

For example, *Leishmania* mutants of p28 (not included in the set of 49 mutants, but published in Beneke et al. (2019b)), which in *Chlamydomonas* is contained in

IDA heavy chains a, c and d and required for their assembly (Kamiya et al., 1991; LeDizet and Piperno, 1995), have been shown to not result in curling. But increased flagellar curling rates are observed especially in IDA f/I1 component mutants, including *ΔIC138* (40.4% ±8.0% curling), *ΔIC140* (54.3% ±6.0% curling), *ΔCFAP44* (21.7% ±10.3% curling), *ΔDHC 1α* (39.4% ±1.2% curling) and *ΔDHC 1β* (35.8% ±5.1% curling) (Figure 29 A and B).

A much stronger variation in flagellar curling was found between different *Leishmania* CPC mutants. *Δhydin* (60.9% ±4.8% curling) deletions showed strong curling, while *ΔPF16* (16.4% ±4.7% curling) and *ΔC1a-34* (15.2% ±1.2% curling) mutants showed much lower rates. In contrast, *ΔPF20* (5.6% ±2.7% curling), *ΔC1a-18* (5.5% ±2.9% curling) and *ΔC1a-86* (curling not observed) deletions showed rates that were not considered as curling (Figure 29 A and B). Data from *C. reinhardtii* suggest that PF20 composes the central pair bridge (Adams et al., 1981; Smith and Lefebvre, 1997), hydin is required for assembly of CPP 2b (Lechtreck and Witman, 2007), and C1a-18, C1a-34 and C1a-86 are major components of CPP C1a (Oda et al., 2014) (Table 1). However, protein blast searches for C1a-18 and C1a-86 could not identify the *Leishmania* orthologue with high confidence (Table 1). For PF16 a clear localization within the flagellar axoneme of *Chlamydomonas* or other organism is missing. But what has been shown is that C1 microtubules of the central apparatus and associated CPPs are missing in demembranated *pf16* mutant flagella (Carbajal-Gonzalez et al., 2013; Dutcher et al., 1984; Loreng and Smith, 2017; Smith and Lefebvre, 1996; Zhao et al., 2019).



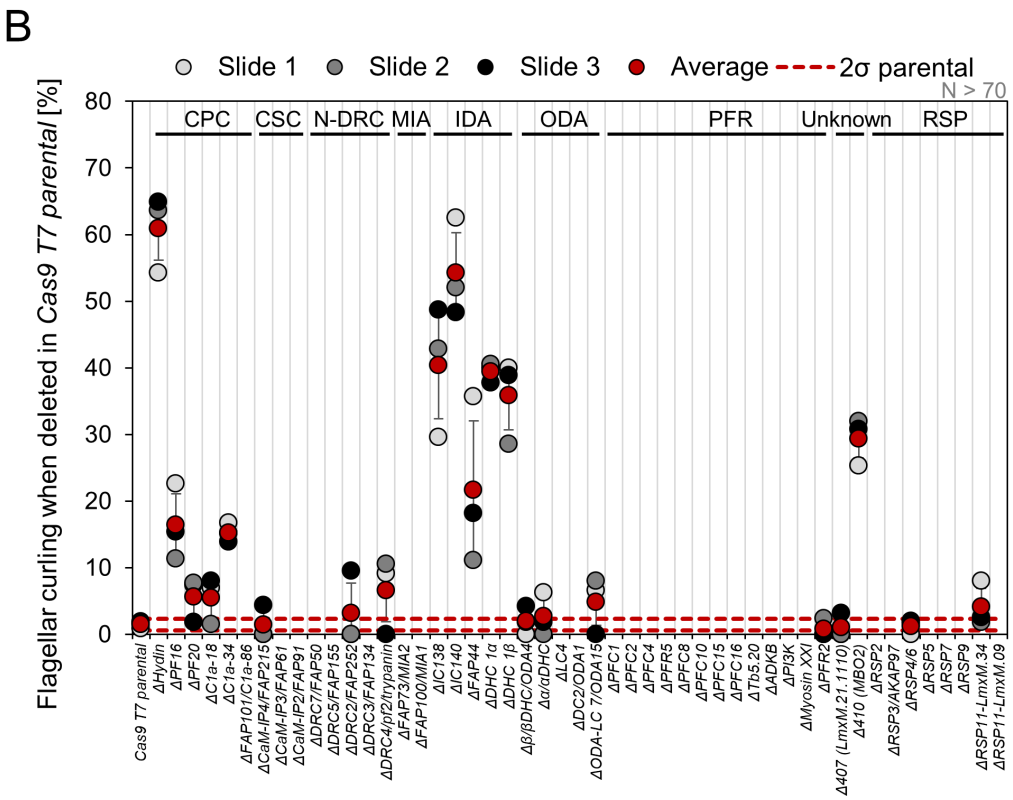


Figure legend on following page.

Figure 29. Loss of IDA and CPC proteins causes flagellar curling.

Flagellar curling rates were assessed in all 49 mutants of the chosen subset when deleting genes in Cas9 T7 parental cells. **(A)** Schematic cross-sectional view of a *Leishmania* promastigote 9 + 2 axoneme with key axonemal structures labelled. Highlighted in blue are components that when subjected to deletion in Cas9 T7 parental cells do not result in more than 15% curling rate and in orange do result in more than 15% curling rate. **(B)** Plot shows raw data for cartoon in **(A)**, with flagellar curling rate [%] on Y axis and mutant on X axis. Structure belonging to each mutant is indicated on the top. Red lines indicate two standard deviations above and below the parental average curling rate. Error bars indicate standard deviation between three counted slides of at least 70 flagella counted flagella in total per cell line.

It is likely that many of these highly conserved CPC orthologues will occupy similar positions in the *Leishmania* axoneme, but there may also be variations still to be uncovered.

Although for most of the other axonemal and extra-axonemal components, including components of the CSC, N-DRC, MIA, ODA, PFR and RSP complex, no curling was observed upon deletion, some showed elevated rates, such as DRC4 (6.5% ±4.7% curling) and ODA15 (4.9% ±3.5% curling) (Figure 29 A and B).

A possible hypothesis to explain higher curling rates in CPC or IDA mutants was that ablation of CPC or IDA components causes the flagellum to “lock” in a curled position and stop the bead propagation. However, whether a curled flagellum is specifically locked at one extreme of bending remains to be tested.

3.2.5 Failed recovery of flagellar curling by ATP reactivation suggest a curled flagellum is locked at one extreme of bending

To test whether curled flagella are locked at one extreme of bending, *Δhydin* mutants were subjected to membrane extraction and ATP reactivation. If a curled flagellum was not locked, membrane extraction would resolve flagellar curling, as it removes ATP which is needed for dynein activity and consequently for flagellar bending. ATP reactivation would then trigger again curling of *Δhydin* mutant membrane extracted flagella, as dynein motor domains get reactivated.

ΔHydin mutants and Cas9 T7 parental cells were subjected to membrane extraction in LRP buffer (Mukhopadhyay and Dey, 2016) supplemented with 0.1% and 1.0% Triton-X100. Extracted flagella were then reactivated using 1.5 mM ATP. A control with no additional Triton-X100 and ATP was performed. While this yielded paralyzed membrane extracted flagella in Cas9 T7 parental cells, which could then be reactivated with ATP, it did not have any effect on flagellar curling rates in

Δhydin mutants. Even detergent concentrations of up to 1% Triton-X100 did not resolve flagellar curling (Figure 30 A, Supplementary: Video files 1). This allows us to speculate that curled flagella are locked at one extreme of bending, as withdrawal of ATP does not relax the curl. One possible explanation is that flagella fail to reverse the flagella beat during progression of the normal flagellar waveform down the flagellum. Alternatively, the mutant axoneme might be collapsed.

However, further evidence for failure of flagellar beat progression can be found in the static behavior of flagellar curls in *Δhydin* mutants. When tracking *Δhydin* mutant cells in culture, 82.6% of flagellar curls (52.6% of all counted flagella) did not resolve their curl over 30 minutes observation time. The same static behavior was observed for mutant flagella that did not curl. 82.4% of normal flagella (30.4% of all counted flagella) remained static and did not form flagellar curls (Figure 30 B).

It remains unclear why ~60% of *Δhydin* mutant flagella curl and ~40% not. A possible hypothesis is that the measured rate of curling is only a “snapshot” of a flagellar curling cycle, which can be slower or quicker. Depending on the damage in the mutant axoneme, the curling rate might be therefore higher or lower between different mutants. Tracking several mutants with different curling rates over a long time in culture might be therefore an interesting follow-up experiment.

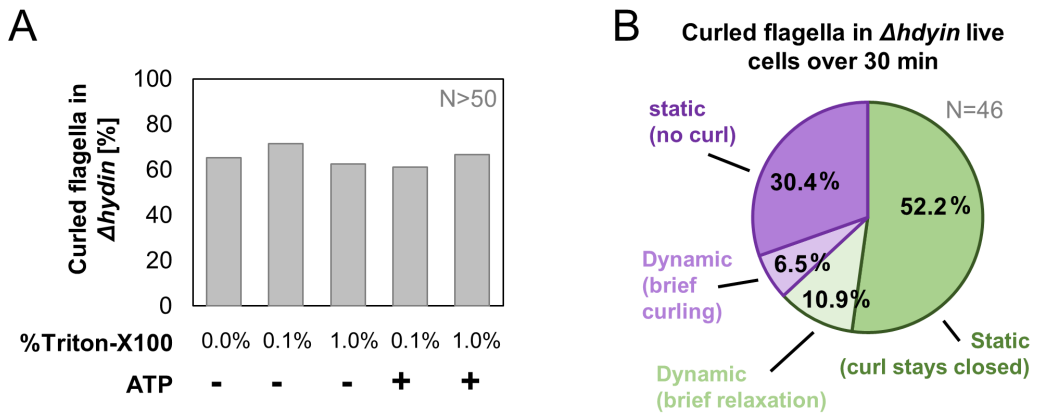


Figure 30. ATP reactivation and live cell video microscopy of *Δhydin* mutants.

(A) Curling rates for *Δhydin* mutant flagella during the ATP reactivation assay as described in the main text. Y axis: percentage of flagellar curling. X axis: used percentage of Triton-X100 and indication whether ATP reactivated or not. At least 50 flagella have been counted under each condition. Caputed videos during the assay are attached (Supplementary: Video files 1) **(B)** Live *Δhydin* mutant cells have been tracked over 30 minutes and the dynamics of flagellar curling were determined. 46 flagella were assessed in total.

3.2.6 A subset of secondary deletions resolves flagellar curling

Since ATP reactivation did not resolve flagellar curling, it was next tested whether suppressor mutations would resolve curling. As there are multiple examples of motility mutants in *C. reinhardtii* and *T. brucei* that can be compensated by secondary mutations (see section 3.3.3), the hypothesis was that additional deletions of at least ODA, IDA, RSP and CPC components will relax the curling formation, as they are required for flagellar bending.

The *Δhydin* mutant was chosen as a parental cell line for co-deletions because it had the highest curling rate. First, the genome of the *Δhydin* mutant was fully sequenced to confirm the precise deletion of both hydin alleles (Figure 31 A, B and C). Then, the subset of 49 genes (Table 2 and Figure 27) were targeted for deletion in *Δhydin* mutants and deletions were confirmed for 45 of 49 mutants with a diagnostic PCR as described in chapter 2 (Figure 32). Next, flagellar curling was evaluated in a blinded manner (Figure 33 A and B).

The *Δhydin*/N-DRC and *Δhydin*/MIA double mutants showed curling rates higher than 40%. Thus, these secondary mutations did not resolve the curls, but perhaps slightly reduced it. Two out of three tested CSC mutants did not suppress curling rates compared to *Δhydin* alone. Interestingly, deletion of FAP91 in the *Δhydin* mutant resolved flagellar curling (only 0.5% ±0.7% curling in the dual mutant), while *ΔFAP91* on its own showed a paralyzed phenotype without flagellar curling (Figure 28 C). None of the 13 proposed PFR components from Portman et al. (2009) (Table 2) or the two unknown proteins, MBO2 and LmxM.21.1110, relaxed flagellar curling upon deletion in *Δhydin* mutants. However, depleting the entire PFR structure by deleting the PFR2 array resolved flagellar curling (11.8% ± 3.2%).

A



B

	ΔHydin	*Cas9 T7 parental*
ORF blast query	**Aligned number of reads to ORF query**	
Hydin	1	>750
Blasticidin	0	0
Puromycin	0	0
Neomycin	587	0

Figure 31. Illumina sequencing of *Δhydin* mutants.

(A) Bam files of *Δhydin* mutants and *L. mexicana* Cas9 T7 parental cells were viewed in IGV genome browser (Robinson et al., 2011) (as described in Figure 17) at the hydin locus. **(B)** Blast queries of the hydin ORF and blasticidin-S deaminase, puromycin N-acetyltransferase and neomycin-kanamycin phosphotransferase resistance gene against *Δhydin* mutants and *L. mexicana* Cas9 T7 parental cells were performed. Only one read in *Δhydin* mutants was identified to align to the hydin ORF. This read was found to be unpaired with low quality scores.

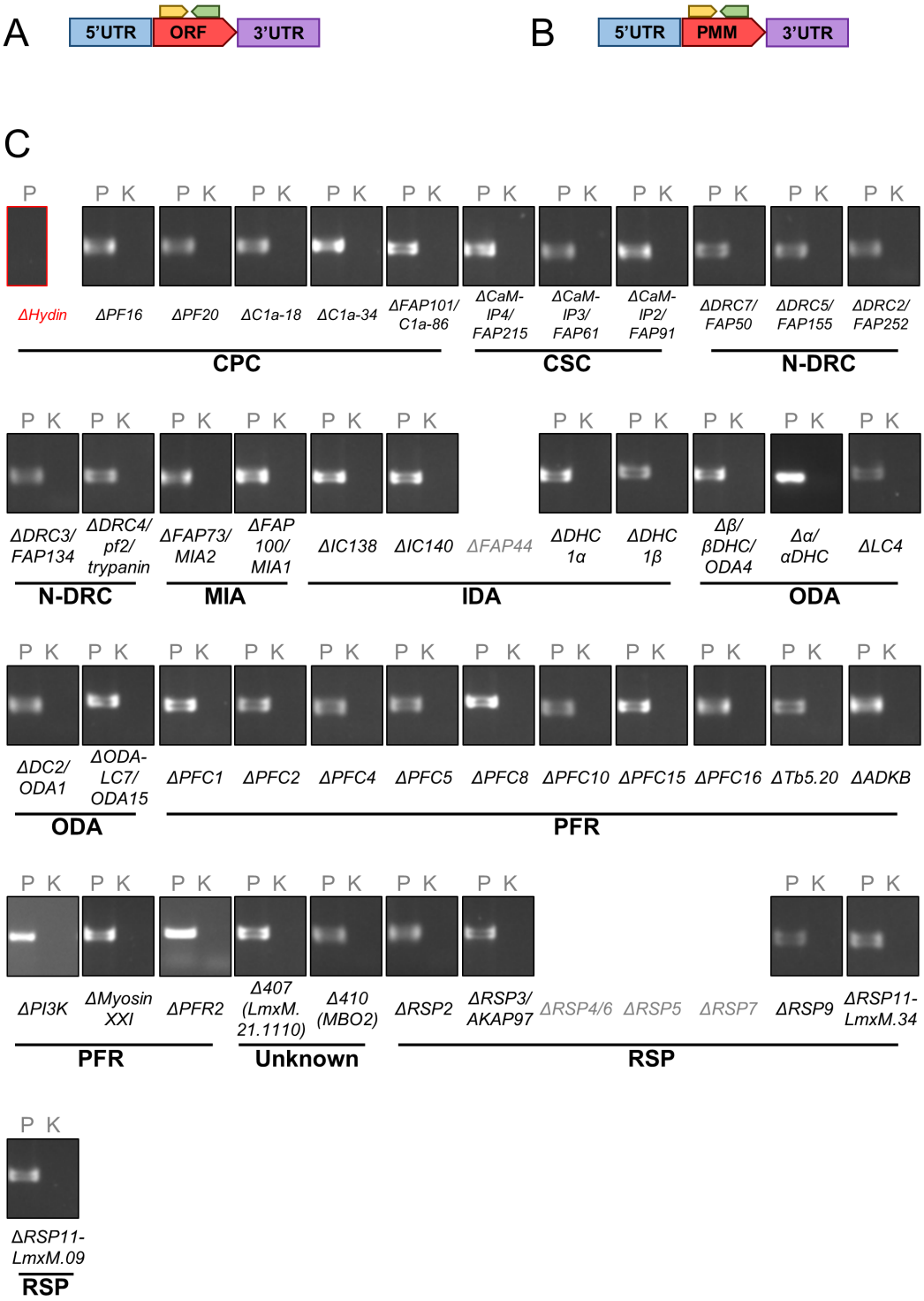


Figure legend on following page.

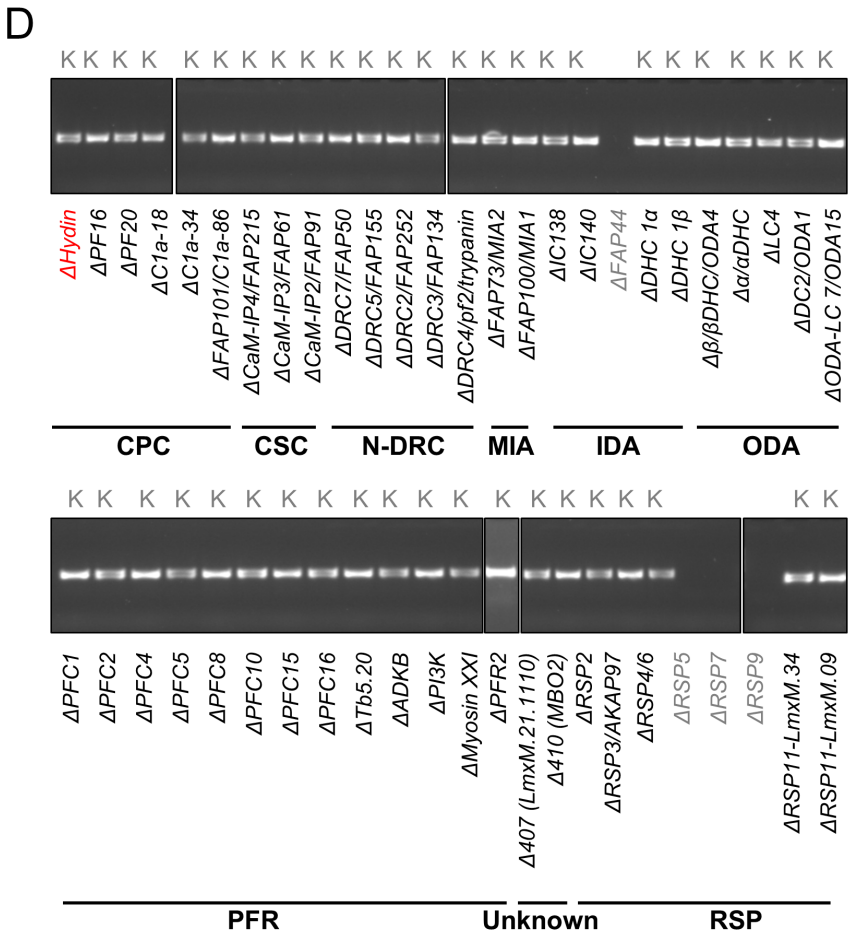
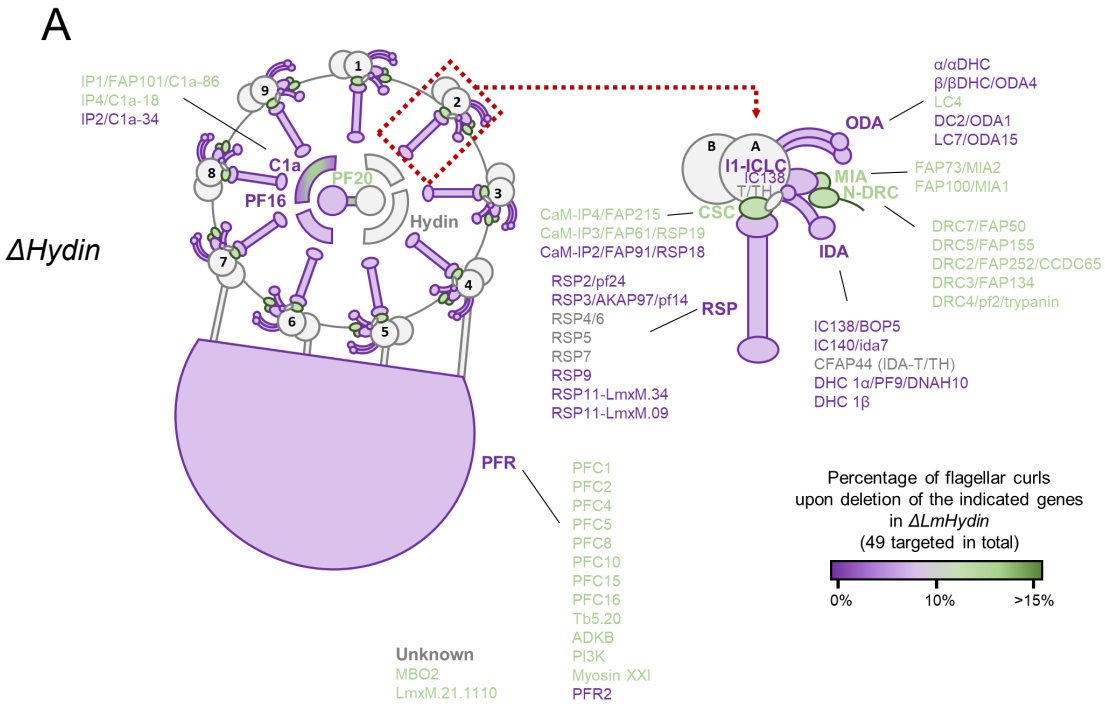


Figure 32. PCR verification of deletions in *Δhydin* mutants.

Cartoons illustrate PCRs for null mutant verification for additional knockouts in *Δhydin* mutants: **(A)** amplification of a fragment of the target gene ORF and **(B)** amplification of a fragment of the PMM ORF. **(C)** Amplicons from the target gene ORF run on agarose gel. Each ORF was amplified in the putative knockout cell line (K) and the parental *Δhydin* mutant cell line (P). No PCR products were obtained in mutant cell lines for amplification of the target gene ORF. Fainter bands below the target gene amplicon are likely to be primer dimers. The gene names indicate the target gene ORF and are part of a structural component in the axoneme as highlighted. For genes marked in grey no agarose gels are shown as transfections failed. Absence of *Δhydin* ORF is highlighted in red. **(E)** Presence of the PMM ORF PCR product was confirmed for all putative knockout cell lines, unless they failed transfections as highlighted in grey.



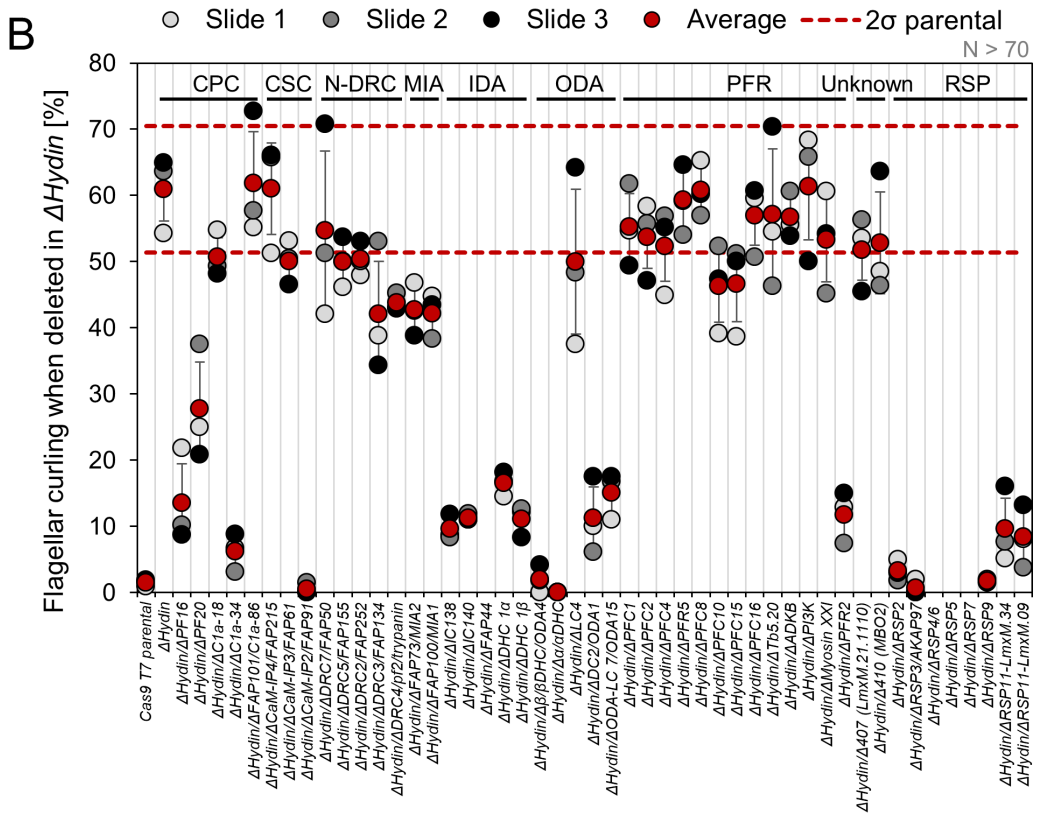


Figure legend on following page.

Figure 33. A subset of secondary deletions resolves flagellar curling.

Flagellar curling rates were assessed in all 49 mutants of the chosen subset when deleting genes in *Δhydin* mutant cells. **(A)** Schematic cross-sectional view on *Leishmania* promastigote 9 + 2 axoneme with key axonemal structures labelled. Highlighted in purple are components that when subjected to deletion in *Δhydin* mutants do resolve flagellar curling (less than 15% curling rate) and in green do not resolve flagellar curling (more than 15% curling rate). **(B)** Plot shows raw data for cartoon in **(A)**, with flagellar curling rate [%] on Y axis and mutant on X axis. Structure belonging to each mutant is indicated on the top. Red lines indicate two standard deviations above and below the *Δhydin* mutant parental average curling rate. Error bars indicate standard deviation between three counted slides of at least 70 flagella counted flagella in total per cell line.

Deletion of IDA, ODA and RSP components in *Δhydin* mutants also decreased curling significantly (0.7 to 16.5% curling in dual mutants), except for ODA protein LC4-like (Figure 33 A and B). Remarkably, PF20 deletion in *Δhydin* mutants reduced curling to 27.8% ± 7.1%, but it did not resolve curling as much as PF16 (13.6% ±5.9% curling) and C1a-34 (6.2% ± 2.3% curling) deletions. *ΔHydin*/*ΔC1a-18* and *Δhydin*/*ΔC1a-86* deletions did not have any notable effects (Figure 33 A and B).

Overall, this showed that components of a given structure typically behaved similar with respect to resolving or causing flagellar curling, as shown for IDA, ODA and RSP components. However, deletions of PFR2 and FAP91, but most obviously CPC deletions, deviated from this rule.

3.2.7 CPPs influence flagellar curling

CPC mutants, especially CPP deletions, had much more diversity in their curling rates than components of other complexes. High curling rates were observed in *Δhydin* mutants, low rates in *ΔPF16* and *ΔC1a-34* mutants and very low rates in *ΔPF20* mutants (Figure 29 A and B). Additionally, while PF16 and C1a-34 deletions had a strong suppressor effect on flagellar curling in *Δhydin* mutants, PF20 deletion had only a weak effect (Figure 29 A, B, Figure 33 A and B).

The “central pair-spoke axis” or “distributor model” (see section 3.1.3) postulates that dynein arms are controlled through their adjacent radial spokes and CPPs (Omoto et al., 1999). The central pair does not rotate in *Leishmania* flagella (Gadelha et al., 2006). Hence, the hypothesis was that the diversity in curling rates of CPC mutants may be an effect of dynein misregulation, as defective or missing

CPPs are unable to transmit dynein regulation signals via their adjacent RSPs. To test this hypothesis, *ΔPF16* mutants were next subjected to co-deletions.

Loss of the PF16 ORF in null mutants was previously confirmed by Illumina sequencing (chapter 2, Figure 17 A and B). Thus, *ΔPF16* null mutants were subjected to additional deletion of C1a-34, ODA α and β heavy chains, as well as IDA f/I1 β heavy chain and co-deletions were compared to hydin co-deletions. Deletions were confirmed by diagnostic PCRs (Figure 34 A - D).

Interestingly, while PF16 and C1a-34 deletions on their own yielded similar curling rates (15-16% curling), co-deletion of both genes resulted in an increase of flagellar curling and did not resolve it (22.9% ±4.5% curling in *ΔPF16/ΔC1a-34*). This is in contrast to reduced curling in *Δhydin* mutants by co-deletion with C1a-34 or PF16 (from 60.9% ±4.8% in *Δhydin* to 6.2% ±2.3% in *Δhydin/ΔC1a-34* and 13.6% ±5.9% in *Δhydin/ΔPF16*) (Figure 35 A – G). Moreover, curling in *ΔPF16* mutants can be suppressed to levels of Cas9 T7 parental flagella by deleting the ODA α or β heavy chain, which is similar to ODA α or β co-deletions with hydin (Figure 33 A and B). By contrast, deletion of IDA f/I1 β in *ΔPF16* mutants does not suppress curling (Figure 35 A – G), although it did suppress curling in *Δhydin* mutants (Figure 33 A and B). However, this difference might be explained by a dominant effect of the IDA f/I1 β mutation itself when performed in a *ΔPF16* mutant background (Figure 29 A and B).

Since in *Chlamydomonas* C1a-34 and PF16 are associated with the C1 microtubule and hydin with the C2 microtubule these findings allow to speculate that CPPs might selectively control dynein activity. Meaning projections on the C1 microtubule control indirectly dynein activity on microtubule 7 to 9 and thus fail to do so in *ΔPF16*, *ΔC1a-34* and *ΔPF16*/*ΔC1a-34* mutants.

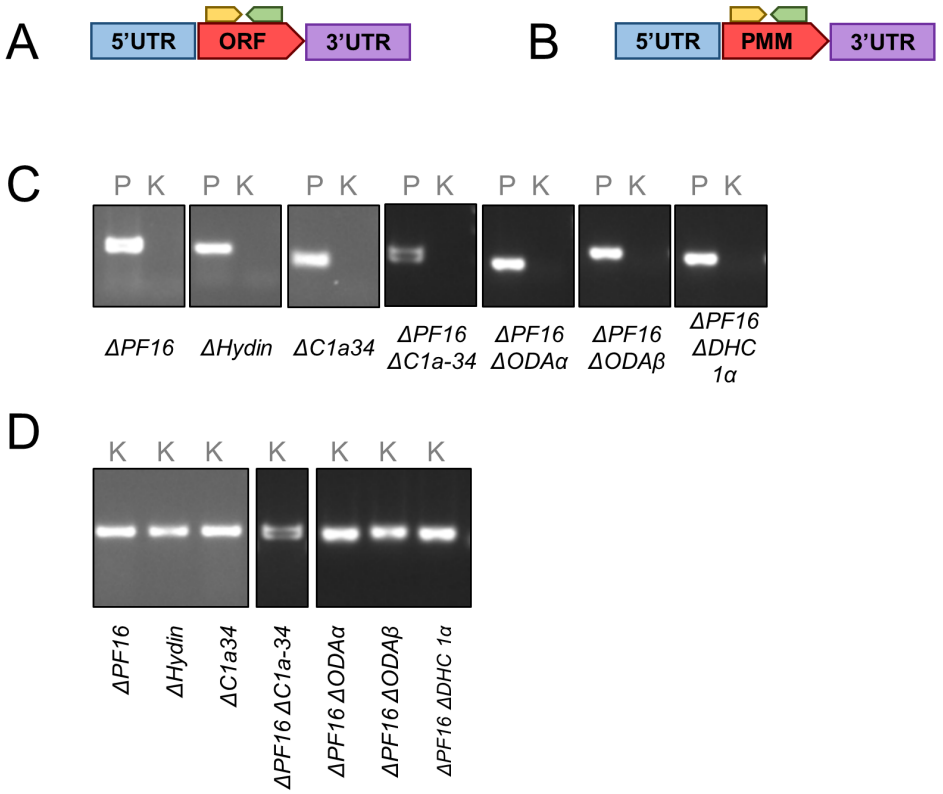
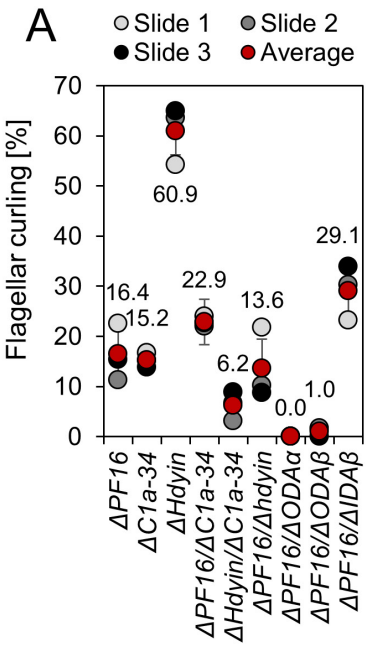


Figure 34. Mutant verification of CPC single and co-deletions.

Cartoons illustrate PCRs for null mutant verification: **(A)** amplification of a fragment of the target gene ORF and **(B)** amplification of a fragment of the PMM ORF. **(C)** Amplicons from the target gene ORF run on agarose gel. Each ORF was amplified in the putative knockout cell line (K) and the parental cell line (P), which is the Cas9 T7 parental cell line for *ΔPF16*, *Δhydin* and *ΔC1a-34* mutants and the *ΔPF16* mutant cell line for shown co-deletions. No PCR products were obtained in mutant cell lines for amplification of the target gene ORF. The gene names, in case of co-deletions the name that is not *ΔPF16*, indicate the target gene ORF. **(D)** Presence of the PMM ORF PCR product was confirmed for all putative knockout cell lines.



B

Tukey's multiple comparisons test:
* P ≤ 0.05; ** P ≤ 0.01; *** P ≤ 0.001; **** P ≤ 0.0001

	ΔPF16	ΔC1a-34	ΔHdyin	ΔPF16/ ΔC1a-34	ΔHdyin/ ΔC1a-34	ΔPF16/ ΔHdyin	ΔPF16/ ΔODAα	ΔPF16/ ΔODAβ	ΔPF16/ ΔIDAβ
ΔPF16			****				**	**	*
ΔC1a-34			****				**	*	*
Δhdyin	****	****		****	****	****	****	****	****
ΔPF16/Δ C1a-34			****		**		****	***	
ΔHdyin/ ΔC1a-34			****	**					****
ΔPF16/Δ hdyin			****				*	*	**
ΔPF16/ ΔODAα	**	**	****	****		*			****
ΔPF16/ ΔODAβ	**	*	****	***		*			****
ΔPF16/ ΔIDAβ	*	*	****		****	**	****	****	

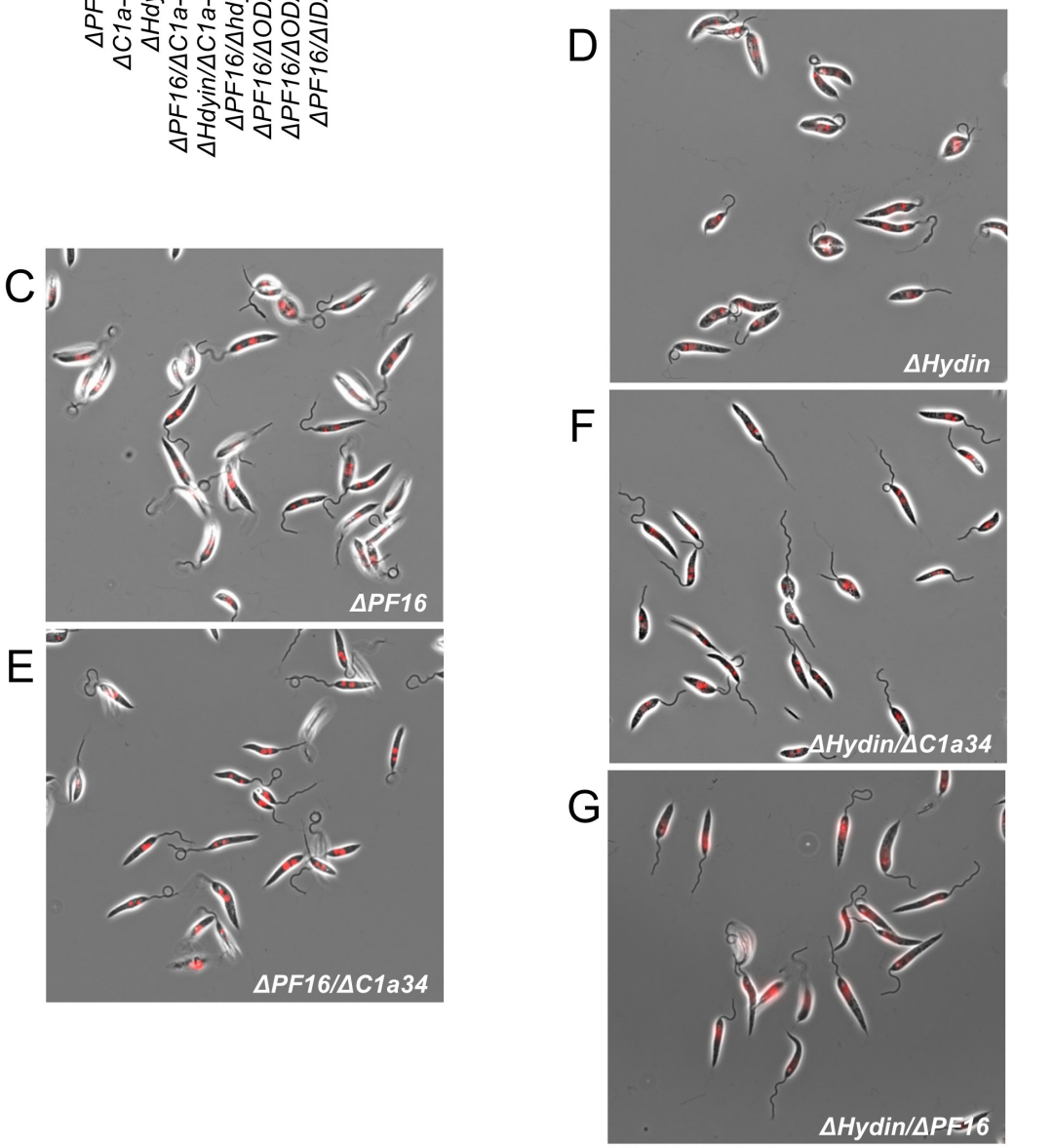


Figure legend on following page.

Figure 35. Analysis of CPC co-deletions.

(A) Curling rates of co-deletions, as plotted on X axis, were determined over three microscopy slides (grey) and their average curling rate [%] was measured (red, value indicated), as plotted on Y axis. Error bars indicate standard deviation between counted slides. **(B)** Curling rates were used to perform Tukey's multiple comparison test. **(C) – (G)** Micrographs for selected measured cell lines support counted curling rates and show merged phase (grey) and fluorescent channels (Hoechst DNA stain in red) for mutant cell lines.

Meanwhile, projections on the C2 microtubule control dynein activity on microtubule 2 to 4 and are unable to fullfil this role in *Δhydin* mutants. This is why curls can be resolved by additional disruption of the dynein regulation on the opposite side of the axoneme, as shown in co-deletions of *ΔPF16/Δhydin* and *ΔC1a-34/Δhydin* (Figure 35 A – G). The flagellum is straight because of resulting balanced dynein forces.

This would well fit with the central pair-spoke axis model of dynein regulation. Ideally this model would be tested next by co-deletions of other CPPs on microtubule C1 and C2. However, it remains difficult to find other appropriate CPP candidates by blast search from *Chlamydomonas* in *Leishmania*. If dyneins are selectively regulated by CPPs, this would cause the flagellum to curl in opposite directions in *ΔPF16* and *Δhydin* mutants. To test this, and determine the direction of curling, both mutants and Cas9 T7 parental cells were membrane extracted and subjected to SEM next.

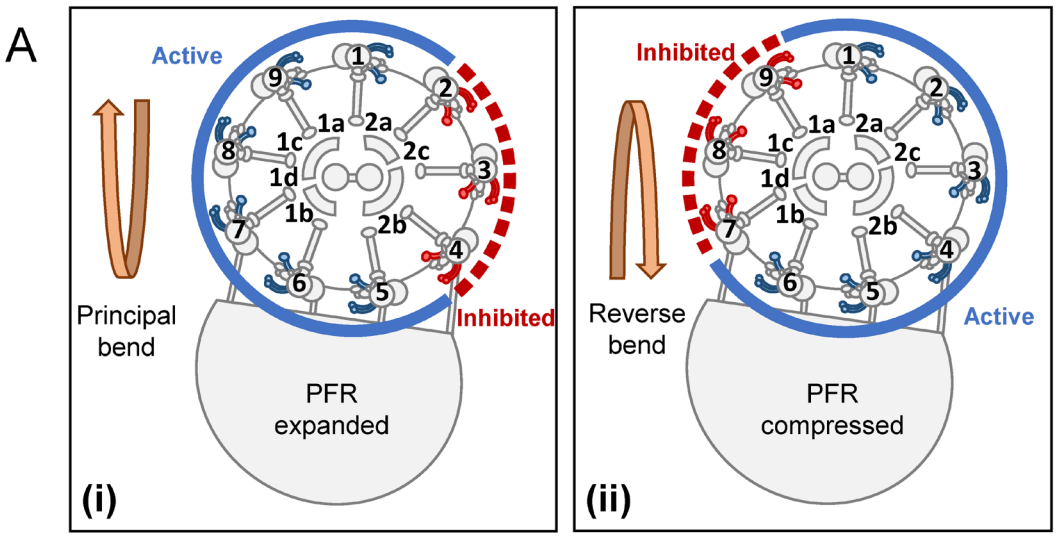
3.2.8 The paraflagellar rod may determine the flagellar curl direction

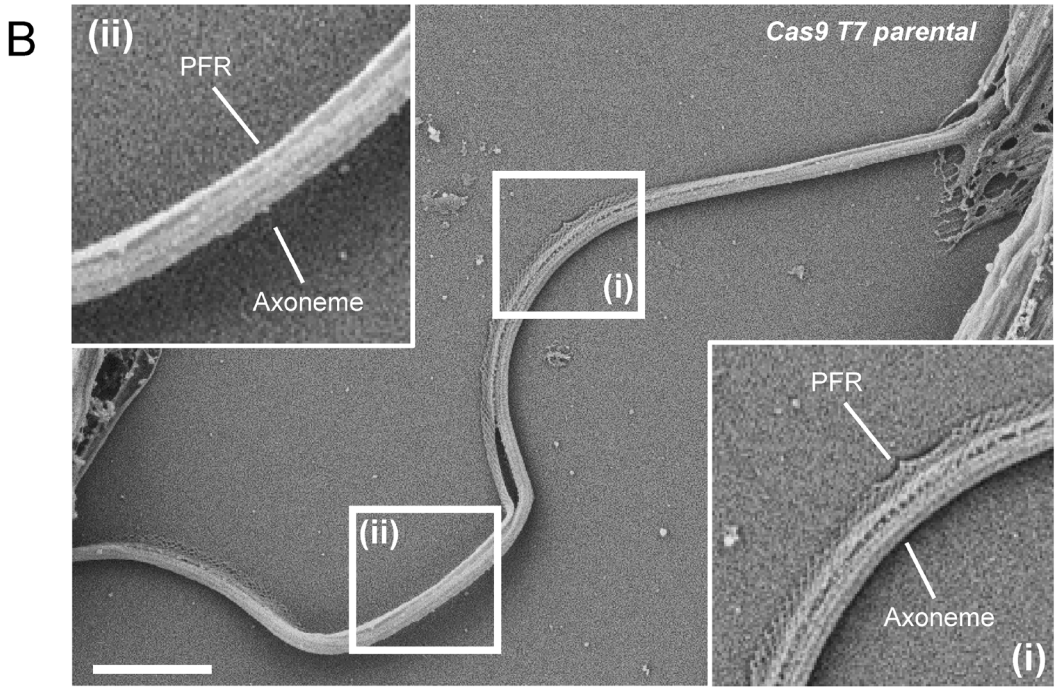
SEM of demembranated cells allows tracking of the position of the PFR relative to the axoneme. The PFR is always in the same orientation when it leaves the flagellar pocket (Wheeler et al., 2016), which means PFR tracking also enables determining the direction of the flagellar curl when the flagellum exits the pocket. Lin and Nicastro (2018) recently proposed a model in which flagellar bending is regulated by periodic inhibition of dyneins 7 to 9 and 2 to 4 (Figure 23 A – D). If this model applies for regulation of flagellar bending in *Leishmania*, this would suggest that the PFR is being expanded during the principal bend and compressed during the reverse bend (Figure 36 A). To provide evidence for such a model in

Leishmania, membrane extracted Cas9 T7 parental cells were imaged and flagellar bends analyzed. In 86.1% of 36 counted bends the PFR was either running along the inside or outside of a bent axoneme, meaning the PFR was either expanded or compressed, which may be supported by its lattice like structure (Figure 36 B and C). This is consistent with the regulatory mechanism proposed for sea urchin sperm and *C. reinhardtii* (Lin and Nicastro, 2018). Assuming that dyneins in *Leishmania* also follow the switch-inhibition model, this would suggest that deletion of C2 CPPs would prevent any inhibition of dynein activity on microtubule 2 to 4, while deletion of C1 CPPs would prevent any inhibition of dynein activity on microtubule 7 to 9. Consequently, *ΔPF16* and *ΔC1a-34* mutant flagella would be stuck in a permanent principal bend, while *Δhydin* mutant flagella are preferentially in their reverse bend position.

To test whether *Δhydin* and *ΔPF16* mutants have different curling directions, SEM of membrane extracted mutants were analyzed next. Contrary to predictions based on the above model, there was no significant difference between *Δhydin* and *ΔPF16* mutants. In 10 out of 10 *Δhydin* mutant flagella and 10 out of 11 *ΔPF16* mutant flagella the PFR was found to be on the outside of the flagellar curl, while the axoneme was running on the inside of the curl (Figure 37 A and B).

Overall, this data is consistent with a physical or a signaling role of the PFR. A physical role of the PFR could especially explain why PFR2 deletions resolve flagellar curling in *Δhydin* mutants (Figure 33 A and B). However, since the *ΔPFR2* mutants lack almost the entire PFR structure (Maga et al., 1999), this too, could affect putative signaling functions. Furthermore, these results propose a model, in which the PFR forces a flagellar curl in a mutant into a permanent principal bend. Consequently, the PFR might function as an “amplifier” for the principal bend.





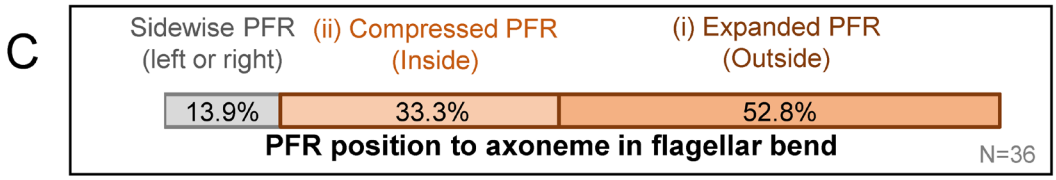


Figure legend on following page.

Figure 36. Position of the paraflagellar rod within bent flagella.

(A) Power stroke conformation model from Lin and Nicastro (2018), shown in Figure 23, has been applied for the *Leishmania* flagellum. **(i)** Dyneins on doublet 5 – 1 are in active pre power stroke conformation during the principal bend, thus expanding the PFR. **(ii)** Dyneins on doublet 1 – 6 are active and 7 – 9 are inhibited during the reverse bend, thus compressing the PFR. **(B)** Scanning electron micrograph of fixed Cas9 T7 parental cells, showing a **(i)** principal bend and **(ii)** reverse bend as proposed in **(A)**. The PFR and axoneme are labelled in magnifications, showing the **(i)** principal bend and **(ii)** reverse bend. Scale bar is 1 µm. **(C)** Categories of PFR position relative to axoneme have been counted in scanning electron micrographs for **(i)** expanded PFR, **(ii)** compressed PFR and sidewise running PFR. The proportion of all three categories in 36 counted bends are shown.

