Genetic dissection of a *Leishmania* flagellar proteome demonstrates requirement for directional motility in sand fly infections

Running Head: *L. mexicana* flagellar mutant screen

Authors

Tom Beneke[1], François Demay[2], Edward Hookway[3], Nicole Ashman[1], Heather Jeffery[1], James Smith[1], Jessica Valli[1], Tomas Becvar[4], Jitka Myskova[4], Tereza Lestinova[4], Shahaan Shafiq[1,5], Jovana Sadlova[4], Petr Volf[4], Richard Wheeler[6,1] and Eva Gluenz[1*]

Affiliations

[1]Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, OX1 3RE, UK

[2] University of Lille 1, Cité Scientifique, 59655 Villeneuve d'Ascq, France

[3] Research Department of Pathology, University College London, 72 Huntley Street, London, WC1E 6JD

[4] Department of Parasitology, Faculty of Science, Charles University, Vinicna 7, Prague 2, 128 44, Czech Republic

[5] School of Life Sciences, Oxford Brookes University, Gipsy Lane, Oxford, OX3 0BP, UK

[6] Peter Medawar Building for Pathogen Research, Nuffield Department of Medicine, University of Oxford, South Parks Road, Oxford OX1 3SY, UK

*Corresponding author: eva.gluenz@path.ox.ac.uk

Abstract

The protozoan parasite *Leishmania* possesses a single flagellum, which is remodelled during the parasite's life cycle from a long motile flagellum in promastigote forms in the sand fly to a short immotile flagellum in amastigotes residing in mammalian phagocytes. This study examined the protein composition and *in vivo* function of the promastigote flagellum. Protein mass spectrometry and label free protein enrichment testing of isolated flagella and deflagellated cell bodies defined a flagellar proteome for *L. mexicana* promastigote forms

(available via ProteomeXchange with identifier PXD011057). This information was used to generate a CRISPR-Cas9 knockout library of 100 mutants to screen for flagellar defects. This first large-scale knockout screen in a *Leishmania* sp. identified 56 mutants with altered swimming speed (52 reduced and 4 increased) and defined distinct mutant categories (faster swimmers, slower swimmers, slow uncoordinated swimmers and paralysed cells, including aflagellate promastigotes and cells with curled flagella and disruptions of the paraflagellar rod). Each mutant was tagged with a unique 17-nt barcode, providing a simple barcode sequencing (bar-seq) method for measuring the relative fitness of *L. mexicana* mutants *in vivo*. In mixed infections of the permissive sand fly vector *Lutzomyia longipalpis*, paralysed promastigotes and uncoordinated swimmers were severely diminished in the fly after defecation of the bloodmeal. Subsequent examination of flies infected with a single mutant lacking the central pair protein PF16 showed that these paralysed promastigotes did not reach anterior regions of the fly alimentary tract. These data show that *L. mexicana* need directional motility for successful colonisation of sand flies.

Author Summary

Leishmania are protozoan parasites, transmitted between mammals by the bite of phlebotomine sand flies. Promastigote forms in the sand fly have a long flagellum, which is motile and used for anchoring the parasites to prevent clearance with the digested blood meal remnants. To dissect flagellar functions and their importance in life cycle progression, we generated here a comprehensive list of >300 flagellar proteins and produced a CRISPR-Cas9 gene knockout library of 100 mutant *Leishmania*. We studied their behaviour *in vitro* before examining their fate in the sand fly *Lutzomyia longipalpis*. Measuring mutant swimming speeds showed that about half behaved differently compared to the wild type*:* a few swam faster, many slower and some were completely paralysed. We also found a group of uncoordinated swimmers. To test whether flagellar motility is required for parasite migration from the fly midgut to the foregut from where they reach the next host, we infected sand flies with a mixed mutant population. Each mutant carried a unique tag and tracking these tags up to nine days after infection showed that paralysed and uncoordinated *Leishmania* were rapidly lost from flies. These data indicate that directional swimming is important for successful colonisation of sand flies.

Introduction

Eukaryotic flagella / cilia are complex multifunctional organelles conserved from protists to humans [1]. Protists use flagella for swimming, feeding, cell-to-cell communication, adherence to substrates and morphogenesis [2]. Single-celled organisms, most commonly among them the green algae *Chlamydomonas reinhardtii,* have served as important model organisms to study molecular mechanisms of ciliogenesis and ciliary function [3], spurred on by the recognition that ciliary defects cause human genetic disorders collectively termed "ciliopathies" [4]. The eukaryotic flagellum is a complex, highly structured organelle and dissection of the molecular mechanisms underpinning its diverse functions requires detailed knowledge of its component parts. Proteomic studies of isolated flagella or axonemes from diverse species typically identified at least 300 distinct proteins [5-9] and phylogenetic profiling identified a set of 274 evolutionarily conserved ciliary genes [10]. All of these datasets comprise many "hypothetical" proteins still awaiting functional characterisation in addition to well-characterised core components of the microtubule axoneme, associated motor proteins and regulatory complexes.

Insights into conserved ciliary biology have helped elucidation of flagellar function in eukaryotic microbes, with a particular focus on human pathogens [11,12]. Among these, flagella have been most extensively studied in the causative agent of African trypanosomiasis, *Trypanosoma brucei* [13], which uses flagellar motility for locomotion and immune evasion [14] and exhibits close spatio-temporal coordination between flagellum assembly and cell morphogenesis during division [15]. The *T. brucei* bloodstream form is particularly sensitive to the loss of flagellar function [6,16], highlighting a potential Achilles' heel that might be exploitable for new anti-parasitic treatments.

The *Leishmania* flagellum is also a multi-functional organelle, which undergoes striking structural changes during the parasite's life cycle [17-19]. Amastigote forms proliferating in mammalian macrophages possess a short sensory-type 9+0 microtubule axoneme, which is remodelled to a canonical long motile 9+2 axoneme during differentiation to promastigote forms, which live in blood-feeding phlebotomine sand flies (Diptera: Psychodidae). In the fly, nectomonad promastigote forms attach via their flagella to the microvilli of the posterior midgut [20] to protect the parasites from being cleared during defecation of remnants of the blood meal. In the oesophageal valve, broad haptomonad forms attach to the cuticular lining via their flagellar tips, forming hemidesmosomes [20]. These life cycle descriptions [21] imply that periods of attachment must be followed by migration to more anterior regions of the alimentary tract and the propulsive function of the *Leishmania* flagellum is presumed to drive this forward migration but this has not been directly tested.

95 To enable a detailed genetic dissection of flagellar functions and mechanisms in *Leishmania*,
96 we defined here a flagellar proteome for motile *L. mexicana* promastigotes. We used new
97 CRISPR-Cas9 genome editing methods [22] to generate a *Leishmania* knockout library of
98 100 mutants, over half of which showed altered swimming speed. We also developed a
99 barcode sequencing (bar-seq) protocol to test the fitness of mutants in the permissive sand
100 fly vector *Lutzomyia longipalpis*. This study identified new genes required for flagellar motility
101 and shows that whilst culture-form promastigotes tolerated loss of the flagellum, paralysed
102 mutants and uncoordinated swimmers failed to colonise sand flies indicating that directional
103 flagellar motility is required for completion of the parasite's life cycle.