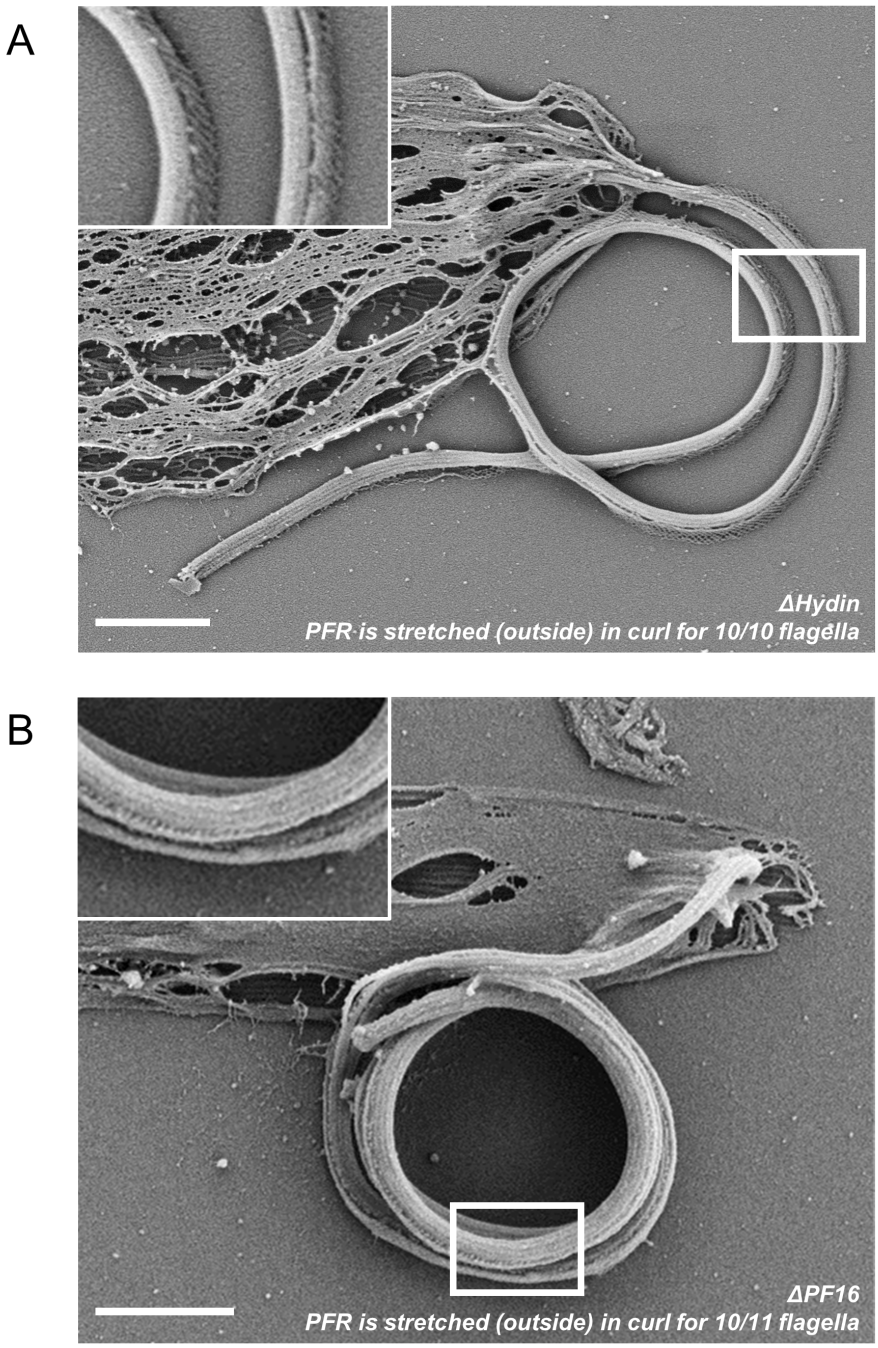


Figure 37. Position of the PFR in *ΔPF16* and *Δhydin* flagellar curling mutants. Scanning electron micrographs of detergent extracted cells, showing the PFR structure on the outer side of a flagellar curl in **(A)** *Δhydin* and **(B)** *ΔPF16* mutants, which was found to be the case for 10 out of 11 *ΔPF16* flagella and for 10 out of 10 *Δhydin* flagella. Magnifications are indicated with white boxes. Scale bars are 1 µm.

3.3 Discussion

3.3.1 A model for flagellar curling gives new insights into the regulation of flagellar bending

The aim of the work shown in this chapter was to study the regulation of flagellar bending by dissecting the flagellar curling phenotype.

The results propose the following flagellar bending model for *Leishmania*: If in a curling mutant, such as *ΔPF16*, *Δhydin* or *ΔC1a-34* CPP mutants, dyneins are asymmetrically mis-regulated, the PFR becomes dominant, as dyneins are imbalanced in their generated force (Figure 38 A). In this case, the CPPs fail to distribute and alternate the signal to the RSPs correctly on both sides of the axoneme. It is likely that the PFR in this case does not actually generate a bending force like dyneins, but instead curls in one direction by itself due to its lattice-like structure.

However, if the dynein mis-regulation is balanced with equal forces on both sides of the axoneme, the physical contribution of the PFR is over-turned by dynein activity and not dominant, causing the flagellum to be straight (Figure 38 B). Consequently, if dyneins in *Leishmania* follow also the switch-inhibition model, which needs to be confirmed with cryo-ET reconstruction studies, this would suggest that C2 CPPs control dynein activity through inhibition on microtubule 2 to 4, while C1 CPPs can inhibit dynein activity on microtubule 7 to 9.

Overall this suggests a model in which the CPPs control the periodic inhibition of dyneins and in which the PFR supports the principal bend.

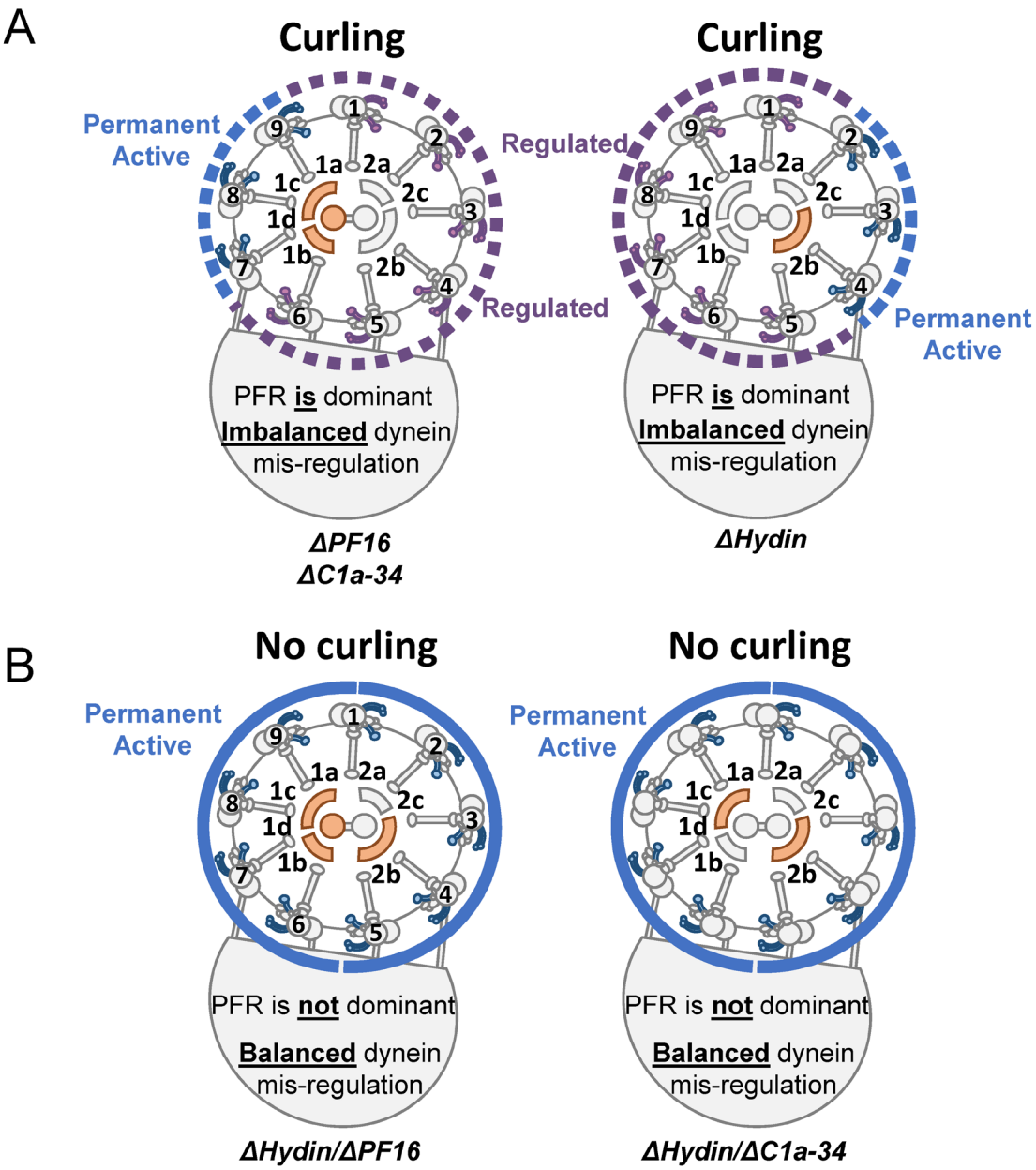


Figure legend on following page.

Figure 38. Proposed model for flagellar curling and flagellar bend regulation. Model for flagellar curling. Ablated axoneme component highlighted in orange. **(A)** In *ΔPF16* and *ΔC1a-34* mutant flagella, dyneins on doublets 7 – 9 are not regulated, as C1 associated CPPs fail to interact with their RSP units. Thus, dyneins on doublets 7 – 9 are permanently active, while doublets 1 – 6 remain regulated. In *Δhydin* mutants the opposite effect can be observed. Since dyneins on the opposite side of the axoneme are mis-regulated, they do not generate a directional force, thus allowing the PFR structure to determine the curling direction. **(B)** In *Δhydin/ΔPF16* and *Δhydin/ΔC1a-34* co-deletions CPPs on both sides of the axoneme are defective and fail to communicate signals further to RSPs. Since dyneins remain unregulated they become permanently active generating equal forces on either side of the axoneme. The resulting balanced dynein force overturns any dominant effects of the PFR, thus resulting in a straight flagellum.

3.3.2 IDA f/I1 components and CPPs are vital for progression of the flagellar beat in *Leishmania*

This flagellar bend regulation model would imply that CPPs and also other structures that are involved in flagellar curling, such as IDA f/I1 components, are essential for a normal flagellar beat cycle in *Leishmania*.

Membrane extracted flagellar axonemes normally bend and move with supplementation of ATP (Lindemann and Rikmenspoel, 1972). Thus, a failure of ATP reactivation suggests a disruption of the beat cycle (Lesich et al., 2014). The formation of curls was not changed in demembranated *Δhydin* mutant flagella, whether supplemented with or without ATP. Additionally, flagellar curls in hydin deletions seemed to behave mostly static over time, suggesting that the flagellum is locked at its current state of bending when the beat cycle is severely damaged. Mutant flagella that occasionally are able to change their curl formation might demonstrate a less severe form of a disrupted flagellar beat cycle. Overall, this emphasizes that the flagellar curling phenotype reflects a disrupted flagellar beat cycle.

In line with this is also that a higher proportion of curling restricts motility. However, *ΔPACRGA* mutants seem to be an exception from this rule (Figure 28 A and B). Mutants have relatively high curling rates, but their swimming speed is only slightly reduced compared to Cas9 T7 parental cells. The RNAi knockdown of PACRG A and B in *T. brucei* showed occurrences of missing doublets that increased towards the distal end of the flagellum (Dawe et al., 2005). Since the flagellar beat in *Leishmania* is mainly initiated from tip-to-base, mostly starting at the distal end but also in the mid-flagellum (Gadelha et al., 2007), this might suggest that curling generally in all mutants results from failed waveform progression upon beat

initiation. Therefore, the axonemal distal defect in *ΔPACRGA* mutants results in an atypical high curling rate for a mild motility phenotype and for other curling mutants there remains a correlation between decreasing swimming speed and directionality and increasing flagellar curling. Thus the curling phenotype provides a unique opportunity to study flagellar components that are required for progression of the normal flagellar waveform.

ΔHydin mutants have the highest flagellar curling rate. However, the question arises, why only about ~60% of mutant flagella curl. Excluding technical reasons in preparation of microscopy slides, this may reflect only the amount of damage caused to the regulation machinery or this may reflect that only certain beat types can result in a flagellar curl. Analysis of *L. mexicana* Cas9 T7 cells showed that about 50% of cells in a population have an uncoordinated/static beat type, while remaining flagellar beat types are either constant (25%), interrupted (24%) or ciliary/reverse beats (1%) (Edwards et al., 2018). Since the amount of damage to the regulation machinery and the type of flagellar beat might change depending on the mutant phenotype, this may also explain why some mutants have lower and some higher curling rates. Also possible is that the measured rate of curling is only a “snapshot” of a flagellar curling cycle and actually all flagella curl when observed long enough.

Analysis of all 49 null mutants showed that only CPC deletions and IDA deletions exhibit the curling phenotype (Figure 28 C, A and B; Figure 29 A and B).

However, IDA components IC138, IC140, CFAP44, LAX28 and IDA f/I1 α and β heavy chains deleted here are all associated with the IDA f/I1 complex, which is believed to be a regulator of all IDA subspecies within the 96 nm repeat and further described in chapter 4. Since in *Chlamydomonas* IDA f/I1 heavy chains have been

shown to switch their power stroke conformations between microtubules 2 to 4 and 7 to 9 (Lin and Nicastro, 2018) it is likely that curling from these IDA mutants reflects a regulatory IDA defect, which disrupts propagation of the flagellar waveform. But, it remains to be tested whether deletion of other IDA subspecies, including a, b, c, d, e and g result also in curling. Ideally, this would be tested by deleting each IDA heavy chain in *Leishmania*. However, it is unclear how many single headed IDAs are actually present in *Leishmania* and which gene encodes them (Wheeler et al., 2015a; Wickstead and Gull, 2007). While this would require cryo-ET reconstruction studies on IDAs of *Leishmania*, previous knockouts in *L. mexicana* of p28, a light chain required in *Chlamydomonas* for IDA heavy chain a, c and d assembly (Kamiya et al., 1991; LeDizet and Piperno, 1995), may have already shown that other IDA subspecies do not cause curling. Only 2.1% of *Δp28* mutants were shown to exhibit the curling phenotype in Beneke et al. (2019b). If p28 in *Leishmania* functions similar to p28 in *Chlamydomonas*, this indicates that mutants of IDAs would not automatically exhibit the curling phenotype. This suggests that the curling phenotype in IDA f/I1 mutants reported here is not a result of missing power strokes from deleted IDA heavy chains, it is a result of IDA f/I1 mediated dynein mis-regulation.

Dynein mis-regulation may be also especially reflected by CPC deletions, including *ΔPF20*, *ΔC1a-34*, *ΔPF16* and *Δhydin* mutants. Electron microscopy of PF16, PF20 and hydin mutants in *C. reinhardtii* might explain observed differences in curling rates. While 98% of imaged *pf16* mutant flagella (Dutcher et al., 1984) and 90% of imaged hydin RNAi knockdown flagella (Lechtreck and Witman, 2007) still had both central pair microtubules, only 34% of imaged *pf20* mutant flagella (Adams et al., 1981) still had the two central pair microtubules. Importantly, these numbers

were generated from flagellar samples that were not demembranated and prepared in a similar way for TEM. Although RNAi knockdowns of PF16, PF20 and hydin in *T. brucei* have also been performed, this quantification is missing in Branche et al. (2006), Dawe et al. (2007) and Ralston et al. (2006). But these studies reported that RNAi knockdowns in *T. brucei* resulted all in aberrant orientations of the central pair microtubules or missing or partly missing central pair microtubules, which was also observed for *L. mexicana ΔPF16* deletions (Beneke et al., 2017).

C. reinhardtii data suggests that *pf20* mutants, which should be deficient for the central pair bridge, do not have central pair microtubules in the majority of flagella, which is in contrast to *pf16* and *hydin* mutants that may have only missing C1 and C2 CPPs (Dutcher et al., 1984; Lechtreck and Witman, 2007). However, it should be noted that the precise PF16 localization and missing CPPs in *pf16* mutant flagella remain to be identified (Carbajal-Gonzalez et al., 2013; Dutcher et al., 1984; Loreng and Smith, 2017; Smith and Lefebvre, 1996; Zhao et al., 2019). This may initially suggest that PF20 must have the highest curling rates of those three deletions. But if CPPs function like a distributor to provide a local signal to the radial spokes that selectively activates subsets of dynein arms (Omoto et al., 1999), a missing CPP on only one side of the axoneme may result in much higher flagellar curling rates than two missing central pair microtubules. This is also demonstrated in *Chlamydomonas* paralyzed *pf19* mutant flagella, in which the CPC is absent. Dyneins in these flagella were found to be in active pre-power stroke conformations causing them to be straight due to balanced dynein forces on both sides of the axoneme (Lin and Nicastro, 2018) (Figure 23 A and D). However, a specific CPP deletion would only disrupt the dynein regulation on one side of the axoneme,

consequently resulting in imbalanced dynein forces. Thus, CPPs may control the propagation of the flagellar waveform by regulating the periodic switching of dynein activity.

3.3.3 A possible model in which CPPs control the periodic switching of dynein activity

These observations strongly support the central pair-spoke axis or distributor model for flagellar bend regulation in *Leishmania*. Meaning, CPPs distribute the inhibitory signal for flagellar bend regulation further to their adjacent radial spokes to control dynein activity. However, it remains unknown whether such a model would require a signaling interaction between RSPs and CPPs or whether also a purely mechanistic interaction is possible.

The idea that flagellar motility is regulated through mechano-chemical transduction via the radial spokes and central apparatus is not new (Smith and Yang, 2004). Probably the strongest evidence to-date for this model can be found in a study on *Chlamydomonas pf6* mutants. *Pf6* mutants miss the CPP C1a and cannot swim. They instead twitch on the bottom of the culture dish (Dutcher et al., 1984; Oda et al., 2014), but their motility phenotype can be rescued by tagging the spokehead proteins RSP3, RSP4 and RSP6 at their C-terminus in *pf6* mutants. Rescue of mutant phenotypes has been shown by using three different protein tags (hemagglutinin, biotin carboxyl carrier protein, and green fluorescent protein). In contrast, tagging these RSPs at their N-terminus in *pf6* mutants does not restore motility, indicating that a connection between RSPs and CPC at the right position is vital for regulation of the flagellar beat. The authors conclude that there must be mechanosignaling between central apparatus and RSPs via the CPPs that controls axonemal dynein activity (Oda et al., 2014). This also supports a more mechanistic

model of RSP CPP interaction because it is purely about the resulting gap between CPP and RSP in CPP mutants that can be filled with any given macromolecule to restore dynein regulation. There are at least two other studies in *Chlamydomonas* that also support a mechanosignaling model for dynein regulation via the RSPs and CPPs. Nakazawa et al. (2014) showed that RSPs pull outer doublet microtubules towards the axonemal center to enable signaling from CPPs to RSPs, while DiPetrillo and Smith (2011) suggested that CPP C1d transmits regulatory signals to the dynein arms via interactions with the RSPs.

A possible signal cascade from CPC via RSPs towards IDAs and ODAs can be also observed in other secondary mutations in *Δhydin* mutants, as co-deletions that resolve curling are also likely to restore balanced dynein forces (Figure 33 A and B). Thus, it is not surprising that all five successfully deleted RSP components resolve curling in *Δhydin* mutants because the additional RSP deletion transforms the mis-regulation of dyneins on one side of the axoneme into a uniformly balanced dynein mis-regulation on all sides of the axoneme. FAP91, a CSC component that resolves curling, might have the same effect because of its direct interaction with RSP3 (Dymek and Smith, 2007). Similar conclusions may also be made for all IDA f/I1 complex deletions as they control IDA activity and flagellar bending, as further discussed in chapter 4. ODA deletions in *Δhydin* mutants resolve curling as their power-stroke conformational changes define the direction of the flagellar bend. Meaning, if ODAs are absent in the flagellum, there will be no bend (Lin and Nicastro, 2018) and therefore no curl.

Interestingly, deletion of two CSC components, all N-DRC and all MIA components does not suppress the curling phenotype, indicating that individual CSC, N-DRC and MIA components may not uniformly control dynein activity in *Leishmania*. This

is especially unexpected for N-DRC components, as paralyzed RSP and CPC mutant phenotypes in *Chlamydomonas* have been shown to be rescued by additional N-DRC deletions (Huang et al., 1982). However, this N-DRC suppressor mutation was later shown to affect assembly of N-DRC 3, 4, 5, 6 and 7 (Heuser et al., 2009), whereas only single N-DRC mutations in *Δhydin* mutants were performed here in this chapter. Additionally, RNAi of DRC4, also referred to as trypanin, was also shown to rescue PF16 knockdown phenotypes in *T. brucei* (Ralston et al., 2006). However, motility was only partly restored. This agrees with the finding that deletion of DRC3 and 4, as well as MIA1 and MIA2 slightly decreased curling in *Δhydin* mutants (Figure 33 A and B).

While these findings are all consistent with the central pair-spoke axis model for flagellar bend regulation, a purely mechanistic connection of CPP and RSP would also agree with curvature models. However, it remains unclear how the beat is initiated in these proposals. Initially contrary to the central pair-spoke axis model is also that in *Chlamydomonas* the rotation of the CPC follows the flagellar beat rather than determines its direction (Mitchell and Nakatsugawa, 2004). However, these findings might just reflect that there is a short delay after the inhibitory signal from CPPs reaches dyneins, thus resulting in first a dynein driven bend and then second the central pair rotation. Also initially opposing the central pair-spoke axis model is the found aberrant orientation of the central pair in PF20, PF16 and hydin CPP mutants in *T. brucei* and *L. mexicana* (Beneke et al., 2017; Dawe et al., 2007; Ralston et al., 2006). However, also here this can be explained and well fitted into the central pair-spoke axis model because if the central pair in CPP mutants rotates, a missing CPP will in any orientation in a fully symmetrically axoneme fail to selectively control dyneins via its adjacent RSP.

Taken together these data suggest that CPPs might actually regulate flagellar bending, potentially by controlling the periodic switching of dynein activity. However, experimental data is missing to conclude that the interaction with RSPs is a purely mechanistic one or involves also signaling transduction pathways.

3.3.4 The paraflagellar rod amplifies the principal bend

Scanning electron micrographs of demembranated Cas9 T7 parental cells, *ΔPF16* mutants and *Δhydin* mutants indicate that in *Leishmania* the PFR is likely to additionally contribute to flagellar bending by supporting especially the principal bend (Figure 36 and Figure 37). This is in line with previous findings, showing that one of the main PFR components, PFR2, is not essential for motility *per se* but is required for normal flagellar wave frequency and beat amplitude in *L. mexicana* and *T. brucei* (Bastin et al., 1998; Santrich et al., 1997). Additionally, it has been proposed before that the PFR may act as a biomechanical spring that stores and transmits energy through its lattice-like structure during a normal flagellar beat cycle (Hughes et al., 2012).

Analysis of flagellar bends in Cas9 T7 parental cells shown here support also that the bending direction is perpendicular to the PFR (Figure 36 B and C). The main flagellar bending direction in *L. mexicana* and *T. brucei* is thought to be perpendicular to the planar arrangement of the central pair microtubules (Gadelha et al., 2007; Gadelha et al., 2006; Koyfman et al., 2011; Rodriguez et al., 2009; Wheeler, 2017). Since the PFR is fixed relative to the axoneme (Gull, 1999), the bending direction is also perpendicular to the PFR. A study on *T. brucei* suggests also that the axoneme drives the expansion and compression of the PFR (Koyfman et al., 2011).

However, the finding that in *ΔPF16* mutants and *Δhydin* mutants the PFR is only expanded and found to be on the outside of the flagellar curl, forcing paralyzed *ΔPF16* mutant and *Δhydin* mutant flagella to curl into the same direction (Figure 37 A and B), supports a model in which the PFR supports especially the principal bend. The PFR may have mechanical properties that might help to amplify a dynein-induced bend. However, also a signaling role or an indirect role in controlling dynein activity (e.g. by modulating ATP levels within the flagellum) of the PFR remains a possible explanation for the flagellar curl to persist. But, since curl formation was not altered in demembranated mutant flagella a signaling or an indirect controlling role of the PFR may be less likely, as use of a detergent would be expected to remove signaling messengers.

Interestingly, curling-like formations can be also seen in flagella of other cells with extra-axonemal structures and surprisingly, even if they are symmetrically positioned within the flagellum. For example, curling can be observed in selected human sperm phenotypes. Compound-heterozygous mutations and frameshift mutations in CFAP44 and CFAP43 have been shown to result in curling rates of up to 39%, which is referred to as coiled flagella by the authors (Tang et al., 2017). Null mutants of *L. mexicana* CFAP44, which is part of the IDA f/I1 tether/tether head complex (T/TH) and further discussed in chapter 4, exhibited also the curling phenotype shown here. Contrary, there has been no curling phenotype reported for *Chlamydomonas fap44* mutant axonemes (Fu et al., 2018; Kubo et al., 2018), potentially as *Chlamydomonas* lacks extra-axonemal structure. Therefore, this might demonstrate that the symmetrically arranged extra-axonemal outer dense fibers and fibrous sheath in mammalian sperm contribute similar to flagellar bend regulation as the PFR does.

Since extra-axonemal structures seem to be the main cause for the flagellar curl to persist, it is not surprising that curling has not been observed in *Chlamydomonas* for CPC and IDA f/I1 mutants before. For example hydin in *Chlamydomonas* and lung epithelial cell flagella, which both lack major extra-axonemal structures, have been studied extensively (Davy and Robinson, 2003; Lechtreck et al., 2008; Lechtreck and Witman, 2007; Olbrich et al., 2012), but curling for mutants has not been reported. However, hydin RNAi knockdowns in *Chlamydomonas* and *hy3/hy3* mouse mutants cause flagella to stall between effective and recovery stroke, indicating mis-regulation of switching dynein activity (Lechtreck et al., 2008; Lechtreck and Witman, 2007). Curled flagella have also not been described for PF16, hydin or PF20 RNAi knockdowns in *T. brucei* (Branche et al., 2006; Dawe et al., 2005; Dawe et al., 2007; Ralston et al., 2006). The attachment of the flagellum to the cell body presumably prevents formation of curls even though a PFR structure is present. This hypothesis could be tested in double RNAi mutants, targeting Fla1 for flagellar detachment and hydin to induce curling.

Taken together, these results are consistent with a flagellar bending model for *Leishmania* in which the CPPs and IDA f/I1 components possibly control the periodic inhibition of dyneins and in which the PFR would eventually support the principal bend. However, it remains unclear whether flagellar bending regulation via the PFR, CPPs and IDA f/I1 components relies on purely mechanical interactions, signaling cascades or both, mechanosignaling pathways.

CHAPTER 4

LAX28 IS REQUIRED FOR ASSEMBLY OF THE INNER DYNEIN ARM F/L1 AND TETHER/TETHER HEAD COMPLEX WITHIN THE *LEISHMANIA* FLAGELLUM

4 LAX28 is required for assembly of the IDA f/I1 and T/TH within the *Leishmania flagellum*

Results described in this chapter have been deposited as a manuscript in bioRxiv (Beneke et al. (2019a), bioRxiv 782888; doi:https://doi.org/10.1101/782888) and proteomics data has been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD014077. LAX28 was tagged by Katherine Banecki during her undergraduate research project in the lab, supervised by Eva Gluenz and me. Katherine Banecki also generated the *LAX28::mNG/ΔCFAP44* mutant within this project. Preparation of thin sections for electron microscopy was carried out by Sophia Fochler with help from Errin Johnson and Raman Dhaliwal (Dunn School Bioimaging Facility). Mass spectrometry analysis of protein samples was carried out by Svenja Hester at the Advanced Proteomics Facility at the Sir William Dunn School of Pathology. All other data and figures presented in this chapter are wholly my own work unless stated otherwise.

4.1 Introduction

4.1.1 Chapter summary and aim

As the model in chapter 3 predicts that dynein activity is regulated by proteins that cause curling when subjected to deletion, the aim of the work presented in this chapter was to characterize the IDA f/I1 tether and tether head (T/TH) complex (1) because it is in a central hub for dynein regulation and (2) because the main component of the T/TH complex, CFAP44, shows increased curling rates when

subjected to deletion. Additionally, the T/TH complex is not very well characterized yet and there seems to be quite a lot of divergence between T/TH proteins in different species.

To gain new insights into the role of the T/TH complex in flagellar motility and the regulation of flagellar bending, the strategy was to identify proteins associated with *L. mexicana* CFAP44 and study their function. To identify T/TH complex associated proteins, flagellar skeletons from CFAP44 null mutants were isolated and their proteomic composition was compared to the one of control cells.

This identified a previously uncharacterized flagellar protein LAX28 (*Leishmania* axonemal protein, 28 kDa, LmxM.24.1310). Further characterization of LAX28, CFAP44 and the IDA f/I1 intermediate chain IC140 showed mutual dependencies for flagellar localization and requirement of all three proteins for normal flagellar motility.

Interestingly, LAX28 has a potential orthologue in humans called TEX47, which is predominantly expressed in testis (Fagerberg et al., 2014) and was identified in the human sperm tail proteome (Amaral et al., 2013), indicating a possible role in sperm motility.

4.1.2 Regulation and function of IDAs

In the cilia model organism *Chlamydomonas* seven IDA subspecies have been identified (dyneins a, b, c, d, e, f/I1 and g). However, it remains to be shown in how far each IDA subspecies has its own role in flagellar function and motility. They may substitute each other's function or may have completely diverse roles (Kato-Minoura et al., 1997; Kubo et al., 2018; Perrone et al., 1998). Direct regulators that control the conformational changes that IDAs undergo have been identified. The

CSC (Dymek et al., 2011), the MIA (Yamamoto et al., 2013) and the N-DRC complex (Heuser et al., 2009; Huang et al., 1982; Oda et al., 2015; Ralston et al., 2006) have all been shown to control the function of IDAs. As discussed in chapter 3, there are also several lines of evidence that RSPs (Barber et al., 2012; Curry et al., 1992; Diener et al., 1993; Ralston et al., 2006; Williams et al., 1989; Yang et al., 2006; Yang et al., 2004) and the CPC (Adams et al., 1981; Dawe et al., 2007; Dutcher et al., 1984; Lechtreck and Witman, 2007; Oda et al., 2014) influence activity of IDAs.

Interestingly, the IDA f/I1 complex is the only IDA that contains two dynein heavy chains (DHC 1α, DHC 1β) and requires an intermediate chain-light chain (ICLC) complex for incorporation into the flagellar axoneme (Heuser et al., 2012; Perrone et al., 1998; Viswanadha et al., 2014). As highlighted in Kubo et al. (2018) this IDA f/I1 ICLC complex contains five light chains in *Chlamydomonas* (LC7a, LC7b, LC8, Tctex1 and Tctex2b), one accessory subunit FAP120 and two intermediate chains IC140 and IC138 (DiBella et al., 2004a; DiBella et al., 2004b; Harrison et al., 1998; Hendrickson et al., 2004; Ikeda et al., 2009; Myster et al., 1997; Myster et al., 1999; Perrone et al., 1998; Piperno, 1990; Porter et al., 1992; Smith and Sale, 1991; Toba et al., 2011). Both intermediate chains have been shown to act either as a regulator or as an assembly factor for the IDA f/I1 complex. IC140 has been shown to preassemble with both heavy chains in the cytoplasm before being transported by IFT proteins to the distal end of a growing flagellum (Viswanadha et al., 2014), while IC138 has been proven to be an important phosphorylation switch of IDA f/I1, regulating the beat amplitude, sliding velocities between microtubules and the speed of bend propagation (Hendrickson et al., 2004; VanderWaal et al., 2011).

4.1.3 The IDA f/I1 tether and tether head complex (T/TH)

Very recently a new interactor of IDA f/I1 has been added to the growing list of IDA regulators. Genetic, biochemical and cryo-ET reconstruction studies in *Chlamydomonas reinhardtii*, *Tetrahymena* and sea urchin sperm have identified FAP44 and its paralogue FAP43 as being part of the IDA f/I1 T/TH complex. Both proteins were shown to be required for assembly of the T/TH complex, but not for the entire IDA f/I1 complex, and needed for regulating conformational changes of IDAs, as well as transferring their activity into microtubule sliding motion (Fu et al., 2018; Kubo et al., 2018; Urbanska et al., 2018). The T/TH has a unique location in close proximity to the start and end of the 96 nm repeat unit, building a connection that allows interaction with the tail of IDA d (Fu et al., 2018). A close interaction of the T/TH complex with IDA f/I1 dynein heads provides a platform for microtubule sliding by linking the A-tubule to both IDA f/I1 heads (Heuser et al., 2012).

FAP43 and FAP44 are part of a core group of 50 genes conserved in organisms with motile flagella (Baron et al., 2007). Studies on their human orthologues CFAP43 and CFAP44 have been linked to non-syndromic male infertility (Krausz et al., 2015; Okutman et al., 2018; Tang et al., 2017) and RNAi mediated loss-of-function studies in *T. brucei* have confirmed a role in motility for protozoan parasites (Coutton et al., 2018). In *L. mexicana* CFAP43 and CFAP44 mutants exhibit a strong decrease in swimming speed and directionality, as well as the flagellar curling phenotype (Beneke et al., 2019b). While CFAP44 and CFAP43 are highly conserved proteins with homologues in *Chlamydomonas*, *Tetrahymena* and humans (Baron et al., 2007), additional components of the T/TH complex have been identified to be more species-specific. FAP244 was shown to be part of the T/TH complex but a homologue in *Tetrahymena* and higher organisms (Fu et al.,

2018) has not been identified. Fap57p has also been proposed to be part of the T/TH complex (Urbanska et al., 2018), as well as less conserved potential T/TH candidates MOT7, FAP102, and Cre10.g452250 (Kubo et al., 2018).

Since the T/TH complex is in a central position for regulating IDA activity and flagellar curling, it was of interest to use quantitative proteomics of flagellar skeleton preparations of *ΔCFAP44* mutants and control cells to identify other T/TH proteins in *Leishmania*.

4.2 Results

4.2.1 Identification of proteins missing from *ΔCFAP44* mutant salt extracted axonemes

To identify potential interactors of the T/TH complex, several attempts were made to solubilize CFAP44 and perform IP experiments. A combination of salt and octylglucoside detergent extraction procedures were tested using a CFAP44::eYFP tagged reporter cell line. A broad range of salt concentrations, including up to 2 M LiCl, 2 M $CaCl_2$, 3.2 M KCl and 4 M NaCl was found to be insufficient to disassociate CFAP44 from the *Leishmania* axoneme (Figure 39). Since CFAP44 was identified to be stably associated with the *Leishmania* axoneme, the strategy was changed and label free quantitative proteomics of *ΔCFAP44* mutants (Beneke et al., 2019b) and parental Cas9 T7 cell line (Beneke et al., 2017) was used to identify proteins that depend on CFAP44 for their axonemal localization. A widely utilized flagellar skeleton isolation protocol in *T. brucei* (Robinson et al., 1991) was adapted for use in *L. mexicana*. To enrich for flagellar skeletons, extracts were additionally purified by running them over a sucrose gradient (Figure 40 A, B and C), as this was previously shown to yield flagella with little cell body contamination (Beneke et al., 2019b).

Liquid chromatography tandem mass spectrometry analysis of flagellar skeleton preparations of *ΔCFAP44* and control cells (Figure 40 C (iv)) identified a total of 1162 proteins (Figure 40 D).

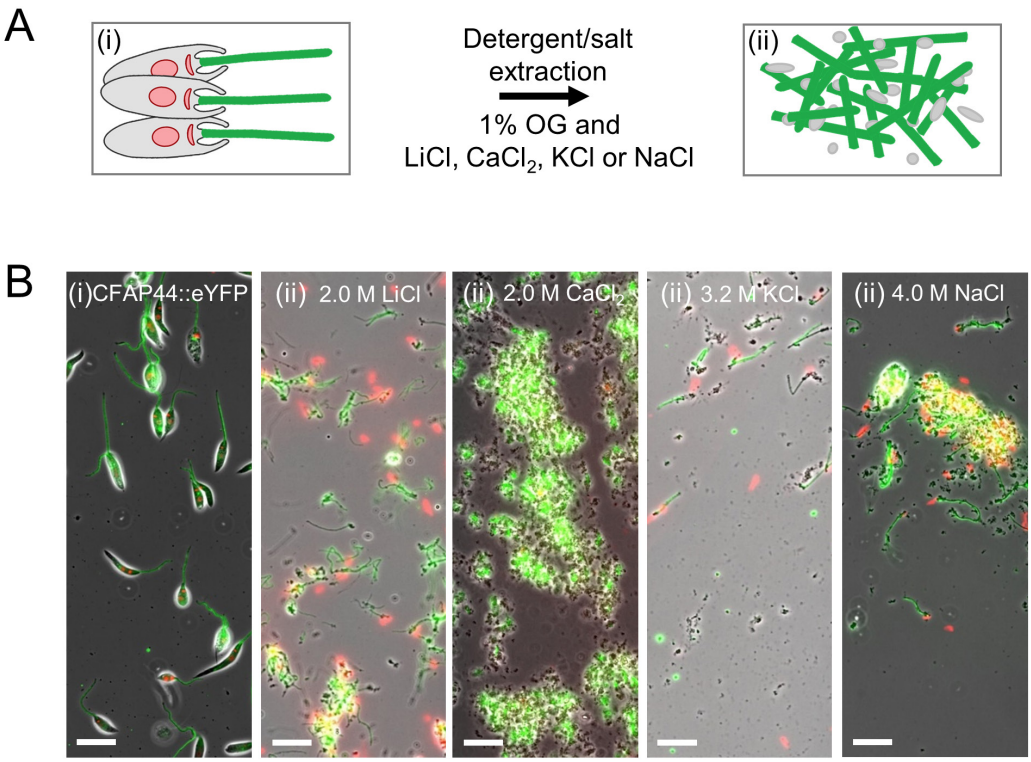
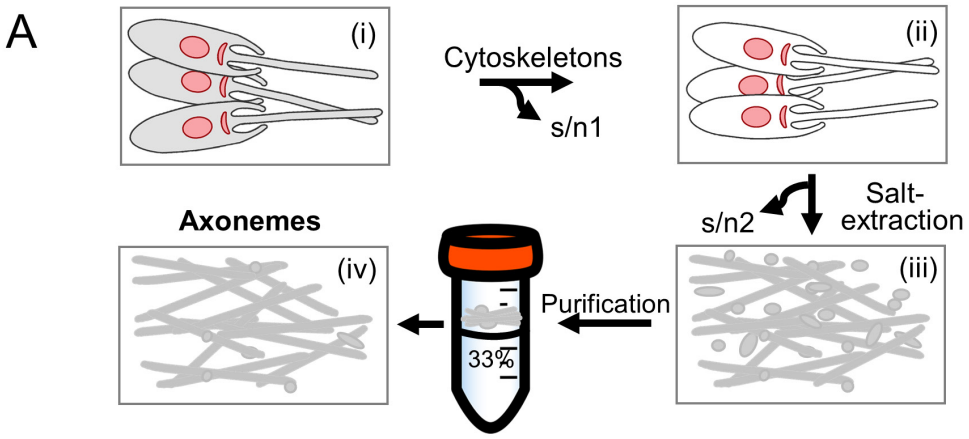
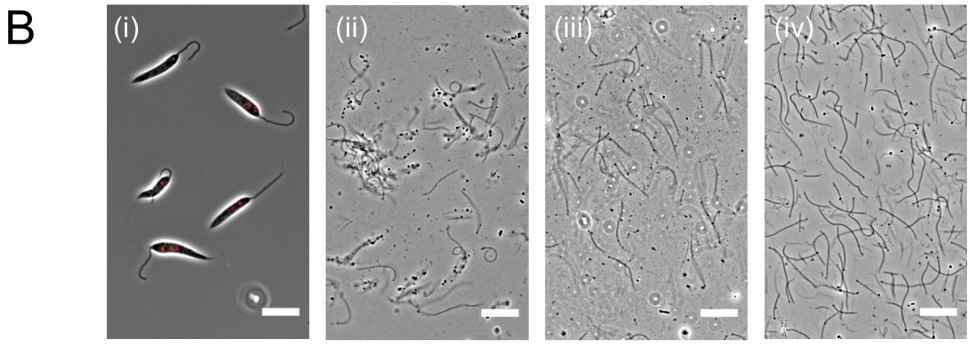
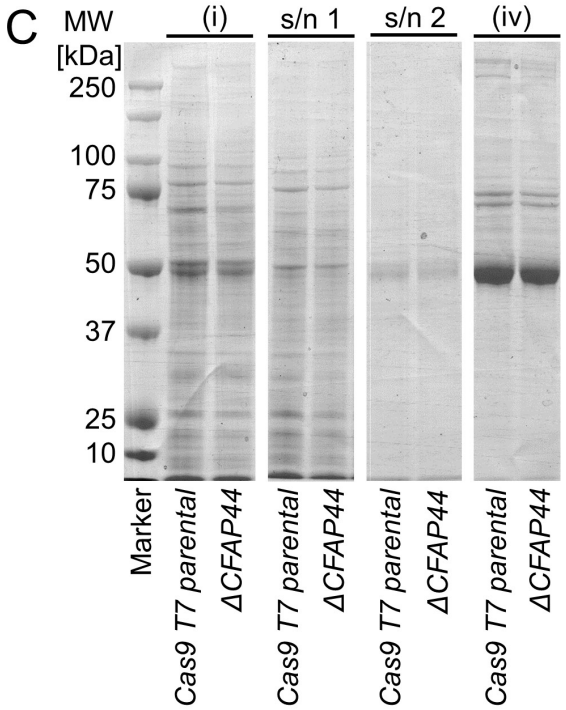


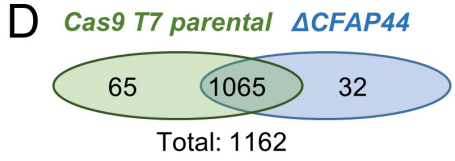
Figure 39. Solubilisation of CFAP44.

(A) Overview of attempts to solubilize CFAP44 in *L. mexicana*. *CFAP44::eYFP* cells were detergent/salt extracted using octylglycoside (1% (w/v) final conc.) and various salts. **(B)** Micrographs show merged phase and Hoechst DNA stain (red) for each isolation stage (i and ii) depicted in **(A)**. **(i)** *CFAP44::eYFP L. mexicana* cells before detergent/salt extraction and **(ii)** cells after detergent/salt extraction using either 2 M LiCl, 2 M $CaCl_2$, 3.2 M KCl or 4 M NaCl. Scale bars is 10 µm.









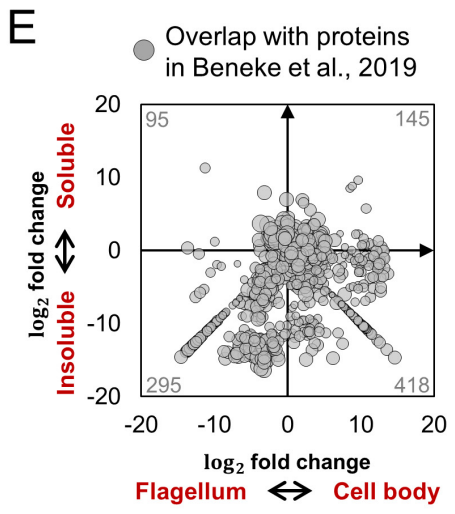


Figure legend on following page.

Figure 40. Salt extracted flagellar axonemes of *L. mexicana.*

(A) Overview of the salt extraction flagellar isolation protocol. Percentage sucrose concentration (w/v) and isolation of supernatants (s/n) are indicated. **(B)** Micrographs show merged phase and Hoechst DNA stain (red) for each isolation stage (i-iv) depicted in **(A)**. **(i)** *L. mexicana* cells before detergent extraction, **(ii)** cells after detergent extraction, **(iii)** axonemes after salt extraction and **(iv)** isolated salt extracted axonemes after differential centrifugation. Scale bars is 10 µm. **(C)** Protein gel stained with Coomassie Blue. Numbers on the left indicate molecular weight in kDa. Lane **(i)** and **(s/n 1)** were loaded with $4 \cdot 10^6$ cell equivalent. Loaded protein amounts for lane **(i)** and **(iv)** were as follows: Cas9 T7 parental **(i)** 2.02 µg and **(iv)** 0.74 µg, *ΔCFAP44* **(i)** 2.56 µg and **(iv)** 0.82 µg. 4 µg of each sample in lane **(iv)** were submitted for mass spectrometry analysis. **(D)** Venn diagram shows total number of all detected proteins (≥ 2 peptides detected, p-value > 0.95). **(E)** Total number of all detected proteins was overlapped with SINQ quantified proteins detected in Beneke et al. (2019b) and shown in Figure 6 (grey). The protein composition of isolated flagella (F) and deflagellated cell bodies (CB), as well as their further separation in detergent soluble (S) and insoluble (I) components, was analysed. Plotted on the X-axis is the log2 fold-change of the spectral index (C_S+C_I / F_S+F_I) and on the Y-axis the log2 fold change of the spectral index (F_S+C_S / F_I+C_I).

Proteins identified in the mutant and parental samples overlapped well with insoluble fractions analysed in Beneke et al. (2019b) (Figure 40 E) and included well characterised flagellar proteins, such as CPC, N-DRC, IDA, ODA, RSP and PFR proteins. Since the basal body remains attached to purified salt extracted axonemes (Robinson et al., 1991) a large number of proteins associated with the basal body and intraflagellar transport were also identified.

Label-free normalized spectral index quantitation method SINQ (Trudgian et al., 2011) was used to determine protein enrichments between both cell lines. 65 proteins were exclusively identified in the Cas9 T7 parental control cells, while 32 proteins were exclusively identified in *ΔCFAP44* mutants.

Proteins that were exclusively identified in *ΔCFAP44* mutants were generally low abundant proteins (2 or 3 detected peptides), with the exception of 48 detected peptides of beta tubulin (LmxM.08.1230) in *ΔCFAP44* mutants. However, only 1 out of 48 peptides was unique (spectral count 1), showing that proteins exclusively identified in *ΔCFAP44* mutants are likely to represent technical limitations of mass spectrometry analysis.

Analysis of proteins missing from *ΔCFAP44* mutants confirmed loss of CFAP44 itself and of its paralogue CFAP43. This confirms previous observations made for deletions and mutations on the T/TH complex in *Chlamydomonas* and *Tetrahymena* (Fu et al., 2018; Kubo et al., 2018; Urbanska et al., 2018). Interestingly, the IDA f/I1 intermediate chain IC140 was also absent in *ΔCFAP44* mutants and the IDA beta and alpha heavy chain were significantly reduced in *ΔCFAP44* mutants (α -55,6%; β -61.3%), while ODA, CPC, RSP and the main component of the PFR, PFR2, showed only little changes (Table 3).

		Salt extracted axonemes of *Cas9 T7 parental* vs *ΔCFAP44* in this study			Flagellar promastigote proteome of *L. mexicana* (Beneke et al., 2019a)			
GeneID	Protein	Unique peptides	SINQ spectral count		SINQ spectral count			
			Cas9 T7 parental	*ΔCFAP44*	Cell body insoluble	Cell body soluble	Flagellar insoluble	Flagellar soluble
LmxM.14.1430	CFAP44	38	47	0	5	0	77	0
LmxM.30.3150	CFAP43	31	38	0	9	0	107	0
LmxM.27.1630	IC140	3	3	0	0	0	14	0
LmxM.24.1310	LAX28	5	7	0	1	0	8	0
LmxM.32.2630	IC138	Not detected			0	0	19	0
LmxM.33.3880	DHC 1α	90	99	42	78	4	256	7
LmxM.23.1310	DHC 1β	62	75	29	85	7	238	5
LmxM.13.1650	αDHC/ODA11 (LmβDHC)	121	131	118	201	60	433	143
LmxM.25.0980	γDHC/ODA2 (LmαDHC)	121	135	132	226	45	546	108
LmxM.16.1430	PFR2	68	294	341	153	83	340	155
LmxM.28.0050	DRC2	25	20	28	4	0	28	0
LmxM.29.1810	Hydin	6	4	4	5	0	104	2
LmxM.20.1400	PF16	20	32	32	9	0	36	3
LmxM.18.1010	LC7	Not detected			4	1	9	6
LmxM.27.0520	RSP3	17	19	22	3	1	27	1
LmxM.24.0840	p28	5	5	4	5	0	14	0

Table 3. Quantitative proteomics of *ΔCFAP44* mutant derived flagellar axonemes.

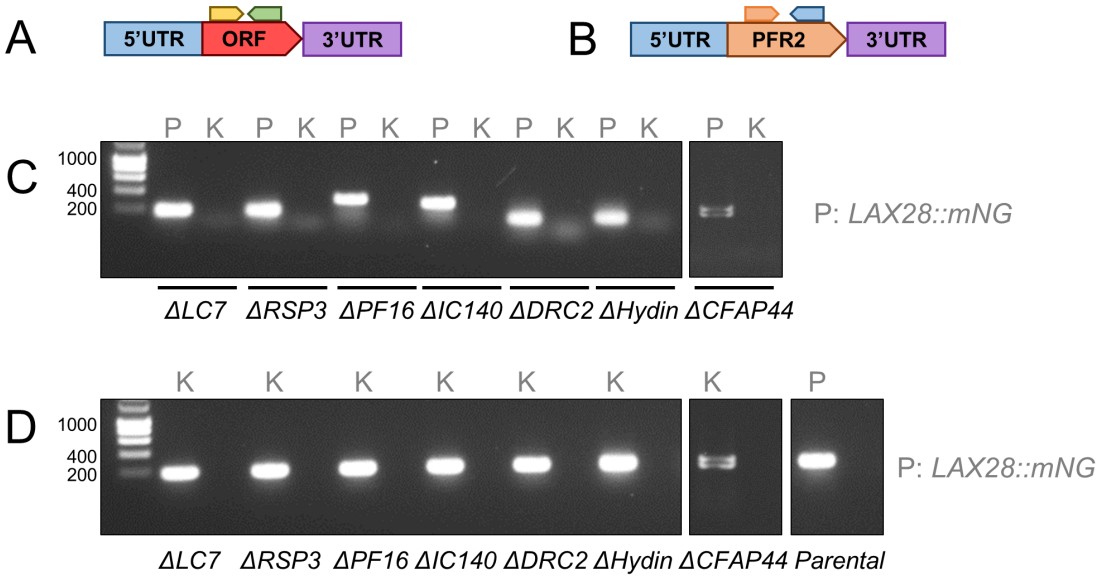
Table shows the spectral index of SINQ quantified detected proteins in this study and in Beneke et al. (2019b). Unique peptides are shown only for data from salt extracted axonemes.

Screening of other axonemal proteins against the flagellar proteome defined in Beneke et al. (2019b) showed that LmxM.24.1310 was not only detected exclusively in control cells (Table 3), but also highly enriched in the promastigote flagellum of *L. mexicana* (Beneke et al., 2019b). LmxM.24.1310, here named LAX28 (*Leishmania* axonemal protein, 28 kDa), is a hypothetical protein of 254 amino acids, with no defined domains and had not previously been linked to the T/TH complex or IDA function.

4.2.2 Axonemal localisation of LAX28 depends on CFAP44 and IC140

Proteomic analysis predicted that LAX28 depends on the presence of CFAP44. To provide further evidence for this dependency and test whether LAX28 is depending on other axonemal proteins, LAX28 was *in situ* tagged with mNeonGreen (mNG) and subjected to deletion of either CFAP44, IC140, DRC2, Hydin, LC7, PF16 or RSP3. Deletion of the ORF was confirmed by PCR verification (Figure 41 A, B, C and D).

While *LAX28::mNG* localized to the axoneme of the promastigote flagellum of *L. mexicana*, it dis-localized in *ΔCFAP44* and *ΔIC140* mutants. No *LAX28::mNG* fluorescent signal was detected in the flagellum of *ΔCFAP44* and *ΔIC140* mutants, but instead a strong signal in the cell body was observed. In contrast other deletions such as *ΔDRC2*, *ΔHydin*, *ΔLC7*, *ΔPF16* or *ΔRSP3* did not affect the localization of *LAX28::mNG* (Figure 42 A). This confirmed the observed loss of LAX28 in *ΔCFAP44* flagellar protein samples (Table 3) and showed that IC140 deletion affects LAX28 localization in a similar way.



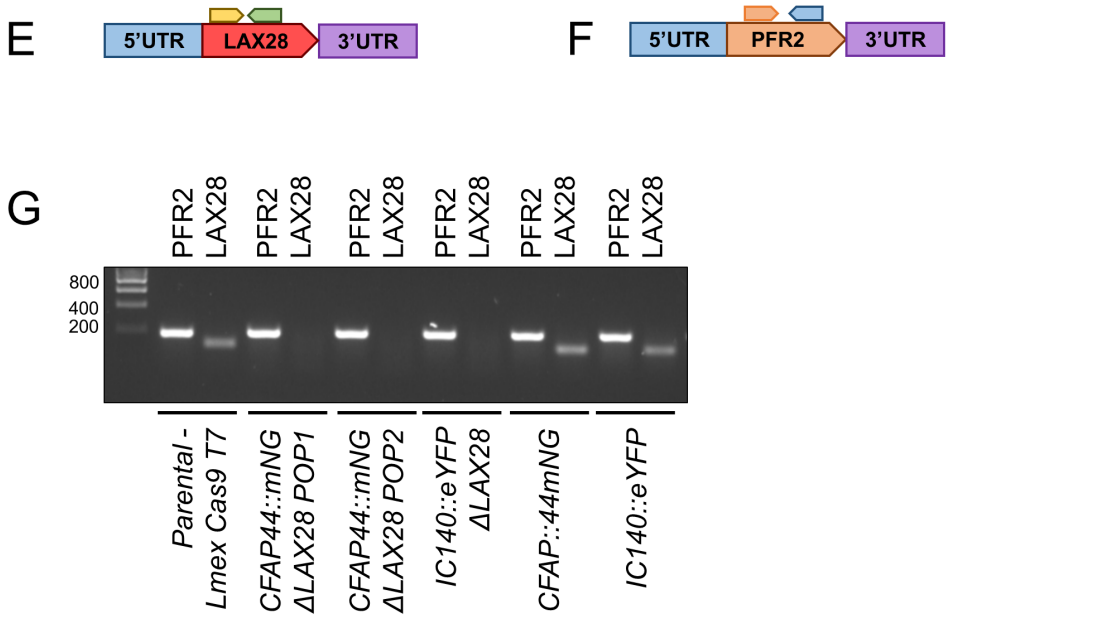


Figure legend on following page.

Figure 41. Null mutant verification.

(A - D) Cartoons illustrate PCRs for null mutant verification for knockouts in Figure 42 A: **(A)** amplification of a fragment of the target gene ORF, **(B)** amplification of a fragment of the PFR2 ORF. **(C)** Amplicons from the target gene ORF run on agarose gel. Each ORF was amplified in the putative knockout cell line (K) and the parental *LAX28::mNG* cell line (P). (**D**) Presence of the PFR2 ORF PCR product was confirmed for all transfected cell lines and the parental *LAX28::mNG* cell line (P).

(E - G) Cartoons illustrate PCRs for null mutant verification for knockouts in Figure 42 B, C, and Figure 46: **(E)** amplification of a fragment of the LAX28 ORF, **(F)** amplification of a fragment of the PFR2 ORF. **(G)** PCR products from amplification of a fragment of the LAX28 ORF and of the PFR2 ORF are run on agarose gel for parentals of *L. mexicana* Cas9 T7, *CFAP44::mNG* and *IC140::eYFP*, as well as for *ΔLAX28* mutants of *CFAP44::mNG* and *IC140::eYFP*. Presence of the PFR2 ORF PCR product was confirmed as in (**D**). *ΔLAX28* mutants of *CFAP44::mNG* were produced in parallel twice, but were indistinguishable in their phenotype. Population 1 (POP1) was taken for all experiments shown in this study.

For **(C)** and **(G)** no PCR products were obtained in mutant cell lines for amplification of the target gene ORF. Fainter bands below the target gene amplicon are likely primer dimers. The gene names indicate the target gene ORF.

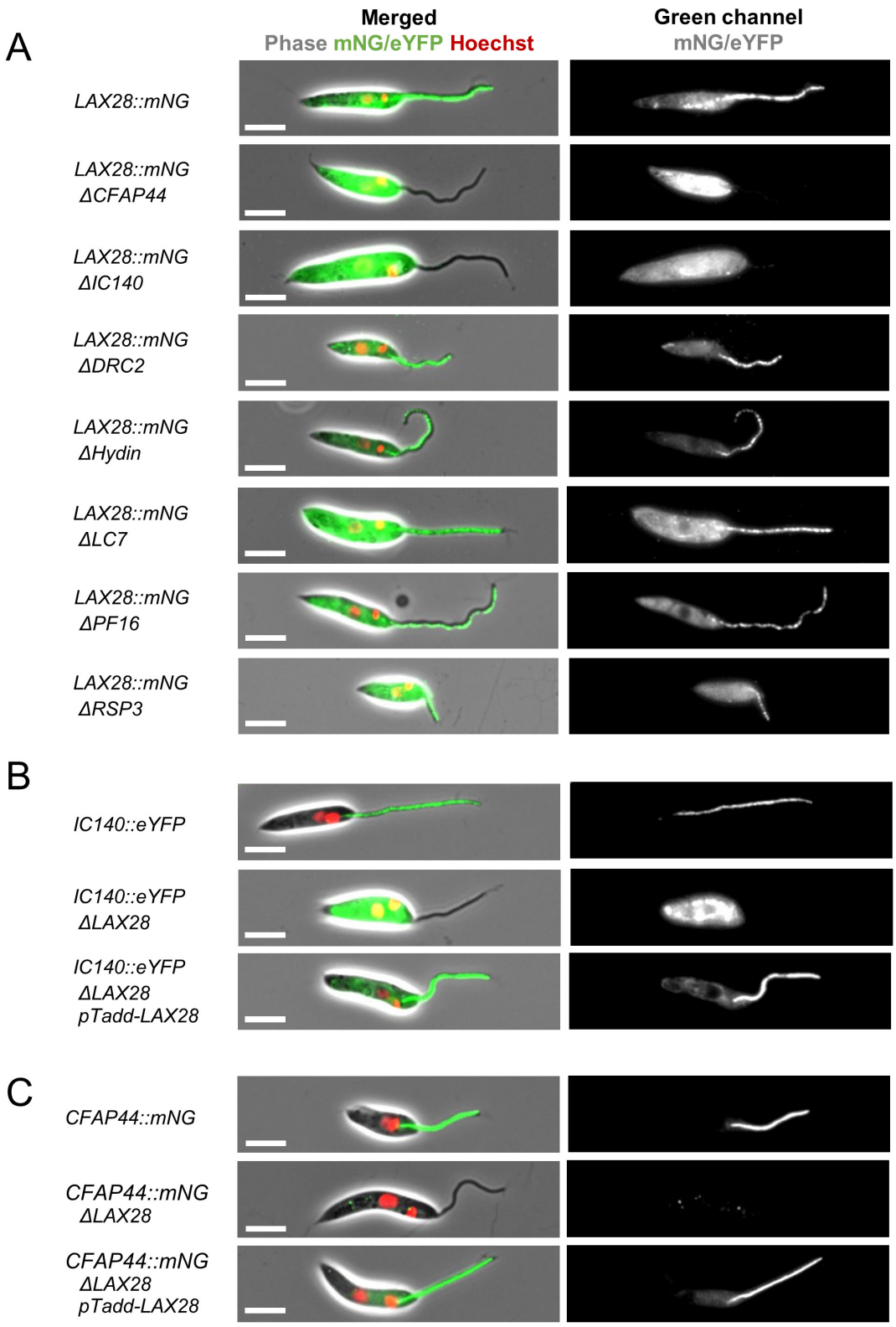


Figure legend on following page.

Figure 42. Testing localization dependencies for LAX28.

Fluorescence micrographs showing **(A)** LAX28 tagged with mNG and mutant cell lines of LAX28 tagged cells (green, *LAX28::mNG* signal), **(B)** IC140 tagged with eYFP and *ΔLAX28* mutant and addback cell lines of IC140 tagged cells (green, *IC140::eYFP* signal) and **(C)** CFAP44 tagged with mNG and *ΔLAX28* mutant and addback cell lines of CFAP44 tagged cells (green, *CFAP44::mNG* signal). In all panels red, Hoechst-stained DNA and scale bar, 5 μm.

4.2.3 Axonemal localisation of CFAP44 and IC140 depends on LAX28

Since these results showed that the axonemal localization of LAX28 depends on presence of CFAP44 and IC140, it was next tested whether the localization of CFAP44 and IC140 also depend on LAX28. LAX28 was subjected to deletion in a cell line expressing either *CFAP44::mNG* or *IC140::eYFP* (Figure 42 B and C). The deletion of LAX28 was confirmed by PCR (Figure 41 E, F and G) and an addback plasmid was transfected to allow episomal expression of LAX28 in null mutants (Figure 42 B and C). While *CFAP44::mNG* and *IC140::eYFP* both localized to the axoneme, they each dis-localized in *ΔLAX28* mutants, showing no signal in the promastigote flagellum. Generated addback cell lines restored both flagellar protein localizations (Figure 42 B and C). Interestingly, deletion of *ΔLAX28* in *IC140::eYFP* cells gave strong cell body signals, while the fluorescent signal of *CFAP44::mNG* in *ΔLAX28* mutants fall below background. This could be either caused by a technical artefact due to differences between the mNG and eYFP tag or it may suggest differences in turnover of unassembled CFAP44 or IC140 proteins, which is discussed in section 4.3.1.

Western blot analysis of the Myc epitope tag, contained within the *CFAP44::mNG* and *IC140::eYFP* fusion protein, showed comparable band intensities between the tagging, deletion and addback mutants (Figure 43).

Taken together these results indicate that LAX28 is required for assembly of CFAP44 and IC140 into the axoneme. To gather further evidence for a potential IDA f/I1 LAX28 association, flagellar ultrastructures of *ΔLAX28* and parental cells were compared next by using TEM. Multiple cross sections of the extracellular part of the promastigote flagellum were imaged and average rotated as previously described (Gadelha et al., 2006; Wheeler et al., 2015a).

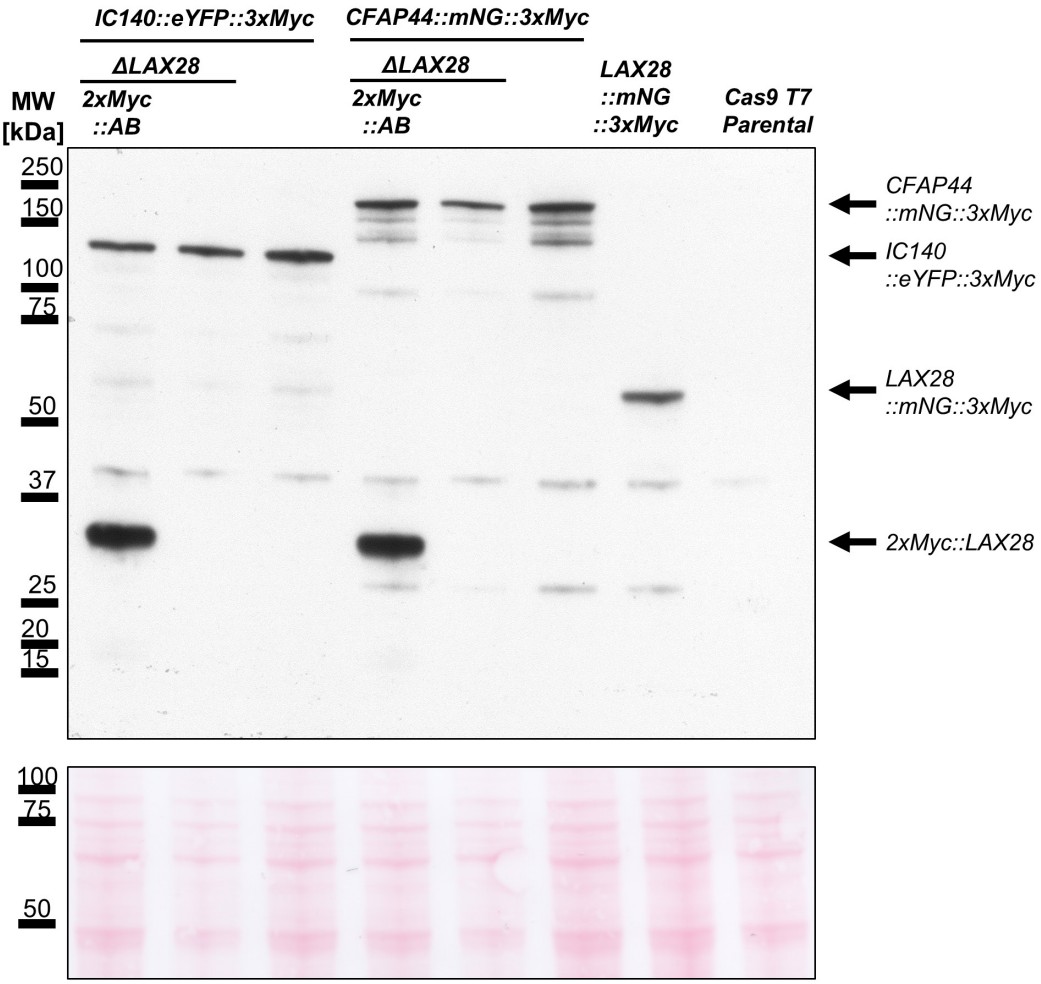


Figure 43. Western blot analysis.

Western blots of parental *L. mexicana Cas9 T7*, *LAX28::mNG*, *CFAP44::mNG* and *IC140::eYFP*, as well as *ΔLAX28* mutants of *CFAP44::mNG*, *IC140::eYFP* and their addbacks. Upper panel: blots probed with anti-myc antibody. Signal of myc epitope tagged proteins (*CFAP44::mNG* and *IC140::eYFP*) stays present in *ΔLAX28* cell lines. The addback of *ΔLAX28* mutations is detected by using a double myc epitope tag. Arrows on the right indicate fusion protein detected by using myc antibody. Molecular weights (MW) are indicated on the left in kDa. Lower panel: Ponceau red stained membrane.

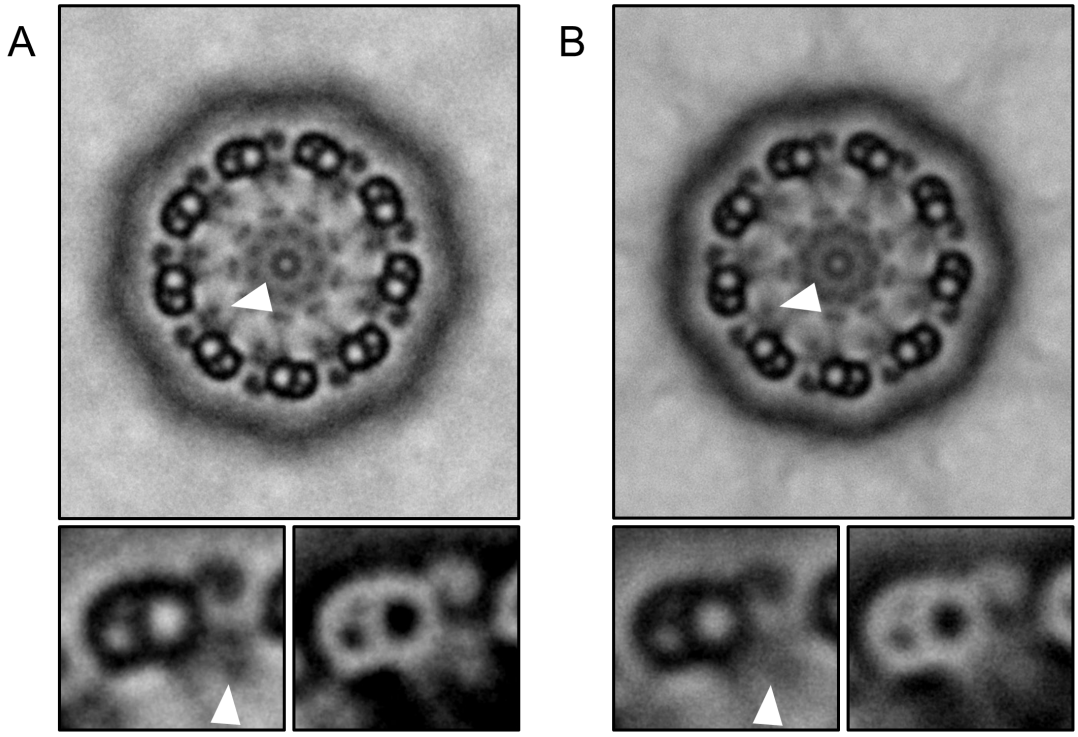


Figure 44. TEM of *ΔLAX28* mutants.

Rotational averages of transmission electron micrographs. 9 rotational averages are stack-averaged for each **(A)** *L. mexicana* Cas9 T7 parental cells and **(B)** *ΔLAX28* mutants. Arrows point to the position of the IDA.

A clear lack of electron density at the location of the IDAs was observed (Figure 44 A and B), suggesting that *LAX28* is indeed affecting IDA localization within the flagellar axoneme.

These results indicate mutual dependencies between LAX28, IC140 and CFAP44. This raised the questions whether LAX28 is required for assembly of T/TH and IDA proteins into the axoneme, or whether the absence of LAX28 renders these structures instable. *Leishmania,* like trypanosomes, offers the opportunity to examine flagella that are being actively assembled alongside the cell's older flagella (Wheeler et al., 2011). Therefore, *LAX28::mNG/ΔCFAP44*, *LAX28::mNG/ΔIC140*, *CFAP44::mNG/ΔLAX28* and *IC140::eYFP/ΔLAX28* mutant cell lines, their parental tagging and addback cell lines, were imaged at different cell cycle stages. Cell cycle stages included the end of G_1 or the beginning of S phase (1K 1N 1F), the beginning of new flagellum growth (1K 1N 2F) and the initiation of cytokinesis (2K 2N 2F) (Figure 45 and Figure 46). This allowed tracking of LAX28/CFAP44 or LAX28/IC140 dependencies in a growing new flagellum. If loss of LAX28 only destabilizes and not affects assembly of the CFAP44 and IC140 protein, as well as their associated structures like the T/TH and IDA f/I1 complex, one might expect to find a stronger fluorescent signal in a growing new flagellum compared to an old one. However, no difference of fluorescent signal over the cell cycle was identified in *ΔLAX28*, *ΔCFAP44* and *ΔIC140* mutants (Figure 45 and Figure 46), meaning there was no visual difference in a growing new flagellum compared to their corresponding old flagellum. This suggests that loss of LAX28 prevented assembly of CFAP44 or IC140 into the axoneme (Figure 45), as well as CFAP44 and IC140 deletions prevented assembly of LAX28 into the axoneme (Figure 46).

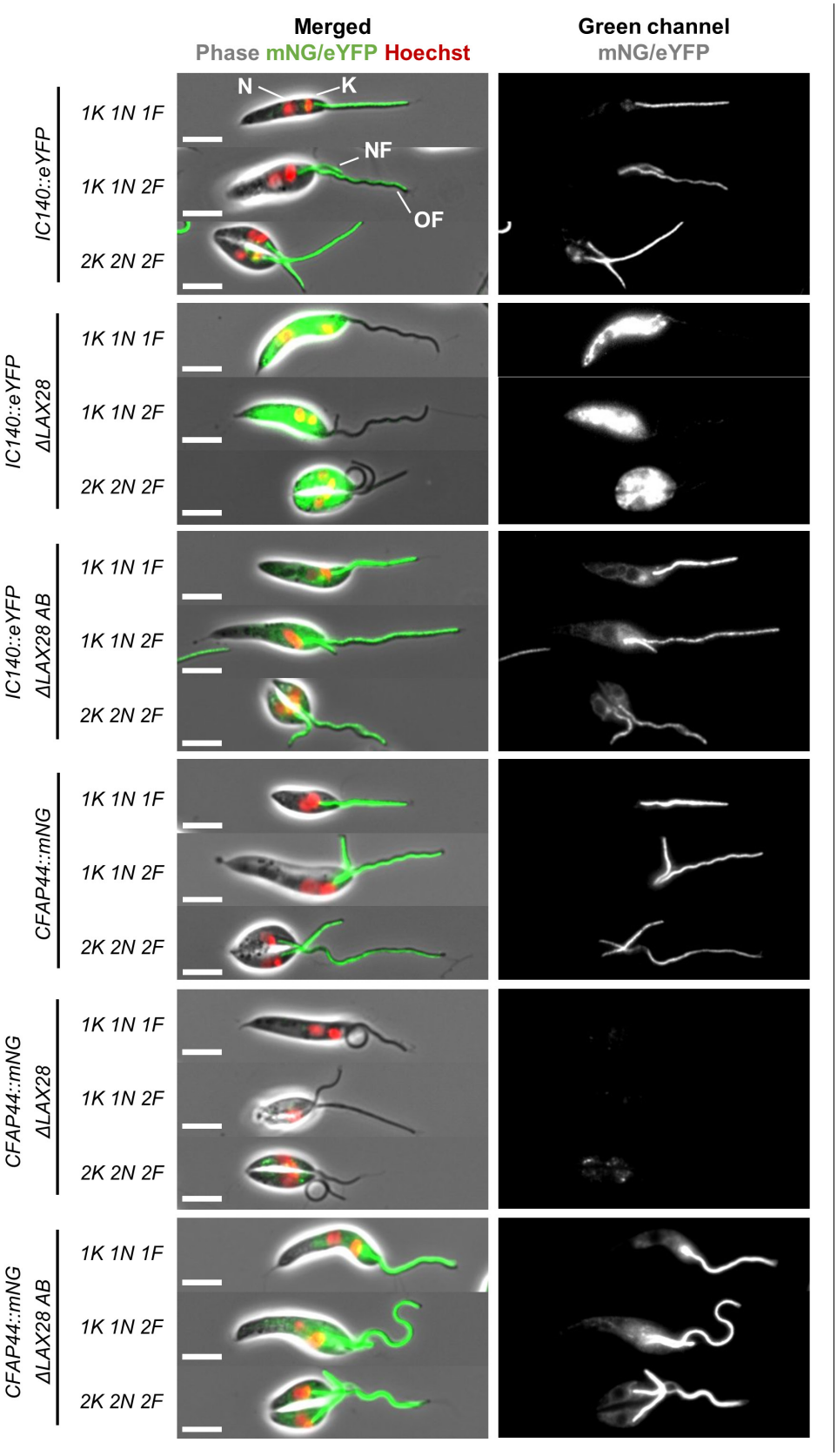


Figure legend on following page.

Figure 45. Tracking IC140::eYFP and CFAP44::mNG signal over the *Leishmania* cell cycle in *ΔLAX28* null mutants.

Fluorescence micrographs showing IC140 tagged with eYFP, CFAP44 tagged with mNG, their *ΔLAX28* mutant cell lines and addbacks (AB). Cell lines are shown in three cell cycle stages during division highlighted by: K (kinetoplast), N (nucleus), F (flagellum), NF (new flagellum) and OF (old flagellum). Green, *IC140::eYFP* or *CFAP44::mNG* signal; red, Hoechst-stained DNA. Scale bar is 5 μm.

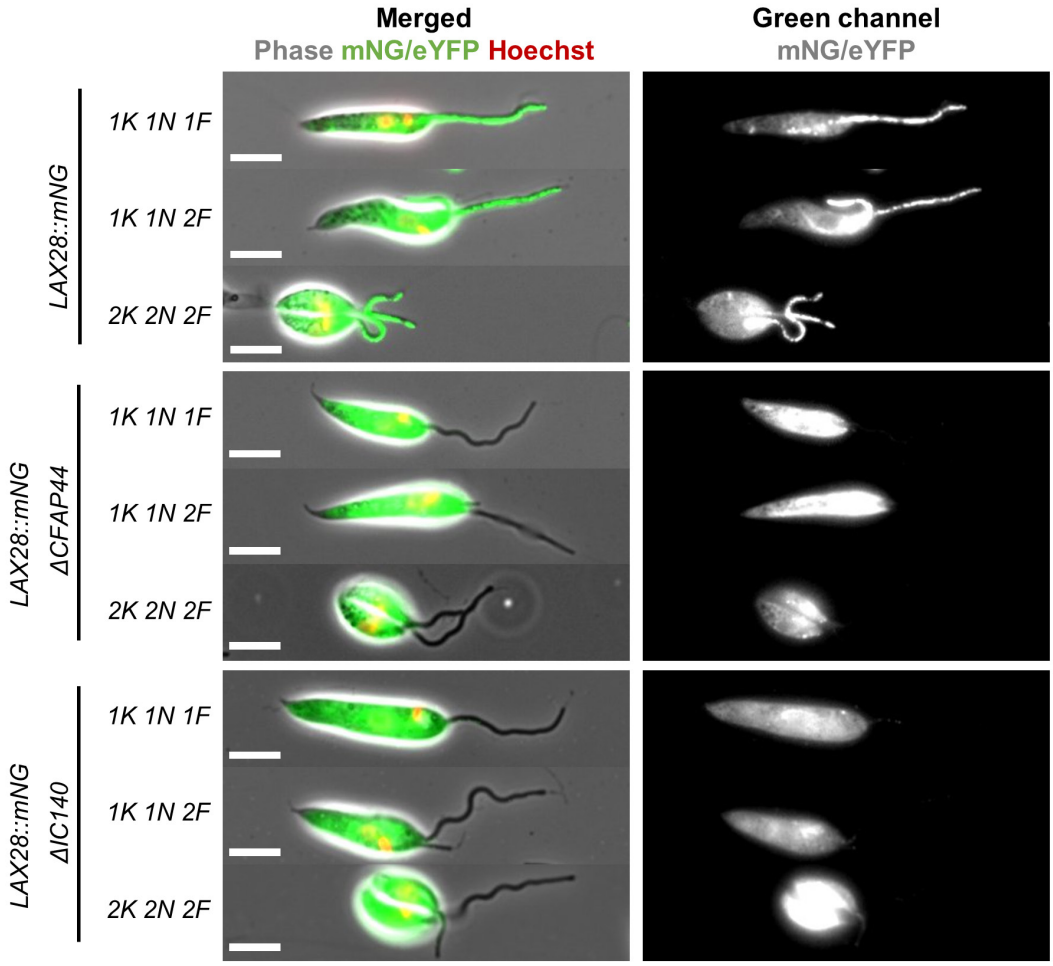


Figure 46. Tracking LAX28::mNG signal over the *Leishmania* cell cycle in *ΔIC140* and *ΔCFAP44* null mutants.

Fluorescence micrographs showing LAX28 tagged with mNG, their mutant cell lines of *ΔCFAP44* and *ΔIC140* and addbacks (AB). Cell lines are shown in three cell cycle stages during division highlighted by: K (kinetoplast), N (nucleus) and F (flagellum). Green, LAX28::mNG signal; red, Hoechst-stained DNA. Scale bar is 5 µm.

In *Chlamydomonas* IC140 was previously shown to be required for assembly of the IDA f/I1 complex (Perrone et al., 1998; Viswanadha et al., 2014) and FAP44 shown to be required for assembly of the T/TH complex (Fu et al., 2018; Kubo et al., 2018; Urbanska et al., 2018). Since LAX28 is required for axonemal localization of IC140 and CFAP44, this suggests that LAX28 is needed for assembly of the T/TH and IDA f/I1 complex.

4.2.4 LAX28 is an inner arm dynein component required for flagellar function

Based on proteomic analysis, protein localisations and mutual dependencies described above, mutants of CFAP44, IC140 and LAX28 should exhibit similar phenotypes linked to impaired IDA f/I1 function. To test this, phenotypes of *ΔCFAP44*, *ΔCFAP43*, *ΔIC140* and *ΔLAX28* were analyzed by measuring swimming speed and directionality of mutants using motility particle tracking analysis (Wheeler, 2017). Phenotypic analysis of mutants included also counting the proportion of flagellar curls within the cell population as defined in Figure 26. While *ΔCFAP44*, *ΔCFAP43* and *ΔIC140* mutants were characterized before (Beneke et al., 2019b), *ΔLAX28* mutants were newly generated as described in chapter 2 and ORF deletion was confirmed by Illumina sequencing (Figure 47). An addback cell line of *ΔLAX28* was generated, allowing constitutive expression from an episome of LAX28.

All four mutants showed relatively high flagellar curling rates compared to the parental cell line and *ΔLAX28* addback cells (Figure 48 A). Motility tracking analysis showed also that these mutants have reduced swimming speed and directionality and cluster closely together when plotting speed vs directionality (Figure 48 B and C). The swimming speeds measured were *ΔLAX28* 3.7 µm/s ±

0.3 µm/s, *ΔCFAP44* 2.9 µm/s ± 0.03 µm/s, *ΔCFAP43* 2.7 µm/s ± 0.1 µm/s and *ΔIC140* 2.9 µm/s ± 0.2 µm/s (Figure 48 B and C). By contrast *ΔLAX28* addback cells expressed normal swimming behaviors, similar to parental cells (Figure 48 B and C).

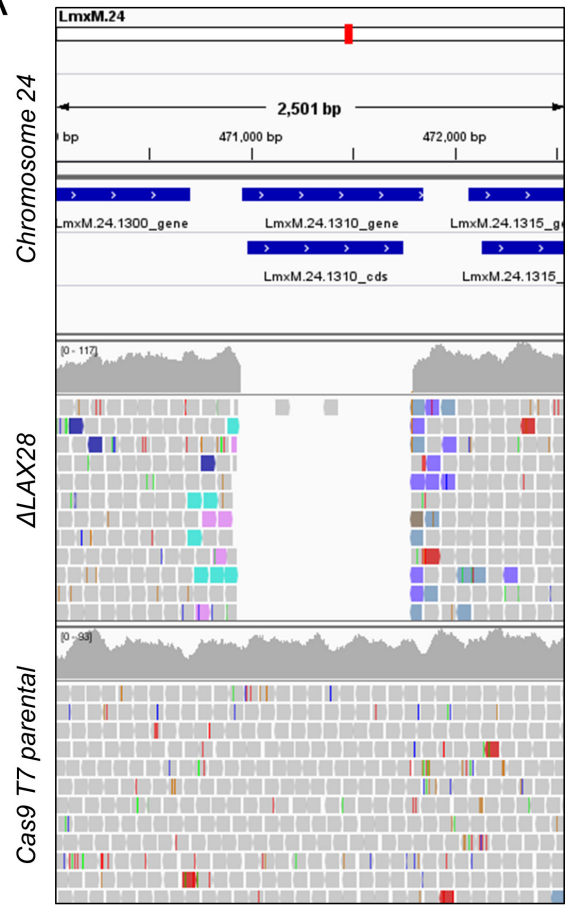
As described in chapter 3, flagellar curling was previously also observed for deletion of CPC proteins PF16 (Beneke et al., 2017) and hydin (Beneke et al., 2019b). *ΔPF16* and *Δhydin* mutants are paralyzed and were shown to have a mean swimming speed of 2.0 – 2.3 µm/s (*ΔPF16* 2.0 µm/s ± 0.1 µm/s; *Δhydin* 2.3 µm/s ± 0.01 µm/s) (Beneke et al., 2019b) (Figure 48 B and C), which is lower than the T/TH and IDA f/I1 mutants measured here. Swimming properties of the IDA f/I1 and T/TH mutants are also different from previously reported uncoordinated motility defects, such as *ΔMBO2* mutants (Beneke et al., 2019b). These results show that LAX28 is essential for normal flagellar motility and indicate that the loss of flagellar motility in *ΔLAX28* knockout cells is most likely due to IDA f/I1 deficiency, affecting possibly assembly of the T/TH and IDA f/I1 complex.

4.2.5 LAX28 shows similarity to human testis expressed protein TEX47

LmxM.24.1310 (LAX28) is clearly conserved across kinetoplastids, but there is only low sequence identity of 31% between the syntenic orthologs from *L. mexicana* (LmxM.24.310) and *T. brucei* (Tb927.8.6920) (Figure 49 A). While simple protein BLAST did not identify any homologs in higher eukaryotes, Position-Specific Iterated BLAST (PSI BLAST; (Altschul et al., 1997)) search identified the homologue TEX47 in humans after running three blast iterations, showing low identity (16%) between both proteins (Figure 49 B). The reciprocal blast search ranked LmxM.24.1310 as third highest-scoring hit in *L. mexicana*, while

LmxM.05.0890, a cleavage and polyadenylation specificity factor subunit, was ranked as highest. Further analysis showed that TEX47 is predominantly expressed in human testis (Fagerberg et al., 2014) and was identified in the human sperm tail proteome (Amaral et al., 2013). Interestingly, MOT7 identified in *Chlamydomonas reinhardtii* and predicted to be as well a possible T/TH complex component (Kubo et al., 2018) identifies TEX47 as PSI BLAST hit, but also with low sequence similarity (24% sequence identity). However, PSI BLAST failed to identify MOT7 when LmxM.24.1310 was used as query sequence or vice versa. While amino acid sequence similarity between LAX28, TEX47 and MOT7 is low, Phyre^2 modelling (Kelley et al., 2015) predicts similar 3D folds of an FAD-binding protein domain of the BLUF family (Gomelsky and Klug, 2002) for all three proteins. Even though a proper genetic characterization of TEX47 remains to be performed these preliminary data point to a possible role for TEX47 in human sperm motility.

A



B

ORF blast query	*ΔLAX28* Aligned number of reads to ORF query	*Cas9 T7 parental* Aligned number of reads to ORF query
LAX28	2	>750
Blasticidin	248	0
Puromycin	356	0
Neomycin	0	0

Figure 47. Illumina sequencing of *ΔLAX28* null mutants.

(A) Bam files of *ΔLAX28* mutants and *L. mexicana* Cas9 T7 parental cells were viewed in IGV genome browser (Robinson et al., 2011) (as described in Figure 17). In *ΔLAX28* mutants one set of paired reads aligned to the LAX28 locus, which is likely to represent cross contamination during genome sample preparation of 22 genomes at a time. No other technical explanation (e.g. error probed demultiplexing or poor read quality) was found. **(B)** Blast queries of the LAX28 and blasticidin-S deaminase, puromycin N-acetyltransferase and neomycin-kanamycin phosphotransferase resistance gene against *ΔLAX28* mutants and *L. mexicana* Cas9 T7 parental cells were performed. Since *ΔLAX28* was generated using blasticidin-S deaminase and puromycin N-acetyltransferase resistance genes, reads aligned in *ΔLAX28* mutants. Two reads of LAX28 were identified in *ΔLAX28* mutants. These two reads were those illustrated in **(A)** within the LAX28 locus.

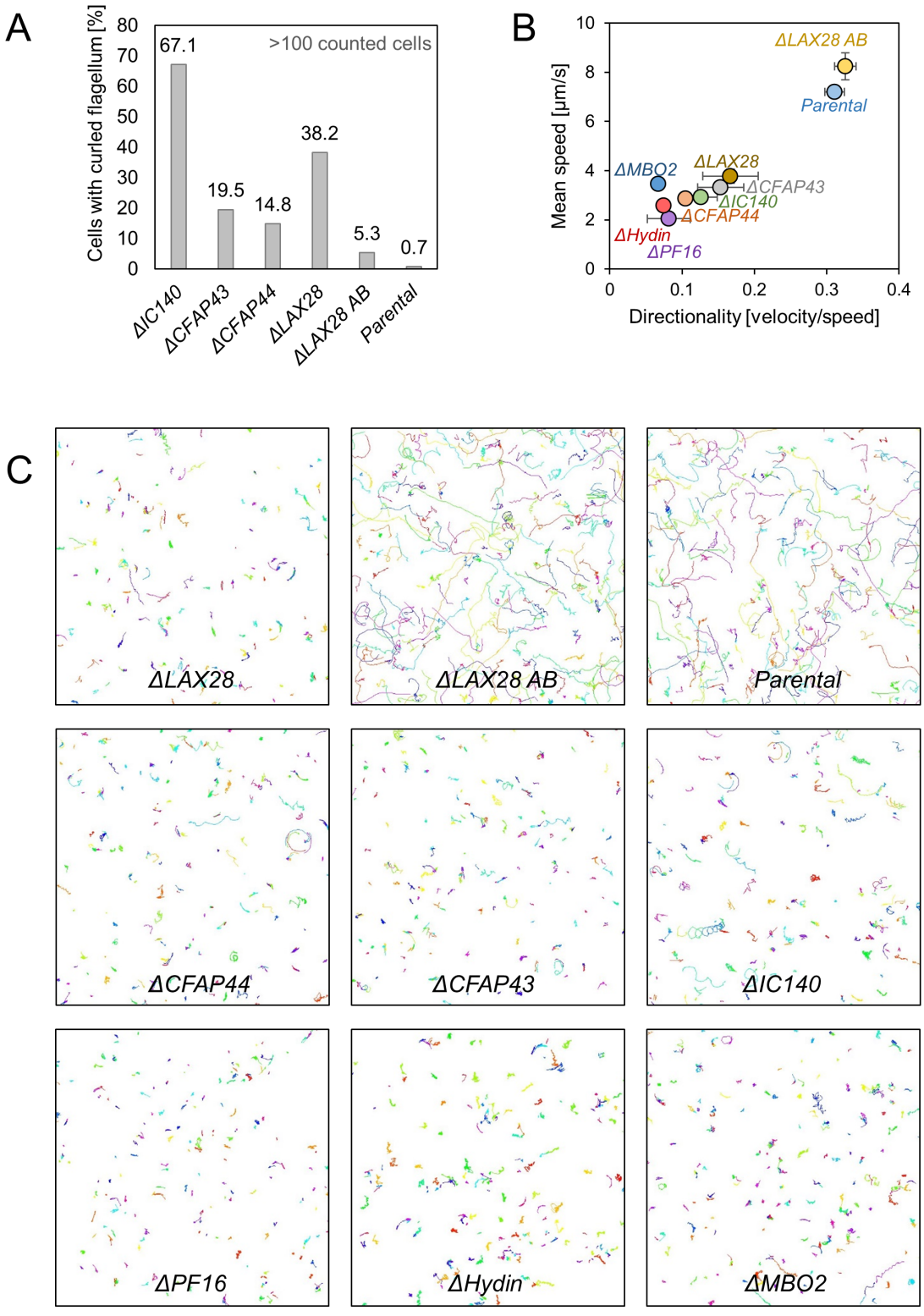


Figure legend on following page.

Figure 48. Flagellar curling and motility analysis of LAX28.

(A) Histogram showing the proportion of cells with curled flagella in *ΔIC140*, *ΔCFAP43*, *ΔCFAP44* and *ΔLAX28* mutants, the parental *L. mexicana* Cas9 T7 cell line and *ΔLAX28* addback (AB) cells. The flagellar curling rate is indicates above. **(B)** Deletion cell lines generated in Beneke et al., 2019 and the *ΔLAX28* mutant and its addback generated in this study were particle motility tracked (Wheeler, 2017) to measure their mean swimming speed and directionality (the ratio of velocity to speed) in triplicates. The average triplicate swimming speed was plotted against the average triplicate directionality of each mutant. Error bars represent the standard deviation of the three replicates. As grouped in Beneke et al., 2019, three of four main mutant phenotype clusters are represented here: paralysed (*ΔPF16*, *ΔHydin*), uncoordinated (*ΔMBO2*) and slow (*ΔIC140*, *ΔCFAP43*, *ΔCFAP44* and *ΔLAX28*). **(C)** Swimming paths used for extracting mean swimming speed and directionality are shown for cell lines in **(B)**. 200 tracks are shown for each cell line.

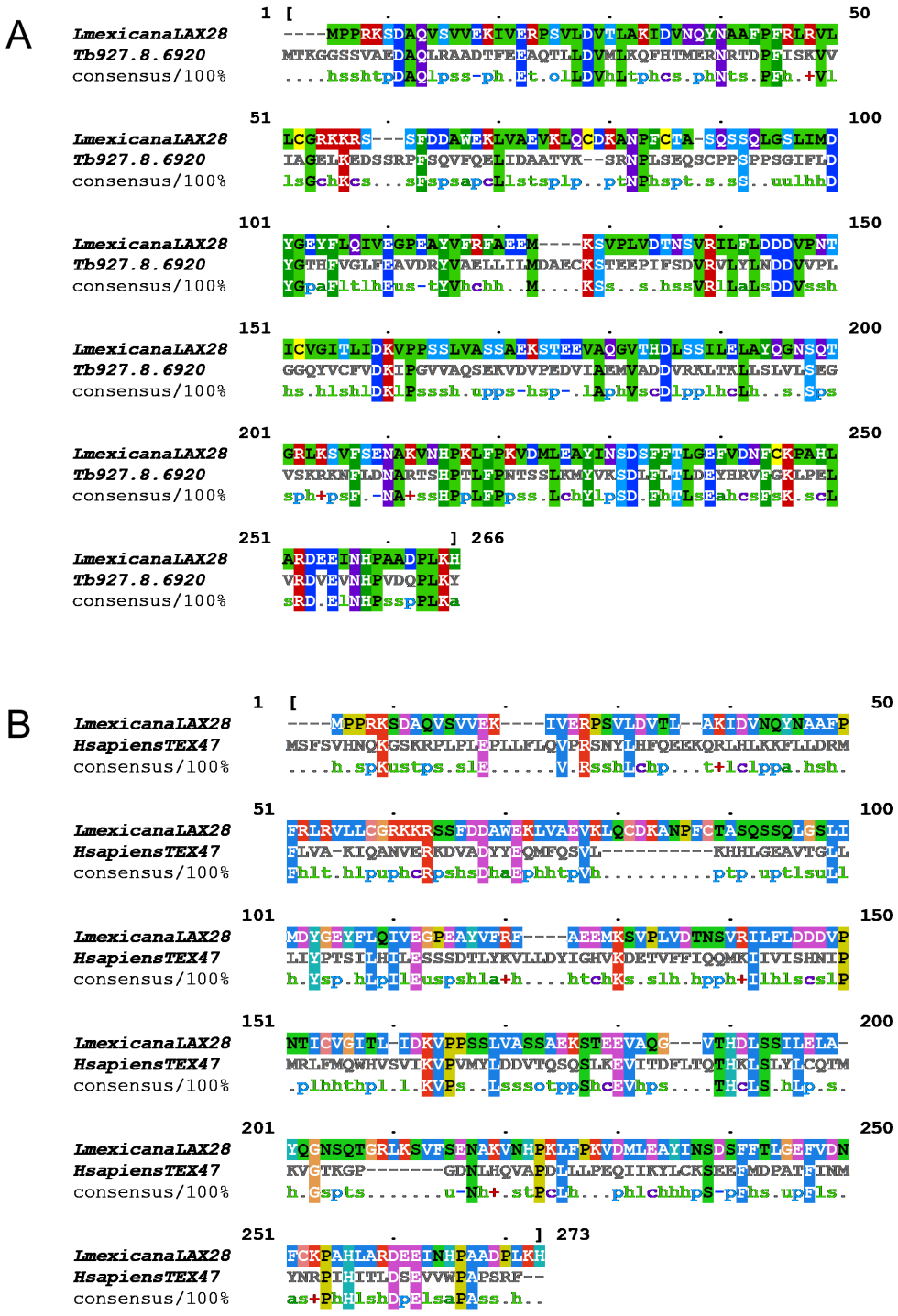


Figure 49. Protein alignments of LmxM.24.1310 (LAX28).

Clustal omega DNA alignment (Sievers et al., 2011) viewed by Mview (Brown et al., 1998) of **(A)** LAX28 and its syntenic orthologue in *T. brucei* Tb927.8.6920 and **(B)** LAX28 and its potential orthologue in *H. sapiens* TEX47.

4.3 Discussion

4.3.1 LAX28 is an IDA component required for assembly of IDA f/I1

The scope of the work presented in this chapter was to identify new T/TH complex interactors in *Leishmania* in order to find new possible regulators of flagellar curling and IDA activity. Using fluorescent reporters showed that LAX28 is an axonemal protein, dependent for its flagellar localization on the presence of CFAP44 and IC140. Interestingly, axonemal localization of CFAP44 and IC140 also depends on the presence of LAX28. This indicates that LAX28 is required for incorporation of the IDA f/I1 and its T/TH complex within the 96 nm repeat. Consequently, the electron density is reduced at the position of the IDA in TEM micrographs of *ΔLAX28*. Tracking *in situ* tagged IC140 and CFAP44 proteins in *ΔLAX28* mutants through the cell cycle provided further evidence for a role of LAX28 in IDA f/I1 assembly (Figure 45 and Figure 46).

Since IC140 does not localize to the flagellum in *ΔLAX28* mutants (Figure 42 B and Figure 46) it is likely that the axoneme is deficient for the IDA f/I1 complex. In *Chlamydomonas* IC140 was shown to be required for assembly of both IDA f/I1 heavy chains (DHC 1α, DHC 1β) and the IDA f/I1 ICLC complex (Perrone et al., 1998; Viswanadha et al., 2014). IC140 binds directly to both microtubules and plays a critical role in assembly, axonemal targeting and regulation of the I1 dynein complex (Hendrickson et al., 2013). IC140 and both heavy chains form a 20S complex and preassemble in the cytoplasm before being transported by IFT proteins to the distal end of a growing flagellum (Viswanadha et al., 2014). If a similar mechanisms exists in *Leishmania*, the strong cell body signal in *IC140::eYFP* cells in the absence of LAX28 (Figure 45 B) may show the

preassembled IDA f/I1 20S complex that can't be incorporated at the growing flagellar tip. Interestingly, knockout of IC140 or CFAP44 in *LAX28::mNG* cells results in a very similar cell body signal localization (Figure 42 A and Figure 46), which is in contrast to knockout of LAX28 in *CFAP44::mNG* cells where no signal above background remains (Figure 42 C and Figure 45). If technical variations between using a mNeonGreen tag and an eYFP tag are excluded, this could indicate that CFAP44 is not assembled into the IDA f/I1 20S complex in the cytoplasm or is subject to more rapid turnover in cells that lack LAX28.

Further evidence of a functional link between LAX28, IC140 and CFAP44 proteins in *Leishmania* is reflected by the observed reduced swimming speed and directionality of mutants, as well as observed increased flagellar curling rate in mutants compared to parental cells. While measured swimming speed and directionality is very similar between all three mutants, flagellar curling rates differ. Although *ΔLAX28* cells (38% curling) exhibit a stronger curling phenotype than *ΔCFAP44* mutants (14.8% curling), they do not fully phenocopy the *ΔIC140* cells (67% curling). While the cause of this remains elusive, one possible explanation for this discrepancy may be that depletion of LAX28 only reduces IC140 protein to levels below the detection limit of the microscope. Consequently, *ΔLAX28* mutants might exhibit lower curling rates than *ΔIC140* mutants, as small amounts of IC140 protein in the *ΔLAX28* mutant axoneme may remain. Alternatively, the loss of LAX28 may reflect a mixture of the observed *ΔIC140* and *ΔCFAP44* phenotype or depletion of LAX28 affects also other IDA-related structures in addition to IDA I1, which counteract the effect of IC140 loss, resulting in an intermediate rate of curling. Taken together this suggest that LAX28 is part of the T/TH and IDA f/I1 complex and required for its incorporation within the flagellar axoneme (Figure 50).

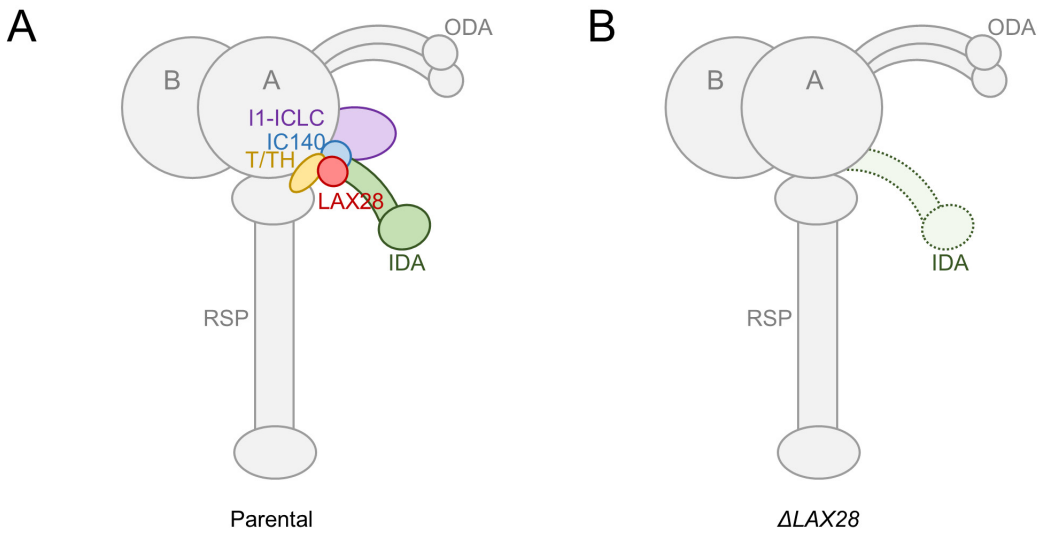


Figure 50. Model: LAX28 is required for assembly of the IDA f/I1 complex within the flagellum.