104 # Results

105 ## Defining the promastigote flagellar proteome

106 To enable a systematic genetic dissection of flagellar functions we sought to isolate *L.*
107 *mexicana* promastigote flagella comprising the axoneme, extra-axonemal structures and the
108 surrounding membrane for subsequent analysis by protein mass spectrometry (MS).
109 Mechanical shearing in the presence of 75 mM Ca^{2+} successfully separated cells into flagella
110 (F) and deflagellated cell bodies (CB) (Figure 1A, B). Subsequent centrifugation on sucrose
111 gradients allowed isolation of F and CB fractions with little cross-contamination: the CB
112 fraction contained only 2.03% (±0.69%) isolated flagella and the F fractions contained 0.56%
113 (±0.15%) deflagellated cell bodies (S1 Figure). Isolated flagella still retained their membrane:
114 First, examination of F fractions by transmission electron microscopy (TEM) confirmed that
115 most axonemes were bounded by a membrane (S2 Figure) and second, tracking an
116 abundant promastigote flagellar membrane protein, the small myristoylated protein 1 (SMP-1,
117 [23]) tagged with enhanced green fluorescent protein (eGFP) showed that it remained
118 associated with isolated flagella (Figure 1). Analysis of the SMP-1::eGFP signal also
119 facilitated flagellar length measurements in whole cells, F and CB fractions, which showed
120 that flagella were separated from the cell body near the exit point from the flagellar pocket.
121 The average break point was 2.7 µm distal to the base of the flagellum. The length of the
122 isolated flagella was similar to those on intact cells, indicating that isolated flagella remained
123 in one piece, with little fragmentation (S3 Figure). Two independently prepared sets of F and
124 CB fractions were separated into detergent soluble (s) and insoluble fractions (i), yielding four
125 fractions, F_S, F_I, CB_S and CB_I (Figure 1C). All four fractions for both replicates were analysed
126 by liquid chromatography tandem mass spectrometry (MS), which detected a total of 2711
127 distinct proteins (Figure 1D). Enrichment of detected proteins between biological replicates

128 correlated well (Pearson's r > 0.72, Spearman's r_s > 0.83, S4 Figure). To discover proteins
129 enriched in each of the four fractions, we used a label-free normalized spectral index
130 quantitation method (SINQ, [24]; S1, S2 and S3 Table) to generate a SINQ enrichment plot
131 (Figure 2A). The promastigote flagellar proteome, defined as proteins enriched in F vs. CB
132 fractions consisted of 701 unique proteins detected in at least one MS run; 352 of these were
133 enriched in F vs. CB fractions in both MS runs.

134 **Comparison with existing datasets validates flagellar proteome**

135 To validate the data, we mapped well-characterised flagellar proteins onto the enrichment
136 data plot (Figure 2A). Axonemal, paraflagellar rod (PFR), flagellar tip and flagellar membrane
137 proteins mapped to the F_I and F_S quadrants. Basal body, FAZ and tripartite attachment
138 complex (TAC) proteins mapped to the CB_I and CB_S quadrants because F fractions
139 contained exclusively the cell-external portion of the flagellum. Intraflagellar transport (IFT)
140 proteins clustered around the midpoint of the plot, indicating their abundance was similar in
141 the F and CB fractions, which is consistent with their known dynamic association with the
142 flagellar basal body and axoneme.
143 We also found substantial overlaps between *L. mexicana* proteins in the F_I quadrant and
144 proteins detected in previously published flagellar proteomes of *L. donovani* and *T. brucei*
145 (S5 Figure A-D). However, *L. mexicana* proteins in the F_S quadrant showed only a moderate
146 overlap with reported soluble *T. brucei* flagellar proteins (S5 Figure C). We designed a
147 website (www.leishgedit.net/leishgedit_db) for interactive browsing of proteins in the
148 enrichment plots shown in Figures 2 and S5.
149 Proteins with predicted trans-membrane domains (TMD) were predominantly detected in the
150 detergent soluble fractions (S5 Figure F). Overall, TMD proteins were underrepresented in
151 the proteome compared to their frequency in the genome (Chi-squared test, $p < 0.0001$), as
152 were proteins smaller than 10 kDa (S6 Figure), most likely due to well-known technical
153 limitations of protein MS [25]. Although ribosomal proteins were detected in individual F
154 fractions, the enrichment plot clustered them around the midpoint, with many enriched in the
155 cell body fractions (S5 Figure E). Our simple strategy of testing for enrichment thus
156 successfully filtered out likely contaminating proteins from the promastigote flagellar
157 proteome, as recently observed for enrichment of other cytoskeleton structures in *T. brucei*
158 [26,27].
159 Interestingly, a comparison of these proteomics data with *L. mexicana* RNA-seq data from
160 promastigotes and amastigotes [28] showed that proteins enriched in the flagellar fractions
161 were significantly more likely to have higher RNA abundance in promastigotes vs.
162 amastigotes, compared to proteins detected in the cell body fraction (Figure 2B; Chi-squared

163 test, $p < 0.0001$). This is consistent with the disassembly of the motile axoneme during
164 differentiation from promastigotes to amastigotes [17]. Whilst on a global scale transcript
165 levels correlate poorly with protein abundance in *Leishmania* spp. [29] these data indicate
166 that modulation of mRNA levels is a key regulatory step in *Leishmania* flagellar biogenesis
167 and differentiation from a 9+2 to a 9+0 flagellum.