Model for LAX28 mediated assembly of the T/TH and IDA f/I1 complex. Schematic cross-sectional view on the outer microtubule doublets (A and B microtubule), RSP, ODA, IDA, T/TH and IDA f/I1 complex, as well as the proposed position of LAX28. **(A)** LAX28 (red) is required for assembly of the T/TH (yellow) and IDA f/I1 complex (violet and blue). All IDAs are fully functional when LAX28 is present. **(B)** In *ΔLAX28* mutants the T/TH and IDA f/I1 complex are not assembled within the *Leishmania* axoneme. Taken the findings in this chapter and phenotypic analysis in *Chlamydomonas* together suggest that also both heavy chains (DHC 1α, DHC 1β) of the IDA f/I1 complex are absent. It remains unknown how the other six IDA subspecies (dyneins a, b, c, d, e and g) are affected (green, dashed line).

4.3.2 IDA f/I1 may control activity of remaining IDA subspecies

However, the reduction of electron density signal at the location of IDAs in *ΔLAX28* mutants (Figure 44) is likely to reflect only an IDA f/I1 assembly defect and remaining signal might demonstrate that other IDA heavy chains within the 96 nm repeat are assembled. This aligns with previous findings. In *Chlamydomonas* each IDA is independently targeted to its location within the 96 nm repeat, causing gaps if assembly of one IDA fails. This gap will be reflected as reduced electron density in thin sections, but not complete absence of signal (Bui et al., 2012; Heuser et al., 2012; Piperno, 1990).

If functional units of IDAs remain and IDAs can substitute for each other one might expect that this gap at the IDA f/I1 position within the 96 nm repeat would result in reduced motility rather than paralysis. However, previous cryo-ET reconstruction studies suggested that IDA f/I1 is in an unique position to be an important regulatory hub for all other IDAs (Heuser et al., 2012), controlled through IC138 (Hendrickson et al., 2004) and other regulatory complexes such as the MIA complex (Yamamoto et al., 2013). Since dyneins are now believed to be controlled by inhibition (Lin and Nicastro, 2018), failure of IDA f/I1 assembly and/or its regulators may cause remaining IDA subspecies to switch to a permanent “on-stage”.

This is why *ΔCFAP43*, *ΔCFAP44*, *ΔIC140* and *ΔLAX28* show reduced swimming speed and directionality compared to parental cells, but are still more mobile than paralyzed central pair mutants. Consequently, IDA f/I1 may control the activity of remaining IDA subspecies, but does not completely abrogate their function when absent.

While it is clear that LAX28 is required for motility of the *Leishmania* flagellum, without using cryo-ET for 3D reconstruction of *ΔLAX28* mutant and parental axonemes little can be concluded for assembly defects of the other IDA heavy chains a, b, c, e, g and d.

4.3.3 The IDA f/I1 and T/TH complex shows diversity across flagellated eukaryotes

In *Chlamydomonas* and *Tetrahymena* the T/TH complex depends on the presence of the IDA f/I1 complex, but the IDA f/I1 complex does not depend on the presence of the T/TH complex (Fu et al., 2018; Kubo et al., 2018; Urbanska et al., 2018). Data presented in this chapter, indicates that this seems to be different to *Leishmania* since mass spectrometry data suggests absence of IC140 in salt extracted flagellar preparations of *ΔCFAP44* mutants and also a reduction of both IDA f/I1 α and β heavy chains (Table 3). In *Chlamydomonas* the T/TH complex is likely to be anchored to the axoneme through their C-terminal coiled-coil domain and might directly interact with the I1 α and β motor domains, IDA d and the base of radial spoke 3 (Fu et al., 2018). It is understood that the T/TH complex requires also dimerization of both paralogues CFAP43 and CFAP44 to be assembled, which agrees with loss of CFAP43 in the mass spectrometry analysis of here shown *ΔCFAP44* mutant axonemes in *Leishmania* (Table 3). However, it is not understood of what the so called “starlike morphology with four-to-five connections to surrounding axonemal structures” (Fu et al., 2018) consists. Components of these structures may remain unknown or may have been already partly identified. Since data for the T/TH and IDA f/I1 assembly shown here for *Leishmania* differ from published data in *Chlamydomonas* it is possible that surrounding structures show species-specific context.

Chlamydomonas proteins Cre10.g452250 (Kubo et al., 2018) and FAP244 were identified to be part of the T/TH complex in *Chlamydomonas*, but have no homologue in *Tetrahymena* and higher organisms (Fu et al., 2018). In contrast C/FAP44, C/FAP43 and also Fap57p (Urbanska et al., 2018) seem to be well conserved across flagellated eukaryotes, including kinetoplastids. Fap57p has been identified to be a homologue of CMF6 and CMF5 in *T. brucei* (Baron et al., 2007) and a partial CMF6 deletion mutant (CMF6 ORF still detectable by PCR after deletion) showed reduced swimming speed previously (Beneke et al., 2019b). Interestingly, the partial CMF6 deletion mutant showed high similarities with swimming patterns of *ΔLAX28*, *ΔCFAP44* and *ΔIC140* knockouts (Beneke et al., 2019b). Additionally, C/FAP44, C/FAP43, Fap57p and FAP244 share all a WD40 repeat at the N-terminal region and a coiled-coil domain at the C-terminus (Fu et al., 2018; Urbanska et al., 2018). Also C/FAP43 and FAP244 can functionally compensate for each other (Kubo et al., 2018). LAX28 and MOT7 (Kubo et al., 2018) share some similarities, but the sequence identity is too low to conclude homology with confidence. Following the blast analysis in this chapter, there is low homology between LAX28 and TEX47, as well as between MOT7 and TEX47. LAX28, TEX47 and MOT7 do not show signatures of known domains, but may be able to adopt a similar 3D fold, as Phyre^2 modelling predicted a beta-sheet with 4-5 strands and 2-3 alpha helices for all three proteins, as well as similarities to FAD-binding protein domains of the BLUF family. BLUF domains were first identified in *Euglena* (Iseki et al., 2002) and bacteria (Masuda and Bauer, 2002) and have been linked to sensing blue-light (Gomelsky and Klug, 2002). Interestingly, the blue-light sensing protein identified in *Euglena gracilis* has been characterised as a photoactivated adenylyl cyclase, regulating flagellar motility by controlling

intraflagellar levels of cyclic AMP (Iseki et al., 2002). *Leishmania* is unlikely to respond to light, but adenylyl cyclases are subject for studying cytokinesis, evasion of the host immune response and social motility (Makin and Gluenz, 2015). A role in light-sensing for MOT7 seems more likely, since *Chlamydomonas* clearly undergoes phototaxis (Bennett and Golestanian, 2015). Whether LAX28 is required only for IDA assembly or whether it has another functional role in the axoneme is currently unknown.

While the precise phylogenetic relationship, if any, between LmxM.24.1310, TEX47 and MOT7 remains to be established, a detailed genetic characterization of TEX47 in infertile men should be followed up.

Compound-heterozygous mutations and frameshift mutations in CFAP44 and CFAP43 have been connected to male infertility (Tang et al., 2017). To our knowledge mutations in TEX47 have not yet been linked to male infertility. However, the predominately expression in human testis of TEX47 (Fagerberg et al., 2014), its presence in the human sperm tail proteome (Amaral et al., 2013) and the data shown here for LAX28 suggests a role in sperm motility. Analysis of TEX47 in humans based on the findings in this chapter, therefore may have possible implementations in mapping a further gene to male infertility.

CHAPTER 5

DISSECTING THE REQUIREMENT OF FLAGELLAR MOTILITY FOR SURVIVAL *IN VIVO* BY USING A CRISPR-CAS9 BAR-SEQ STRATEGY

5 Dissecting the requirement of flagellar motility for survival *in vivo* by using a CRISPR-Cas9 bar-seq strategy

Data from the bar-seq sand fly infection screen and individual sand fly infections shown in this chapter were published in Beneke et al. (2019b). Jessica Valli created *ΔBBS2* mutants used for individual sand fly infections and Edward Hookway developed primers for the Illumina amplicon sequencing strategy. Sand fly infections were carried out by Petr Volf's laboratory at Charles University in Prague, while mouse infections were performed by Jeremy Mottram's laboratory at the University of York. All other data and figures presented in this chapter are wholly my own work unless stated otherwise.

5.1 Introduction

5.1.1 Chapter summary and aim

The aim of the work presented in this chapter was to develop a bar-seq mutant screen that enables a dissection of flagellar function *in vivo*. In particular, it was of interest to analyze whether flagellar motility is required for survival in sand flies and mice.

A robust bar-seq phenotyping pipeline was developed and tested *in vitro* and *in vivo* in two pilot experiments. Using the bar-seq approach, twenty-one mutants from a set of 100 mutants (see section 3.1.5) were barcoded and passaged through sand flies and mice. The results showed different requirements for flagellar "functions" in different life cycle stages: (1) flagellar motility and directional swimming are essential for the colonisation of sand flies but dispensable in mouse

infections and (2) flagellar assembly is essential for both mouse and fly infections, but dispensable for *in vitro* growth. Furthermore, preliminary data from mouse infections suggest that disruption of flagellar protein trafficking (*ΔBBS2* and *ΔKHARON1* mutants) reduces virulence, thus supporting a model in which the flagellum contributes to amastigote survival within macrophages. The requirement of flagellar-driven directional motility for colonialization in sand flies was subsequently confirmed in fly infections with paralyzed *ΔPF16* and uncoordinated swimming *ΔMBO2* mutants. However, mouse infections with homogenous populations of single mutants remain to be performed to confirm findings in the bar-seq assay.

Overall, these data demonstrate that the bar-seq assay is a powerful approach for studying flagellar function *in vivo* and that the *Leishmania* flagellum is required for survival throughout the *Leishmania* life cycle.

5.1.2 The *Leishmania* flagellum and its role in parasite-host and parasite-vector interactions

Much remains unclear about the *Leishmania* flagellum and its role in parasite-host and parasite-vector interactions. Previous studies in the closely related parasite *T. brucei* suggested that the flagellum is required for both, survival in flies and in mice. Bloodstream forms are sensitive to the loss of flagellar function (Alsford et al., 2011; Broadhead et al., 2006; Ralston et al., 2006) and immotile procyclics are defecated from their vector (Rotureau et al., 2014).

However, it remains unknown whether a functional *Leishmania* amastigote flagellum is required, like the *T. brucei* bloodstream flagellum, for survival in the mammalian host (see introduction section 1.2.3).

Morphological and life cycle related differences between *T. brucei* and *Leishmania* parasites indicate that *Leishmania* amastigote forms have developed different mechanisms for survival within their mammalian host. For example, *T. brucei* has its flagellum attached along the cell body, while in *Leishmania* the flagellum is mostly extracellular with a very short flagellar attachment zone (FAZ), a structure that connects the flagellum with the cell body cytoskeleton (Sunter and Gull, 2016; Wheeler et al., 2016). Furthermore, *T. brucei* bloodstream form parasites are motile, while *Leishmania* amastigotes are immotile.

Thus, the function of the amastigote flagellum remains particularly elusive. There are several hypotheses regarding how the flagellum may contribute to survival in the mammalian host.

(1) The first proposal suggests that the *Leishmania* flagellum may function as an antenna or secretory organelle for releasing virulence factors in parasite–host interactions. This theory is supported by the 9 + 0 or 9v ultrastructure of the amastigote flagellum, as it resembles very much the structure of sensory cilia (Gluenz et al., 2010b). Additionally, data from *Chlamydomonas* suggests that cilia may well function also as a secretory organelle (Baldari and Rosenbaum, 2010). Also, previous research on avirulent flagellar protein trafficking mutants *ΔKHARON1* and *ΔBBS1* (see introduction 1.1.3 for more details) (Price et al., 2013; Tran et al., 2013; Tran et al., 2015) may also support the interpretation that the amastigote flagellum is required for releasing virulence factors to promote parasite survival within the mammalian host.

(2) Another hypothesis is that not the amastigote flagellum is required for survival in the host, but instead the motile flagellum of an invading metacyclic parasite is needed for host invasion. Forestier et al. (2011) showed that live promastigotes

oscillate their flagellar beat shortly before invading their host cells and invade macrophages in a polarized manner, having their flagellum engulfed first. The later was also suggested in previous studies (Chang, 1978; Zenian et al., 1979). The location of the parasite thereby spatiotemporally correlated with the lysosomal-associated membrane protein 1 (LAMP-1) in macrophages, indicating lysosomal activity on the macrophage plasma membrane (Forestier et al., 2011). Since dead parasites, used as a control in this study, failed to be phagocytosed by macrophages the authors proposed that *Leishmania* parasites must injure the macrophage surface via their flagellum to promote phagocytosis (Forestier et al., 2011). Contrary to this, there have been reports that *Leishmania* parasites enter their host cells with their posterior ends first (Courret et al., 2002) or with both ends (Chang, 1979; Pearson et al., 1983; Rittig et al., 1998). This model also remains highly controversial as it needs to be tested with appropriate controls, for example by using immotile mutants instead of dead parasites.

(3) A third hypothesis is that the amastigote flagellum interacts with the flagellar pocket to control the release of virulence factors within the parasitophorous vacuole. The flagellar pocket is a specialized invagination in the plasma membrane from which the flagellum emerges and is believed to be the only site where kinetoplastids can do endo-and exocytosis (Field and Carrington, 2009; Gull, 2003; McConville et al., 2002). Hence, it is highly likely that virulence factors are released though the flagellar pocket. Consequently, the amastigote flagellum might function to support the structure of the pocket and to restrict the access to the pocket for endo-and exocytosis (Gluenz et al., 2010a).

Meanwhile, the role of the long motile promastigote flagellum is less controversial. Two clear functions have been attributed to the promastigote flagellum: (1) attachment and (2) motility.

(1) Following their escape from the peritrophic matrix, the promastigotes attach to the abdominal midgut by inserting their flagellum into the microvilli (Killick-Kendrick et al., 1974), protecting them from defecation after a blood meal (see section 1.2.2 for complete life cycle description). Additionally, in later sand fly infection stages they also have been observed to attach as haptomonad form with their flagellar tips to the alimentary tract in the anterior parts of the fly. Attachment is conferred by formation of a hemidesmosomes, a complex that connects two attached membranes through a junction (Killick-Kendrick et al., 1974).

(2) In-between both periods of attachment promastigotes have been observed to be free swimming, for example in the cardia lumen, suggesting that parasites migrate by motility towards anterior parts of the fly (Walters et al., 1989).

Overall, it remains unclear whether flagellar motility is required throughout the parasite life cycle and especially whether flagellar driven motility is required to migrate through the sand fly. It also remains to be shown what drives the migration process of parasites from posterior to anterior parts in the fly and whether the promastigote flagellum thereby senses signals during migration and metacyclogenesis, as previously suggested (Di-Blasi et al., 2015).

5.1.3 The principles of bar-seq experiments

To understand the diverse functions of the flagellum during the different life cycle stages of *Leishmania*, it can be informative to study a large cohort of mutants

(individually and in competition). For this, bar-seq methods can be applied to track the phenotype of several mutants at once.

In a typical bar-seq assay each mutant is tagged with a unique barcode to allow phenotyping of several mutants in a pooled population. The barcode might consist of a short DNA sequence optimized for Illumina sequencing or also a sequence that is actually required for targeting the gene itself. For the bar-seq analysis mutants can be either pooled after they have been individually generated or mutants are already generated as pool, e.g. if the system requires only a single target vector to silence (e.g. RNAi) or delete (e.g. some CRISPR methods) a gene of interest. The barcoded mutant pool is then subjected to a time course experiment under different conditions. Using next-generation sequencing the relative abundance of each barcode within the entire pool can be calculated, allowing e.g. to determine the *in vitro* or *in vivo* fitness phenotype of each mutant within the pool.

Genome scale bar-seq screens have been already pursued in multiple organisms, including yeast (Robinson et al., 2014; Smith et al., 2009; Smith et al., 2010), *Plasmodium berghei* (Bushell et al., 2017; Gomes et al., 2015), *T. brucei* (Alsford et al., 2011) and *Toxoplasma gondii* (Sidik et al., 2016). This has led to major discoveries. For example, a novel virulence factor has been identified in *T. gondii* by using a genome wide bar-seq CRISPR library approach in infections of human fibroblasts. This identified ~200 previously uncharacterized fitness-conferring genes of which 16 were further shown to be essential during infection of human cells. It also has led to the discovery of the claudin-like apicomplexan microneme protein invasion factor (CLAMP), which is required for the initiation of infection (Sidik et al., 2016). Another example is the pooled bar-seq knockout screen in

Plasmodium. 2,578 barcoded *Plasmodium berghei* knockout mutants, representing >50% of the genome, were created and phenotyped by measuring their competitive growth rates in mice. The authors of the study found that *P. berghei* requires ~65% of genes for optimal growth, which is the highest proportion reported from any organism where this has been measured so far (Bushell et al., 2017).

These two examples show each a unique strategy of a bar-seq approach for fitness phenotyping of mutants. In total, there are at least four different types of bar-seq libraries, which may be grouped as follows:

(1) bar-seq RNAi libraries, in which RNAi vector inserts are used for both, barcode read out and knockdown of the target gene (Alsford et al., 2011),

(2) bar-seq CRISPR libraries, in which sgRNAs in vectors are used for deletion of the target ORF and also serve as a barcode for the read out of the assay (Sidik et al., 2016),

(3) bar-seq pooled knockout libraries, in which mutants are generated in pools by transfecting barcoded vectors (Bushell et al., 2017; Gomes et al., 2015) and

(4) bar-seq single knockout libraries, in which each mutant is barcoded individually in the process of gene deletion before pooling of mutants ((Han et al., 2010; Smith et al., 2009); possible with LeishGEdit toolbox as described in section 2.2.3).

These bar-seq strategies all have different advantages and disadvantages. For example, while bar-seq approach 4 is much more labour intensive compared to approaches 1 to 3, it allows us to assess each mutant individually to show that the deletion was complete. In contrast, for mutants generated in pools (approaches 1

to 3) there is no actual proof that each specific gene was successfully subjected to a knockdown or knockout, unless individually tested later on.

Another drawback for RNAi libraries is that the bar-seq read out reflects only a part of the target gene as not the entire RNAi vector insert is sequenced, thus increasing the false discovery rate. However, RNAi libraries are inducible, which might be more suitable for studying the loss-of-function phenotype of essential genes. Targeting essential genes remains especially challenging in bar-seq experiments although it is often the desired outcome to identify them. If systems lack an inducible component, transfection replicates are usually used to infer essentiality of a gene. More reliable are additional addback copies introduced prior to the deletion, but these require labour intensive cloning procedures and are usually not scalable approaches.

Additionally, all bar-seq strategies are also limited to the analysis of a few specific phenotypes. For example only phenotypes that result in differences of growth or survival, or enrichment of a sub-population in a particular place from which DNA can be isolated can be assessed with bar-seq methods.

However, overall the chosen bar-seq strategy often relies on available tools in the screened organism rather than a list of pro and contra for various strategies. For example most *Leishmania* spp. lack the RNAi machinery (Ullu et al., 2004) and since CRISPR-Cas9 mediated DSBs are mainly repaired by SSA (Zhang and Matlashewski, 2019) (see section 2.1.4), the use of a single vector based genome-wide CRISPR library is unlikely to work in *Leishmania* (see discussion in 6.2).

Therefore, the use of a bar-seq strategy, in which mutants are generated individually and barcoded in the process is an attractive alternative to other bar-

seq approaches. The LeishGEdit toolbox can be easily adapted for barcoding of individual mutants (as shown in section 2.2.3).

This kind of bar-seq knockout approach can be particularly useful for dissecting the *in vivo* function of many flagellar mutants already studied *in vitro* in the Gluenz lab (as described in section 3.1.5), thus helping to understand the diverse functions of the *Leishmania* flagellum in different life cycle stages. For example, the motility analysis of the first 100 mutants of the current CRISPR-Cas9 knockout screen in the Gluenz lab (as described in section 3.1.5) identified different *in vitro* swimming behaviours among mutants. But it remains unclear whether there are any *in vivo* consequences of these flagellar defects.

Therefore, to gain insight into the biological significance of motility in the mammalian and insect host a selection of mutants that represented a broad range of flagellar function and swimming phenotypes (recently published in Beneke et al. (2019b)) was chosen and analysed by using a bar-seq strategy (Figure 51 B). A rapid *in vitro* and *in vivo* screening assay to test fitness of these diverse flagellar mutants in sand flies and in mouse models was developed.

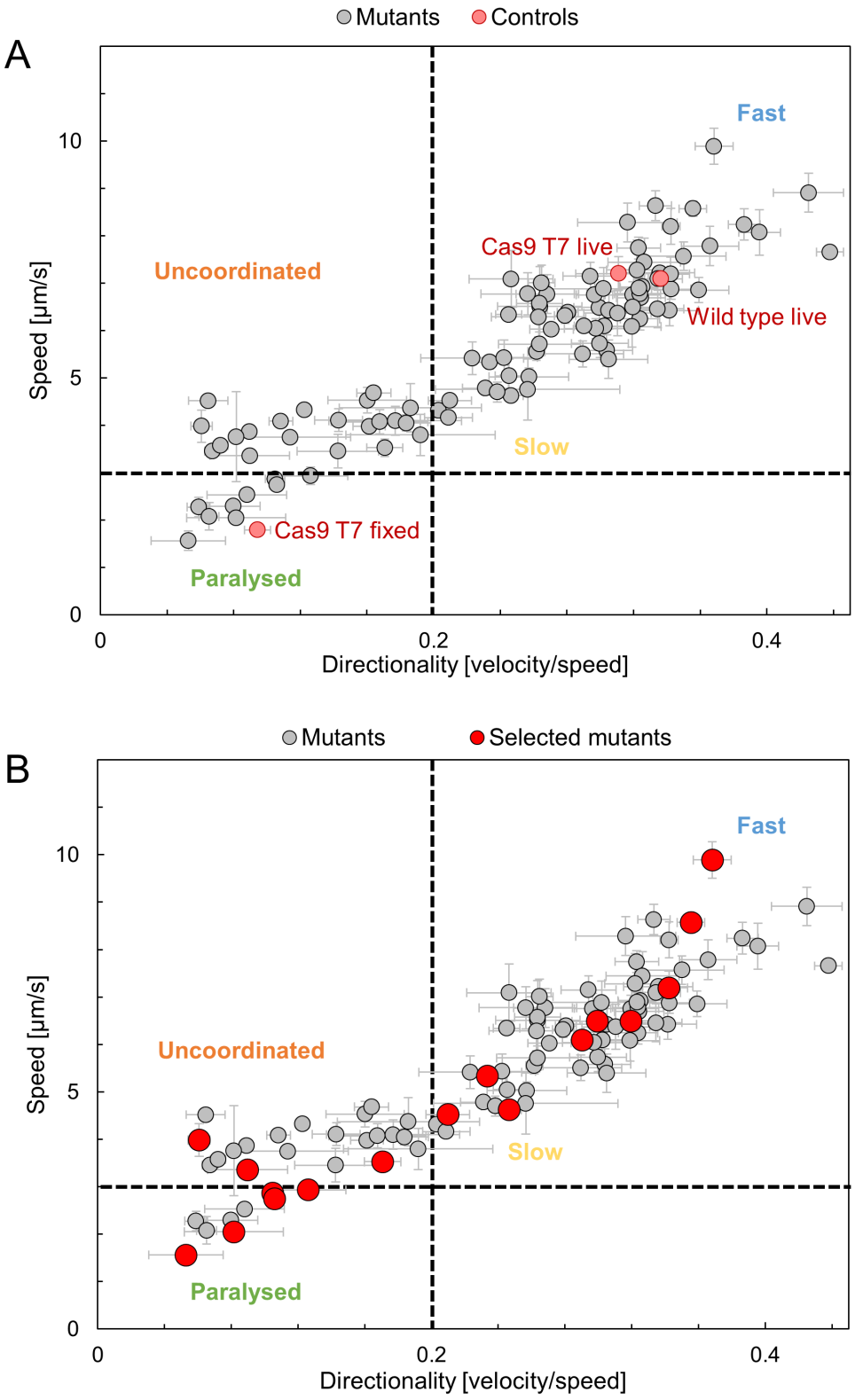


Figure legend on following page.

Figure 51. Motility analysis of mutants.

Figure drawn by Tom Beneke based on data from Beneke et al. (2019b). Average speed [µm/s] on Y axis has been plotted against directionality [velocity/speed] on X axis by measuring swimming parameters as described in Wheeler (2017). Mutants are classified to be either paralyzed, uncoordinated, fast or slow swimming. **(A)** Measurements of all 100 mutants in Beneke et al. (2019b) (grey) and control cells, including live and fixed Cas9 T7 parental cells and live wildtype cells (red). **(B)** All 100 mutants measured in Beneke et al. (2019b) as in **(A)** with highlighted selected mutants that have been subjected to a bar-seq mutant screen (red).

5.2 Results

5.2.1 A bar-seq strategy allows pooled mutant screens

To enable a dissection of the role of the flagellum in parasite–host and parasite–vector interactions, a bar-seq strategy for *Leishmania* was developed. As described in chapter 2, a 17 nt barcode and a 20 nt constant region were added to the existing 30 nt HF and 20 nt pT-pPLOT-pPOT primer binding site in the upstream forward primer. Each barcode has 40-60% GC content, there at least 3 MM between barcodes and barcodes have no blast hit in *Lutzomyia longipalpis*, in *Mus musculus* and in *L. mexicana*. Using the additional 20 nt constant region and 20 nt pT-pPLOT-pPOT primer binding site adjacent to the 17 nt barcode allows reading out barcode abundance by using Illumina amplicon sequencing strategies. Edward Hookway developed long primers that contain flow cell binding sites for Illumina sequencing and additional indices to multiplex multiple samples (see 7.6.1). Using only a single PCR reaction to amplify barcodes from isolated genomic DNA allows rapid and high-throughput library preparation to read out barcode abundance (Figure 52).

To test whether a bar-seq experiment would be actually suitable to simultaneously determine the growth “fitness” of several mutants in a pooled population, a pilot *in vitro* bar-seq experiment was performed first. 20 control cell lines (SBL1-20) were tagged with a unique 17 nt barcode in their small subunit (SSU) ribosomal RNA locus (Figure 53 A) to have a barcoded cell line in the pool that is close to the parental cell line (constructed by Tomas Becvar, under my supervision). Sanger sequencing of the integrated barcode cassette showed that barcodes and Illumina primer binding sites were inserted as desired in the locus (Figure 53 B and C).

Next, all 20 control promastigote cell lines were pooled in equal proportions and cultured *in vitro* over 48 hours, growing from early log phase to late log phase. Genomic DNA was extracted at the start and after 24 and 48 hours in culture. Reading out barcode abundance showed that the proportion of a single cell line within the pool, reflected as normalized read, did not significantly change over time for any cell line within the pool (Figure 54). Consequently, mutants were not randomly selected during the bar-seq assay.

Next, it was tested whether the read out of barcode abundance also reflects the actual abundance of a mutant within the pool. Therefore, all 20 control cell lines were pooled in different proportion and cultured again for 48 hours in log phase culture. Strikingly, the barcode abundance correlated well with cell numbers added to the pool, thus showing that not only presence or absence of a barcode can be detected, but also its relative abundance (Figure 54).

To test the bar-seq assay performance *in vivo*, a pilot sand fly infection assay was set up next. In addition to the 20 barcoded control cell lines, a *ΔBBS2* mutant cell line and a paralyzed *ΔPF16* mutant cell line were barcoded and the presence of the correct barcode cassette in the locus was confirmed by Sanger sequencing as before (Figure 53 D). Both mutants and all 20 control cell lines were then pooled in equal proportions shortly before infections and late log phase promastigote mutant pools were used to infect *L. longipalpis* (sand fly infections were carried out by Jovana Sadlova). Infections were performed in triplicate with 150 sand flies each and after two, six and nine days post blood meal, 50 sand flies per time point per replicate were isolated and subjected to DNA extraction.

Using Illumina sequencing the relative abundance of barcodes over time was determined and used to calculate “mutant fitness”. In most bar-seq *in vivo* assays

it is not possible to extract multiple DNA samples from one fly or animal. Consequently, a mutant pool cannot be tracked over time in one subject. Therefore, the mutant fitness was calculated by comparing each time point against the time point before infection (start of assay) and not the previous time point.

Using this analysis showed that *ΔPF16* mutants had a strong fitness phenotype in sand flies (Figure 55 A and B), which is expected since motility has been proposed before to be required for migration within sand flies (see section 5.1.2). Meanwhile, all 20 control cell lines and the *ΔBBS2* mutant cell line did not significantly change over time when using Turkey's multiple comparison test (Figure 55 B). However, a decrease of the *ΔBBS2* mutant barcode could be still observed (Figure 55 A). Overall, this showed that the bar-seq assay can be a powerful tool for phenotyping several mutants simultaneously when performing sand fly infections.

To show that also the single library step Illumina amplicon sequencing strategy is a reliable way for library preparation and reading out barcode abundance, p5 and p7 indexed primers used for library preparation (Table 6) were shuffled between sand fly infection samples and Illumina sequencing was repeated (except for samples "9 days post blood meal", as DNA amount was insufficient for a technical replicate). Both technical replicates correlated well, thus showing that the read out was robust in the developed pipeline (Figure 55 C).

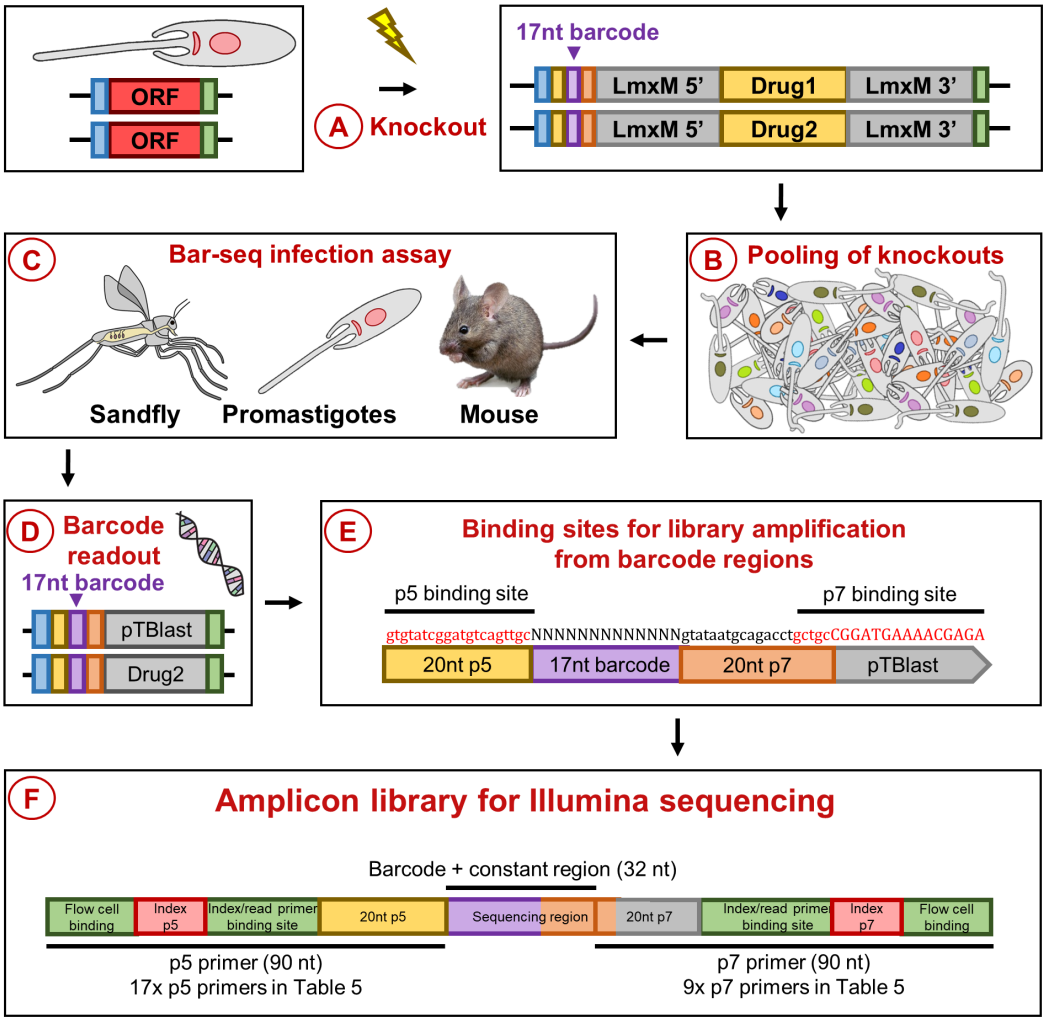
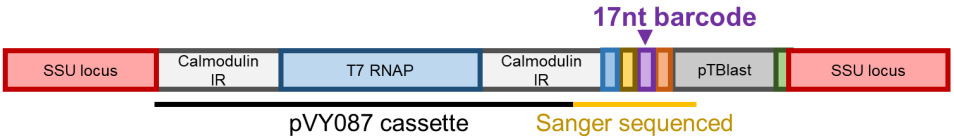


Figure 52. Overview of bar-seq strategy for mutant fitness phenotyping in pools.

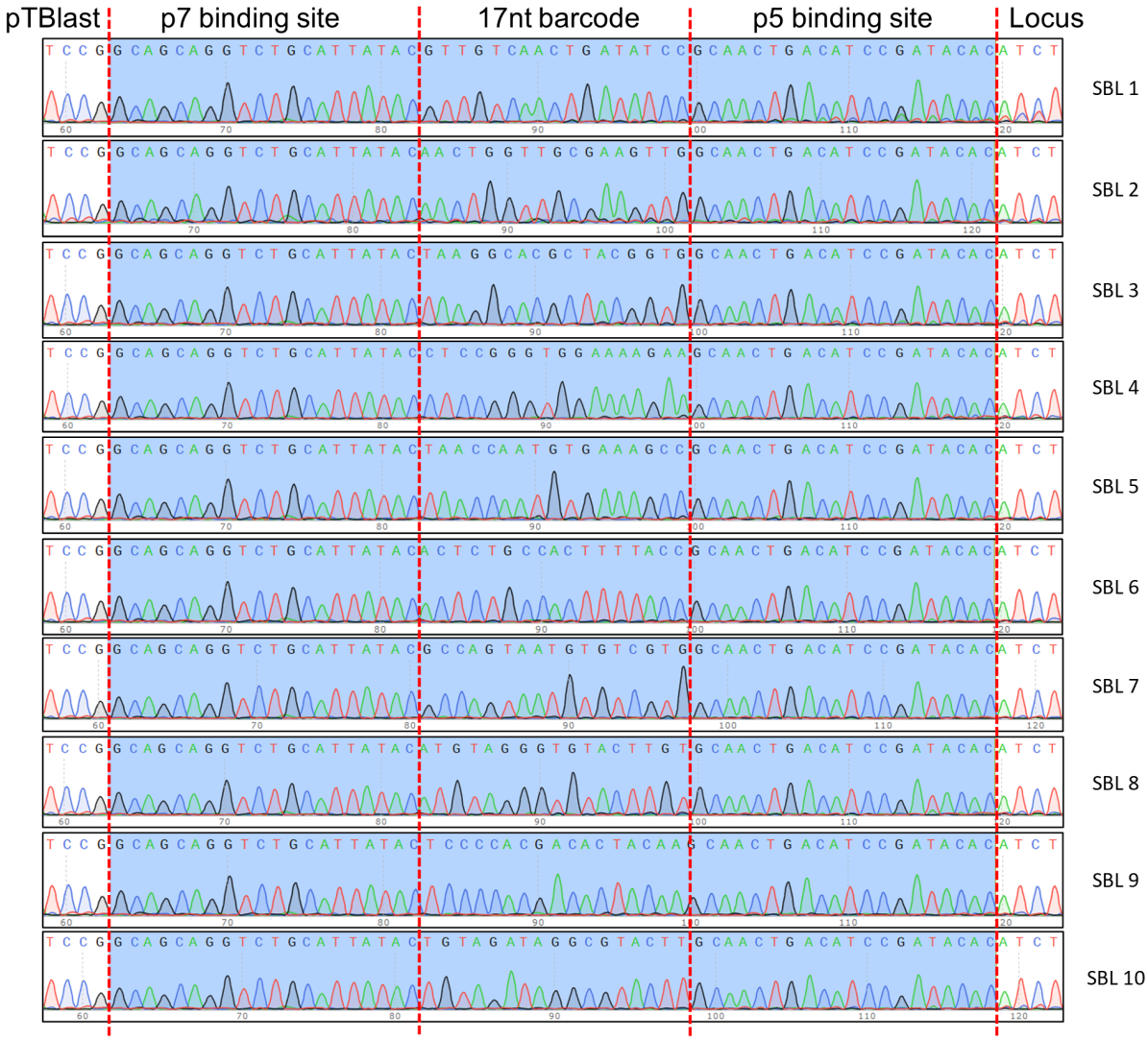
Figure made by Tom Beneke, adapted from Beneke et al. (2019c).Strategy for bar-seq phenotyping. **(A)** Cas9 T7 parental cells are subjected for double allele deletion using two different drug selectable markers. A 17 nt barcode (purple), surrounded by two constant regions (yellow and orange) is inserted into the target locus by using donor DNA with HFs (blue and green). **(B)** Barcoded mutants are pooled and **(C)** passaged through an *in vitro* or *in vivo* screen for example in mice or sand flies. **(D)** Barcode abundance is read out by **(E)** amplifying barcodes using their surrounding constant regions with two long primers (p5 and p7 primers), containing **(F)** flow cell binding sites and additional indices for Illumina sequencing, allowing to multiplex dozens of samples for sequencing.

A

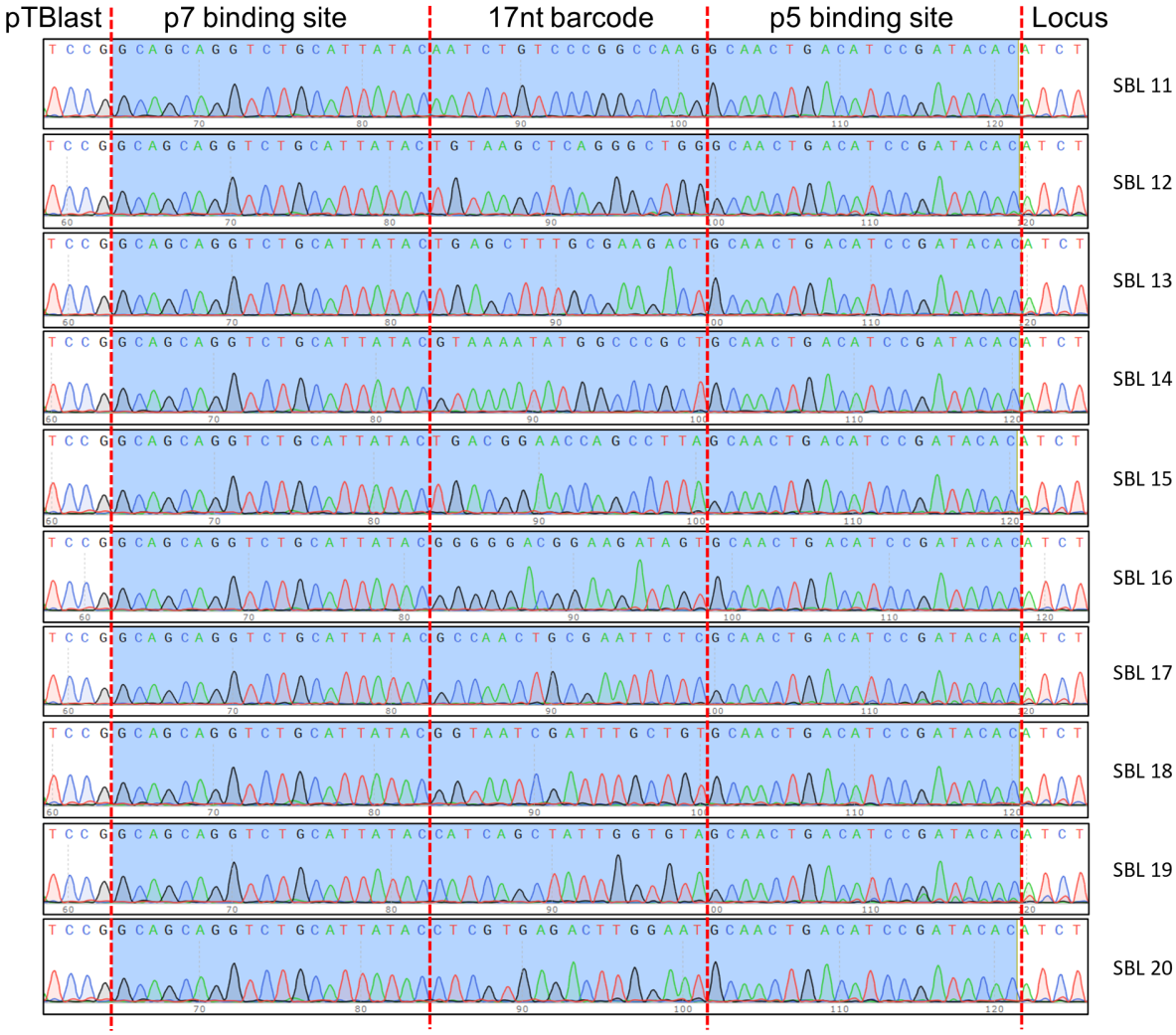


Barcode location for SBL1-20

B



C



D

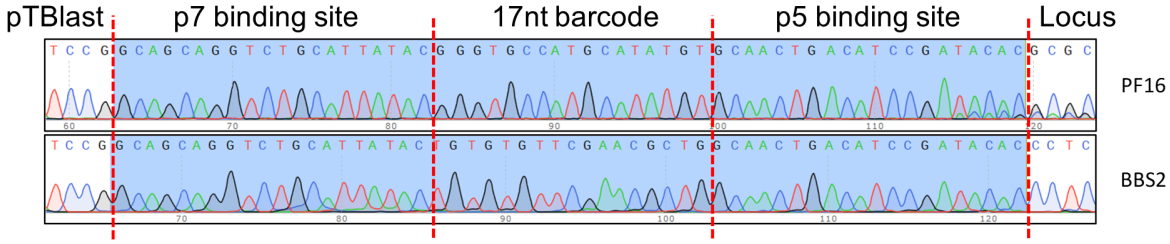


Figure legend on following page.

Figure 53. Sanger sequencing of barcoded mutants and control cell lines.
(A) Location of barcode in control cell lines. The 17 nt barcode, including two constant regions is inserted in the integrated pVY087 cassette. Highlighted in yellow is the region that has been Sanger sequenced. **(B)**, **(C)** and **(D)** DNA sequencing chromatograms are shown for SBL1-20, *ΔPF16* and *ΔBBS2* mutants. Highlighted are sequences that align to the pTBlast cassette, p7 binding site, 17nt barcode, p5 binding site and the locus of the integrated construct.

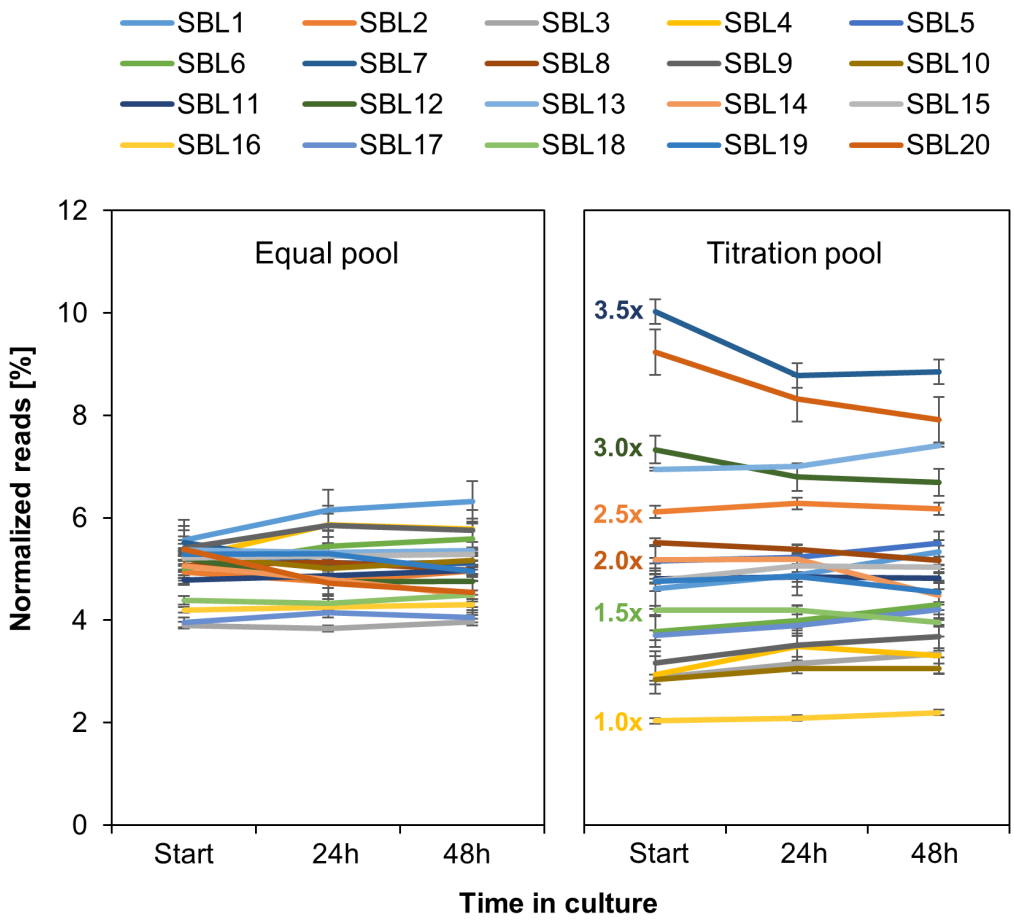


Figure 54. Testing the *in vitro* bar-seq screen.

Barcoded control cell lines (SBL1-20) have been pooled and subjected to a bar-seq *in vitro* screen. Y axis: Normalized reads, meaning the relative abundance of each barcode within the entire pool. X axis: DNA isolation time points. The average of biological triplicates is shown. Error bars show standard deviation between triplicates. Left panel: Cell lines have been pooled equally. Right panel: Cell lines were pooled by using different cell numbers. Numbers on left indicate how much more cells of a cell line compared to the bottom cell line (1.0x) were used.

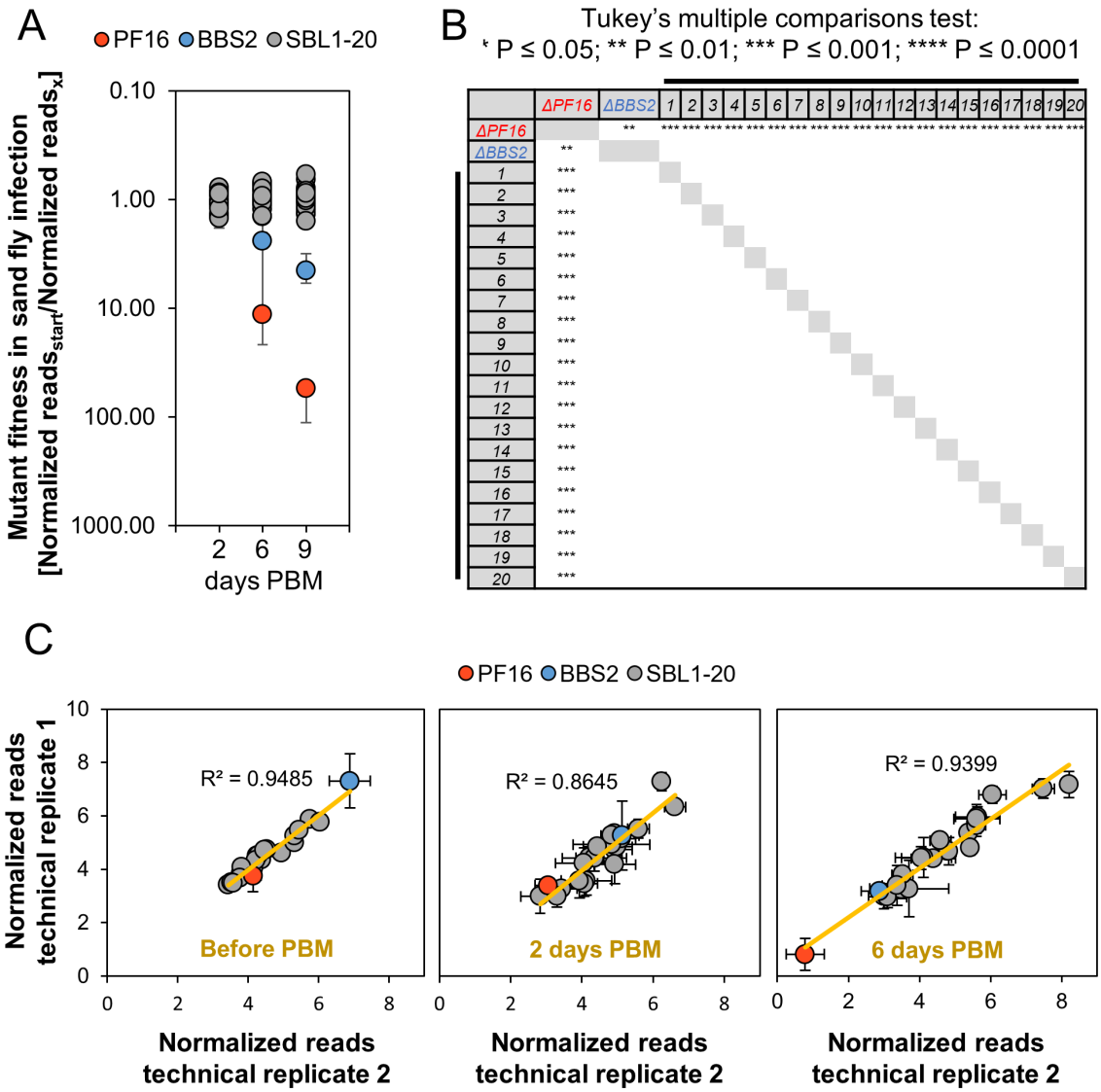


Figure 55. Pilot bar-seq *in vivo* screen in sand flies.

(A) Barcoded cell lines have been pooled and subjected to a bar-seq *in vivo* sand fly screen. Y axis: mutant fitness calculated by normalized reads at the start of the assay divided by normalized reads at each DNA isolation time point. X axis: DNA isolation time points post blood meal (PBM). The average of biological triplicates is shown. Error bars show standard deviation between triplicates. **(B)** Mutant fitness from **(A)** was used to perform Tukey's multiple comparison test for *ΔBBS2*, *ΔPF16* and SBL1-20. **(C)** DNA samples from the bar-seq assay have been library prepped again with shuffled index primers. Y-axis: Normalized reads from the bar-seq assay in **(A)**. X-axis: Normalized reads from the technical sequencing replicate. The R squared value of the linear trendline (yellow) is shown.

5.2.2 Setting up a bar-seq *in vivo* and *in vitro* fitness screen

Next a subset of 21 genes (all published in Beneke et al. (2019a)) was chosen to study flagellar function by using the bar-seq fitness screen. These mutants were representative for all discovered flagellar phenotype categories (Figure 25 and Figure 51 B) and were selected to study the consequence of flagellar-associated defects for survival in mice, as well as persistence and migration defects in the fly. Phenotype categories included mutants that: (1) were aflagellate (*ΔIFT88*), (2) have a short flagellum (*ΔARL-3A*), (3) are paralyzed with a flagellum of normal length (*ΔPF16*), (4) are slow swimming parasites (*ΔCFAP43*, *ΔCFAP44*, *ΔIC140*), (5) are slow swimming parasites with little directionality (ΔdDC2, ΔRSP4/6, *ΔN-DRC2*), (6) have a shorter PFR or a disrupted PFR structure (*ΔLmxM.21.1110*, *ΔFM458*), (7) are fast swimming parasites (*ΔLC4-like*, *ΔLmxM.18.1090*) and (8) have no obvious defect in promastigote culture, but are associated with the flagellar membrane (*ΔLmxM.23.1020*, *ΔLmxM.17.0870*) or are involved in flagellar protein trafficking (*ΔBBS2*, *ΔKHARON1*). Thus this subset included genes that have a role in motility (IFT88, PF16, CFAP43, CFAP44, IC140, LmxM.21.1110, FM458, LC4-like, LmxM.18.1090), flagellar protein trafficking (ARL-3A, BBS2, KHARON1) and possibly also sensing, meaning reacting to environmental stimuli, as it has been proposed for protein kinase A regulatory subunit FM458 (Oberholzer et al., 2011) and cyclic nucleotide specific phosphodiesterases, such as LmxM.18.1090 (Oberholzer et al., 2007b), before.

Furthermore, barcoded mutants defective in LPG synthesis and defective in the pathway leading to mannose activation for synthesis of LPG and other glycoconjugates were generated. This included *ΔLPG1* (Beneke et al., 2017; Ryan et al., 1993), phosphomannose isomerase (*ΔPMI*) (Garami and Ilg, 2001b),

phosphomannomutase (*ΔPMM*) (Garami et al., 2001) and GDP-mannose pyrophosphorylase (*ΔGDP-MP*) (Garami and Ilg, 2001a) mutants. On the one hand these mutants serve as controls in sand fly and mouse infections, as an intact LPG coat has been shown to be essential for parasite attachment to the sand fly midgut (Jecna et al., 2013; Sacks et al., 2000) and as the mannose pathway has been proposed to be required for survival of amastigote parasites *in vivo* and *in vitro* (Garami and Ilg, 2001a; Garami and Ilg, 2001b; Garami et al., 2001; Kedzierski et al., 2006; Stewart et al., 2005). On the other hand the role of LPG as a potential virulence factor for survival in the mammalian host has been controversial, as *ΔLPG1 L. major* knockouts showed reduced virulence (Spath et al., 2003) and *ΔLPG1 L. mexicana* knockouts showed normal virulence (Ilg, 2000) (see discussion 5.3.2). Therefore, testing these mutants in a bar-seq pool may shed new light on the contentious role of LPG1 for survival in mice.

All 21 mutants were barcoded by using the LeishGEdit toolbox for replacing the target ORF with two barcoded drug selectable marker cassettes and absence of the target ORF was confirmed by using a diagnostic PCR strategy (all as described in chapter 2). Additionally, five control cell lines from above (SBL1 – 5) were included to have a barcoded version in the pool that is close to the parental cell line (Figure 56 A, B and C).

Next, all barcoded cell lines (21 knockouts and 5 control cell lines) were subjected to *in vivo* and *in vitro* bar-seq screens. For the bar-seq promastigote *in vitro* assay parasites were pooled in equal proportions and observed over 96 hours. However, for bar-seq *in vivo* assays parasites were divided into four sub-pools according to their *in vitro* growth rates. This approach was chosen to ensure that all cell lines reached the same density for infections. Each sub-pool was grown for 48 hours

until they reached late log phase for sand fly infection or for 96 hours until they reached late stationary phase for mouse infection. Sub-pools were then pooled in equal proportions just before the infection. Genomic DNA of pooled mutants was isolated at the start of *in vitro* growth assays or before infecting sand flies or mice and then again isolated at various time points (Figure 57). Using Illumina sequencing the relative abundance of barcodes over time was determined and used to calculate mutant fitness as above.

Measuring the mutant fitness in a pool of promastigotes *in vitro* over 96 hours (Figure 57 A) showed that four cell lines became significantly depleted: *ΔIFT88*, *ΔPMM*, *ΔPMI* and *ΔGDP-MP* (Figure 58 A). The barcoded control cell lines and *ΔLmxM.21.1110* grew slightly, but not significantly quicker in promastigote culture (Figure 58 A). It is likely that the barcoded control cell line grew more rapidly because cells express fewer drug resistance markers than other cell lines. It remains unclear why the abundance of *ΔLmxM.21.1110* mutants increases.

As a proof of principle for the bar-seq assay, fitness phenotypes were compared to measured doubling times of these cell lines (carried out by James Smith) (Figure 58 B). Cell lines depleted in the bar-seq assay, such as *ΔIFT88*, *ΔPMM*, *ΔPMI* and *ΔGDP-MP*, had also much longer doubling times. Additionally, the barcoded control cell lines had shorter doubling times, thus correlating with their observed increase in the bar-seq assay. This showed that mutant fitness measured in the bar-seq assay correlated well with the actual growth phenotype.

ΔDRC2 was not detected in the pool by Illumina sequencing. Sanger sequencing of the integrated barcode cassette in the DRC2 locus showed that the p5 primer binding site of the pTBlast cassette had 8 nt MM (Figure 59) (likely introduced in oligo synthesis). Therefore, *ΔDRC2* was omitted from the subsequent analysis.

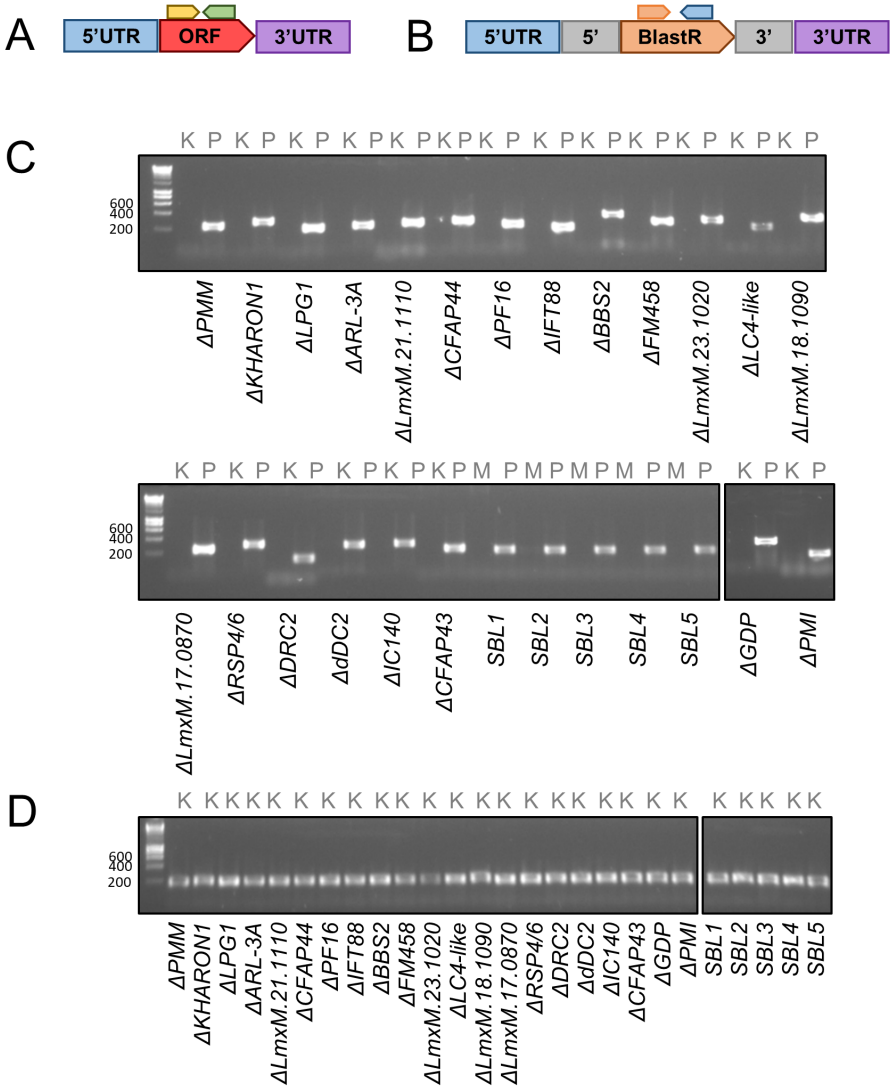


Figure 56. Null mutant PCR validation.

Cartoons illustrate PCRs for null mutant verification for knockouts used in bar-seq screen: **(A)** amplification of a fragment of the target gene ORF and **(B)** amplification of a fragment of the blasticidin-S deaminase ORF. **(C)** Amplicons from the target gene ORF run on agarose gel. Each ORF was amplified in the putative knockout cell line (K) and the parental Cas9 T7 cell line (P). No PCR products were obtained in mutant cell lines for amplification of the target gene ORF. Fainter bands below the target gene amplicon are likely primer dimers. The gene names indicate the target gene ORF. **(D)** Presence of the blasticidin-S deaminase ORF PCR product was confirmed for all transfected cell lines.

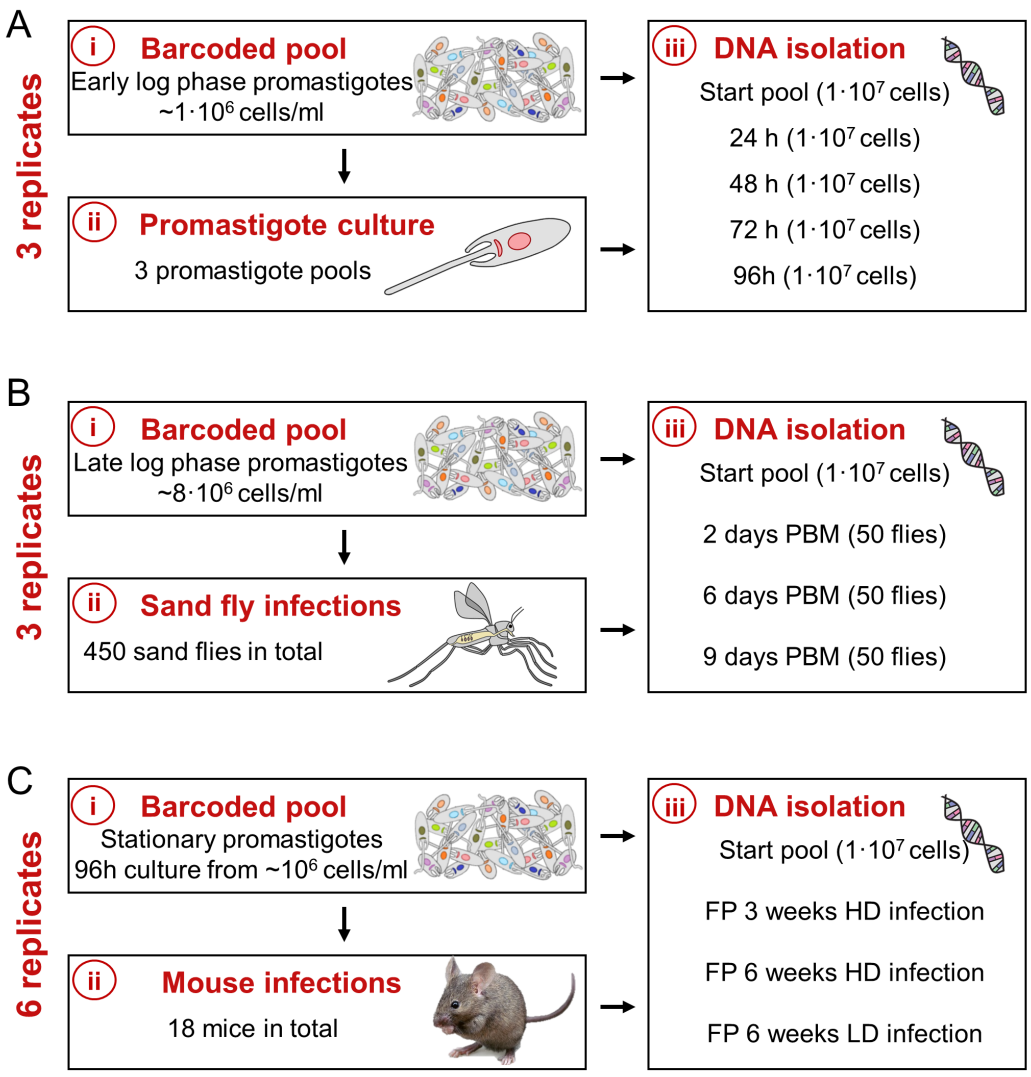


Figure 57. Overview for *in vitro* and *in vivo* bar-seq screen.

(i) Barcoded promastigote pools are subjected to **(ii)** bar-seq assays *in vivo* or *in vitro* by **(iii)** isolating DNA over a time course experiment with biological replicates as indicated. **(A)** Early log phase promastigote pools are cultured over 96 hours and DNA is isolated every 24 hours. **(B)** Late log phase promastigote pools are used for sand fly infections and DNA is isolated over 9 days post blood meal (PBM). **(C)** Stationary promastigote pools are used for infecting mice in high dose (HD, $2{\cdot}10^6$ parasites) and low dose (FD, $2{\cdot}10^5$ parasites) infections. DNA from footpad infections is harvested at 3 and 6 weeks post infections.

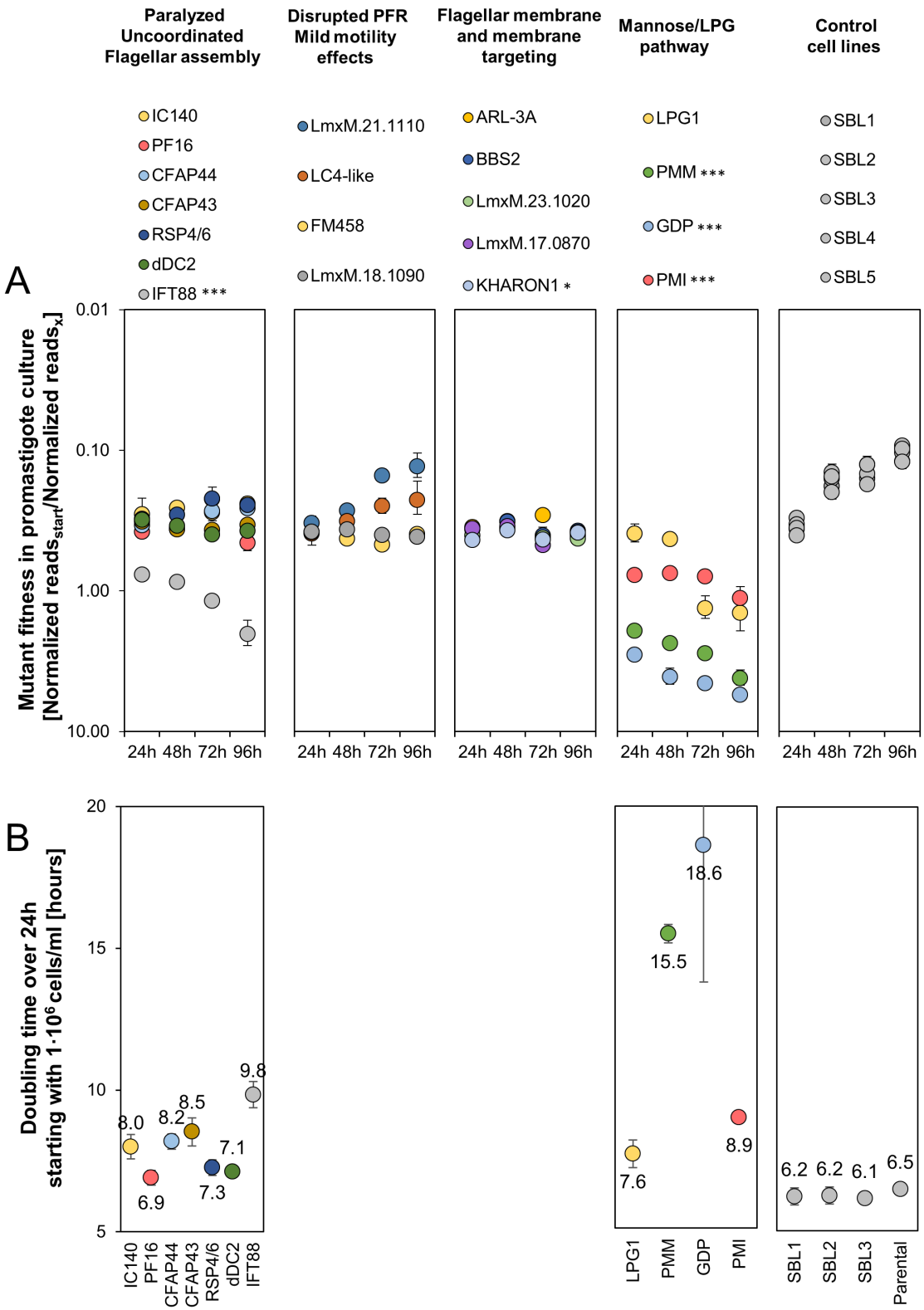


Figure legend on following page.

Figure 58. Promastigote growth mutant analysis.

Figure made by Tom Beneke, taken from Beneke et al. (2019b). Each panel shows mutant classes associated with swimming or other, such as pathway related, defects. Barcoded control cell lines are shown on the right. **(A)** Barcoded mutants have been pooled and subjected to a bar-seq *in vitro* screen as shown in Figure 57 A. Y axis: mutant fitness calculated by normalized reads at the start of the assay divided by normalized reads at each DNA isolation time point. X axis: DNA isolation time points. The average of biological triplicates is shown. Error bars show standard deviation between triplicates. Student's t-test between each mutant and control lines is indicated by asterisks above. **(B)** Y axis: doubling time was measured in log phase culture in triplicates. The average is shown and indicated by data labels. Error bars show standard deviation between triplicates. X axis: cell line used for doubling time measurement. Parental refers to Cas9 T7 parental cell line.



Figure 59. Sanger sequencing the integrated barcode cassette in *ΔDRC2* mutants.

The upstream homology flank and the adjacent barcode cassette in the DRC2 locus have been Sanger sequenced. The desired annotation and the actual sequence in the pTBlast cassette is indicated. Sanger sequencing chromatogram is shown. Highlighted in red is the mutation in the p5 primer binding site.

5.2.3 A bar-seq fitness screen in sand flies identifies proteins required for colonialization of sand flies

Next the *in vivo* mutant fitness phenotype of each barcoded cell line was determined by infecting *L. longipalpis* with a pooled mutant population and isolating DNA of pools at two, six and nine days post infection (Figure 57 B) (sand fly infections were carried out by Jovana Sadlova).

As expected, *ΔLPG1*, *ΔPMM*, *ΔGDP-MP* and *ΔPMI* mutants were all rapidly depleted over time, indicating that LPG and other glycoconjugates are required for colonialization of sand flies. Phosphoglycan-containing glycoconjugate surface molecules are believed to have a protective role during blood meal digestion in the sand fly (Sacks et al., 2000). They are also believed to be essential for parasite attachment to the sand fly midgut (Jecna et al., 2013). This would explain why all four mutants become depleted already two days post blood meal (Figure 60).

Paralyzed, slow swimming, uncoordinated and aflagellated mutants were also out-competed during sand fly infections. Especially *ΔIFT88*, which fails to assemble a flagellum, was rapidly depleted over time (Figure 60). The proportion of other mutants, such as *ΔPF16*, *ΔCFAP43*. *ΔCFAP44*, *ΔRSP4/6*, *ΔdDC2* and *ΔIC140* decreased less quickly. The paralyzed mutant *ΔPF16*, which has a normal flagellar length, decreases most among these mutants. *ΔKHARON1*, which is associated with flagellar protein trafficking (Tran et al., 2013), and *ΔARL-3A*, which do have shorter flagella (Beneke et al., 2019b), are also significantly depleted over time (Figure 60). However, both these mutants show only mild mutant phenotypes in sand flies compared to other mutants that are cleared from *L. longipalpis* infections. In contrast, mutants of genes associated with the flagellar membrane (*ΔLmxM.23.1020*, *ΔLmxM.17.0870*) or flagellar protein trafficking (*ΔBBS2*) and

mutants that have a PFR-associated defect (*ΔLmxM.21.1110*, *ΔFM458*) or are fast swimming parasites (*ΔLC4-like*, *ΔLmxM.18.1090*) do not show a defect in sand fly infections (Figure 60). This is also found for barcoded control cells, which seem to slightly increase, even though by student's t-test not significantly, similar to their determined *in vitro* growth rate (Figure 58). Taken together, these results indicate that directional motility and glycoconjugate synthesis related pathways are essential for colonialization of the sand fly (Table 4).

While this shows that bar-seq assays of pooled mutant populations are clearly a powerful technique to phenotype multiple mutants simultaneously, they do not give the resolution that infections with a single defined parasite line provide. It is possible that different phenotypes are observed in a bar-seq assay with mixed infections compared to an infection with a genetically homogenous population. Additionally, individual sand fly infections give anatomical resolution and allow to determine whether an immotile mutant fails to migrate to anterior parts of the fly gut. Therefore, individual sand fly infections were pursued next and carried out by Petr Volf's lab. This allowed also to confirm that immotile mutants fail to migrate to anterior parts of the fly gut.

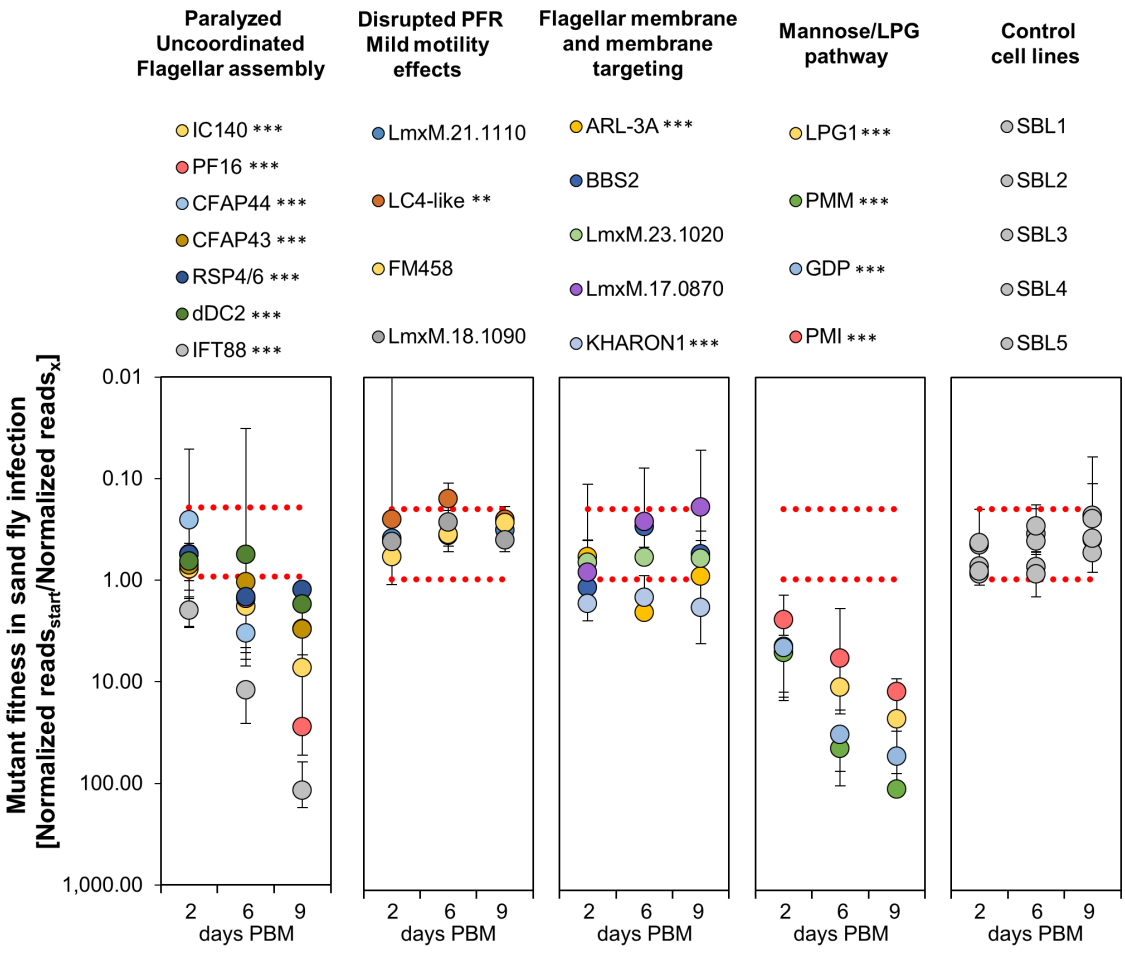


Figure 60. Bar-seq *in vivo* screen in sand flies.

Figure made by Tom Beneke, taken from Beneke et al. (2019b). Barcoded mutants have been pooled and subjected to a bar-seq *in vivo* screen in sand flies as shown in Figure 57 B. PBM: post blood meal.

Panels are arranged as in Figure 58 and show mutant classes associated with swimming or other, such as pathway related, defects. Red dotted lines indicate two standard deviations above the average of control cell lines. Y axis: mutant fitness calculated by normalized reads at the start of the assay divided by normalized reads at each DNA isolation time point. X axis: DNA isolation time points. The average of biological triplicates is shown. Error bars show standard deviation between replicates. Student's t-test between each mutant and control lines is indicated by asterisks above.

5.2.4 Paralysed and uncoordinated motility phenotypes are cleared from sand flies

For individual sand fly infections three representative motility mutants, each having a distinct motility phenotype, were chosen. Paralysed *ΔPF16* mutants, uncoordinated swimming *ΔMBO2* mutants and motile *ΔBBS2* mutants along with their parental and addback cell lines (*ΔBBS2* mutant used for sand fly infection were created by Jessica Valli (Valli, 2017)) were used for infection of *L. longipalpis* (Figure 61 and Figure 62). *ΔBBS2* mutants, which lack a component of the BBSome complex (Nachury et al., 2007) (see section 1.1.3), Cas9 T7 parental cells, *L. mexicana* wild type cells and *ΔBBS2* addbacks showed high infection rates of more than 60% two and six days post blood meal. This is in contrast to *ΔPF16* and *ΔMBO2* mutants, which are unable to complete their life cycle within their vector. Mutants are less infective from the beginning and only 1.6% of *ΔPF16* infected sand flies and 7.5% of *ΔMBO2* infected sand flies were still infected at six days post blood meal, while more than 60.0% of sand flies remain infected when colonialized with parental, wild type or addback cell lines (Figure 61 A).

ΔBBS2 mutants and parental control cell lines localized to the abdominal midgut after two days post blood meal and colonized the entire sand fly midgut and stomodeal valve by six days post blood meal. However, *ΔPF16* and *ΔMBO2* mutants failed to migrate to anterior parts of the sand fly midgut and were restricted to the abdominal midgut in the few sand flies they colonized (Figure 61 B). Addbacks for *ΔPF16* and *ΔMBO2* mutants rescued the respective phenotype.

These results confirm findings made in bar-seq *in vivo* assays (Figure 62 A and B) and strongly support a model in which *Leishmania* flagellar-driven directional motility is essential for migration and completion of the life cycle within the fly.

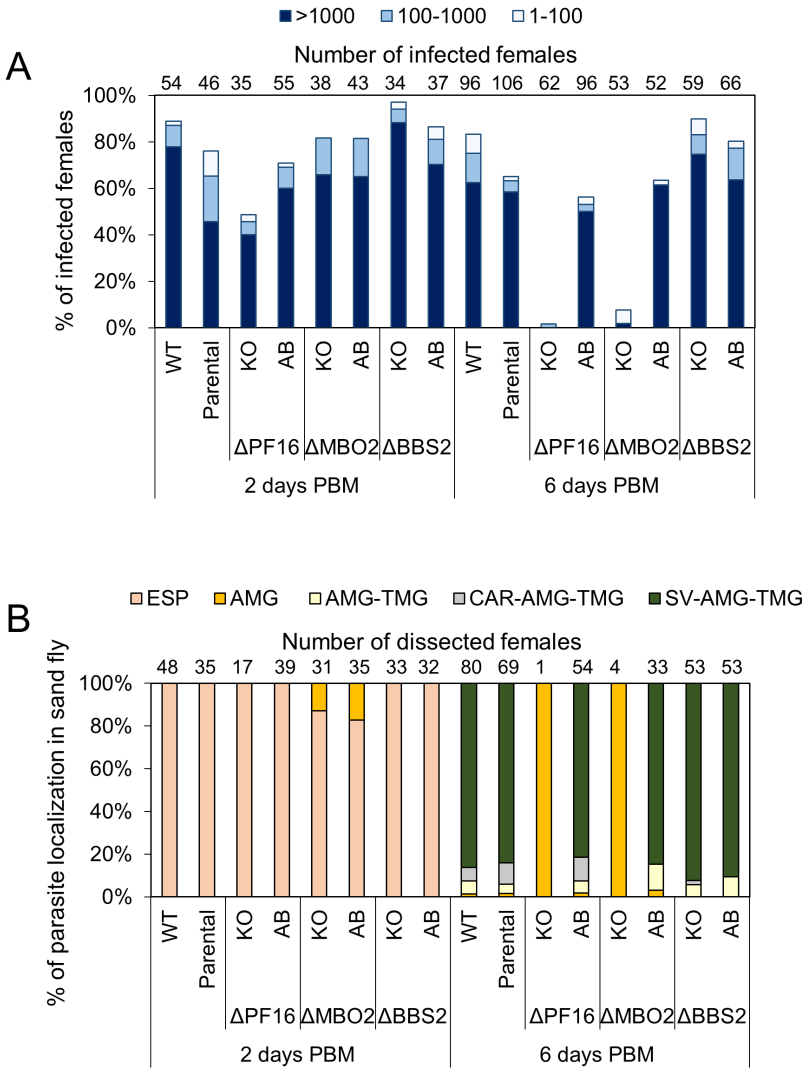


Figure 61. Development of *ΔPF16*, *ΔMBO2* and *ΔBBS2* mutants in *L. longipalpis* infections.

Figure made by Tom Beneke, taken from Beneke et al. (2019b). Infection rates and localisation of *L. mexicana* mutant and control cell lines have been determined in *L. longipalpis* infections. KO: knockout; AB: addback; WT: wildtype. X axis bottom: tested mutants and parental cell lines at two and six days post blood meal (PBM). X axis top: Numbers of **(A)** infected or **(B)** dissected females in the group. Y axis: percentage of **(A)** infected females or **(B)** localization within sand flies. **(A)** Parasites per infected sand flies have been measured: dark-blue: heavy; blue: middle; light-blue: light. **(B)** Localisation of parasites in flies have been determined. ESP: endoperitrophic space; AMG: abdominal midgut; TMG: thoracic midgut; CAR: cardia; SV: stomodeal valve.

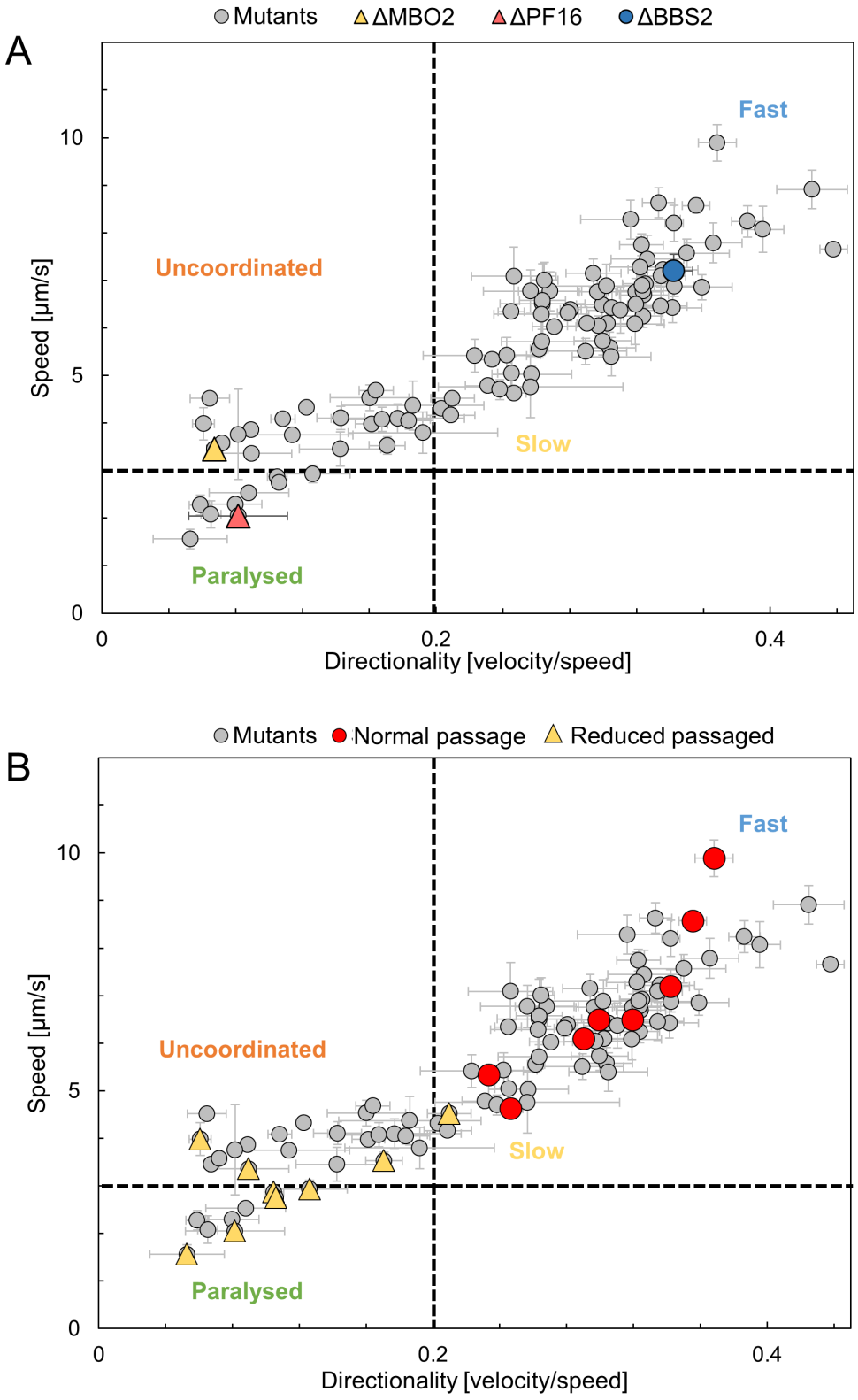


Figure legend on following page.

Figure 62. Mapping mutants with reduced passage in sand flies with their motility phenotype.

Figure drawn by Tom Beneke based on data from Beneke et al. (2019b). Average speed [µm/s] on Y axis has been plotted against directionality [velocity/speed] on X axis and mutants are classified to be either paralyzed, uncoordinated, fast or slow swimming as in Figure 51. Measurements of all 100 mutants in Beneke et al. (2019b) are shown in grey. Additionally highlighted are **(A)** *ΔPF16*, *ΔMBO2* and *ΔBBS2* mutants that have been tested *in vivo* in individual sand fly infections (reduced passaged: triangle; normal passage: circle) and **(B)** all mutants that were included in the bar-seq sand fly *in vivo* assay (reduced passage: yellow triangles; normal passage: red circles).

5.2.5 A bar-seq fitness screen in mice identifies proteins required for infection

To test what consequence flagellar defects have in mouse infections the same barcoded mutants used in the *in vitro* assay and the sand fly infections above (see section 5.2.2 and 5.2.3) were used for infections of female BALB/c mice. Therefore, a population of pooled mutant promastigotes in late stationary phase was injected in the left foot pad using either a high dose (HD, $2 \cdot 10^6$ cells) or a low dose (LD, $2 \cdot 10^5$ cells) and DNA from the left foot pad was isolated from euthanized mice after three and six weeks post infection (Figure 57 C) (mouse infections and DNA isolation of pools were carried out by Carolina Catta Preta in Jeremy Mottram's laboratory and are described in section 7.5.3).

Results showed that the controls persisted over the three and six week infection time. When comparing the relative barcode abundance from late stationary promastigotes shortly before infection to three and six weeks post infection no difference between all five control cell lines was observed, demonstrating that the concept of the bar-seq screen in mice works (Figure 63).

Barcodes of mutants with PFR-associated defects and mutants of flagellar membrane localized proteins were also not decreased within the pool over time. Remarkably, the same result was found for paralyzed, uncoordinated, slow swimming and fast swimming mutants (Figure 63).

In contrast, *ΔBBS2* and *ΔKHARON1* mutants were rapidly depleted during mouse infections (Figure 63). BBS2 is part of the BBSome complex (see section 1.1.3), which is highly conserved across ciliated eukaryotes (Hodges et al., 2010) and expected to play a role in flagellar protein trafficking (Liu and Lechtreck, 2018). However, this has not been directly demonstrated in *Leishmania* yet. Meanwhile,

KHARON1 has been shown in *L. mexicana* to be required for flagellar targeting of the glucose transporter 1 in promastigotes. While this transporter is not expressed in amastigotes, it is believed that KHARON1 may also be involved in trafficking of other proteins (Tran et al., 2013). Therefore, both proteins, BBS2 and KHARON1, are expected to play a role in flagellar protein trafficking (see section 1.1.3). This allows to speculate that avirulence of *ΔBBS2* and *ΔKHARON1* mutants is related to flagellar protein trafficking defects.

While this is consistent with previous findings, mutant phenotypes displayed from mannose pathway and LPG synthesis mutants do only partly agree with previous literature. *ΔPMM* and *ΔGDP-MP* are both rapidly depleted during mouse infections (Figure 63) which is in line with previous findings of *in vitro* infections with *ΔPMM L. major* mutants (Kedzierski et al., 2006), *in vitro* and *in vivo* infections with *ΔPMM L. mexicana* mutants (Garami et al., 2001) and *in vitro* and *in vivo* infections with *ΔGDP-MP L. mexicana* mutants (Garami and Ilg, 2001a; Stewart et al., 2005). However, *ΔPMI L. mexicana* mutants have been shown to be avirulent in *in vitro* and *in vivo* infections (Garami and Ilg, 2001b), although bar-seq data presented here suggests that PMI is not required for survival in mice. Bar-seq analysis of mouse infections also suggest that *ΔLPG1 L. mexicana* mutants are infection competent (Figure 63), which agrees with observations made by Ilg (2000) in *L. mexicana*. It is, however, in contrast to findings from Spath et al. (2003), which showed that *ΔLPG1 L. major* mutants are avirulent. However, this might be explained by species related differences between *L. major* and *L. mexicana*. Also, to be fully comparable with these studies, findings from bar-seq mouse infections have to be repeated individually with a single defined parasite line rather than in a

pool. For example the LPG deficiency in *L. mexicana* mutants may be compensated by other surrounding LPG-carrying parasites in the pool.

Interestingly, *ΔIFT88* mutants are avirulent (Figure 63). IFT proteins are required for flagellar assembly (see section 1.1.3) and fail to assemble the flagellar axoneme beyond the flagellar pocket (Sunter et al., 2018). Therefore, avirulent and aflagellate *ΔIFT88* mutants give strong evidence that components of the *Leishmania* flagellum might be actually required for parasite invasion or survival in the mammalian host. Meanwhile, motility seems not to be essential for survival in the mammalian host (Table 4).

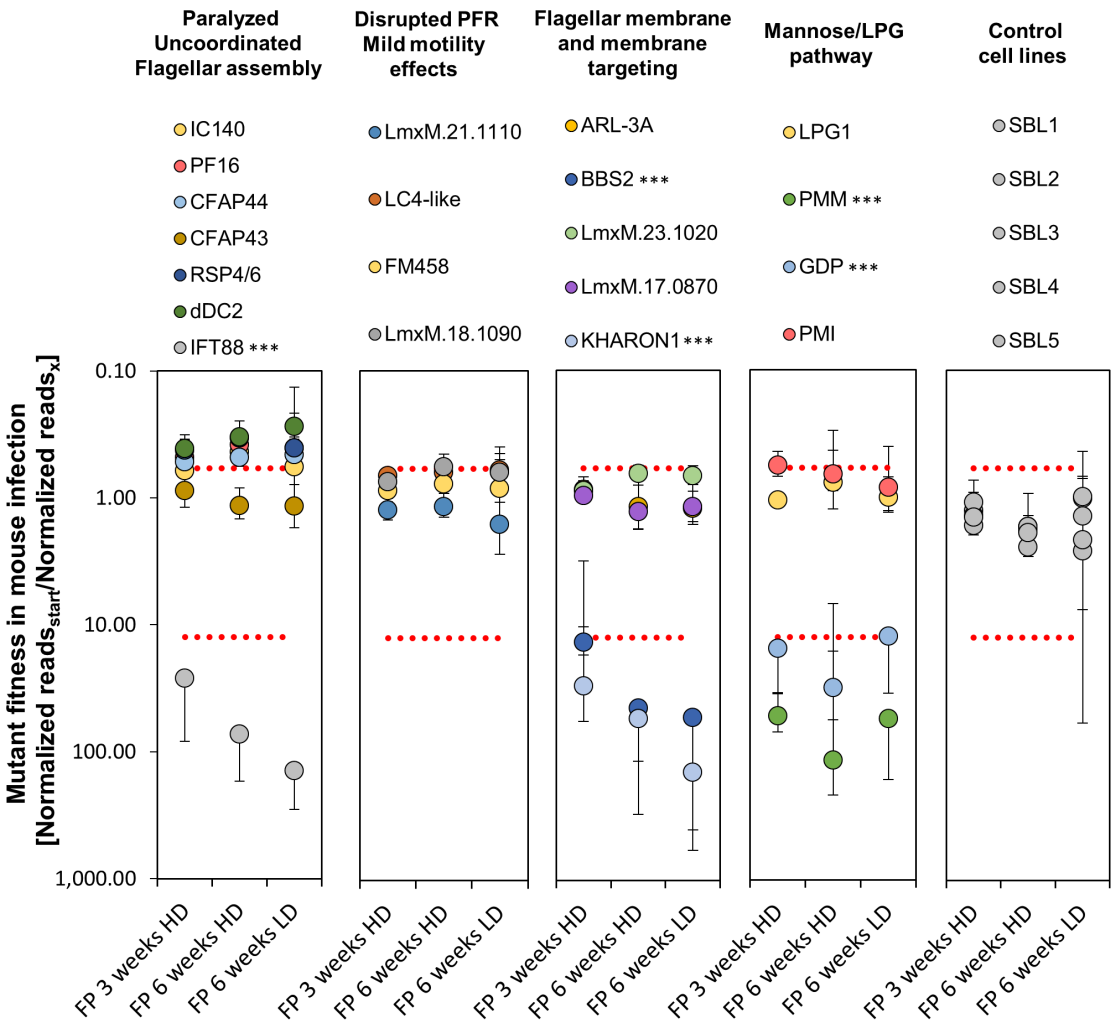


Figure 63. Bar-seq *in vivo* screen in mice.

Barcoded mutants have been pooled and subjected to a bar-seq *in vivo* screen in mice as shown in Figure 57 C. FP: footpad infection; LD: low dose infection ($2 \cdot 10^5$ parasites); HD: high dose infection ($2 \cdot 10^6$ parasites). The figure legend from Figure 60 applies, except that the average of biological sextuplicate is shown here.

Motility	Other phenotypes	Reduced fitness phenotype in mice	Reduced fitness phenotype in sand fly
Paralyzed	Curling		*ΔPF16*
	No flagellum	*ΔIFT88*	*ΔIFT88*
Uncoordinated	Curling		*ΔMBO2*
			ΔRSP4/6, ΔdDC2
Slow	Curling		*ΔIC140, ΔCFAP44, ΔCFAP43*
	PFR integrity		
	Short PFR/flagellum		
Fast			
Normal	Mannose/LPG pathway linked	*ΔGDP, ΔPMM*	*ΔLPG1, ΔPMI, ΔGDP, ΔPMM*
	Flagellar membrane and membrane targeting	*ΔBBS2, ΔKHARON1*	

Table 4. *Leishmania* genes identified to be required for colonialization of the sand fly and infection of mouse.

Findings in Figure 60 and Figure 63 have been summarized. Phenotype categories are primary based on swimming behaviors and then secondary on additional morphological features or other, such as pathway related, defects. *In vivo* mutant fitness phenotypes in mice and sand flies (white columns) are matched to these primary and secondary phenotypes (grey columns).

5.3 Discussion

5.3.1 Flagellar-driven directional motility in *Leishmania* is essential for colonialization of the vector

The aim of the work shown in this chapter was to study the requirement of flagellar function, in particular flagellar motility, for survival of *Leishmania* parasites in the mouse and sand fly. A bar-seq assay was developed and proven to be a powerful approach to phenotype multiple mutants simultaneously. Using a pool of mutants to infect sand flies is only one of many possible assays that can be transformed into a high-throughput approach with this strategy.

However, bar-seq results have to be interpreted carefully. For example, parasites in sand flies may form aggregates through glycoconjugate surface molecules. Consequently, motile mutants that swim towards the stomodeal valve could drag immotile parasites from the same mutant pool along. This might be especially the case in sand flies with high infection rates (Gossage et al., 2003; Rogers, 2012; Rogers et al., 2002). Surprisingly, there is no evidence for this in the bar seq-data. *ΔPF16* mutants are cleared in individual sand fly infection assays, as well as in sand flies infected with pooled *Leishmania* mutants (Figure 60 and Figure 61). Similar observations can be made for uncoordinated swimming mutants within the pool. *ΔdDC2* and *ΔRSP4/6* decrease significantly over time in sand fly infections, which is also found for uncoordinated swimming *ΔMBO2* mutants in individual sand fly infections (Figure 60, Figure 61 and Figure 62). Overall, there seems to be a correlation with the degree of loss of motility and increasing failure to survive in sand flies. Consequently, there must be substantial competition within the fly that allows to outcompete unfit parasites. Also, parasites are likely to undergo a strong

selection progress, in which for example parasites are lost that cannot attach to the sand fly gut.

However, it is unclear whether *Leishmania* experience a bottleneck in sand flies. Testing for a bottleneck would require following a mixed infection of barcoded “wild type” *Leishmania* parasites, similar to the control cell lines used here, through different compartments in the sand fly. If there is at any point a compartment found with a selected subset of the infecting population, this could indicate a bottleneck. Using eight barcoded cell lines, this kind of approach was used to show that *T. brucei* parasites are selected when passing the foregut in tetse flies (Oberle et al., 2010). Since the bar-seq approach used here only measures presence of mutant DNA within the entire fly, the issue of bottlenecks couldn’t be address.

Clearance of *ΔPF16* and *ΔMBO2* mutants in sand fly infections suggest that flagellar-driven directional motility is needed for colonialization of the sand fly. *ΔPF16* mutants are paralyzed due to a central pair defect (Beneke et al., 2017), but still possess a long flagellum with a low curling rate (Beneke et al., 2019b). *ΔMBO2* mutants do have a slightly shorter flagellum (Beneke et al., 2019b), but their observed failure to complete their life cycle in the fly is likely to be explained by their uncoordinated movement possibly initiated through defective waveforms, which have been also reported in *Chlamydomonas* MBO2 mutant flagella (Segal et al., 1984). Both mutants are found in the endoperitrophic space, which isolates the blood meal in the peritrophic matrix, at two days post blood meal. But at six days they are cleared from more than 93% of dissected sand flies or only colonize the abdominal midgut. Thus parasites fail to migrate to the posterior parts of the sand fly gut, proving a migration defect. Flagellar-driven directional motility as a requirement for completion of the life cycle in sand flies seems to be well in line

with previous reports in *T. brucei*, which showed that deletion of *DNAI1* is required for parasites to reach the tsetse fly foregut. *DNAI1* is an ODA intermediate chain and required for ODA assembly (Rotureau et al., 2014).