168 Selection of candidates for a motility mutant screen

169 Many of the proteins detected in the F fractions had orthologs in previously defined flagellar
170 and ciliary proteomes yet lacked any functional characterisation. Arguably, endowing cells
171 with motility is the primary function of the promastigote flagellum and we took advantage of
172 our high-throughput CRISPR-Cas9 toolkit [22] to identify proteins required for motility and
173 subsequently study the phenotypes of the mutant *Leishmania*. In our knockout (KO) library
174 (S4 Table) we included 19 highly conserved axonemal proteins involved in the regulation of
175 flagellar beating, three intraflagellar transport (IFT) proteins, 60 flagellar proteins with
176 transcript enrichment in promastigotes [28] and eight additional soluble and four insoluble
177 flagellar proteins. Twenty of the selected proteins were detected in the promastigote flagellar
178 proteome but have to our knowledge not been linked to flagella before. We also made
179 deletion mutants for two genes implicated in membrane protein trafficking, *BBS2* and
180 *Kharon1*. Flagellar localisation of a subset of proteins was independently examined by
181 generating cell lines expressing proteins tagged with a fluorescent protein at the N- and/or C-
182 terminus (S7 Figure).
183 Orthofinder [30] was used to generate genome-wide orthologous protein sequence families
184 using genome sequences of 33 ciliated and 15 non-ciliated species from across eukaryotic
185 life, including *L. mexicana* and *T. brucei* (S5 Table). Twenty-two proteins were kinetoplastid-
186 specific (*L. mexicana* and *T. brucei*), 30 were conserved specifically in ciliated organisms and
187 23 widely conserved across eukaryotes whilst the remainder showed no clear pattern. In the
188 following, we refer to genes of unknown function by their GeneID from TriTrypDB.org [31]
189 and where we identified named orthologs we used these gene names.

190 Screening CRISPR-Cas9 knockout mutants for motility defects

191 The target genes were then deleted as described previously [22]. To facilitate high-
192 throughput generation of knockout (KO) cell lines, PCR reactions and transfections were
193 performed in 96-well plates. Analysis of drug-resistant transfectants by PCR confirmed loss
194 of the target ORF and integration of the drug-resistance gene in 94 of 98 cell lines (S8
195 Figure). This 96% success rate highlights the power of our gene deletion strategy. The
196 reason for the presence of the target ORF in the remaining four cell lines was not further

197 investigated, but was confirmed by diagnostic PCR of two independently isolated samples of
198 genomic DNA from the relevant mutants.
199 The flagellar mutants generated in this study, the previously generated paralysed cell line
200 Δ*PF16* [22], the parental line *L. mex* Cas9 T7, and wild type promastigotes were subjected
201 to motility assays using dark field microscopy to track the swimming behaviour of cells and
202 measure swimming speed and directionality as previously described [32]. Parental cells
203 immobilised though formaldehyde fixation were also measured. More than half of all mutant
204 lines showed a significant deviation from the normal average swimming speed measured for
205 the parental cell line and wild type controls (Figure 3A, B): 52 (53.6%) mutants showed a
206 significant reduction in speed and 4 (4.1%) swam faster (Student's t-test, $p<0.005$; Figure 3A,
207 S4 Table). Plotting mean swimming speed against mean directionality (velocity/speed)
208 shows broad groups of mutants (Figure 3B): Those which are paralysed, slower swimmers,
209 slow uncoordinated swimmers, faster swimmers and a single mutant that had faster and
210 more directional swimming (Δ*LmxM.36.3620*). The mechanistic contribution to swimming
211 behaviour remains to be clarified for many proteins in this set. Loss of flagellar waveform
212 modulators would cause altered motility patterns, and this is exemplified by two mutants in
213 this set: the Δ*dDC2* mutant, which lacks the outer dynein arm docking complex protein dDC2
214 and can perform a ciliary beat but no flagellar beat [33] clusters with the uncoordinated
215 group. By contrast, Δ*LC4-like*, which lacks a distal regulator of outer dynein arms and spends
216 more time doing a flagellar beat at a higher beat frequency [33], was among the faster
217 swimmers.
218 The most severe loss of motility was observed in three cell lines that had no visible external
219 flagellum (Figure 4); all of these were deletions of conserved intraflagellar transport (IFT)
220 proteins (Δ*IFT122B*, Δ*IFT139* and Δ*IFT88*). Ablation of the central pair (CP) protein hydin
221 also resulted in almost complete paralysis, comparable to the deletion of the CP protein
222 PF16 [22].
223 In a subset of paralysed or slow-swimming uncoordinated mutants (Figure 3C) we noted that
224 the flagella tended to be in a curled rather than straight conformation. Δ*hydin* mutants had
225 the highest proportion of curled-up flagella (62.6%, Figure 4 and S9 Figure) while fewer than
226 1% of flagella were curled-up in the parental cell line and many other slow swimming mutants
227 (S9 Figure). A high proportion (>10%) of curled-up flagella was also found in four paralysed
228 KO lines (inner dynein arm intermediate chain protein mutant Δ*IC140*, 57%; Δ*PF16*, 14%;
229 tether and tether head complex protein mutants Δ*CFAP44*, 15% and Δ*CFAP43*, 19%) and
230 three uncoordinated KO lines (Δ*MBO2*, 26%; nexin-dynein regulatory complex protein mutant
231 Δ*DRC4*, 13%; Δ*LmxM.33.0560*, 12%). The curls were observed in aldehyde fixed cells as