From data shown here, it remains unclear whether flagellar length has a major role in attachment to the sand fly gut. Aflagellate *ΔIFT88* mutants are rapidly cleared from sand flies, but *ΔARL-3A* mutants, which do have a short flagellum, are only slightly decreased (Figure 60). Overexpression of GTP-locked ADP-ribosylation factor-like protein 3A (ARL-3A) in *L. amazonensis* contributed to clearance of parasites from *L. longipalpis* infections. But it is unclear whether this was due to a shortened flagellum or through other alterations within mutants that contribute to the phenotype in flies (Cuvillier et al., 2003). Further evidence for independence of flagellar length for sand fly infections can be found in *L. braziliensis* isolates from patients that fail to assemble a full length flagellum upon promastigote differentiation, but are still able to colonize *L. longipalpis* (Zauli et al., 2012).

Furthermore, it remains unclear in how far the PFR contributes to flagellar attachment to the sand fly gut. *ΔFM458* mutants possess a shorter PFR and *ΔLmxM.21.1110* have a fragmented PFR, but both mutants do not decrease over time (Figure 60).

Overall, it remains to be further investigated how the PFR and flagellar length contribute to parasite–vector interactions. It might be interesting to investigate whether complete absence of PFR results in failure to colonialize sand flies. The PFR in promastigotes almost equals the axoneme in its dimensions. Hence, the PFR could be required for flagellar attachment in the sand fly gut, as it enlarges the flagellar surface.

Decrease of *ΔLPG1*, *ΔPMM*, *ΔGDP-MP* and *ΔPMI* mutants over time in the sand fly bar-seq assay (Figure 60) are also in line with previous observations (Jecna et al., 2013; Pimenta et al., 1994), showing that surface glycoconjugates are essential for colonizing sand flies. Mutants are deficient for LPG and mannose-containing glycoconjugate surface molecules, thus being sensitive for the blood meal digesting hostile environment in the fly gut (Sacks et al., 2000).

Rather surprisingly, the *ΔKHARON1* mutant, which fails to mediate targeting of flagellar localized glucose transporter 1 in *L. mexicana* promastigotes (Tran et al., 2013), are only slightly reduced and *ΔBBS2* mutants, which have been shown to be deficient for flagellar membrane protein trafficking in *Chlamydomonas* and metazoa (Jin et al., 2010), does not have a phenotype in sand flies (Figure 60 and Figure 61). Additionally, null mutants of flagellar membrane proteins LmxM.17.0870 and LmxM.23.1020 did not decrease over time (Figure 60), altough previous observations point towards a role of the flagellar membrane for attachement. For example the flagellar membrane of nectomonads have been shown to swell up in order to confer better attachment to the microvilli (Killick-Kendrick et al., 1974). Thus, the flagellar membrane may be re-organized to allow flagellar membrane proteins or other flagellar surface molecules to better interact with the sand fly gut.

Further interaction between parasite and vector might also be expected from *Leishmania* proteins involved in cyclic nucleotide signal transduction pathways. These pathways have been linked to environmental sensing with regulation of flagellar beat patterns (Mukhopadhyay and Dey, 2016; Porter and Sale, 2000) and have been shown to be essential for migration of *T. brucei* parasites within their vector (Shaw et al., 2019). However, knockouts of phosphodiesterase

LmxM.18.1090 and protein kinase A regulatory subunit FM458 remained as abundant as controls post 9 days blood meal (Figure 60).

Overall, these results may allow two possible hypotheses for the function of the promastigote flagellum during the life cycle stages in the vector. Directional motility of parasites within the fly is required (1) to coordinate the attachment via the flagellum to the microvilli within the sand fly gut and/or (2) to migrate through the fly and reach the stomodeal valve.

Individual sand fly infections with *ΔPF16* and *ΔMBO2* mutants, and also results from the bar-seq screen, allow to conclude a migration defect within the fly, but no conclusions can be made on a role of flagellar attachment to the sand fly gut.

5.3.2 The *Leishmania* amastigote flagellum has a role in parasite survival in the mammalian host

Interestingly, the bar-seq screen data from mouse infections draws a different conclusion for the role of motility for host invasion and amastigote survival in the mouse. Mutants that possess a short flagellum (*ΔARL-3A*), paralysed mutants (*ΔPF16*), slow swimming mutants (*ΔCFAP43*, *ΔCFAP44*, *ΔIC140*), slow uncoordinated swimming mutants (ΔdDC2, ΔRSP4/6), mutants that have a shorter PFR or a disrupted PFR structure (*ΔLmxM.21.1110*, *ΔFM458*), and mutants that are fast swimming parasites (*ΔLC4-like*, *ΔLmxM.18.1090*), do not decrease during mouse infections over time. This suggests that motility is not required for parasite survival within the mammalian host. This is in line with the morphological features of promastigotes and amastigotes. Once parasites are fully differentiated from motile promastigotes to immotile amastigotes motility becomes dispensable. Also PFR-associated proteins are expected to be dispensable for host survival, as the PFR is absent in *Leishmania* amastigotes.

However, these findings are in contrast to the proposal that the parasite oscillates its flagellum to injure macrophages and promote phagocytosis (Forestier et al., 2011). Furthermore, slow swimming *L. mexicana* FAZ5 mutants are avirulent in mice, but can infect marcophages and initial macrophage infection rates are actually higher in mutants than in wild type cells. The authors concluded that phagocytosis was additionally promoted by mutant parasites as they were unable to move away effectively from the macrophages (Sunter et al., 2019). While this indicates that *Leishmania* motility is not essential for macrophage invasion, it may support the flagellar-mediated macrophage surface injury model to promote phagocytosis (Forestier et al., 2011). If parasites are immotile, but still able to beat their flagellum, which is observed in FAZ5 mutants (Sunter et al., 2019), they would effectively damage the macrophage surface even more as *Leishmania* mutants are constantly lying on their host cells, thus promoting more parasite uptake.

However, the bar-seq assay measures only presence or absence of mutant DNA. Hence, it does not distinguish whether a mutant fails during: (1) host invasion, (2) *in vivo* differentiation from promastigotes to amastigotes, (3) parasitophorous vacuole formation, (4) surviving intracellularly or (5) *in vivo* growth. Another drawback of the bar-seq assay is that mutants may compensate for each other within the pool. For example, if flagellated motile promastigotes injure cells to promote phagocytosis (Forestier et al., 2011), then motile cells might still be able to fulfil this role for immotile cells if they are in the same pool.

Therefore, conclusions form the bar-seq experiments have to be independently tested in experiments where mice are infected with homogenous populations of single mutants to exclude compensatory effects.

While this remains to be further investigated, the nature of *ΔIFT88*, *ΔBBS2* and *ΔKHARON1* mutant phenotypes strongly supports a model in which the flagellum contributes to amastigote survival within macrophages (Figure 63).

Aflagellate *ΔIFT88* mutants decrease ~100 times over six weeks infection. However, it remains unclear whether this is due to a growth defect which has been also observed *in vitro* (Figure 58) or whether this is caused by failed interaction between parasite and host due to the absence of the flagellum.

ΔBBS2 and *ΔKHARON1* mutants provide much better evidence that the flagellum is required for amastigote survival, as both mutants decrease over time in mice, but do not exhibit a growth defect *in vitro* as promastigotes (Figure 58 and Figure 63). *ΔBBS1* mutants (*ΔBBS2* mutants have not been individually tested yet) in *L. major* and *ΔKHARON1* mutants in *L. mexicana* have been both shown to be avirulent in their mammalian host. *ΔKHARON1* mutants have a cytokinesis defect as intracellular amastigotes and are unable to divide within THP-1 marcophages and BALB/c mice (Tran et al., 2015). *ΔBBS1* mutants are able to infect murine BALB/c peritoneal macrophages, but are avirulent in BALB/c mice (Price et al., 2013). Only little is known about how the flagellar protein composition changes in these mutants. However, recently a putative calflagin (LmxM.16.0920), which localizes to the flagellum, has been shown to mislocalize in *L. mexicana ΔBBS3* mutants (unpublished, FHS project by Charlotte Laycock supervised by me and Eva Gluenz, project initiated by Jessica Valli). Although additional experiments are required to confirm this, this would be the first evidence of a protein trafficking role for the BBSome in *Leishmania*.

Results for *ΔLPG1* mutants, which remain at a similar proportion within the pool, are in contrast to the hypothesis that the LPG coat protects the parasite from

complement-mediated lysis immediately after injection into the host. Additionally, it is in contrast to the proposal that the LPG promotes a delay of phagosome maturation, allowing metacyclics to differentiate fully into amastigotes (reviewed in Franco et al. (2012)). However, if the parasite is injected in mice in a too high dose, some parasites might just escape this immune response by chance even though they lack the LPG surface coat. Once intracellular and fully differentiated to amastigotes, LPG is strongly reduced or undetectable and thus not required for survival in the host (Bahr et al., 1993; Glaser et al., 1991; McConville and Blackwell, 1991; Moody et al., 1993; Turco and Sacks, 1991). Therefore, the not changing barcode abundance of *ΔLPG1* mutants in mouse infection pools over time might just reflect the limitation of the bar-seq assay as more cells have to be injected for the assay and as other cells may compensate for the loss of LPG when invading macrophages.

This complementation effect may also explain the different results for infection competent *ΔPMI* mutants here in the bar-seq screen (Figure 63) and previously reported avirulent *ΔPMI* mutants in individual mouse infections (Garami and Ilg, 2001b). However, since *ΔPMM* and *ΔGDP-MP* did show a loss of fitness in mice (Figure 63), which is in line with previous reports (Garami and Ilg, 2001a; Garami et al., 2001; Kedzierski et al., 2006; Stewart et al., 2005), a complementation effect for *ΔPMI* mutants seems unlikely. It is more likely that the loss of PMI can be compensated by PMM and GDP-MP, but not vice versa. PMI converts fructose-6-phosphate to mannose-6-phosphate, which is then the substrate for a PMM and GDP-MP glycoconjugate biosynthesis pathway in *Leishmania*. However, mannose-6-phosphate can be also synthesized from alternative pathways, allowing to compensate the loss of PMI if for example mannose is available (Pomel

et al., 2012). Therefore, the phenotype of *ΔPMM* and *ΔGDP-MP* mutants in mice is likely to be related to the deficiency of producing other PG-containing surface glycoconjugates, which are expressed at high levels in amastigotes, shielding *Leishmania* from reactive oxygen and nitrogen species produced by the host to kill the invading parasite (reviewed in Franco et al. (2012)).

5.3.3 The *Leishmania* flagellum is a multifunctional organelle

Overall, these data emphasize that the *Leishmania* flagellum is a multi-functional organelle. It is clearly required for directional motility in the sand fly, potentially to mediate attachment to the sand fly gut and clearly to enable migration within the fly to posterior parts. However, it seems to have a complete different function within the mammalian host. Motility does not seem to be required for host invasion, differentiation or survival. Whether the flagellum may serve as a potential antenna in parasite-host interactions as previously proposed (Gluenz et al., 2010b) remains to be shown.

The bar-seq assay developed and applied here is a powerful and effective tool allowing for the first time large phenotyping assays of *Leishmania* mutants *in vivo* and *in vitro*. But mutant analysis has to be carried out carefully as mutants may mask each other's phenotypes in a pooled population and should therefore always be followed up individually.

CHAPTER 6

CONCLUSIONS AND OUTLOOK

6 Conclusions and Outlook

6.1 Reflection of research strategies for project aims

This thesis had three main project aims: (1) developing a high-throughput CRISPR-Cas9 toolbox that allows to dissect flagellar function *in vitro* and *in vivo*, (2) using this toolbox to dissect the mechanism and function of conserved flagellar proteins, in particular to understand how the flagella beat is regulated and (3) using this toolbox in combination with a bar-seq approach to study whether flagellar driven motility is essential for survival within the vector and the mammalian host.

The development of a CRISPR-Cas9 high-throughput toolbox was achieved through five key components: (1) using short HFs for integration of donor DNA, (2) using T7 RNAP for *in vivo* transcription from short sgRNA template DNA constructs, (3) implementing the pT-pPLOT modular plasmid system, (4) enabling barcoding of mutants for bar-seq phenotyping and (5) developing an online tool for primer design. Since this system is cloning - and *in vitro* transcription – free, it can be easily scaled up and used to dissect flagellar function *in vivo* and *in vitro*.

This toolbox has since been employed to start a knockout screen of more than 500 potential flagellar protein coding genes, which had been identified in a proteomic analysis of the promastigote flagellum (Beneke et al., 2019b). Since the flagellar knockout screen produced defined flagellar mutants on a scale that is unprecedented in other model organisms, it was invaluable in addressing the second overarching aim: to study the regulation of flagellar bending in *Leishmania*. A systematic dissection of the curling phenotype provided supporting evidence for the previously proposed central pair-spoke axis model (Omoto et al., 1999). Especially CPP and IDA f/I1 proteins were found to be involved in the regulation of

flagellar bending in *Leishmania*. Additionally, a model specific to *Leishmania* was proposed, in which the PFR supports flagellar bending by amplifying the principal bend.

Furthermore, studying the regulation of flagellar bending showed that LAX28 is required for IDA f/I1 assembly within the *Leishmania* axoneme. Interestingly, LAX28 has a potential orthologue in humans TEX47, which underlines once again the potential of *Leishmania* as a model organism for studying motile cilia and flagella.

The LeishGEdit toolbox was also irreplaceable to address the third overarching aim: to dissect the requirement of *Leishmania* motility for survival *in vivo*. Bar-seq experiments, involving sand fly and mouse infections, showed that directional motility is required for survival in the sand fly, but dispensable for survival in the mammalian host. Meanwhile, flagellar assembly is essential throughout the entire life cycle of *Leishmania*.

Overall, this thesis made new discoveries about the molecular machinery that powers and regulates flagellar motility, as well as the consequences of flagellar defects, in particular loss of motility, for survival in the vector and host.

6.2 The future of gene editing in *Leishmania*

The results in the chapters above demonstrate the precision and efficiency of the LeishGEdit toolbox. This system is scalable to a 96 well plate format, as it does not involve any cloning or *in vitro* transcription procedures compared to other gene editing systems available for kinetoplastids (see section 2.1.4). Together with the bar-seq approach for phenotyping mutants in pools *in vivo* and *in vitro* this is a suitable strategy not only for dissecting the function of more than 500 potential

flagellar protein coding genes, this is a valid approach for a genome-wide knockout screen in *Leishmania*.

A genome-wide knockout screen would be incredibly beneficial to study any aspect of *Leishmania* biology, especially if coupled with genome-wide protein tagging projects, such as Tryptag which is currently pursued in *T. brucei* (Dean et al., 2017). This would not only accelerate the functional dissection of all *Leishmania* proteins, but also to determine their cellular localization. However, this type of screen would be limited by two factors: (1) the possibility of targeting essential genes, meaning genes that are required for survival under certain conditions and (2) the assay used for determining the phenotype, for example bar-seq experiments can only be used to measure certain phenotypes in pools (e.g. growth defects, survival in mouse, colonialization of sand flies, as discussed in section 5.1.3).

Studying the loss-of-function phenotype of essential genes is particularly challenging, as most CRISPR-Cas9 gene editing systems available for kinetoplastids are not inducible. Meanwhile, RNAi systems are well established in *T. brucei* and allow for inducible gene knockdowns by using promotors that are regulated by a tetracycline repressor protein (Wirtz and Clayton, 1995). Thus, first steps towards inducible CRISPR systems have been made in *T. brucei* by using tetracycline inducible promoters to control expression of Cas9 (Rico et al., 2018). But an inducible gene editing system in *Leishmania* does not exist yet. While tetracycline inducible promoters have been also shown to be fully functional for conditional gene expression in *Leishmania* promastigotes (Kraeva et al., 2014; Kushnir et al., 2005; Yan et al., 2002; Yan et al., 2001; Yao et al., 2007), these systems are not widely used in the *Leishmania* community. However, a tetracycline

inducible system could be combined with the LeishGEdit toolbox. This would enable to generate conditional null mutants by introducing a tetracycline inducible episomal addback copy of the target ORF first and then deleting the target ORF in its genomic locus. It is also likely that other available inducible or conditional systems will be combined with CRISPR approaches for *Leishmania*. For example, the DiCre system (Duncan et al., 2016) and the Forced Plasmid Shuffle system (Murta et al., 2009) (see section 2.1.2) can be easily combined with the LeishGEdit toolbox. The DiCre system may be even repurposed to induce expression of Cas9 and T7 RNAP. Both genes could be integrated in reverse into the genome and "loxP-flopped" upon induction. Another approach might be to fuse the destabilization domain to a protein of interest to allow for induced protein degradation (Banaszynski et al., 2006). For example, one allele can be deleted while the other allele is *in situ* tagged with a destabilization domain. The destabilization domain has been effectively used in *Leishmania* (Madeira da Silva et al., 2009; Wheeler et al., 2015a), but my own attempts to use the destabilization domain for knockdown of IFT88 or to regulate expression of Cas9 and T7 RNAP by using "loxP-flop" sites have not yielded satisfying results (data not shown). Other future options for Cas9 induction may be also to use split-Cas9, meaning the Cas9 protein is dimerized from two monomers through rapamycin induction. However, further developments are required to reduce un-induced dimerization of split-Cas9 monomers (Zetsche et al., 2015).

A functional inducible knockout system that uses CRISPR for gene editing would be also an important step towards a genome-wide CRISPR-Cas9 knockout screen that relies only on sgRNAs and does not require additional donor DNA for effective gene editing. For example, for mammalian cell lines pooled libraries of CRISPR

sgRNAs are available, like the TKOv3 library (Hart et al., 2017), that can be used to transfect more than 70,000 sgRNAs in a single transfection and target more than 18,000 genes, thus nearly targeting every human gene at once. These kinds of genome-wide CRISPR library approaches have also already been used in Apicomplexan parasites, such as *T. gondii*, showing the potential of this strategy (Sidik et al., 2016) (see section 5.1.3). However, *T. gondii* has the great advantage of having a haploid genome, making it easy to manipulate (Black and Boothroyd, 2000), as only one gene copy needs to be targeted and as there isn't a second allele to provide a repair template. This is in contrast to mostly diploid *Leishmania* parasites (Figure 20). To-date no CRISPR tool that is capable of producing homogenous populations of specific mutants without additional co-transfection of donor DNA, additional drug selection or sub-cloning has been reported to function for genome-wide knockout screens in kinetoplastids. This may be related to a diploid genome organisation in *Leishmania*, but it may be also especially linked to the overall low DSB repair efficiency in *Leishmania*, resulting in cell death for more than half of CRISPR-Cas9 created DSBs (Zhang and Matlashewski, 2019) (see further details in section 2.1.4).

Although the bar-seq LeishGEdit knockout approach would enable a genome-wide knockout screen, it would be much more labour intensive to produce barcoded mutants individually for a pooled analysis than generating mutants already in pools. Thus, alternative strategies for genome-wide loss-of-function screens, including for example CRISPR targeting RNA enzymes like CasRx (Konermann et al., 2018), should be further explored.

6.3 The *Leishmania* flagellum as a model for motile cilia and flagella gives new insight into flagellar bending

Motility is not only an essential feature for *Leishmania* promastigotes within the fly, it is an important feature for many other ciliated or flagellated cells. Defects in cilia or flagella in humans cause ciliopathies. Modelling these defects in genetically modifiable organisms, such as *Leishmania*, are vital to study flagellar function and the cause of these diseases. Motility profiles of mutants missing specific axonemal and extra-axonemal flagellar proteins shown here, the analysis of the flagellar curling phenotype and the on-going CRISPR-Cas9 knockout screen of >500 potential flagellar protein coding genes in the Gluenz lab, of which parts were presented here, demonstrate the potential of *Leishmania* as a model organism for studying motile cilia and flagella.

However, many other model organisms (as listed in section 3.1.5) for studying cilia and flagella exist and each of them may have advantages that other systems don't have. For example, *Drosophila* is an excellent model for studying cilia and flagella in multicellular organisms (Zur Lage et al., 2019) because compared to other metazoan models, such as mice or zebrafish, ciliary and centrosome mutants are not embryonic lethal (Basto et al., 2006). Additionally, flies have many diverse cilia and flagella, including 9 + 2, 9 + 1 and 9 + 0 cilia with or without dynein arms (reviewed in Jana et al. (2016)). Also *Drosophila* sperm flagella assemble without need of IFT (Han et al., 2003) making it an ideal model organism to study other IFT-independent flagellar assembly pathways.

Despite these advantages, the classic model organism to-date is *C. reinhardtii* with many discoveries made over the past 60 years. Thus, it is questionable whether using of *Leishmania* for studying motile cilia and flagella remains a side project to research on the *in vivo* function of the *Leishmania* flagellum or whether the

Leishmania flagellum as a model system may pose advantages over other systems for the discovery of general mechanisms of flagellar motility.

As reviewed in Sasso et al. (2018) and Harris (2001), *C. reinhardtii* has been on the forefront for many years to explore the function of motile cilia and flagella. Main landmark discoveries have been made. For example, research in *C. reinhardtii* identified how IFT proteins maintain and assemble flagella (Kozminski et al., 1993). Furthermore, the flagellum and its many components have been detailed dissected, allowing to model mutations that cause ciliopathies, such as PCD (Pazour et al., 2000). Using *Chlamydomonas* as a laboratory species has several advantages. Cell culture is easy and cheap, but most importantly, *Chlamydomonas* reproduces sexually allowing to combine mutations from different strains into a single haploid strain (Harris, 2001; Sasso et al., 2018). Haploid genome organization also allows to introduce quick and easy genetic modifications (Cheng et al., 2017). More recently, CRISPR-Cas9 has been also successfully explored for genetic editing in *C. reinhardtii* (Baek et al., 2016; Shin et al., 2016). Additionally, there is now also the first whole genome indexed library of mapped insertion mutants for *C. reinhardtii* available, covering 83% of nuclear protein-coding genes (Li et al., 2019). If this library is linked with clever motility analysis screens, this will once more enhance flagellar research in *Chlamydomonas*. However, traditionally most flagellar protein coding genes have been discovered by isolation of mutants from *C. reinhardtii* mutation libraries, created by using UV-light (Levine, 1960). These mutations are then mapped back onto the genome, which has clearly broadened our understanding of many biological processes, but which sometimes leaves open questions about in how far a mutation that is causing an indel can be compared to a deletion.

This is a new key advantage of the CRISPR-Cas9 screen in *Leishmania* presented here, where not only a single mutation or an insertion into an ORF is linked to the observed phenotype, but instead the entire ORF is deleted and replaced by drug resistance markers. Still, there are many examples of single mutations that are well comparable with the deletion of the same gene. For example hydin knockdowns in *Chlamydomonas* and *T. brucei*, as well as *hy3/hy3* mouse mutants and complete deletion of hydin shown here, all cause paralyzed flagella (Dawe et al., 2007; Lechtreck et al., 2008; Lechtreck and Witman, 2007). However, mutants created by UV-light induced indel mutations require labor intensive follow-up experiments to check if any other locus was altered. Thus, loss-of-function approaches to study flagellar function in *Leishmania* and *T. brucei* are much more precise and scalable. Additionally, in *T. brucei* RNAi can be used to study flagellar function. However, the flagellum in *T. brucei* is attached to the entire cell body, making it less ideal to study flagellar bend regulation.

The other key advantage for *Leishmania* as a model organism for research on motile flagella and cilia is that the promastigotes are able to survive *in vitro* without assembling a flagellum. Aflagellate *ΔIFT88* mutants shown here and previously published work on *ΔIFT140* mutants (Sunter et al., 2018) demonstrates that promastigotes are viable *in vitro* without having an assembled flagellum. They only have a slight growth defect in promastigote culture, making the *Leishmania* flagellum an ideal system to study its components, flagellar assembly and flagellar length regulation. Additionally, *Leishmania* mutants of key axonemal structures, such as IDA, ODA, RSP and CPC, express similar mutant phenotypes as previously characterized *Chlamydomonas* mutations. They are also well comparable to observed phenotypes for the respective orthologous gene in

humans, which are often linked to PCDs. Work shown here on IDA f/I1 interactors and the newly characterized LAX28 flagellar gene with a potential orthologous in humans indicate that *Leishmania* can be well used to map more potential genetic causes to PCD. Only 25% of infertile men can be linked to a genetic cause so far (reviewed in Neto et al. (2016)). Thus, it remains vital to characterize flagellar components with orthologues in humans further in model organisms, such as *Leishmania*.

Meanwhile, the PFR in *Leishmania* flagella also allows research questions in the context of extra-axonemal flagellar structures, which are missing from *Chlamydomonas* flagella. Data shown here indicates that the PFR may contribute to flagellar bending by amplifying the principal bend. In how far this is relevant for other flagella with extra-axonemal structures or for *Leishmania in vivo*, especially within the vector, remains to be shown.

Besides linking more mutations in flagellar protein coding genes to PCD, the research focus of the cilia and flagella field seems to remain heavily on understanding the regulation of flagellar bending, in particular how the periodic switch between of dynein activity is controlled, and how cilia and flagella assembly is regulated and functions. The analysis of the flagellar curling phenotype shows that *Leishmania* can be well used to study flagellar bending and to draw conclusions on currently competing dynein regulation models. Data shown here indicates that regulation of flagellar bending in *Leishmania* follows the central pair-spoke axis model, with CPPs and IDA f/I1 associated proteins as key regulatory components. As discussed in chapter 3, there are many lines of evidence that this applies also for other eukaryotic flagella. However, further functional analyses through genetic, biochemical and especially cryo-ET reconstruction studies are

required to conclude on competing models for dynein regulation in *Leishmania* and more broadly.

6.4 General conclusion

This thesis project has led to the development of a high-throughput CRISPR-Cas9 gene editing and bar-seq screening toolbox for kinetoplastids (here termed LeishGEdit), allowing scalable functional analyses that have the potential to develop further into genome-wide knockout screens. Using LeishGEdit has allowed three major discoveries: (1) regulation of flagellar bending in *Leishmania* seems to follow the central pair-spoke axis model with an additional contribution of the PFR to amplify the principal bend, (2) LAX28 was identified as a novel IDA f/I1 assembly factor, mapping potentially a new ciliopathy associated gene in humans and (3) flagellar driven directional motility is essential for colonialization of the sand fly vector, but dispensable in the mammalian host. Thus, research presented here has strengthen *Leishmania* as a model organism for motile cilia and flagella. Meanwhile, the role of a promastigote to amastigote differentiating flagellum and amastigote flagellum remains elusive, but the tools developed here enable large screens that have the potential to shed new light onto this subject.

CHAPTER 7

MATERIALS AND METHODS

7 Materials and Methods

7.1 Cell culture

Promastigote-form *L. mexicana* (WHO strain MNYC/BZ/62/M379) were grown at 28°C in M199 medium (Life Technologies) supplemented with 2.2 g/L $NaHCO_3$, 0.005% haemin, 40 mM 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7.4 and 10% FCS. Media were supplemented with the relevant selection drugs: 32 µg/ml Hygromycin B, 20 µg/ml Puromycin Dihydrochloride, 5 µg/ml Blasticidin S Hydrochloride, 40 µg/ml G-418 Disulfate, 50 µg/ml Nourseothricin Sulfate and 25 µg/ml Phleomycin. *T. brucei SmOx P9* and *T. brucei SmOx B4* were cultured as described in Beneke et al. (2017).

A CASY model TT cell counter (Roche Diagnostics) with a 60 µm capillary was used to measure cell culture densities. Particles with a pseudo diameter below 2.0 µm were excluded.

7.2 LeishGEdit system and CRISPR mutants

7.2.1 Construction of pT knockout, pPLOT tagging and Cas9 T7 RNAP expression plasmids

pT knockout, pPLOT tagging and Cas9 T7 RNAP expression (pTB007, pTB011) plasmids were constructed as described in Beneke et al. (2017). pTB008 was constructed for expression of T7 RNAP by amplifying the calmodulin 5'UTR and the adjacent T7 RNAP ORF in pVY087 (Kraeva et al., 2014) and cloning it into B6182 (Madeira da Silva et al., 2009) using *Kas* I and *Xba* I cloning sites. All plasmid maps are available under www.LeishGEdit.net.

7.2.2 CRISPR-Cas9 gene knockouts and tags

Gene deletions and tags were essentially done as described in Beneke et al. (2017). For chapter 3 knockouts were generated on a 96 well plate as described in Beneke et al. (2019b). Primers were obtained from www.LeishGEdit.net and LeishGEdit v4.2 search results (Supplementary: File 2). Knockouts were generated in populations transfecting two amplified pT plasmids donor DNA constructs. If needed mutants were generated using only one drug resistance marker by transfecting only one pT donor DNA construct and isolating clones by limiting dilution. Tagged cell lines were generated by isolating clones of pPLOT transfected parasites. Tagged clones were then subjected to deletion in populations.

7.2.3 Diagnostic PCR for knockout verification

Genomic DNA of drug-selected mutants was isolated post transfection with the DNeasy Blood & Tissue Kit (Qiagen) after at least one 1:10 and one 1:100 dilution. Extracted genomic DNA was subjected to diagnostic PCRs to test for the presence of the target gene ORF in putative KO lines and the parental cell line as previously shown in (Beneke and Gluenz, 2019; Beneke et al., 2017) and using primers designed by Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012) as shown in Beneke et al. (2019b) and here (Supplementary: File 3). To show presence of DNA a second PCR reaction was performed to amplify the ORF of PFR2, blasticidin-S deaminase or PMM (Table 5). SBL barcoded parental cell lines were verified by amplifying the streptothricin acetyltransferase ORF (Table 5).

Amplified fragment of ORF	Primer	Sequence 5' to 3'
PFR2	Forward	CGCAGAAGGAGAAGAGCGAG
	Reverse	GTTGTACACGGACAGCTCCA
Blasticidin-S deaminase	Forward	GTATAATGCAGACCTGCTGC
	Reverse	CACCCTCATTGAAAGAGCAACGG
PMM	Forward	CTTTGACGTTGATGGCACCCTC
	Reverse	TGCGGTGAAACTCTTTGCCATC
SBL	Forward	TACCGGAAGGATTACATCTCG
	Reverse	GCGAAAAAGTGGGCACTAAGC

Table 5. Primer sequences used for diagnostic PCR to show presence of DNA.

7.2.4 Addback construction

PF16, MBO2 and BBS2 addbacks were constructed by cloning the ORF into the *Eco* RI and *Spe* I sites of pRM005 (Beneke et al., 2017). For other addbacks pRM005 was further modified to pTadd by cloning a multi-cloning restriction site in-between the *Eco* RI and *Spe* I sites. The LAX28 addback was made by cloning the ORF into *Spe* I and *Nco* I cloning sites of pTadd using primers that allow the fusion of two myc epitope tags to the N-terminus of the addback ORF: F: 5'-TTAGCAACTAGTATGGAACAGAAGCTGATCAGCGAAGAAGACCTGGAGCAA AAGCTCATTAGCGAGGAGGACCTCATGCCTCCGCGCAAGTCAGA-3' and R: 5'- TTAGCACCATGGGCGCGGGTTCATTTTATCGT-3'. 5 µg of circular plasmid DNA was transfected as described previously (Beneke et al., 2017) to allow episomal expression of *PF16*, *MBO2, BBS2* or *2xMyc::LAX28*. Drug resistance cells were selected using 25 µg/ml phleomycin.

7.3 Microscopy

7.3.1 Light microscopy

Leishmania parasites were prepared as described in (Wheeler et al., 2015b) and imaged live, with a Zeiss Axioimager.Z2 microscope with a 63× numerical aperture (NA) 1.40 oil immersion objective and a Hamamatsu ORCA-Flash4.0 camera at the ambient temperature of 25–28°C. Micrographs were taken all with a 3,000 ms exposure time for green fluorescent channels. For video microscopy of *Δhydin* mutants, 15 µl cell culture at approximately $1 \cdot 10^6$ cells/ml was placed on a 200 mesh sizes gold grid (Agar Scientific) embedded on a glass slide in a 250-µm deep chamber with a 1.5 mm cover slip and imaged using phase illumination with a 20×

NA 0.3 objective and a Hamamatsu ORCA-ER camera on an Olympus CKX41 microscope at the ambient temperature of 25–28°C.

7.3.2 Fixation and cytoskeleton preparation for TEM and SEM

For TEM and SEM, cells were prepared with a chemical fixation protocol similar to Hoog et al. (2010). Briefly, cells were fixed for TEM with 2.5% glutaraldehyde and 4% paraformaldehyde in M199 culture medium or for SEM with 2.5% glutaraldehyde in M199 culture medium for 2 hours at room temperature and then stored in the fridge overnight. Cells for TEM samples were fixed in culture flasks, while cells for SEM samples were fixed on coverslips in 24 well plates.

For SEM cytoskeleton preparations, cells were demembranated prior to fixation. Briefly, cells were collected, washed three times in phosphate buffered saline (PBS), settled on 24 well plates, washed with 0.1% NP-40 in PBS, washed three times in PBS and then fixed in 2.5% glutaraldehyde in PBS.

7.3.3 TEM

Fixed cells were washed six times for 10 min in 0.1 M PIPES buffer at pH 7.2. Wash four of six was supplement with 50 mM glycine. Cells were embedded in 4% low melting point agarose and incubated in 1% osmium tetroxide and 1.5% potassium ferrocyanide in 0.1 M PIPES buffer at 4 °C for 1 hour in darkness. Samples were then washed five times with ddH_2O for 5 min each time and stained with 0.5% uranyl acetate in darkness at 4°C overnight. Samples were dehydrated, embedded in epoxy resin, sectioned and on-section stained as described previously (Hoog et al., 2010). Electron micrographs were captured on a Tecnai 12 TEM (FEI) with an Ultrascan 1000 CCD camera (Gatan).

7.3.4 SEM

Fixed cells on coverslips were washed three times for 5 min in PBS and incubated in 1% osmium tetroxide in PBS at 4°C for 1 hour in darkness. Samples were then washed three times with ddH_2O for 5 min each time and dehydrated. Sample drying was completed by critical point drying using a Tousimis autosamdri-815. Coverslips were then sputter gold coated for 60 s. Samples were imaged at ~20,000-35,000x on a JEOL JSM-6390 SEM with an Everhart-Thornley secondary electron detector using an EHT target of 2 kV and a 20 µm aperture.

7.3.5 Image processing

All micrographs were processed using Fiji (Schindelin et al., 2012). To allow comparison between fluorescent signals in tagged cell lines in chapter 4, the same settings were used to display the green fluorescence channel. Channel settings in Fiji were set for CFAP44::mNG tagged cell lines 7,000 to 20,000, for IC140::eYFP tagged cell lines 6,000 to 11,000 and for LAX28::mNG tagged cell lines 5,000 to 12,000. Settings were identical for the tagged cell line and additional knockouts and addbacks on top of the tagged cell lines.

For average rotation of axonemes in TEM images, nine sections from each cell line were average rotated as previously described (Gadelha et al., 2006; Wheeler et al., 2015a). The resulting images were aligned using "Align image by line ROI" function in Fiji (Schindelin et al., 2012). A stack was generated from these aligned images and signals were averaged by using the “Z projection” function.

7.3.6 Motility analysis

Motility analysis was performed as previously described in Beneke et al. (2019b) using the method from Wheeler (2017). Directionality ([Velocity/Speed]) and mean speed was measured in triplicates, taking samples from cell cultures at a density of approximately $6 \cdot 10^6$ cells/ml. 5 µl of cell culture was placed on a glass slide in a 250-µm deep chamber covered with a 1.5 cover slip and imaged using darkfield illumination with a 10× NA 0.3 objective and a Hamamatsu ORCA-Flash4.0 camera on a Zeiss Axioimager.Z2 microscope at the ambient temperature of 25–28°C.

7.4 ATP reactivation assay

For the ATP reactivation assay $1 \cdot 10^7$ cells of the *L. mexicana* Cas9 T7 parental cell line and *Δhydin* mutant cell line were collected by centrifugation (800*g*, 3 minutes), washed three times in 500 µl PBS and resuspended in 100 µl PBS. Then 5 µl of the cell suspension was transferred onto a glass slide in a 250-µm deep chamber. Subsequently, 60 µl LRP buffer (24 mM glutamine, 140 mM sucrose, 20 mM Tris base, 1.5 mM $MgSO_4$, 2 mM EGTA, 0.25 mM EDTA, 1 mM DTT, pH 7.8 titrated with KOH, 0 – 1% Triton X-100, Mukhopadhyay and Dey (2016)) was added. Finally, 0.9 µl of 100 mM ATP (Thermo Fisher) were added and the suspension mixed. Each step was recorded by video microscopy (Supplementary: Video files 1) at the ambient temperature of 25–28°C as described above (section 7.3.1).

7.5 Bar-seq pooling assays and infections

7.5.1 Pooling of mutants

For the pilot bar-seq *in vitro* screen (section 5.2.1) mutants were pooled at a density of $1 \cdot 10^5$ cells/ml and grown for 48 hours. For the pilot bar-seq sand fly infection screen (section 5.2.1), mutants were pooled at a density of $8 \cdot 10^6$ cells/ml and immediately used for sand fly infections.

For the bar-seq *in vitro* growth assay (section 5.2.2) mutants were pooled at a density of $1 \cdot 10^6$ cells/ml and grown for 96 hours. The pool was divided into three replicates. DNA was extracted at least from $5 \cdot 10^6$ cells per time point and per replicate using DNeasy Blood & Tissue Kit (Qiagen). Mutant growth rates recorded *in vitro* were taken to divide mutants into four sub-pools. For sand fly infections (section 5.2.3) sub-pools were seeded at $5 \cdot 10^6$ cells/ml (pool 1), $3 \cdot 10^5$ cells/ml (pool 2), $1 \cdot 10^5$ cells/ml (pool 3) or $8 \cdot 10^4$ cells/ml (pool 4), respectively, and grown for 48 hours. For mouse infections (section 5.2.5) sub-pools were seeded at $6 \cdot 10^6$ cells/ml (pool 1), $3 \cdot 10^6$ cells/ml (pool 2), $1.5 \cdot 10^6$ cells/ml (pool 3) or $1 \cdot 10^6$ cells/ml (pool 4), respectively, and grown for 96 hours. The sub-pools were then mixed in equal proportions just before mouse or sand fly infection. Using different starting densities of sub-pools was chosen as a strategy to ensure that each mutant in the pool would be equally represented in the pool when infections are carried out in late log phase (for sand fly infection) or stationary phase (for mouse infection). Mutants for pool 1 were: *ΔPMM*, *ΔGDP*; for pool 2: *ΔIFT88*, *ΔARL-3A*, *ΔLPG1*, *ΔPMI*; for pool 3: *ΔIC140*, *ΔPF16*, *ΔCFAP44*, *ΔCFAP43*, *ΔRSP4/6*, *ΔdDC2*, *ΔLmxM.21.1110*, *ΔLC4-like*, *ΔFM458*, *ΔLmxM.18.1090*, *ΔLmxM.23.1020*, *ΔLmxM.17.0870*, *ΔBBS2*, *ΔKHARON 1* and for pool 4: SBL1-5.

7.5.2 *Lutzomyia longipalpis* infections, dissection and DNA extraction

Pooled populations, *L. mexicana* MNYC/BZ/62/M379 wild type, *L. mexicana* Cas9 T7 parental, *ΔPF16*, *ΔBBS2* and *ΔMBO2*, as well as their addback cell lines were used for sand fly infections. Infection of *Lutzomyia longipalpis* (from Jacobina, Brazil) with parasites and subsequent fly dissection as well as DNA extraction was carried out by Petr Volf's laboratory as described in Beneke et al. (2019b).

7.5.3 BALB/c infections and DNA extraction

Animal work was carried out according to the Animals (Scientific Procedures) Act of 1986, United Kingdom, was approved by the University of York Animal Welfare and Ethical Review Body (AWERB) committee and was carried out by Jeremy Mottram's laboratory (University of York). 18 female 4 – 6 weeks old BALB/c mice were subcutaneously infected at the left footpad with stationary promastigote pools using either $2 \cdot 10^6$ parasites (high dose) or $2 \cdot 10^5$ parasites (low dose) in 40 µl of sterile PBS. Footpad swelling was recorded (Castanys-Munoz et al., 2012) and increased generally from 1.8 mm to 2.5 mm over 6 weeks. After three and six weeks mice were euthanized by cervical dislocation and the infected left footpad was dissected from heel to toes avoiding bones and nails. Samples were immediately weighted, transferred to sterile Eppendorf tubes and kept at 20°C overnight. DNA was extracted in the following according to the manufacture instructions for tissue using DNeasy Blood & Tissue Kit (Qiagen).

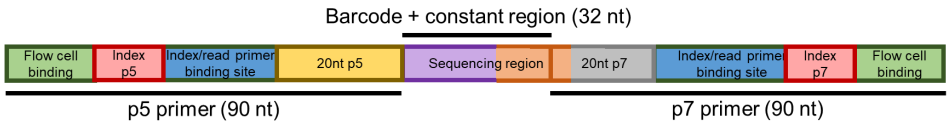
7.6 DNA sequencing

7.6.1 Library preparation for bar-seq samples

Extracted DNA from pools was prepared for Illumina sequencing using a single step library approach as described in Beneke et al. (2019b). p5 and p7 primers used for MiSeq and NextSeq amplicon sequencing contained flow cell binding sites for Illumina sequencing and additional indices to multiplex multiple samples (Figure 52). Indices were derived from Illumina Nextera (indices 501-517) and TruSeq (indices RPI1-RPI48) indexing kits (Table 6).

7.6.2 Library preparation for genome samples and Illumina sequencing

Leishmania genomic DNA of Cas9 T7 parental cells (before and after mouse passaged), *ΔPKAC1* (LmxM.34.4010), *ΔLmxM.08_29.1000*, *ΔCD047* (LmxM.09.0210), *Δhydin*, *ΔPF16* and *ΔLAX28* was prepared using Illumina TruSeq Nano DNA Library kit according to the manufacturer's instructions. The final sequencing pool was quantified by qPCR using NEB Library Quant Kit and library size was determined using Agilent High Sensitivity DNA Kit on a 2100 Bioanalyzer instrument. The final library was multiplexed with bar-seq library samples and other sequencing samples. To ensure high quality scores, the proportion of the bar-seq samples in the final sequencing pool was not allowed to be higher than 60%. Sequencing was performed using either a MiSeq kit v3 loaded with 8 pM (2x75 cycles, 6 and 8 cycles index read) or a NextSeq 500/550 High Output Kit v2.5 loaded with 1.8 pM (2x75 cycles, 6 and 8 cycles index read).



	Index number	Primer sequence 5' to 3'	Index for NextSeq 5' to 3'	Index for MiSeq 5' to 3'
p5 primer (Nextera index, Illumina)	502	AATGATACGGCGACCACCGAGATCTACACCTCTCTATACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	ATAGAGAG	CTCTCTAT
	503	AATGATACGGCGACCACCGAGATCTACACTATCCTCTACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	AGAGGATA	TATCCTCT
	504	AATGATACGGCGACCACCGAGATCTACACAGAGTAGAACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	TCTACTCT	AGAGTAGA
	505	AATGATACGGCGACCACCGAGATCTACACGTAAGGAGACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	CTCCTTAC	GTAAGGAG
	506	AATGATACGGCGACCACCGAGATCTACACACTGCATAACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	TATGCAGT	ACTGCATA
	507	AATGATACGGCGACCACCGAGATCTACACAAGGAGTAACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	TACTCCTT	AAGGAGTA
	508	AATGATACGGCGACCACCGAGATCTACACCTAAGCCTACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	AGGCTTAG	CTAAGCCT
	510	AATGATACGGCGACCACCGAGATCTACACCGTCTAATACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	ATTAGACG	CGTCTAAT
	511	AATGATACGGCGACCACCGAGATCTACACTCTCTCCGACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	CGGAGAGA	TCTCTCCG
	513	AATGATACGGCGACCACCGAGATCTACACTCGACTAGACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	CTAGTCGA	TCGACTAG
	515	AATGATACGGCGACCACCGAGATCTACACTTCTAGCTACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	AGCTAGAA	TTCTAGCT
	516	AATGATACGGCGACCACCGAGATCTACACCCTAGAGTACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	ACTCTAGG	CCTAGAGT
	517	AATGATACGGCGACCACCGAGATCTACACGCGTAAGAACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	TCTTACGC	GCGTAAGA
	518	AATGATACGGCGACCACCGAGATCTACACCTATTAAGACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	CTTAATAG	CTATTAAG
	520	AATGATACGGCGACCACCGAGATCTACACAAGGCTATACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	ATAGCCTT	AAGGCTAT
	521	AATGATACGGCGACCACCGAGATCTACACGAGCCTTAACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	TAAGGCTC	GAGCCTTA
	522	AATGATACGGCGACCACCGAGATCTACACTTATGCGAACACTCTTT CCCTACACGACGCTCTTCCGATCTGTGTATCGGATGTCAGTTGC	TCGCATAA	TTATGCGA
p7 primer (TruSeq index, Illumina)	40-2	CAAGCAGAAGACGGCATACGAGATTCTGAGGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	CTCAGA	CTCAGA
	41-2	CAAGCAGAAGACGGCATACGAGATGTCGTCGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	GACGAC	GACGAC
	42-2	CAAGCAGAAGACGGCATACGAGATCGATTAGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	TAATCG	TAATCG
	43-2	CAAGCAGAAGACGGCATACGAGATGCTGTAGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	TACAGC	TACAGC
	44-2	CAAGCAGAAGACGGCATACGAGATATTATAGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	TATAAT	TATAAT
	45-2	CAAGCAGAAGACGGCATACGAGATGAATGAGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	TCATTC	TCATTC
	46-2	CAAGCAGAAGACGGCATACGAGATTCGGGAGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	TCCCGA	TCCCGA
	47-2	CAAGCAGAAGACGGCATACGAGATCTTCGAGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	TCGAAG	TCGAAG
	48-2	CAAGCAGAAGACGGCATACGAGATTGCCGAGTGACTGGAGTTCCTT GGCACCCGAGAATTCCATCTCGTTTTCATCCGGCAGC	TCGGCA	TCGGCA

Table legend on following page.

Table 6. Custom designed p5 and p7 primers for Illumina amplicon sequencing.

Table shows p5 and p7 oligo sequences custom designed by Edward Hookway for single library step Illumina amplicon sequencing. Complete amplicon fragment for Illumina sequencing is shown on top, consisting of p5 primer sequence, the 17nt barcode and 15nt constant region, and the p7 primer sequence. Highlighted for p5 and p7 primer sequences are MiSeq/NextSeq flow cell binding sites (green), multiplex index sequence (red), index read primer binding site (blue) and overlap to constant region in pTBlast cassette (for p5 yellow; for p7 gray). Both index sequence columns on right indicate index read orientation for NextSeq and MiSeq, which should be used respectively for Illumina sequencing sample sheets.

7.6.3 Sanger sequencing

For Sanger sequencing two primers, spanning the integration site, were PCR amplified and reactions ran on a 2% agarose gel. The band was excised from the gel and purified using the QIAquick Gel Extraction Kit (Qiagen). While the forward primer used for the amplification varied depending on the target locus, the reverse primer sequence depended on the used pT plasmid (either pTBlast, pTNeo or pTPuro). Primers for amplification of fragments and primers used for Sanger sequencing are listed in Table 7. Sanger sequencing was performed by Source BioScience.

7.7 Computing

7.7.1 Barcode counting

Barcodes were counted as described in Beneke et al. (2019b). Briefly, MiSeq or NextSeq raw files were de-multiplexed using bcl2fastq (Illumina) and barcodes in de-multiplexed samples of bar-seq assays were counted using a bash script (Supplementary: Script 4). De-multiplexing and barcode counting were both performed with 0 nt MM in-between indexed samples or barcodes. Counted barcodes were normalized for each sample by calculating their abundance relative to all 25 barcodes in the pool. The “fitness” of each mutant in the pool was calculated by dividing normalized counts in the pooled population at the start of a bar-seq assay by normalized counts at the relevant time point in the assay (e.g. before feeding flies by 2 days PBM).

A

Amplified fragment of ORF	Primer	Sequence 5' to 3'
PF16	Forward	GACGTCGCTGCAGAAGTTATCCT
BBS2	Forward	GTCTCACCCGCTACCTGTACAG
DRC2	Forward	TCTCGCTGCTTCACTCACTACC
SBL	Forward	TGCACATTGACTCTCCACAGC
pTNeo	Reverse	GCCCGGTTCTTTTTGTCAAGAC
pTPuro	Reverse	GTTGACCCAGATCGACACATTGA
pTBlast	Reverse	CCGTTGCTCTTTCAATGAGGGTG

B

Sanger sequencing of amplicon	Sequence 5' to 3'
pTNeo	CTGTTCCTGTCGTACGAAAG
pTPuro	GCGTTGTGAAGCCGCGGATA
pTBlast	AGACAGGCATGCCTTGACTG

Table 7. Primer sequences for Sanger sequencing.

(A) Forward and reverse primer sequences used to amplify fragments. **(B)** Primer sequences used for Sanger sequencing.

7.7.2 Bam file assembly

De-multiplexed genome samples were ran through a bash wrapper script (Supplementary: Script 5), in which they were first assembled using Burrows-Wheeler Aligner (Li and Durbin, 2009), then sorted and indexed using Samtools (Li et al., 2009) and then viewed using IGV genome browser (Robinson et al., 2011).

Samtools (Li et al., 2009) was also used to determine total reads per chromosome and genome, allowing to calculate read coverage per chromosome.

7.7.3 Automated design of sgRNA and primers using LeishGEdit v4.2

For the design of sgRNA target sequences a 130 nt window upstream of the start codon or downstream of the stop codon was searched for the highest scoring sgRNA and chosen based on the CCTop scoring pipeline (Stemmer et al., 2015). Additionally, selected sgRNA sequences (23 nt, including the protospacer adjacent motif NGG) were verified by counting their number of perfect matches within the genome. The 30nt HFs of upstream forward primer (#1) and downstream reverse primer (#5) were designed to be immediately adjacent to the sgRNA target and PAM sequence. The script is written as bash script and can be run standalone in a single (Supplementary: Script 1) or batch mode (Supplementary: Script 2). The output of LeishGEdit v4.2 for all TritrypDB genomes (for DBv41) is attached (Supplementary: File 2).

7.7.4 BLASTp searches

BLASTp searches were performed under https://blast.ncbi.nlm.nih.gov/Blast.cgi using default settings. For reciprocal best protein blast searches NCBI sequence

IDs of *Chlamydomonas reinhardtii* (taxid:3055) were ran against organism *Leishmania mexicana* MHOM/GT/2001/U1103 (taxid:929439) and vice versa. E-values and reciprocal best protein blast search results were recorded using non-redundant protein sequence database version 5 (searches in database from February 2019 until May 2019).

7.8 Proteomics

7.8.1 Western blot

Western blots for detection of FLAG (Cas9), T7 RNAP and myc were done as described in Beneke et al. (2017). Antibody and buffer dilutions are listed in Table 8.

7.8.2 Salt extracted axonemes of *ΔCFAP44* mutants and parentals

A widely utilized isolation protocol for *T. brucei* flagellar skeletons (Robinson et al., 1991) was adapted for use in *L. mexicana*. Cells were first non-ionic detergent extracted in 1% octylglucoside to yield whole cell cytoskeletons and then salt extracted on ice using 1 M NaCl to depolymerize subpellicular microtubules. To avoid proteolytic degradation, all procedures were performed on ice or at 4°C during centrifugations. $1 \cdot 10^9$ *L. mexicana ΔCFAP44* mutants and *L. mexicana* Cas9 T7 parental cells were collected at 800*g* for 15 min and resuspended in 2.5 ml PBS, containing a protease inhibitor cocktail [final concentration, 1x Halt Protease Inhibitor Cocktail (Thermo Fisher), 500 µM Phenylmethylsulfonyl fluoride (PMSF) and 5 mM EDTA]. 50 µl of cell suspension was isolated and supplemented with 5 µl 20% SDS solution (Figure 40 (i)).

Antibody	Blocking buffer	1st antibody dilution	Wash buffer	2nd antibody
Anti - T7 RNAP, Novagen 70566	TBST* + 1% gelatin	1:2,500 in TBST*	TBST*	Anti - mouse 1:20,000 TBST* + 1% milk (Marvel)
Anti - myc, Millipore 05-724	TBST* + 5% Milk (Marvel)	1:3,000 in TBST* + 5% Milk (Marvel)	TBST* + 5% Milk (Marvel)	Anti - mouse 1:20,000 TBST* + 5% milk (Marvel)
Anti - FLAG M2, Sigma F3165	TBST* + 5% milk (Marvel)	1:25,000 in TBST* + 5% milk (Marvel)	TBST* + 5% milk (Marvel)	Anti - mouse 1:20,000 TBST* + 5% milk (Marvel)

*TBST: Tris - buffered saline with 0.05% Tween 20

Table 8. Anitbodies and buffers used in Western blots.

Table taken from Beneke et al. (2017).

The remaining cell suspension was centrifuged again and resuspended in 2.5 ml 10 mM PIPES [10 mM NaCl, 10 mM piperazine-N,N′-bis(2-ethanesulfonic acid, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 0.32 M sucrose, adjusted to pH 7.2], containing a protease inhibitor cocktail and octylglycoside (1% (w/v) final conc.). Cells were left on ice for 5 min and centrifuged for 10 min at 2,000*g*. 50 µl of supernatant was isolated and supplemented with 5 µl 20% SDS solution (Figure 40 (s/n 1)). 1.25 ml cold 10 mM PIPES buffer, containing a protease inhibitor cocktail, octylglycoside (1% (w/v) final conc.) and 1 M NaCl was added to the pellet and vortexed for 60 sec. After incubation for 30 min on ice, the solution was centrifuged at 12,200*g* for 20 min. 50 µl of supernatant was isolated and supplemented with 5 µl 20% SDS solution (Figure 40 (s/n 2)). The pellet was resuspended in 1 ml cold 10 mM PIPES buffer, containing a protease inhibitor cocktail and 0.32 M sucrose. The sample was loaded on top of a sucrose-bed containing one layer of 10 mM PIPES with 33% w/v sucrose [10 mM NaCl, 10 mM piperazine-N,N′-bis(2-ethanesulfonic acid, 1 mM $CaCl_2$, 1 mM $MgCl_2$, adjusted to pH 7.2 with 0.96M sucrose] and centrifuged at 800*g* for 15 min. While the top layer was centrifuged again at 12,200*g* for 20 min the remaining sucrose-bed was discarded. After centrifugation the supernatant was discarded and the pellet resuspended in 100 µl PBS containing a protease inhibitor cocktail and 2% SDS (Figure 40 (iv)). Proteins were quantified using BCA assay. For Coomassie gels samples were mixed with 4x Laemmli buffer (1x final concentration) and heated at 60°C for 10 minutes. For proteomic analysis ~4 µg of protein (~50 µl) of the final fraction (Figure 40 (iv)) were analysed by liquid tandem mass spectrometry.

7.8.3 Proteomics

Mass spectrometry analysis was carried out by Svenja Hester at the Advanced Proteomics Facility at the Sir William Dunn School of Pathology and Biochemistry Department. Protein samples were prepared using filter-aided sample preparation (FASP) digestion (Wisniewski et al., 2009) (Vicacon500, Sartorius, VN01H02 10kDA). The FASP filter were washed with 200 µl 0.1 % trifluoroacetic acid (TFA) in 50 % acetonitrile (ACN) (14,300*g*, 10min) and ~4 µg of protein was loaded onto the filter. Samples were denatured with 200 µl 8 M urea in 100 mM TEAB (triethylammonium bicarbonate) for 30 min at RT, reduced with 10 mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride) for 30 min at RT and alkylated with 50 mM CAA (chloroacetamide) for 30min at RT in the dark. FASP columns were centrifuged (14,300*g*, 10min) and washed first multiple times with 200 µl 6 M urea in 50 mM TEAB until no more bubbles form and then washed twice with 200 µl 1 M urea in 50 mM TEAB. FASP columns were centrifuged again. Samples were digested with 200 ng trypsin in 300 µl 50 mM TEAB overnight at 37°C. Columns were centrifuged and flow through was kept. Columns were washed with 200 µl 0.1 % TFA and with 200 µl 50 % ACN in 0.1 % TFA. The flow through was collected from both washes. Samples were dried in a SpeedVac, resuspend in 50 µl 5% formic acid and 5% DMSO and then trapped on a C18 PepMap100 pre-column (300µm i.d. x 5mm, 100 Å, Thermo Fisher Scientific) using solvent A (0.1% Formic Acid in water) at a pressure of 500 bar and separated on an Ultimate 3000 UHPLC system (Thermo Fischer Scientific) coupled to a QExactive mass spectrometer (Thermo Fischer Scientific) for proteomic analysis.

7.8.4 Proteomic analysis

MS-data were analysed as previously described (Beneke et al., 2019b). Briefly, MS-data was converted from .RAW to .MGF file using ProteoWizard and uploaded to the Central Proteomics Facilities Pipeline (CPFP Trudgian et al. (2010)). Protein lists were generated by using CPFP meta-searches against the predicted *L. mexicana* proteome (gene models based on (Fiebig et al., 2015b)), followed by label-free SINQ quantification. Proteomics data has been deposited to the ProteomeXchange Consortium via PRIDE (Vizcaino et al., 2016) (dataset: PXD014077).

7.8.5 Solubilisation of CFAP44

Solubilisation of CFAP44 was tested using a reporter cell line, expressing *in situ* tagged *CFAP44::eYFP* fusion protein. $1 \cdot 10^7$ cells were collected at 800*g* for 5 min and washed once in PBS. Cells were pelleted again and resuspended in 100 µl of 10 mM PIPES buffer (as above), containing a protease inhibitor cocktail (as above) and octylglycoside (1% (w/v) final conc.), as well as either 2 M LiCl, 2 M $CaCl_2$, 3.2 M KCl or 4 M NaCl. Lower concentrations were also tested including: 0.01 M, 0.1 M, 0.25 M, 0.5 M and 0.75 M NaCl; 1 M LiCl; 0.25 M, 0.5 M, 0.75 M, 1.0 M and 1.5 M $CaCl_2$. Cell suspensions were vortexed for 60 sec and incubated for 30 min on ice. The solution was centrifuged at 17,000*g* for 2 min and the pellet washed once in PBS. Cells were pelleted again, resuspended in 10 mM PIPES buffer and pipetted onto a glass slide for viewing under a microscope.

8 References

Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A. et al. (2017). RNA targeting with CRISPR-Cas13. *Nature* **550**, 280-284.

Adams, G. M., Huang, B., Piperno, G. and Luck, D. J. (1981). Central-pair microtubular complex of Chlamydomonas flagella: polypeptide composition as revealed by analysis of mutants. *J Cell Biol* **91**, 69-76.

Afzelius, B. (1959). Electron microscopy of the sperm tail; results obtained with a new fixative. *J Biophys Biochem Cytol* **5**, 269-78.

Afzelius, B. (1976a). [What happens when the cilia remain immobile?]. *Lakartidningen* **73**, 509-10.

Afzelius, B. A. (1976b). A human syndrome caused by immotile cilia. *Science* **193**, 317-9.

Alsford, S., Turner, D. J., Obado, S. O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C. and Horn, D. (2011). High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res* **21**, 915-24.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-402.

Amaral, A., Castillo, J., Estanyol, J. M., Ballesca, J. L., Ramalho-Santos, J. and Oliva, R. (2013). Human sperm tail proteome suggests new endogenous metabolic pathways. *Mol Cell Proteomics* **12**, 330-42.

Baek, K., Kim, D. H., Jeong, J., Sim, S. J., Melis, A., Kim, J. S., Jin, E. and Bae, S. (2016). DNA-free two-gene knockout in Chlamydomonas reinhardtii via CRISPR-Cas9 ribonucleoproteins. *Sci Rep* **6**, 30620.

Bahr, V., Stierhof, Y. D., Ilg, T., Demar, M., Quinten, M. and Overath, P. (1993). Expression of lipophosphoglycan, high-molecular weight phosphoglycan and glycoprotein 63 in promastigotes and amastigotes of Leishmania mexicana. *Mol Biochem Parasitol* **58**, 107-21.

Baldari, C. T. and Rosenbaum, J. (2010). Intraflagellar transport: it's not just for cilia anymore. *Curr Opin Cell Biol* **22**, 75-80.

Banaszynski, L. A., Chen, L. C., Maynard-Smith, L. A., Ooi, A. G. and Wandless, T. J. (2006). A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* **126**, 995-1004.

Barak, E., Amin-Spector, S., Gerliak, E., Goyard, S., Holland, N. and Zilberstein, D. (2005). Differentiation of Leishmania donovani in host-free system: analysis of signal perception and response. *Mol Biochem Parasitol* **141**, 99-108.

Barber, C. F., Heuser, T., Carbajal-Gonzalez, B. I., Botchkarev, V. V., Jr. and Nicastro, D. (2012). Three-dimensional structure of the radial spokes reveals heterogeneity and interactions with dyneins in Chlamydomonas flagella. *Mol Biol Cell* **23**, 111-20.

Barnes, R. L. and McCulloch, R. (2007). Trypanosoma brucei homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases. *Nucleic Acids Res* **35**, 3478-93.

Baron, D. M., Ralston, K. S., Kabututu, Z. P. and Hill, K. L. (2007). Functional genomics in Trypanosoma brucei identifies evolutionarily conserved components of motile flagella. *J Cell Sci* **120**, 478-91.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-12.

Basmaciyan, L., Azas, N. and Casanova, M. (2019a). A potential acetyltransferase involved in Leishmania major metacaspase-dependent cell death. *Parasit Vectors* **12**, 266.

Basmaciyan, L., Jacquet, P., Azas, N. and Casanova, M. (2019b). A novel hydrolase with a pro-death activity from the protozoan parasite Leishmania major. *Cell Death Discov* **5**, 99.

Bassett, A. and Liu, J. L. (2014a). CRISPR/Cas9 mediated genome engineering in Drosophila. *Methods* **69**, 128-36.

Bassett, A. R. and Liu, J. L. (2014b). CRISPR/Cas9 and genome editing in Drosophila. *J Genet Genomics* **41**, 7-19.

Bastin, P., Sherwin, T. and Gull, K. (1998). Paraflagellar rod is vital for trypanosome motility. *Nature* **391**, 548.

Basto, R., Lau, J., Vinogradova, T., Gardiol, A., Woods, C. G., Khodjakov, A. and Raff, J. W. (2006). Flies without centrioles. *Cell* **125**, 1375-86.

Bates, P. A. (2007). Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies. *Int J Parasitol* **37**, 1097-106.

Bayly, P. V. and Wilson, K. S. (2015). Analysis of unstable modes distinguishes mathematical models of flagellar motion. *J R Soc Interface* **12**.

Beneke, T. (2014). Proteomic analysis of the flagellar membrane composition of Leishmania mexicana. *Research Project at Brandenburg University of Technology Cottbus Senftenberg and University of Oxford*.

Beneke, T., Banecki, K., Fochler, S. and Gluenz, E. (2019a). LAX28 is required for assembly of the inner dynein arm I1 and tether/tether head complex in the Leishmania flagellum. *bioRxiv*.

Beneke, T., Demay, F., Hookway, E., Ashman, N., Jeffery, H., Smith, J., Valli, J., Becvar, T., Myskova, J., Lestinova, T. et al. (2019b). Genetic dissection of a Leishmania flagellar proteome demonstrates requirement for directional motility in sand fly infections. *PLoS Pathog* **15**, e1007828.

Beneke, T., Demay, F., Wheeler, R. J. and Gluenz, E. (2019c). Isolation of Leishmania promastigote flagella. *Methods Mol Biol* **Trypanosomatids**.

Beneke, T. and Gluenz, E. (2019). LeishGEdit: A Method for Rapid Gene Knockout and Tagging Using CRISPR-Cas9. *Methods Mol Biol* **1971**, 189-210.

Beneke, T., Madden, R., Makin, L., Valli, J., Sunter, J. and Gluenz, E. (2017). A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *R Soc Open Sci* **4**, 170095.

Bennett, R. R. and Golestanian, R. (2015). A steering mechanism for phototaxis in Chlamydomonas. *J R Soc Interface* **12**, 20141164.

Berbari, N. F., Lewis, J. S., Bishop, G. A., Askwith, C. C. and Mykytyn, K. (2008). Bardet-Biedl syndrome proteins are required for the localization of G protein-coupled receptors to primary cilia. *Proc Natl Acad Sci U S A* **105**, 4242-6.

Bessen, M., Fay, R. B. and Witman, G. B. (1980). Calcium control of waveform in isolated flagellar axonemes of Chlamydomonas. *J Cell Biol* **86**, 446-55.

Beverley, S. M. and Coburn, C. M. (1990). Recurrent de novo appearance of small linear DNAs in Leishmania major and relationship to extra-chromosomal DNAs in other species. *Mol Biochem Parasitol* **42**, 133-41.

Beverley, S. M., Coderre, J. A., Santi, D. V. and Schimke, R. T. (1984). Unstable DNA amplifications in methotrexate-resistant Leishmania consist of extrachromosomal circles which relocalize during stabilization. *Cell* **38**, 431-9.

Black, M. W. and Boothroyd, J. C. (2000). Lytic cycle of Toxoplasma gondii. *Microbiol Mol Biol Rev* **64**, 607-23.

Branche, C., Kohl, L., Toutirais, G., Buisson, J., Cosson, J. and Bastin, P. (2006). Conserved and specific functions of axoneme components in trypanosome motility. *J Cell Sci* **119**, 3443-55.

Broadhead, R., Dawe, H. R., Farr, H., Griffiths, S., Hart, S. R., Portman, N., Shaw, M. K., Ginger, M. L., Gaskell, S. J., McKean, P. G. et al. (2006). Flagellar motility is required for the viability of the bloodstream trypanosome. *Nature* **440**, 224-7.

Brokaw, C. J. (1971). Bend propagation by a sliding filament model for flagella. *J Exp Biol* **55**, 289-304.

Brokaw, C. J. (1972a). Computer simulation of flagellar movement. I. Demonstration of stable bend propagation and bend initiation by the sliding filament model. *Biophys J* **12**, 564-86.

Brokaw, C. J. (1972b). Flagellar movement: a sliding filament model. *Science* **178**, 455-62.

Brokaw, C. J. (1975). Effects of viscosity and ATP concentration on the movement of reactivated sea-urchin sperm flagella. *J Exp Biol* **62**, 701-19.

Brokaw, C. J. (1985). Computer simulation of flagellar movement. VI. Simple curvature-controlled models are incompletely specified. *Biophys J* **48**, 633-42.

Brown, N. P., Leroy, C. and Sander, C. (1998). MView: a web-compatible database search or multiple alignment viewer. *Bioinformatics* **14**, 380-1.

Bryant, J. M., Baumgarten, S., Glover, L., Hutchinson, S. and Rachidi, N. (2019). CRISPR in Parasitology: Not Exactly Cut and Dried! *Trends Parasitol* **35**, 409-422.

Bui, K. H., Yagi, T., Yamamoto, R., Kamiya, R. and Ishikawa, T. (2012). Polarity and asymmetry in the arrangement of dynein and related structures in the Chlamydomonas axoneme. *J Cell Biol* **198**, 913-25.

Bushell, E., Gomes, A. R., Sanderson, T., Anar, B., Girling, G., Herd, C., Metcalf, T., Modrzynska, K., Schwach, F., Martin, R. E. et al. (2017). Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes. *Cell* **170**, 260-272 e8.

Bussotti, G., Gouzelou, E., Cortes Boite, M., Kherachi, I., Harrat, Z., Eddaikra, N., Mottram, J. C., Antoniou, M., Christodoulou, V., Bali, A. et al. (2018). Leishmania Genome Dynamics during Environmental Adaptation Reveal Strain-Specific Differences in Gene Copy Number Variation, Karyotype Instability, and Telomeric Amplification. *MBio* **9**.

Cabello-Donayre, M., Orrego, L. M., Herraez, E., Vargas, P., Martinez-Garcia, M., Campos-Salinas, J., Perez-Victoria, I., Vicente, B., Marin, J. J. G. and Perez-Victoria, J. M. (2019). Leishmania heme uptake involves LmFLVCRb, a novel porphyrin transporter essential for the parasite. *Cell Mol Life Sci*.

Carbajal-Gonzalez, B. I., Heuser, T., Fu, X., Lin, J., Smith, B. W., Mitchell, D. R. and Nicastro, D. (2013). Conserved structural motifs in the central pair complex of eukaryotic flagella. *Cytoskeleton (Hoboken)* **70**, 101-120.

Castanys-Munoz, E., Brown, E., Coombs, G. H. and Mottram, J. C. (2012). Leishmania mexicana metacaspase is a negative regulator of amastigote proliferation in mammalian cells. *Cell Death Dis* **3**, e385.

Ceccaldi, R., Rondinelli, B. and D'Andrea, A. D. (2016). Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol* **26**, 52-64.

Chakravarty, J. and Sundar, S. (2010). Drug resistance in leishmaniasis. *J Glob Infect Dis* **2**, 167-76.

Chang, K. P. (1978). Leishmania infection of human skin fibroblasts in vitro: absence of phagolysosomal fusion after induced phagocytosis of promastigotes, and their intracellular transformation. *Am J Trop Med Hyg* **27**, 1084-96.

Chang, K. P. (1979). Leishmania donovani: promastigote--macrophage surface interactions in vitro. *Exp Parasitol* **48**, 175-89.

Chang, K. P. and Dwyer, D. M. (1976). Multiplication of a human parasite (Leishmania donovani) in phagolysosomes of hamster macrophages in vitro. *Science* **193**, 678-80.

Cheng, X., Liu, G., Ke, W., Zhao, L., Lv, B., Ma, X., Xu, N., Xia, X., Deng, X., Zheng, C. et al. (2017). Building a multipurpose insertional mutant library for forward and reverse genetics in Chlamydomonas. *Plant Methods* **13**, 36.

Chikne, V., Shanmugha Rajan, K., Shalev-Benami, M., Decker, K., Cohen-Chalamish, S., Madmoni, H., Biswas, V. K., Kumar Gupta, S., Doniger, T., Unger, R. et al. (2019). Small nucleolar RNAs controlling rRNA processing in Trypanosoma brucei. *Nucleic Acids Res* **47**, 2609-2629.

Child, F. M. (1959). The characterization of the cilia of Tetrahymena pyriformis. *Exp Cell Res* **18**, 258-67.

Clayton, C. (2019). Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open Biol* **9**, 190072.

Cohen, E. and Meininger, V. (1987). Ultrastructural analysis of primary cilium in the embryonic nervous tissue of mouse. *Int J Dev Neurosci* **5**, 43-51.

Cole, D. G., Diener, D. R., Himelblau, A. L., Beech, P. L., Fuster, J. C. and Rosenbaum, J. L. (1998). Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in Caenorhabditis elegans sensory neurons. *J Cell Biol* **141**, 993-1008.

Comai, L. and Howell, T. (2009). Barcode Generator 2.8. *GitHub*.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-23.

Cornillie, F. J., Lauweryns, J. M. and Corbeel, L. (1984). Atypical bronchial cilia in children with recurrent respiratory tract infections. A comparative ultrastructural study. *Pathol Res Pract* **178**, 595-604.

Costa, F. C., Francisco, A. F., Jayawardhana, S., Calderano, S. G., Lewis, M. D., Olmo, F., Beneke, T., Gluenz, E., Sunter, J., Dean, S. et al. (2018). Expanding the toolbox for Trypanosoma cruzi: A parasite line incorporating a bioluminescence-fluorescence dual reporter and streamlined CRISPR/Cas9 functionality for rapid in vivo localisation and phenotyping. *PLoS Negl Trop Dis* **12**, e0006388.

Courret, N., Frehel, C., Gouhier, N., Pouchelet, M., Prina, E., Roux, P. and Antoine, J. C. (2002). Biogenesis of Leishmania-harbouring parasitophorous vacuoles following phagocytosis of the metacyclic promastigote or amastigote stages of the parasites. *J Cell Sci* **115**, 2303-16.

Coutton, C., Vargas, A. S., Amiri-Yekta, A., Kherraf, Z. E., Ben Mustapha, S. F., Le Tanno, P., Wambergue-Legrand, C., Karaouzene, T., Martinez, G., Crouzy, S. et al. (2018). Mutations in CFAP43 and CFAP44 cause male infertility and flagellum defects in Trypanosoma and human. *Nat Commun* **9**, 686.

Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J., Joung, J. and Zhang, F. (2017). RNA editing with CRISPR-Cas13. *Science* **358**, 1019-1027.

Croft, S. L. (2001). Monitoring drug resistance in leishmaniasis. *Trop Med Int Health* **6**, 899-905.

Cruz, A. and Beverley, S. M. (1990). Gene replacement in parasitic protozoa. *Nature* **348**, 171-3.

Cruz, A. K., Titus, R. and Beverley, S. M. (1993). Plasticity in chromosome number and testing of essential genes in Leishmania by targeting. *Proc Natl Acad Sci U S A* **90**, 1599-603.

Cunningham, D. D. (1885). The presence of Peculiar parasitic organism in the tissue of a speciment of Delhi Boil. *Sci. Mem. Med. Off Army (India)* **1**, p. 21.

Curry, A. M., Williams, B. D. and Rosenbaum, J. L. (1992). Sequence analysis reveals homology between two proteins of the flagellar radial spoke. *Mol Cell Biol* **12**, 3967-77.

Cuvillier, A., Miranda, J. C., Ambit, A., Barral, A. and Merlin, G. (2003). Abortive infection of Lutzomyia longipalpis insect vectors by aflagellated LdARL-3A-Q70L overexpressing Leishmania amazonensis parasites. *Cell Microbiol* **5**, 717-28.

Cuvillier, A., Redon, F., Antoine, J. C., Chardin, P., DeVos, T. and Merlin, G. (2000). LdARL-3A, a Leishmania promastigote-specific ADP-ribosylation factor-like protein, is essential for flagellum integrity. *J Cell Sci* **113 (Pt 11)**, 2065-74.

Dallai, R., Lupetti, P. and Mencarelli, C. (2006). Unusual axonemes of hexapod spermatozoa. *Int Rev Cytol* **254**, 45-99.

Davy, B. E. and Robinson, M. L. (2003). Congenital hydrocephalus in hy3 mice is caused by a frameshift mutation in Hydin, a large novel gene. *Hum Mol Genet* **12**, 1163-70.

Dawe, H. R., Farr, H., Portman, N., Shaw, M. K. and Gull, K. (2005). The Parkin co-regulated gene product, PACRG, is an evolutionarily conserved axonemal protein that functions in outer-doublet microtubule morphogenesis. *J Cell Sci* **118**, 5421-30.

Dawe, H. R., Shaw, M. K., Farr, H. and Gull, K. (2007). The hydrocephalus inducing gene product, Hydin, positions axonemal central pair microtubules. *BMC Biol* **5**, 33.

de Lima, L. P., Calderano, S. G., da Silva, M. S., de Araujo, C. B., Vasconcelos, E. J. R., Iwai, L. K., Pereira, C. A., Fragoso, S. P. and Elias, M. C. (2019). Ortholog of the polymerase theta helicase domain modulates DNA replication in Trypanosoma cruzi. *Sci Rep* **9**, 2888.

Dean, S., Moreira-Leite, F. and Gull, K. (2019). Basalin is an evolutionarily unconstrained protein revealed via a conserved role in flagellum basal plate function. *Elife* **8**.

Dean, S., Sunter, J., Wheeler, R. J., Hodkinson, I., Gluenz, E. and Gull, K. (2015). A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol* **5**, 140197.

Dean, S., Sunter, J. D. and Wheeler, R. J. (2017). TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol* **33**, 80-82.

Delling, M., Indzhykulian, A. A., Liu, X., Li, Y., Xie, T., Corey, D. P. and Clapham, D. E. (2016). Primary cilia are not calcium-responsive mechanosensors. *Nature* **531**, 656-60.

Di-Blasi, T., Lobo, A. R., Nascimento, L. M., Cordova-Rojas, J. L., Pestana, K., Marin-Villa, M., Tempone, A. J., Telleria, E. L., Ramalho-Ortigao, M., McMahon-Pratt, D. et al. (2015). The flagellar protein FLAG1/SMP1 is a candidate for Leishmania-sand fly interaction. *Vector Borne Zoonotic Dis* **15**, 202-9.

DiBella, L. M., Sakato, M., Patel-King, R. S., Pazour, G. J. and King, S. M. (2004a). The LC7 light chains of Chlamydomonas flagellar dyneins interact with components required for both motor assembly and regulation. *Mol Biol Cell* **15**, 4633-46.

DiBella, L. M., Smith, E. F., Patel-King, R. S., Wakabayashi, K. and King, S. M. (2004b). A novel Tctex2-related light chain is required for stability of inner dynein arm I1 and motor function in the Chlamydomonas flagellum. *J Biol Chem* **279**, 21666-76.

Dick, W. M. (1997). Fungi, flagella and phylogeny. *Mycological Research* **101**, 385-394.

Diener, D. R., Ang, L. H. and Rosenbaum, J. L. (1993). Assembly of flagellar radial spoke proteins in Chlamydomonas: identification of the axoneme binding domain of radial spoke protein 3. *J Cell Biol* **123**, 183-90.

DiPetrillo, C. G. and Smith, E. F. (2011). The Pcdp1 complex coordinates the activity of dynein isoforms to produce wild-type ciliary motility. *Mol Biol Cell* **22**, 4527-38.

Donovan, C. (1903). On the possibility of the occurrence of trypanosomiasis in India. 1903. *Indian Med Gaz* **7**, 478.

Dostalova, A. and Volf, P. (2012). Leishmania development in sand flies: parasite-vector interactions overview. *Parasit Vectors* **5**, 276.

Duncan, S. M., Jones, N. G. and Mottram, J. C. (2017). Recent advances in Leishmania reverse genetics: Manipulating a manipulative parasite. *Mol Biochem Parasitol* **216**, 30-38.

Duncan, S. M., Myburgh, E., Philipon, C., Brown, E., Meissner, M., Brewer, J. and Mottram, J. C. (2016). Conditional gene deletion with DiCre demonstrates an essential role for CRK3 in Leishmania mexicana cell cycle regulation. *Mol Microbiol* **100**, 931-44.

Dupe, A., Dumas, C. and Papadopoulou, B. (2015). Differential Subcellular Localization of Leishmania Alba-Domain Proteins throughout the Parasite Development. *PLoS One* **10**, e0137243.

Dutcher, S. K., Huang, B. and Luck, D. J. (1984). Genetic dissection of the central pair microtubules of the flagella of Chlamydomonas reinhardtii. *J Cell Biol* **98**, 229-36.

Dymek, E. E., Heuser, T., Nicastro, D. and Smith, E. F. (2011). The CSC is required for complete radial spoke assembly and wild-type ciliary motility. *Mol Biol Cell* **22**, 2520-31.

Dymek, E. E. and Smith, E. F. (2007). A conserved CaM- and radial spoke associated complex mediates regulation of flagellar dynein activity. *J Cell Biol* **179**, 515-26.

Eckert, R. and Lux, H. D. (1976). A voltage-sensitive persistent calcium conductance in neuronal somata of Helix. *J Physiol* **254**, 129-51.

Edwards, B. F. L., Wheeler, R. J., Barker, A. R., Moreira-Leite, F. F., Gull, K. and Sunter, J. D. (2018). Direction of flagellum beat propagation is controlled by proximal/distal outer dynein arm asymmetry. *Proc Natl Acad Sci U S A* **115**, E7341-E7350.

el-On, J. and Hamburger, A. D. (1987). Topical treatment of New and Old World cutaneous leishmaniasis in experimental animals. *Trans R Soc Trop Med Hyg* **81**, 734-7.

El-Sayed, N. M., Myler, P. J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E. A., Hertz-Fowler, C. et al. (2005). Comparative genomics of trypanosomatid parasitic protozoa. *Science* **309**, 404-9.

Ernstrom, G. G. and Chalfie, M. (2002). Genetics of sensory mechanotransduction. *Annu Rev Genet* **36**, 411-53.

Fagerberg, L., Hallstrom, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K. et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics* **13**, 397-406.

Fernandez-Prada, C., Sharma, M., Plourde, M., Bresson, E., Roy, G., Leprohon, P. and Ouellette, M. (2018). High-throughput Cos-Seq screen with intracellular Leishmania infantum for the discovery of novel drug-resistance mechanisms. *Int J Parasitol Drugs Drug Resist* **8**, 165-173.

Fiebig, M., Kelly, S. and Gluenz, E. (2015a). Comparative Life Cycle Transcriptomics Revises Leishmania mexicana Genome Annotation and Links a Chromosome Duplication with Parasitism of Vertebrates. *PLoS Pathog* **11**, e1005186.

Fiebig, M., Kelly, S. and Gluenz, E. (2015b). Comparative lifecycle transcriptomics revises *Leishmania mexicana* genome annotation and links a chromosome duplication with parasitism of vertebrates. *PLoS Pathog* **11**, e1005186.

Field, M. C. and Carrington, M. (2009). The trypanosome flagellar pocket. *Nat Rev Microbiol* **7**, 775-86.

Forestier, C. L., Machu, C., Loussert, C., Pescher, P. and Spath, G. F. (2011). Imaging host cell-Leishmania interaction dynamics implicates parasite motility, lysosome recruitment, and host cell wounding in the infection process. *Cell Host Microbe* **9**, 319-30.

Franco, L. H., Beverley, S. M. and Zamboni, D. S. (2012). Innate immune activation and subversion of Mammalian functions by leishmania lipophosphoglycan. *J Parasitol Res* **2012**, 165126.

Fred, E. B. (1933). Antony van Leeuwenhoek: On the Three-hundredth Anniversary of his Birth. *J Bacteriol* **25**, iv 2-18.

Fu, G., Wang, Q., Phan, N., Urbanska, P., Joachimiak, E., Lin, J., Wloga, D. and Nicastro, D. (2018). The I1 dynein-associated tether and tether head complex is a conserved regulator of ciliary motility. *Mol Biol Cell* **29**, 1048-1059.

Fuge, H. (1969). Electron microscopic studies on the intra-flagellar structures of trypanosomes. *J Protozool* **16**, 460-6.

Fukuyama, K. and Epstein, W. L. (1973). Heterogeneous ultrastructure of keratohyalin granules: a comparative study of adjacent skin and mucous membrane. *J Invest Dermatol* **61**, 94-100.

Gadelha, C., Wickstead, B. and Gull, K. (2007). Flagellar and ciliary beating in trypanosome motility. *Cell Motil Cytoskeleton* **64**, 629-43.

Gadelha, C., Wickstead, B., McKean, P. G. and Gull, K. (2006). Basal body and flagellum mutants reveal a rotational constraint of the central pair microtubules in the axonemes of trypanosomes. *J Cell Sci* **119**, 2405-13.

Garami, A. and Ilg, T. (2001a). Disruption of mannose activation in Leishmania mexicana: GDP-mannose pyrophosphorylase is required for virulence, but not for viability. *EMBO J* **20**, 3657-66.

Garami, A. and Ilg, T. (2001b). The role of phosphomannose isomerase in Leishmania mexicana glycoconjugate synthesis and virulence. *J Biol Chem* **276**, 6566-75.

Garami, A., Mehlert, A. and Ilg, T. (2001). Glycosylation defects and virulence phenotypes of Leishmania mexicana phosphomannomutase and dolicholphosphate-mannose synthase gene deletion mutants. *Mol Cell Biol* **21**, 8168-83.

Gasiunas, G., Barrangou, R., Horvath, P. and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579-86.

Ghorbani, M. and Farhoudi, R. (2018). Leishmaniasis in humans: drug or vaccine therapy? *Drug Des Devel Ther* **12**, 25-40.

Gibbons, B. H. and Gibbons, I. R. (1972). Flagellar movement and adenosine triphosphatase activity in sea urchin sperm extracted with triton X-100. *J Cell Biol* **54**, 75-97.

Glaser, T. A., Moody, S. F., Handman, E., Bacic, A. and Spithill, T. W. (1991). An antigenically distinct lipophosphoglycan on amastigotes of Leishmania major. *Mol Biochem Parasitol* **45**, 337-44.

Gluenz, E., Ginger, M. L. and McKean, P. G. (2010a). Flagellum assembly and function during the Leishmania life cycle. *Curr Opin Microbiol* **13**, 473-9.

Gluenz, E., Hoog, J. L., Smith, A. E., Dawe, H. R., Shaw, M. K. and Gull, K. (2010b). Beyond 9+0: noncanonical axoneme structures characterize sensory cilia from protists to humans. *FASEB J* **24**, 3117-21.

Gomelsky, M. and Klug, G. (2002). BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem Sci* **27**, 497-500.

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. (2015). A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. *Cell Host Microbe* **17**, 404-413.

Goodenough, U. W. and Heuser, J. E. (1985). The Chlamydomonas cell wall and its constituent glycoproteins analyzed by the quick-freeze, deep-etch technique. *J Cell Biol* **101**, 1550-68.

Gopfert, M. C. and Robert, D. (2003). Motion generation by Drosophila mechanosensory neurons. *Proc Natl Acad Sci U S A* **100**, 5514-9.

Gossage, S. M., Rogers, M. E. and Bates, P. A. (2003). Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle. *Int J Parasitol* **33**, 1027-34.

Gradilone, S. A., Masyuk, A. I., Splinter, P. L., Banales, J. M., Huang, B. Q., Tietz, P. S., Masyuk, T. V. and Larusso, N. F. (2007). Cholangiocyte cilia express TRPV4 and detect changes in luminal tonicity inducing bicarbonate secretion. *Proc Natl Acad Sci U S A* **104**, 19138-43.

Grewal, J. S., Catta-Preta, C. M. C., Brown, E., Anand, J. and Mottram, J. C. (2019). Evaluation of Clan CD C11 peptidase PNT1 and other Leishmania mexicana cysteine peptidases as potential drug targets. *Biochimie*.

Gull, K. (1999). The cytoskeleton of trypanosomatid parasites. *Annu Rev Microbiol* **53**, 629-55.

Gull, K. (2003). Host-parasite interactions and trypanosome morphogenesis: a flagellar pocketful of goodies. *Curr Opin Microbiol* **6**, 365-70.

Halliday, C., Billington, K., Wang, Z., Madden, R., Dean, S., Sunter, J. D. and Wheeler, R. J. (2019). Cellular landmarks of Trypanosoma brucei and Leishmania mexicana. *Mol Biochem Parasitol* **230**, 24-36.

Han, T. X., Xu, X. Y., Zhang, M. J., Peng, X. and Du, L. L. (2010). Global fitness profiling of fission yeast deletion strains by barcode sequencing. *Genome Biol* **11**, R60.

Han, Y. G., Kwok, B. H. and Kernan, M. J. (2003). Intraflagellar transport is required in Drosophila to differentiate sensory cilia but not sperm. *Curr Biol* **13**, 1679-86.

Harris, E. H. (2001). Chlamydomonas as a Model Organism. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 363-406.

Harrison, A., Olds-Clarke, P. and King, S. M. (1998). Identification of the t complex-encoded cytoplasmic dynein light chain tctex1 in inner arm I1 supports the involvement of flagellar dyneins in meiotic drive. *J Cell Biol* **140**, 1137-47.

Hart, T., Tong, A. H. Y., Chan, K., Van Leeuwen, J., Seetharaman, A., Aregger, M., Chandrashekhar, M., Hustedt, N., Seth, S., Noonan, A. et al. (2017). Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 (Bethesda)* **7**, 2719-2727.

Hendrickson, T. W., Goss, J. L., Seaton, C. A. and Rohrs, H. W. (2013). The IC138 and IC140 intermediate chains of the I1 axonemal dynein complex bind directly to tubulin. *Biochim Biophys Acta* **1833**, 3265-3271.

Hendrickson, T. W., Perrone, C. A., Griffin, P., Wuichet, K., Mueller, J., Yang, P., Porter, M. E. and Sale, W. S. (2004). IC138 is a WD-repeat dynein intermediate chain required for light chain assembly and regulation of flagellar bending. *Mol Biol Cell* **15**, 5431-42.

Heuser, T., Barber, C. F., Lin, J., Krell, J., Rebesco, M., Porter, M. E. and Nicastro, D. (2012). Cryoelectron tomography reveals doublet-specific structures and unique interactions in the I1 dynein. *Proc Natl Acad Sci U S A* **109**, E2067-76.

Heuser, T., Raytchev, M., Krell, J., Porter, M. E. and Nicastro, D. (2009). The dynein regulatory complex is the nexin link and a major regulatory node in cilia and flagella. *J Cell Biol* **187**, 921-33.

Hodges, M. E., Scheumann, N., Wickstead, B., Langdale, J. A. and Gull, K. (2010). Reconstructing the evolutionary history of the centriole from protein components. *J Cell Sci* **123**, 1407-13.

Hodges, M. E., Wickstead, B., Gull, K. and Langdale, J. A. (2012). The evolution of land plant cilia. *New Phytol* **195**, 526-40.

Hoey, D. A., Downs, M. E. and Jacobs, C. R. (2012). The mechanics of the primary cilium: an intricate structure with complex function. *J Biomech* **45**, 17-26.

Hoog, J. L., Gluenz, E., Vaughan, S. and Gull, K. (2010). Ultrastructural investigation methods for Trypanosoma brucei. *Methods Cell Biol* **96**, 175-96.

Housden, B. E., Muhar, M., Gemberling, M., Gersbach, C. A., Stainier, D. Y., Seydoux, G., Mohr, S. E., Zuber, J. and Perrimon, N. (2017). Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. *Nat Rev Genet* **18**, 24-40.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O. et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827-32.

Huang, B., Ramanis, Z. and Luck, D. J. (1982). Suppressor mutations in Chlamydomonas reveal a regulatory mechanism for Flagellar function. *Cell* **28**, 115-24.

Huang, K., Kunkel, T. and Beck, C. F. (2004). Localization of the blue-light receptor phototropin to the flagella of the green alga Chlamydomonas reinhardtii. *Mol Biol Cell* **15**, 3605-14.

Hughes, L. C., Ralston, K. S., Hill, K. L. and Zhou, Z. H. (2012). Three-dimensional structure of the Trypanosome flagellum suggests that the paraflagellar rod functions as a biomechanical spring. *PLoS One* **7**, e25700.

Ihry, R. J., Worringer, K. A., Salick, M. R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C. et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med* **24**, 939-946.

Ikeda, K., Yamamoto, R., Wirschell, M., Yagi, T., Bower, R., Porter, M. E., Sale, W. S. and Kamiya, R. (2009). A novel ankyrin-repeat protein interacts with the regulatory proteins of inner arm dynein f (I1) of Chlamydomonas reinhardtii. *Cell Motil Cytoskeleton* **66**, 448-56.

Ilg, T. (2000). Lipophosphoglycan is not required for infection of macrophages or mice by Leishmania mexicana. *EMBO J* **19**, 1953-62.

Iseki, M., Matsunaga, S., Murakami, A., Ohno, K., Shiga, K., Yoshida, K., Sugai, M., Takahashi, T., Hori, T. and Watanabe, M. (2002). A blue-light-activated adenylyl cyclase mediates photoavoidance in Euglena gracilis. *Nature* **415**, 1047-51.

Ishemgulova, A., Hlavacova, J., Majerova, K., Butenko, A., Lukes, J., Votypka, J., Volf, P. and Yurchenko, V. (2018). CRISPR/Cas9 in Leishmania mexicana: A case study of LmxBTN1. *PLoS One* **13**, e0192723.

Ishikawa, T. (2017). Axoneme Structure from Motile Cilia. *Cold Spring Harb Perspect Biol* **9**.

Ishino, S., Mizukami, T., Yamaguchi, K., Katsumata, R. and Araki, K. (1987). Nucleotide sequence of the meso-diaminopimelate D-dehydrogenase gene from Corynebacterium glutamicum. *Nucleic Acids Res* **15**, 3917.

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. (2005). The genome of the kinetoplastid parasite, Leishmania major. *Science* **309**, 436-42.

Jaganathan, D., Ramasamy, K., Sellamuthu, G., Jayabalan, S. and Venkataraman, G. (2018). CRISPR for Crop Improvement: An Update Review. *Front Plant Sci* **9**, 985.

Jana, S. C., Mendonca, S., Werner, S. and Bettencourt-Dias, M. (2016). Methods to Study Centrosomes and Cilia in Drosophila. *Methods Mol Biol* **1454**, 215-36.

Jecna, L., Dostalova, A., Wilson, R., Seblova, V., Chang, K. P., Bates, P. A. and Volf, P. (2013). The role of surface glycoconjugates in Leishmania midgut attachment examined by competitive binding assays and experimental development in sand flies. *Parasitology* **140**, 1026-32.

Jiang, W., Brueggeman, A. J., Horken, K. M., Plucinak, T. M. and Weeks, D. P. (2014). Successful transient expression of Cas9 and single guide RNA genes in Chlamydomonas reinhardtii. *Eukaryot Cell* **13**, 1465-9.

Jin, H., White, S. R., Shida, T., Schulz, S., Aguiar, M., Gygi, S. P., Bazan, J. F. and Nachury, M. V. (2010). The conserved Bardet-Biedl syndrome proteins assemble a coat that traffics membrane proteins to cilia. *Cell* **141**, 1208-19.

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**, 816-21.

Jones, N. G., Catta-Preta, C. M. C., Lima, A. and Mottram, J. C. (2018). Genetically Validated Drug Targets in Leishmania: Current Knowledge and Future Prospects. *ACS Infect Dis* **4**, 467-477.

Kamiya, R., Kurimoto, E. and Muto, E. (1991). Two types of Chlamydomonas flagellar mutants missing different components of inner-arm dynein. *J Cell Biol* **112**, 441-7.

Kato-Minoura, T., Hirono, M. and Kamiya, R. (1997). Chlamydomonas inner-arm dynein mutant, ida5, has a mutation in an actin-encoding gene. *J Cell Biol* **137**, 649-56.

Kedzierski, L., Malby, R. L., Smith, B. J., Perugini, M. A., Hodder, A. N., Ilg, T., Colman, P. M. and Handman, E. (2006). Structure of Leishmania mexicana phosphomannomutase highlights similarities with human isoforms. *J Mol Biol* **363**, 215-27.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-58.

Killick-Kendrick, R., Molyneux, D. H. and Ashford, R. W. (1974). Leishmania in phlebotomid sandflies. I. Modifications of the flagellum associated with attachment to the mid-gut and oesophageal valve of the sandfly. *Proc R Soc Lond B Biol Sci* **187**, 409-19.

King, S. M. (2016). Axonemal Dynein Arms. *Cold Spring Harb Perspect Biol* **8**.

King, S. M. (2018). Turning dyneins off bends cilia. *Cytoskeleton (Hoboken)* **75**, 372-381.

Kinnamon, S. C. and Reynolds, S. D. (2009). Cell biology. Using taste to clear the air(ways). *Science* **325**, 1081-2.

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z. and Joung, J. K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-5.

Konermann, S., Lotfy, P., Brideau, N. J., Oki, J., Shokhirev, M. N. and Hsu, P. D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* **173**, 665-676 e14.

Koressaar, T. and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289-91.

Kotze, R. C. M., Muller, N., du Plessis, L. and van der Horst, G. (2019). The importance of insect sperm: Sperm ultrastructure of Hermetia illucens (black soldier fly). *Tissue Cell* **59**, 44-50.

Koyfman, A. Y., Schmid, M. F., Gheiratmand, L., Fu, C. J., Khant, H. A., Huang, D., He, C. Y. and Chiu, W. (2011). Structure of Trypanosoma brucei flagellum accounts for its bihelical motion. *Proc Natl Acad Sci U S A* **108**, 11105-8.

Kozminski, K. G., Johnson, K. A., Forscher, P. and Rosenbaum, J. L. (1993). A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proc Natl Acad Sci U S A* **90**, 5519-23.

Kraeva, N., Ishemgulova, A., Lukes, J. and Yurchenko, V. (2014). Tetracycline-inducible gene expression system in Leishmania mexicana. *Mol Biochem Parasitol* **198**, 11-3.

Krausz, C., Escamilla, A. R. and Chianese, C. (2015). Genetics of male infertility: from research to clinic. *Reproduction* **150**, R159-74.

Kubo, T., Hou, Y., Cochran, D. A., Witman, G. B. and Oda, T. (2018). A microtubule-dynein tethering complex regulates the axonemal inner dynein f (I1). *Mol Biol Cell* **29**, 1060-1074.

Kumar, R. and Engwerda, C. (2014). Vaccines to prevent leishmaniasis. *Clin Transl Immunology* **3**, e13.

Kushnir, S., Gase, K., Breitling, R. and Alexandrov, K. (2005). Development of an inducible protein expression system based on the protozoan host Leishmania tarentolae. *Protein Expr Purif* **42**, 37-46.

Kwon, R. Y., Temiyasathit, S., Tummala, P., Quah, C. C. and Jacobs, C. R. (2010). Primary cilium-dependent mechanosensing is mediated by adenylyl cyclase 6 and cyclic AMP in bone cells. *FASEB J* **24**, 2859-68.

Labuhn, M., Adams, F. F., Ng, M., Knoess, S., Schambach, A., Charpentier, E. M., Schwarzer, A., Mateo, J. L., Klusmann, J. H. and Heckl, D. (2018). Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications. *Nucleic Acids Res* **46**, 1375-1385.

Lachaud, L., Bourgeois, N., Kuk, N., Morelle, C., Crobu, L., Merlin, G., Bastien, P., Pages, M. and Sterkers, Y. (2014). Constitutive mosaic aneuploidy is a unique genetic feature widespread in the Leishmania genus. *Microbes Infect* **16**, 61-6.

Lander, N., Chiurillo, M. A. and Docampo, R. (2016). Genome Editing by CRISPR/Cas9: A Game Change in the Genetic Manipulation of Protists. *J Eukaryot Microbiol* **63**, 679-90.

Lander, N., Li, Z. H., Niyogi, S. and Docampo, R. (2015). CRISPR/Cas9-Induced Disruption of Paraflagellar Rod Protein 1 and 2 Genes in Trypanosoma cruzi Reveals Their Role in Flagellar Attachment. *MBio* **6**, e01012.

Lechtreck, K. F., Delmotte, P., Robinson, M. L., Sanderson, M. J. and Witman, G. B. (2008). Mutations in Hydin impair ciliary motility in mice. *J Cell Biol* **180**, 633-43.

Lechtreck, K. F., Johnson, E. C., Sakai, T., Cochran, D., Ballif, B. A., Rush, J., Pazour, G. J., Ikebe, M. and Witman, G. B. (2009). The Chlamydomonas reinhardtii BBSome is an IFT cargo required for export of specific signaling proteins from flagella. *J Cell Biol* **187**, 1117-32.

Lechtreck, K. F. and Witman, G. B. (2007). Chlamydomonas reinhardtii hydin is a central pair protein required for flagellar motility. *J Cell Biol* **176**, 473-82.

LeDizet, M. and Piperno, G. (1995). The light chain p28 associates with a subset of inner dynein arm heavy chains in Chlamydomonas axonemes. *Mol Biol Cell* **6**, 697-711.

Lehane, M. J. (1997). Peritrophic matrix structure and function. *Annu Rev Entomol* **42**, 525-50.

Lehti, M. S. and Sironen, A. (2017). Formation and function of sperm tail structures in association with sperm motility defects. *Biol Reprod* **97**, 522-536.

Leishman, W. B. (1903). On the possibility of the occurrence of trypanosomiasis in India. 1903. *Br Med J* **7**, 1252-1254.

Leprohon, P., Fernandez-Prada, C., Gazanion, E., Monte-Neto, R. and Ouellette, M. (2014). Drug resistance analysis by next generation sequencing in Leishmania. *Int J Parasitol*.

Lesich, K. A., dePinho, T. G., Dionne, B. J. and Lindemann, C. B. (2014). The effects of Ca2+ and ADP on dynein switching during the beat cycle of reactivated bull sperm models. *Cytoskeleton (Hoboken)* **71**, 611-27.

Levine, R. P. (1960). A screening technique for photosynthetic mutants in unicellular algae. *Nature* **188**, 339-40.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-60.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-9.

Li, X., Patena, W., Fauser, F., Jinkerson, R. E., Saroussi, S., Meyer, M. T., Ivanova, N., Robertson, J. M., Yue, R., Zhang, R. et al. (2019). A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat Genet* **51**, 627-635.

Liew, G. M., Ye, F., Nager, A. R., Murphy, J. P., Lee, J. S., Aguiar, M., Breslow, D. K., Gygi, S. P. and Nachury, M. V. (2014). The intraflagellar transport protein IFT27 promotes BBSome exit from cilia through the GTPase ARL6/BBS3. *Dev Cell* **31**, 265-278.

Lin, J., Heuser, T., Song, K., Fu, X. and Nicastro, D. (2012). One of the nine doublet microtubules of eukaryotic flagella exhibits unique and partially conserved structures. *PLoS One* **7**, e46494.

Lin, J. and Nicastro, D. (2018). Asymmetric distribution and spatial switching of dynein activity generates ciliary motility. *Science* **360**.

Lindemann, C. B. (1994). A model of flagellar and ciliary functioning which uses the forces transverse to the axoneme as the regulator of dynein activation. *Cell Motil Cytoskeleton* **29**, 141-54.

Lindemann, C. B. and Lesich, K. A. (2010). Flagellar and ciliary beating: the proven and the possible. *J Cell Sci* **123**, 519-28.

Lindemann, C. B. and Rikmenspoel, R. (1972). Sperm flagellar motion maintained by ADP. *Exp Cell Res* **73**, 255-9.

Liu, P. and Lechtreck, K. F. (2018). The Bardet-Biedl syndrome protein complex is an adapter expanding the cargo range of intraflagellar transport trains for ciliary export. *Proc Natl Acad Sci U S A* **115**, E934-E943.

Lodish, H., Berk, A., Zipursky, S., Matsudaira, P., Baltimore, D. and Darnell, J. (2000). Cilia and Flagella: Structure and Movement. *New York: W. H. Freeman and Company.* **Molecular Cell Biology. 4th edition.**

Loreng, T. D. and Smith, E. F. (2017). The Central Apparatus of Cilia and Eukaryotic Flagella. *Cold Spring Harb Perspect Biol* **9**.

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

Madeira da Silva, L., Owens, K. L., Murta, S. M. and Beverley, S. M. (2009). Regulated expression of the Leishmania major surface virulence factor lipophosphoglycan using conditionally destabilized fusion proteins. *Proc Natl Acad Sci U S A* **106**, 7583-8.

Maga, J. A., Sherwin, T., Francis, S., Gull, K. and LeBowitz, J. H. (1999). Genetic dissection of the Leishmania paraflagellar rod, a unique flagellar cytoskeleton structure. *J Cell Sci* **112 (Pt 16)**, 2753-63.

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J., Charpentier, E., Haft, D. H. et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**, 722-36.

Makin, L. and Gluenz, E. (2015). cAMP signalling in trypanosomatids: role in pathogenesis and as a drug target. *Trends Parasitol* **31**, 373-9.

Malone, A. M., Anderson, C. T., Stearns, T. and Jacobs, C. R. (2007a). Primary cilia in bone. *J Musculoskelet Neuronal Interact* **7**, 301.

Malone, A. M., Anderson, C. T., Tummala, P., Kwon, R. Y., Johnston, T. R., Stearns, T. and Jacobs, C. R. (2007b). Primary cilia mediate mechanosensing in bone cells by a calcium-independent mechanism. *Proc Natl Acad Sci U S A* **104**, 13325-30.

Mannaert, A., Downing, T., Imamura, H. and Dujardin, J. C. (2012). Adaptive mechanisms in pathogens: universal aneuploidy in Leishmania. *Trends Parasitol* **28**, 370-6.

Martel, D., Beneke, T., Gluenz, E., Spath, G. F. and Rachidi, N. (2017). Characterisation of Casein Kinase 1.1 in Leishmania donovani Using the CRISPR Cas9 Toolkit. *Biomed Res Int* **2017**, 4635605.

Masuda, S. and Bauer, C. E. (2002). AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in Rhodobacter sphaeroides. *Cell* **110**, 613-23.

Masyuk, T. V., Masyuk, A. I., Torres, V. E., Harris, P. C. and Larusso, N. F. (2007). Octreotide inhibits hepatic cystogenesis in a rodent model of polycystic liver disease by reducing cholangiocyte adenosine 3',5'-cyclic monophosphate. *Gastroenterology* **132**, 1104-16.

McConville, M. J. and Blackwell, J. M. (1991). Developmental changes in the glycosylated phosphatidylinositols of Leishmania donovani. Characterization of the promastigote and amastigote glycolipids. *J Biol Chem* **266**, 15170-9.

McConville, M. J., Mullin, K. A., Ilgoutz, S. C. and Teasdale, R. D. (2002). Secretory pathway of trypanosomatid parasites. *Microbiol Mol Biol Rev* **66**, 122-54; table of contents.

Merritt, C. and Stuart, K. (2013). Identification of essential and non-essential protein kinases by a fusion PCR method for efficient production of transgenic Trypanosoma brucei. *Mol Biochem Parasitol* **190**, 44-9.

Mitchell, D. R. (2003). Reconstruction of the projection periodicity and surface architecture of the flagellar central pair complex. *Cell Motil Cytoskeleton* **55**, 188-99.

Mitchell, D. R. (2017). Evolution of Cilia. *Cold Spring Harb Perspect Biol* **9**.

Mitchell, D. R. and Nakatsugawa, M. (2004). Bend propagation drives central pair rotation in Chlamydomonas reinhardtii flagella. *J Cell Biol* **166**, 709-15.

Mitchell, D. R. and Sale, W. S. (1999). Characterization of a Chlamydomonas insertional mutant that disrupts flagellar central pair microtubule-associated structures. *J Cell Biol* **144**, 293-304.

Mitchison, H. M. and Valente, E. M. (2017). Motile and non-motile cilia in human pathology: from function to phenotypes. *J Pathol* **241**, 294-309.

Mittra, B., Cortez, M., Haydock, A., Ramasamy, G., Myler, P. J. and Andrews, N. W. (2013). Iron uptake controls the generation of Leishmania infective forms through regulation of ROS levels. *J Exp Med* **210**, 401-16.

Mittra, B., Laranjeira-Silva, M. F., Miguel, D. C., Perrone Bezerra de Menezes, J. and Andrews, N. W. (2017). The iron-dependent mitochondrial superoxide dismutase SODA promotes Leishmania virulence. *J Biol Chem* **292**, 12324-12338.

Moafi, M., Rezvan, H., Sherkat, R. and Taleban, R. (2019). Leishmania Vaccines Entered in Clinical Trials: A Review of Literature. *Int J Prev Med* **10**, 95.

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**, 174-82.

Mojica, F. J., Juez, G. and Rodriguez-Valera, F. (1993). Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. *Mol Microbiol* **9**, 613-21.

Moody, S. F., Handman, E., McConville, M. J. and Bacic, A. (1993). The structure of Leishmania major amastigote lipophosphoglycan. *J Biol Chem* **268**, 18457-66.

Mukhopadhyay, A. G. and Dey, C. S. (2016). Reactivation of flagellar motility in demembranated Leishmania reveals role of cAMP in flagellar wave reversal to ciliary waveform. *Sci Rep* **6**, 37308.

Murta, S. M., Vickers, T. J., Scott, D. A. and Beverley, S. M. (2009). Methylene tetrahydrofolate dehydrogenase/cyclohydrolase and the synthesis of 10-CHO-THF are essential in Leishmania major. *Mol Microbiol* **71**, 1386-401.

Myster, S. H., Knott, J. A., O'Toole, E. and Porter, M. E. (1997). The Chlamydomonas Dhc1 gene encodes a dynein heavy chain subunit required for assembly of the I1 inner arm complex. *Mol Biol Cell* **8**, 607-20.

Myster, S. H., Knott, J. A., Wysocki, K. M., O'Toole, E. and Porter, M. E. (1999). Domains in the 1alpha dynein heavy chain required for inner arm assembly and flagellar motility in Chlamydomonas. *J Cell Biol* **146**, 801-18.

Nachury, M. V., Loktev, A. V., Zhang, Q., Westlake, C. J., Peranen, J., Merdes, A., Slusarski, D. C., Scheller, R. H., Bazan, J. F., Sheffield, V. C. et al.

(2007). A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. *Cell* **129**, 1201-13.

Nakazawa, Y., Ariyoshi, T., Noga, A., Kamiya, R. and Hirono, M. (2014). Space-dependent formation of central pair microtubules and their interactions with radial spokes. *PLoS One* **9**, e110513.

Nauli, S. M., Alenghat, F. J., Luo, Y., Williams, E., Vassilev, P., Li, X., Elia, A. E., Lu, W., Brown, E. M., Quinn, S. J. et al. (2003). Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nat Genet* **33**, 129-37.

Nenarokova, A., Zahonova, K., Krasilnikova, M., Gahura, O., McCulloch, R., Zikova, A., Yurchenko, V. and Lukes, J. (2019). Causes and Effects of Loss of Classical Nonhomologous End Joining Pathway in Parasitic Eukaryotes. *MBio* **10**.

Neto, F. T., Bach, P. V., Najari, B. B., Li, P. S. and Goldstein, M. (2016). Genetics of Male Infertility. *Curr Urol Rep* **17**, 70.

Ngo, H., Tschudi, C., Gull, K. and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. *Proc Natl Acad Sci U S A* **95**, 14687-92.

O'Connell, M. R., Oakes, B. L., Sternberg, S. H., East-Seletsky, A., Kaplan, M. and Doudna, J. A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516**, 263-6.

Oberholzer, M., Bregy, P., Marti, G., Minca, M., Peier, M. and Seebeck, T. (2007a). Trypanosomes and mammalian sperm: one of a kind? *Trends Parasitol* **23**, 71-7.

Oberholzer, M., Langousis, G., Nguyen, H. T., Saada, E. A., Shimogawa, M. M., Jonsson, Z. O., Nguyen, S. M., Wohlschlegel, J. A. and Hill, K. L. (2011). Independent analysis of the flagellum surface and matrix proteomes provides insight into flagellum signaling in mammalian-infectious Trypanosoma brucei. *Mol Cell Proteomics* **10**, M111 010538.

Oberholzer, M., Lopez, M. A., Ralston, K. S. and Hill, K. L. (2009). Approaches for functional analysis of flagellar proteins in African trypanosomes. *Methods Cell Biol* **93**, 21-57.

Oberholzer, M., Marti, G., Baresic, M., Kunz, S., Hemphill, A. and Seebeck, T. (2007b). The Trypanosoma brucei cAMP phosphodiesterases TbrPDEB1 and TbrPDEB2: flagellar enzymes that are essential for parasite virulence. *FASEB J* **21**, 720-31.

Oberle, M., Balmer, O., Brun, R. and Roditi, I. (2010). Bottlenecks and the maintenance of minor genotypes during the life cycle of Trypanosoma brucei. *PLoS Pathog* **6**, e1001023.

Oda, T., Yanagisawa, H. and Kikkawa, M. (2015). Detailed structural and biochemical characterization of the nexin-dynein regulatory complex. *Mol Biol Cell* **26**, 294-304.

Oda, T., Yanagisawa, H., Yagi, T. and Kikkawa, M. (2014). Mechanosignaling between central apparatus and radial spokes controls axonemal dynein activity. *J Cell Biol* **204**, 807-19.

Okutman, O., Rhouma, M. B., Benkhalifa, M., Muller, J. and Viville, S. (2018). Genetic evaluation of patients with non-syndromic male infertility. *J Assist Reprod Genet* **35**, 1939-1951.

Olbrich, H., Schmidts, M., Werner, C., Onoufriadis, A., Loges, N. T., Raidt, J., Banki, N. F., Shoemark, A., Burgoyne, T., Al Turki, S. et al. (2012).

Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry. *Am J Hum Genet* **91**, 672-84.

Omoto, C. K., Gibbons, I. R., Kamiya, R., Shingyoji, C., Takahashi, K. and Witman, G. B. (1999). Rotation of the central pair microtubules in eukaryotic flagella. *Mol Biol Cell* **10**, 1-4.

Omoto, C. K. and Kung, C. (1980). Rotation and twist of the central-pair microtubules in the cilia of Paramecium. *J Cell Biol* **87**, 33-46.

Oumeish, O. Y. (1999). Cutaneous leishmaniasis: a historical perspective. *Clin Dermatol* **17**, 249-54.

Pan, J. and Snell, W. J. (2000). Signal transduction during fertilization in the unicellular green alga, Chlamydomonas. *Curr Opin Microbiol* **3**, 596-602.

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.

Pazour, G. J., Dickert, B. L., Vucica, Y., Seeley, E. S., Rosenbaum, J. L., Witman, G. B. and Cole, D. G. (2000). Chlamydomonas IFT88 and its mouse homologue, polycystic kidney disease gene tg737, are required for assembly of cilia and flagella. *J Cell Biol* **151**, 709-18.

Pearson, R. D., Sullivan, J. A., Roberts, D., Romito, R. and Mandell, G. L. (1983). Interaction of Leishmania donovani promastigotes with human phagocytes. *Infect Immun* **40**, 411-6.

Peng, D., Kurup, S. P., Yao, P. Y., Minning, T. A. and Tarleton, R. L. (2014). CRISPR-Cas9-mediated single-gene and gene family disruption in Trypanosoma cruzi. *MBio* **6**, e02097-14.

Peng, D. and Tarleton, R. (2015). EuPaGDT: a web tool tailored to design CRISPR guide RNAs for eukaryotic pathogens. *Microb Genom* **1**, e000033.

Perrone, C. A., Yang, P., O'Toole, E., Sale, W. S. and Porter, M. E. (1998). The Chlamydomonas IDA7 locus encodes a 140-kDa dynein intermediate chain required to assemble the I1 inner arm complex. *Mol Biol Cell* **9**, 3351-65.

Pimenta, P. F., Modi, G. B., Pereira, S. T., Shahabuddin, M. and Sacks, D. L. (1997). A novel role for the peritrophic matrix in protecting Leishmania from the hydrolytic activities of the sand fly midgut. *Parasitology* **115 (Pt 4)**, 359-69.

Pimenta, P. F., Saraiva, E. M., Rowton, E., Modi, G. B., Garraway, L. A., Beverley, S. M., Turco, S. J. and Sacks, D. L. (1994). Evidence that the vectorial competence of phlebotomine sand flies for different species of Leishmania is controlled by structural polymorphisms in the surface lipophosphoglycan. *Proc Natl Acad Sci U S A* **91**, 9155-9.

Piperno, G. (1990). Functional diversity of dyneins. *Cell Motil Cytoskeleton* **17**, 147-9.

Pomel, S., Rodrigo, J., Hendra, F., Cave, C. and Loiseau, P. M. (2012). In silico analysis of a therapeutic target in Leishmania infantum: the guanosine-diphospho-D-mannose pyrophosphorylase. *Parasite* **19**, 63-70.

Poon, S. K., Peacock, L., Gibson, W., Gull, K. and Kelly, S. (2012). A modular and optimized single marker system for generating Trypanosoma brucei cell lines expressing T7 RNA polymerase and the tetracycline repressor. *Open Biol* **2**, 110037.

Porter, M. E., Power, J. and Dutcher, S. K. (1992). Extragenic suppressors of paralyzed flagellar mutations in Chlamydomonas reinhardtii identify loci that alter the inner dynein arms. *J Cell Biol* **118**, 1163-76.

Porter, M. E. and Sale, W. S. (2000). The 9 + 2 axoneme anchors multiple inner arm dyneins and a network of kinases and phosphatases that control motility. *J Cell Biol* **151**, F37-42.

Portman, N. and Gull, K. (2010). The paraflagellar rod of kinetoplastid parasites: from structure to components and function. *Int J Parasitol* **40**, 135-48.

Portman, N., Lacomble, S., Thomas, B., McKean, P. G. and Gull, K. (2009). Combining RNA interference mutants and comparative proteomics to identify protein components and dependences in a eukaryotic flagellum. *J Biol Chem* **284**, 5610-9.

Pourcel, C., Salvignol, G. and Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653-63.

Praetorius, H. A. and Spring, K. R. (2001). Bending the MDCK cell primary cilium increases intracellular calcium. *J Membr Biol* **184**, 71-9.

Praetorius, H. A. and Spring, K. R. (2003a). Removal of the MDCK cell primary cilium abolishes flow sensing. *J Membr Biol* **191**, 69-76.

Praetorius, H. A. and Spring, K. R. (2003b). The renal cell primary cilium functions as a flow sensor. *Curr Opin Nephrol Hypertens* **12**, 517-20.

Prensier, G., Vivier, E., Goldstein, S. and Schrevel, J. (1980). Motile flagellum with a "3 + 0" ultrastructure. *Science* **207**, 1493-4.

Preston, R. and Saimi, Y. (1990). Ca2+ and motility in Paramecium. In: Bloodgood R, editor. Ciliary and Flagellar Membranes. *Plenum Press*, pp. 173–200.

Prevo, B., Scholey, J. M. and Peterman, E. J. G. (2017). Intraflagellar transport: mechanisms of motor action, cooperation, and cargo delivery. *FEBS J* **284**, 2905-2931.

Price, H. P., Paape, D., Hodgkinson, M. R., Farrant, K., Doehl, J., Stark, M. and Smith, D. F. (2013). The Leishmania major BBSome subunit BBS1 is essential for parasite virulence in the mammalian host. *Mol Microbiol* **90**, 597-611.

Ralston, K. S., Lerner, A. G., Diener, D. R. and Hill, K. L. (2006). Flagellar motility contributes to cytokinesis in Trypanosoma brucei and is modulated by an evolutionarily conserved dynein regulatory system. *Eukaryot Cell* **5**, 696-711.

Raymond, F., Boisvert, S., Roy, G., Ritt, J. F., Legare, D., Isnard, A., Stanke, M., Olivier, M., Tremblay, M. J., Papadopoulou, B. et al. (2012). Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic Acids Res* **40**, 1131-47.

Reiter, J. F. and Leroux, M. R. (2017). Genes and molecular pathways underpinning ciliopathies. *Nat Rev Mol Cell Biol* **18**, 533-547.

Rico, E., Jeacock, L., Kovarova, J. and Horn, D. (2018). Inducible high-efficiency CRISPR-Cas9-targeted gene editing and precision base editing in African trypanosomes. *Sci Rep* **8**, 7960.

Rijal, S., Ostyn, B., Uranw, S., Rai, K., Bhattarai, N. R., Dorlo, T. P., Beijnen, J. H., Vanaerschot, M., Decuypere, S., Dhakal, S. S. et al. (2013). Increasing failure of miltefosine in the treatment of Kala-azar in Nepal and the potential role of parasite drug resistance, reinfection, or noncompliance. *Clin Infect Dis* **56**, 1530-8.

Rittig, M. G., Schroppel, K., Seack, K. H., Sander, U., N'Diaye, E. N., Maridonneau-Parini, I., Solbach, W. and Bogdan, C. (1998). Coiling phagocytosis of trypanosomatids and fungal cells. *Infect Immun* **66**, 4331-9.

Roberts, S. C. (2011). The genetic toolbox for Leishmania parasites. *Bioeng Bugs* **2**, 320-6.

Robinson, D., Beattie, P., Sherwin, T. and Gull, K. (1991). Microtubules, tubulin, and microtubule-associated proteins of trypanosomes. *Methods Enzymol* **196**, 285-99.

Robinson, D. G., Chen, W., Storey, J. D. and Gresham, D. (2014). Design and analysis of Bar-seq experiments. *G3 (Bethesda)* **4**, 11-8.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. and Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol* **29**, 24-6.

Rodriguez, J. A., Lopez, M. A., Thayer, M. C., Zhao, Y., Oberholzer, M., Chang, D. D., Kisalu, N. K., Penichet, M. L., Helguera, G., Bruinsma, R. et al. (2009). Propulsion of African trypanosomes is driven by bihelical waves with alternating chirality separated by kinks. *Proc Natl Acad Sci U S A* **106**, 19322-7.

Rogers, M. E. (2012). The role of leishmania proteophosphoglycans in sand fly transmission and infection of the Mammalian host. *Front Microbiol* **3**, 223.

Rogers, M. E., Chance, M. L. and Bates, P. A. (2002). The role of promastigote secretory gel in the origin and transmission of the infective stage of Leishmania mexicana by the sandfly Lutzomyia longipalpis. *Parasitology* **124**, 495-507.

Rojas, F., Silvester, E., Young, J., Milne, R., Tettey, M., Houston, D. R., Walkinshaw, M. D., Perez-Pi, I., Auer, M., Denton, H. et al. (2019). Oligopeptide Signaling through TbGPR89 Drives Trypanosome Quorum Sensing. *Cell* **176**, 306-317 e16.

Rosenbaum, J. L. and Witman, G. B. (2002). Intraflagellar transport. *Nat Rev Mol Cell Biol* **3**, 813-25.

Rotureau, B., Ooi, C. P., Huet, D., Perrot, S. and Bastin, P. (2014). Forward motility is essential for trypanosome infection in the tsetse fly. *Cell Microbiol* **16**, 425-33.

Ryan, K. A., Garraway, L. A., Descoteaux, A., Turco, S. J. and Beverley, S. M. (1993). Isolation of virulence genes directing surface glycosyl-phosphatidylinositol synthesis by functional complementation of Leishmania. *Proc Natl Acad Sci U S A* **90**, 8609-13.

Sacks, D. L., Modi, G., Rowton, E., Spath, G., Epstein, L., Turco, S. J. and Beverley, S. M. (2000). The role of phosphoglycans in Leishmania-sand fly interactions. *Proc Natl Acad Sci U S A* **97**, 406-11.

Sadlova, J. and Volf, P. (2009). Peritrophic matrix of Phlebotomus duboscqi and its kinetics during Leishmania major development. *Cell Tissue Res* **337**, 313-25.

Sahin, A., Espiau, B., Marchand, C. and Merlin, G. (2008). Flagellar length depends on LdARL-3A GTP/GDP unaltered cycling in Leishmania amazonensis. *Mol Biochem Parasitol* **157**, 83-7.

Sahin, A., Lemercier, G., Tetaud, E., Espiau, B., Myler, P., Stuart, K., Bakalara, N. and Merlin, G. (2004). Trypanosomatid flagellum biogenesis: ARL-3A is involved in several species. *Exp Parasitol* **108**, 126-33.

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.

Sarpal, R., Todi, S. V., Sivan-Loukianova, E., Shirolikar, S., Subramanian, N., Raff, E. C., Erickson, J. W., Ray, K. and Eberl, D. F. (2003). Drosophila KAP interacts with the kinesin II motor subunit KLP64D to assemble chordotonal sensory cilia, but not sperm tails. *Curr Biol* **13**, 1687-96.

Sartori, P., Geyer, V. F., Scholich, A., Julicher, F. and Howard, J. (2016). Dynamic curvature regulation accounts for the symmetric and asymmetric beats of Chlamydomonas flagella. *Elife* **5**.

Sasso, S., Stibor, H., Mittag, M. and Grossman, A. R. (2018). From molecular manipulation of domesticated Chlamydomonas reinhardtii to survival in nature. *Elife* **7**.

Satir, P. and Matsuoka, T. (1989). Splitting the ciliary axoneme: implications for a "switch-point" model of dynein arm activity in ciliary motion. *Cell Motil Cytoskeleton* **14**, 345-58.

Schadeli, D., Serricchio, M., Ben Hamidane, H., Loffreda, A., Hemphill, A., Beneke, T., Gluenz, E., Graumann, J. and Butikofer, P. (2019). Cardiolipin depletion-induced changes in the Trypanosoma brucei proteome. *FASEB J*, fj201901184RR.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-82.

Schneider, A., Sherwin, T., Sasse, R., Russell, D. G., Gull, K. and Seebeck, T. (1987). Subpellicular and flagellar microtubules of Trypanosoma brucei brucei contain the same alpha-tubulin isoforms. *J Cell Biol* **104**, 431-8.

Scholey, J. M. (2008). Intraflagellar transport motors in cilia: moving along the cell's antenna. *J Cell Biol* **180**, 23-9.

Segal, R. A., Huang, B., Ramanis, Z. and Luck, D. J. (1984). Mutant strains of Chlamydomonas reinhardtii that move backwards only. *J Cell Biol* **98**, 2026-34.

Shah, A. S., Ben-Shahar, Y., Moninger, T. O., Kline, J. N. and Welsh, M. J. (2009). Motile cilia of human airway epithelia are chemosensory. *Science* **325**, 1131-4.

Shapiro, T. A. and Englund, P. T. (1995). The structure and replication of kinetoplast DNA. *Annu Rev Microbiol* **49**, 117-43.

Shaw, S., DeMarco, S. F., Rehmann, R., Wenzler, T., Florini, F., Roditi, I. and Hill, K. L. (2019). Flagellar cAMP signaling controls trypanosome progression through host tissues. *Nat Commun* **10**, 803.

Shin, S. E., Lim, J. M., Koh, H. G., Kim, E. K., Kang, N. K., Jeon, S., Kwon, S., Shin, W. S., Lee, B., Hwangbo, K. et al. (2016). CRISPR/Cas9-induced knockout and knock-in mutations in Chlamydomonas reinhardtii. *Sci Rep* **6**, 27810.

Shrivastava, R., Tupperwar, N., Drory-Retwitzer, M. and Shapira, M. (2019). Deletion of a Single LeishIF4E-3 Allele by the CRISPR-Cas9 System Alters Cell Morphology and Infectivity of Leishmania. *mSphere* **4**.

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. (2016). A

Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes. *Cell* **166**, 1423-1435 e12.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J. et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.

Smith, A. M., Heisler, L. E., Mellor, J., Kaper, F., Thompson, M. J., Chee, M., Roth, F. P., Giaever, G. and Nislow, C. (2009). Quantitative phenotyping via deep barcode sequencing. *Genome Res* **19**, 1836-42.

Smith, A. M., Heisler, L. E., St Onge, R. P., Farias-Hesson, E., Wallace, I. M., Bodeau, J., Harris, A. N., Perry, K. M., Giaever, G., Pourmand, N. et al. (2010). Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples. *Nucleic Acids Res* **38**, e142.

Smith, E. F. (2002). Regulation of flagellar dynein by calcium and a role for an axonemal calmodulin and calmodulin-dependent kinase. *Mol Biol Cell* **13**, 3303-13.

Smith, E. F. and Lefebvre, P. A. (1996). PF16 encodes a protein with armadillo repeats and localizes to a single microtubule of the central apparatus in Chlamydomonas flagella. *J Cell Biol* **132**, 359-70.

Smith, E. F. and Lefebvre, P. A. (1997). PF20 gene product contains WD repeats and localizes to the intermicrotubule bridges in Chlamydomonas flagella. *Mol Biol Cell* **8**, 455-67.

Smith, E. F. and Sale, W. S. (1991). Microtubule binding and translocation by inner dynein arm subtype I1. *Cell Motil Cytoskeleton* **18**, 258-68.

Smith, E. F. and Yang, P. (2004). The radial spokes and central apparatus: mechano-chemical transducers that regulate flagellar motility. *Cell Motil Cytoskeleton* **57**, 8-17.

Soares Medeiros, L. C., South, L., Peng, D., Bustamante, J. M., Wang, W., Bunkofske, M., Perumal, N., Sanchez-Valdez, F. and Tarleton, R. L. (2017). Rapid, Selection-Free, High-Efficiency Genome Editing in Protozoan Parasites Using CRISPR-Cas9 Ribonucleoproteins. *MBio* **8**.

Sollelis, L., Ghorbal, M., MacPherson, C. R., Martins, R. M., Kuk, N., Crobu, L., Bastien, P., Scherf, A., Lopez-Rubio, J. J. and Sterkers, Y. (2015). First efficient CRISPR-Cas9-mediated genome editing in Leishmania parasites. *Cell Microbiol* **17**, 1405-12.

Sorokin, S. P. (1968). Reconstructions of centriole formation and ciliogenesis in mammalian lungs. *J Cell Sci* **3**, 207-30.

Spath, G. F., Garraway, L. A., Turco, S. J. and Beverley, S. M. (2003). The role(s) of lipophosphoglycan (LPG) in the establishment of Leishmania major infections in mammalian hosts. *Proc Natl Acad Sci U S A* **100**, 9536-41.

Srivastava, P., Prajapati, V. K., Rai, M. and Sundar, S. (2011). Unusual case of resistance to amphotericin B in visceral leishmaniasis in a region in India where leishmaniasis is not endemic. *J Clin Microbiol* **49**, 3088-91.

States, D. J. and Gish, W. (1994). Combined use of sequence similarity and codon bias for coding region identification. *J Comput Biol* **1**, 39-50.

Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J. and Mateo, J. L. (2015). CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* **10**, e0124633.

Sterkers, Y., Lachaud, L., Bourgeois, N., Crobu, L., Bastien, P. and Pages, M. (2012). Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in Leishmania. *Mol Microbiol* **86**, 15-23.

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. and Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62-7.

Stewart, J., Curtis, J., Spurck, T. P., Ilg, T., Garami, A., Baldwin, T., Courret, N., McFadden, G. I., Davis, A. and Handman, E. (2005). Characterisation of a Leishmania mexicana knockout lacking guanosine diphosphate-mannose pyrophosphorylase. *Int J Parasitol* **35**, 861-73.

Subota, I., Julkowska, D., Vincensini, L., Reeg, N., Buisson, J., Blisnick, T., Huet, D., Perrot, S., Santi-Rocca, J., Duchateau, M. et al. (2014). Proteomic analysis of intact flagella of procyclic Trypanosoma brucei cells identifies novel flagellar proteins with unique sub-localization and dynamics. *Mol Cell Proteomics* **13**, 1769-86.

Sugino, K. and Naitoh, Y. (1982). Simulated Cross-Bridge Patterns Corresponding to Ciliary Beating in Paramecium. *Nature* **295**, 609-611.

Sugrue, P., Hirons, M. R., Adam, J. U. and Holwill, M. E. (1988). Flagellar wave reversal in the kinetoplastid flagellate Crithidia oncopelti. *Biol Cell* **63**, 127-31.

Sun, S., Fisher, R. L., Bowser, S. S., Pentecost, B. T. and Sui, H. (2019). Three-dimensional architecture of epithelial primary cilia. *Proc Natl Acad Sci U S A* **116**, 9370-9379.

Sundar, S., More, D. K., Singh, M. K., Singh, V. P., Sharma, S., Makharia, A., Kumar, P. C. and Murray, H. W. (2000). Failure of pentavalent antimony in visceral leishmaniasis in India: report from the center of the Indian epidemic. *Clin Infect Dis* **31**, 1104-7.

Sunter, J. D. and Gull, K. (2016). The Flagellum Attachment Zone: 'The Cellular Ruler' of Trypanosome Morphology. *Trends Parasitol* **32**, 309-324.

Sunter, J. D., Moreira-Leite, F. and Gull, K. (2018). Dependency relationships between IFT-dependent flagellum elongation and cell morphogenesis in Leishmania. *Open Biol* **8**.

Sunter, J. D., Yanase, R., Wang, Z., Catta-Preta, C. M. C., Moreira-Leite, F., Myskova, J., Pruzinova, K., Volf, P., Mottram, J. C. and Gull, K. (2019). Leishmania flagellum attachment zone is critical for flagellar pocket shape, development in the sand fly, and pathogenicity in the host. *Proc Natl Acad Sci U S A* **116**, 6351-6360.

Szoor, B., Simon, D. V., Rojas, F., Young, J., Robinson, D. R., Kruger, T., Engstler, M. and Matthews, K. R. (2019). Positional Dynamics and Glycosomal Recruitment of Developmental Regulators during Trypanosome Differentiation. *MBio* **10**.

Tang, S., Wang, X., Li, W., Yang, X., Li, Z., Liu, W., Li, C., Zhu, Z., Wang, L., Wang, J. et al. (2017). Biallelic Mutations in CFAP43 and CFAP44 Cause Male Infertility with Multiple Morphological Abnormalities of the Sperm Flagella. *Am J Hum Genet* **100**, 854-864.

Taschner, M. and Lorentzen, E. (2016). The Intraflagellar Transport Machinery. *Cold Spring Harb Perspect Biol* **8**.

Toba, S., Fox, L. A., Sakakibara, H., Porter, M. E., Oiwa, K. and Sale, W. S. (2011). Distinct roles of 1alpha and 1beta heavy chains of the inner arm dynein I1 of Chlamydomonas flagella. *Mol Biol Cell* **22**, 342-53.

Torres-Guerrero, E., Quintanilla-Cedillo, M. R., Ruiz-Esmenjaud, J. and Arenas, R. (2017). Leishmaniasis: a review. *F1000Res* **6**, 750.

Tran, K. D., Rodriguez-Contreras, D., Vieira, D. P., Yates, P. A., David, L., Beatty, W., Elferich, J. and Landfear, S. M. (2013). KHARON1 mediates flagellar targeting of a glucose transporter in Leishmania mexicana and is critical for viability of infectious intracellular amastigotes. *J Biol Chem* **288**, 22721-33.

Tran, K. D., Vieira, D. P., Sanchez, M. A., Valli, J., Gluenz, E. and Landfear, S. M. (2015). Kharon1 null mutants of Leishmania mexicana are avirulent in mice and exhibit a cytokinesis defect within macrophages. *PLoS One* **10**, e0134432.

Trudgian, D. C., Ridlova, G., Fischer, R., Mackeen, M. M., Ternette, N., Acuto, O., Kessler, B. M. and Thomas, B. (2011). Comparative evaluation of label-free SINQ normalized spectral index quantitation in the central proteomics facilities pipeline. *Proteomics* **11**, 2790-7.

Trudgian, D. C., Thomas, B., McGowan, S. J., Kessler, B. M., Salek, M. and Acuto, O. (2010). CPFP: a central proteomics facilities pipeline. *Bioinformatics* **26**, 1131-2.

Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A. J., Le, L. P. et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187-197.

Tull, D., Naderer, T., Spurck, T., Mertens, H. D., Heng, J., McFadden, G. I., Gooley, P. R. and McConville, M. J. (2010). Membrane protein SMP-1 is required for normal flagellum function in Leishmania. *J Cell Sci* **123**, 544-54.

Tull, D., Vince, J. E., Callaghan, J. M., Naderer, T., Spurck, T., McFadden, G. I., Currie, G., Ferguson, K., Bacic, A. and McConville, M. J. (2004). SMP-1, a member of a new family of small myristoylated proteins in kinetoplastid parasites, is targeted to the flagellum membrane in Leishmania. *Mol Biol Cell* **15**, 4775-86.

Turco, S. J. and Descoteaux, A. (1992). The lipophosphoglycan of Leishmania parasites. *Annu Rev Microbiol* **46**, 65-94.

Turco, S. J. and Sacks, D. L. (1991). Expression of a stage-specific lipophosphoglycan in Leishmania major amastigotes. *Mol Biochem Parasitol* **45**, 91-9.

Ubeda, J. M., Raymond, F., Mukherjee, A., Plourde, M., Gingras, H., Roy, G., Lapointe, A., Leprohon, P., Papadopoulou, B., Corbeil, J. et al. (2014). Genome-wide stochastic adaptive DNA amplification at direct and inverted DNA repeats in the parasite Leishmania. *PLoS Biol* **12**, e1001868.

Ullu, E., Tschudi, C. and Chakraborty, T. (2004). RNA interference in protozoan parasites. *Cell Microbiol* **6**, 509-19.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. and Rozen, S. G. (2012). Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**, e115.

Urbanska, P., Joachimiak, E., Bazan, R., Fu, G., Poprzeczko, M., Fabczak, H., Nicastro, D. and Wloga, D. (2018). Ciliary proteins Fap43 and Fap44 interact with each other and are essential for proper cilia and flagella beating. *Cell Mol Life Sci* **75**, 4479-4493.

Valli, J. (2017). The Role of the Leishmania mexicana Amastigote Flagellum in Parasite-Host Interactions. *DPhil, University of Oxford*.

VanderWaal, K. E., Yamamoto, R., Wakabayashi, K., Fox, L., Kamiya, R., Dutcher, S. K., Bayly, P. V., Sale, W. S. and Porter, M. E. (2011). bop5 Mutations reveal new roles for the IC138 phosphoprotein in the regulation of flagellar motility and asymmetric waveforms. *Mol Biol Cell* **22**, 2862-74.

Vasquez, J. J., Wedel, C., Cosentino, R. O. and Siegel, T. N. (2018). Exploiting CRISPR-Cas9 technology to investigate individual histone modifications. *Nucleic Acids Res* **46**, e106.

Vincensini, L., Blisnick, T. and Bastin, P. (2011). 1001 model organisms to study cilia and flagella. *Biol Cell* **103**, 109-30.

Viswanadha, R., Hunter, E. L., Yamamoto, R., Wirschell, M., Alford, L. M., Dutcher, S. K. and Sale, W. S. (2014). The ciliary inner dynein arm, I1 dynein, is assembled in the cytoplasm and transported by IFT before axonemal docking. *Cytoskeleton (Hoboken)* **71**, 573-86.

Vizcaino, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T. et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* **44**, 11033.

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**, 915-8.

Wagner, M. K. and Yost, H. J. (2000). Left-right development: the roles of nodal cilia. *Curr Biol* **10**, R149-51.

Wais-Steider, J. and Satir, P. (1979). Effect of vanadate on gill cilia: switching mechanism in ciliary beat. *J Supramol Struct* **11**, 339-47.

Walker, B. J. and Wheeler, R. J. (2019). High-speed multifocal plane fluorescence microscopy for three-dimensional visualisation of beating flagella. *J Cell Sci* **132**.

Walters, L. L., Chaplin, G. L., Modi, G. B. and Tesh, R. B. (1989). Ultrastructural biology of Leishmania (Viannia) panamensis (=Leishmania braziliensis panamensis) in Lutzomyia gomezi (Diptera: Psychodidae): a natural host-parasite association. *Am J Trop Med Hyg* **40**, 19-39.

Wang, G., Guo, Y., Zhou, T., Shi, X., Yu, J., Yang, Y., Wu, Y., Wang, J., Liu, M., Chen, X. et al. (2013). In-depth proteomic analysis of the human sperm reveals complex protein compositions. *J Proteomics* **79**, 114-22.

Warner, F. D. and Satir, P. (1974). The structural basis of ciliary bend formation. Radial spoke positional changes accompanying microtubule sliding. *J Cell Biol* **63**, 35-63.

Wheeler, R. J. (2017). Use of chiral cell shape to ensure highly directional swimming in trypanosomes. *PLoS Comput Biol* **13**, e1005353.

Wheeler, R. J., Gluenz, E. and Gull, K. (2011). The cell cycle of Leishmania: morphogenetic events and their implications for parasite biology. *Mol Microbiol* **79**, 647-62.

Wheeler, R. J., Gluenz, E. and Gull, K. (2015a). Basal body multipotency and axonemal remodelling are two pathways to a 9+0 flagellum. *Nat Commun* **6**, 8964.

Wheeler, R. J., Gluenz, E. and Gull, K. (2015b). Routes to a 9+0 flagellum: Basal body multipotency and axonemal plasticity. *Nature Communications* **6**, 8964.

Wheeler, R. J., Sunter, J. D. and Gull, K. (2016). Flagellar pocket restructuring through the Leishmania life cycle involves a discrete flagellum attachment zone. *J Cell Sci* **129**, 854-67.

Wickstead, B. and Gull, K. (2007). Dyneins across eukaryotes: a comparative genomic analysis. *Traffic* **8**, 1708-21.

Williams, B. D., Velleca, M. A., Curry, A. M. and Rosenbaum, J. L. (1989). Molecular cloning and sequence analysis of the Chlamydomonas gene coding for radial spoke protein 3: flagellar mutation pf-14 is an ochre allele. *J Cell Biol* **109**, 235-45.

Williams, C. L., McIntyre, J. C., Norris, S. R., Jenkins, P. M., Zhang, L., Pei, Q., Verhey, K. and Martens, J. R. (2014). Direct evidence for BBSome-associated intraflagellar transport reveals distinct properties of native mammalian cilia. *Nat Commun* **5**, 5813.

Wingfield, J. L., Lechtreck, K. F. and Lorentzen, E. (2018). Trafficking of ciliary membrane proteins by the intraflagellar transport/BBSome machinery. *Essays Biochem* **62**, 753-763.

Wirtz, E. and Clayton, C. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. *Science* **268**, 1179-83.

Wisniewski, J. R., Zougman, A., Nagaraj, N. and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat Methods* **6**, 359-62.

Witman, G. B., Plummer, J. and Sander, G. (1978). Chlamydomonas flagellar mutants lacking radial spokes and central tubules. Structure, composition, and function of specific axonemal components. *J Cell Biol* **76**, 729-47.

Woolley, D. M. (1997). Studies on the eel sperm flagellum. I. The structure of the inner dynein arm complex. *J Cell Sci* **110 (Pt 1)**, 85-94.

Woolley, D. M. (2007). A novel motility pattern in quail spermatozoa with implications for the mechanism of flagellar beating. *Biol Cell* **99**, 663-75.

Woolley, D. M. (2010). Flagellar oscillation: a commentary on proposed mechanisms. *Biol Rev Camb Philos Soc* **85**, 453-70.

World Health Organization, W. (2019). Leishmaniasis Fact sheet. *[WWW Document]*.

Xiao, Z. S. and Quarles, L. D. (2010). Role of the polycytin-primary cilia complex in bone development and mechanosensing. *Ann N Y Acad Sci* **1192**, 410-21.

Yamamoto, R., Song, K., Yanagisawa, H. A., Fox, L., Yagi, T., Wirschell, M., Hirono, M., Kamiya, R., Nicastro, D. and Sale, W. S. (2013). The MIA complex is a conserved and novel dynein regulator essential for normal ciliary motility. *J Cell Biol* **201**, 263-78.

Yan, S., Martinez-Calvillo, S., Schnaufer, A., Sunkin, S., Myler, P. J. and Stuart, K. (2002). A low-background inducible promoter system in Leishmania donovani. *Mol Biochem Parasitol* **119**, 217-23.

Yan, S., Myler, P. J. and Stuart, K. (2001). Tetracycline regulated gene expression in Leishmania donovani. *Mol Biochem Parasitol* **112**, 61-9.

Yang, P., Diener, D. R., Yang, C., Kohno, T., Pazour, G. J., Dienes, J. M., Agrin, N. S., King, S. M., Sale, W. S., Kamiya, R. et al. (2006). Radial spoke proteins of Chlamydomonas flagella. *J Cell Sci* **119**, 1165-74.

Yang, P., Yang, C. and Sale, W. S. (2004). Flagellar radial spoke protein 2 is a calmodulin binding protein required for motility in Chlamydomonas reinhardtii. *Eukaryot Cell* **3**, 72-81.

Yang, Y. and Lu, X. (2011). Drosophila sperm motility in the reproductive tract. *Biol Reprod* **84**, 1005-15.

Yao, C., Luo, J., Hsiao, C. H., Donelson, J. E. and Wilson, M. E. (2007). Leishmania chagasi: a tetracycline-inducible cell line driven by T7 RNA polymerase. *Exp Parasitol* **116**, 205-13.

Zauli, R. C., Yokoyama-Yasunaka, J. K., Miguel, D. C., Moura, A. S., Pereira, L., da Silva, I. A., Jr., Lemes, L. G., Dorta, M. L., de Oliveira, M. A., Pitaluga, A. N. et al. (2012). A dysflagellar mutant of Leishmania (Viannia) braziliensis isolated from a cutaneous leishmaniasis patient. *Parasit Vectors* **5**, 11.

Zenian, A., Rowles, P. and Gingell, D. (1979). Scanning electron-microscopic study of the uptake of Leishmania parasites by macrophages. *J Cell Sci* **39**, 187-99.

Zetsche, B., Volz, S. E. and Zhang, F. (2015). A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat Biotechnol* **33**, 139-42.

Zhang, F. (2019). Development of CRISPR-Cas systems for genome editing and beyond. *Quarterly Reviews of Biophysics* **52**, 1-31.

Zhang, W. W., Lypaczewski, P. and Matlashewski, G. (2017). Optimized CRISPR-Cas9 Genome Editing for Leishmania and Its Use To Target a Multigene Family, Induce Chromosomal Translocation, and Study DNA Break Repair Mechanisms. *mSphere* **2**.

Zhang, W. W. and Matlashewski, G. (2015). CRISPR-Cas9-Mediated Genome Editing in Leishmania donovani. *MBio* **6**, e00861.

Zhang, W. W. and Matlashewski, G. (2019). Single-Strand Annealing Plays a Major Role in Double-Strand DNA Break Repair following CRISPR-Cas9 Cleavage in Leishmania. *mSphere* **4**.

Zhang, W. W., McCall, L. I. and Matlashewski, G. (2013). Role of cytosolic glyceraldehyde-3-phosphate dehydrogenase in visceral organ infection by Leishmania donovani. *Eukaryot Cell* **12**, 70-7.

Zhao, L., Hou, Y., Picariello, T., Craige, B. and Witman, G. B. (2019). Proteome of the central apparatus of a ciliary axoneme. *J Cell Biol* **218**, 2051-2070.

Zilberstein, D. and Shapira, M. (1994). The role of pH and temperature in the development of Leishmania parasites. *Annu Rev Microbiol* **48**, 449-70.

Zur Lage, P., Newton, F. G. and Jarman, A. P. (2019). Survey of the Ciliary Motility Machinery of Drosophila Sperm and Ciliated Mechanosensory Neurons Reveals Unexpected Cell-Type Specific Variations: A Model for Motile Ciliopathies. *Front Genet* **10**, 24.

9 Appendix

9.1 List of supplementary material

Supplementary: Script 1. LeishGEdit primer design v4.2 single mode.

Automated primer design bash script to generate all six primers needed for LeishGEdit gene editing. Donor DNA primers, including upstream forward primer (#1), upstream reverse primer (#2), downstream forward primer (#4), downstream reverse primer (#5), contain the 30 nt HF sequence adjacent to the sgRNA target sequence and its PAM site, as well as primer binding sites compatible with pT, pPLOT and pPOT plasmids (pPOT plasmids can be used for CRISPR-Cas9 mediated gene editing in *T. brucei*) (Figure 11 A and B). Additionally, two sgRNA primers are designed, one targeting the 5'UTR of the target gene (#3) and one the 3'UTR (#6). Further information for sgRNAs are given, including the CCTop sgRNA MM counts (potential sgRNA target sites that have up to 2 MM in the first 12 nt upstream of the PAM site or up to 4 MM in the entire sgRNA target sequence), as well as sgRNA perfect match counts within the entire genome.

Supplementary: Script 2. LeishGEdit primer design v4.2 batch mode.

Automated primer design bash script to generate all six primers needed for LeishGEdit gene editing that operates as described in "Supplementary: Script 1", but in batch mode.

Supplementary: Script 3. Primer3 primer search script.

Automated primer design bash script to run Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012) and pipe its output into a column based format that includes: CDS of all given geneIDs, and forward primer and reverse primer that yield amplicons of CDS fragments with 100 to 350 bp lengths. The script can be operated in batch mode.

Supplementary: Script 4. Barcode counter.

An automated bash script for counting barcodes in de-multiplexed bar-seq samples. The script is set to count barcodes with 0 MM over the 17nt barcode, but can be modified if desired to count also barcodes with MM.

Supplementary: Script 5. Bam assembler.

A bash wrapper script to produce bam files of sequenced cell lines. De-multiplexed genome samples are first assembled using Burrows-Wheeler Aligner (Li and Durbin, 2009), then sorted and indexed using Samtools (Li et al., 2009).

Supplementary: File 1. Generated Barcodes.

This file contains 15,489 17nt barcodes for bar-seq experiments. Barcodes were initially generated by using barcode generator 2.8 (Comai and Howell, 2009) and then customized to select barcodes with 40-60% GC content (for unbiased Illumina sequencing), at least 3 MM between barcodes and allowing no blast hit in *Lutzomyia longipalpis*, in *Mus musculus* and in *L. mexicana*.

Supplementary: File 2. LeishGEdit v4.2 output for TritrypDBv41 genomes.

The LeishGEdit v4.2 script (Supplementary: Script 2) has been used to design primers for all currently available TritrypDB genomes (for DBv41).

Supplementary: File 3. Primer3 primer search output.

The Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012) bash wrapper script (Supplementary: Script 3) has been used to design primers for ORF amplification for all currently available TritrypDB genomes (DBv41). Amplicon lengths was set to 100 - 350 bp.

Supplementary: Video files 1. ATP reactivation videos.

L. mexicana Cas9 T7 parental cells and *Δhydin* mutants have been subjected to the ATP reactivation procedure (see section 7.4). Important steps of the protocol have been captured by video microscopy (~15 sec recording time per video). This includes cells in PBS, cells in LRP buffer with 0.1% or 1.0% Triton X-100 and cells in LRP buffer with 0.1% or 1.0% Triton X-100 re-activated with 1.5 mM ATP.