232 well as in live cells in culture, indicating they were not an artefact of microscopy sample
233 preparation. This novel phenotype might be caused by disrupted dynein regulation and
234 warrants further investigation.
235 Null mutants for the major PFR protein PFR2, lacking the paracrystalline PFR lattice
236 structure, are known to have impaired motility [34]. To compare motility of a Δ*PFR2* mutant
237 with other mutants generated in this study, we used CRISPR-Cas9 to delete both allelic
238 copies of the *PFR2* array (*PFR2A*, *PFR2B* and *PFR2C*) and confirmed loss of PFR2
239 expression by Western blot (S10 Figure). This Δ*PFR2* line had slower and less directional
240 swimming compared to the parental cells, clustering with other slow swimming mutants
241 defined in Figure 3B. To test whether gene deletion in other slow swimming mutants had a
242 major disruptive effect on the PFR, which might explain their motility defect, we expressed
243 PFR2::mNG in KO lines and looked for changes to PFR length or loss of PFR integrity
244 (defined as gaps in the PFR2::mNG signal) (Figures 4 and 5). Three mutants had shorter
245 flagella compared to the parental cell line, but the PFR remained proportional to the overall
246 flagellar length and was uninterrupted (Δ*ARL-3A*, Δ*CFAP44*, and Δ*FLAM2*). Six mutants had
247 PFR-specific defects (Figure 5B): a shorter flagellum with a disproportionately shorter PFR
248 (Δ*LmxM.27.0860;* Δ*TTC29;* Δ*LmxM.14.1220*), a normal-length flagellum with a shorter PFR
249 (Δ*FM458*) or a shorter PFR with gaps (Δ*LmxM.21.1110,* 25.3% of all flagella; Δ*MBO2* only
250 4.1% of all flagella). Interestingly, these comparatively subtle alterations to PFR length and
251 integrity reduced swimming speed to similar levels as PFR2 deletion (Figure 3C).
252 We generated 13 add-back cell lines to rescue mutant phenotypes by transfecting episomes
253 containing the deleted ORF. Four complemented mutants fully recovered parental swimming
254 speed (Figure 3) and complemented Δ*CFAP44* and Δ*MBO2* lines showed fewer curled
255 flagella (S9 Figure). Complementation of the other slow swimming mutants resulted in a
256 significant increase in swimming speed close to parental levels (Figure 3) and reduction of
257 curling compared to the KO lines (S9 Figure).
258 Thus, our screen readily identified promastigote mutants with impaired motility and even the
259 most severe phenotype, ablation of flagellar assembly caused by loss of IFT components,
260 was compatible with promastigote survival *in vitro*, in line with earlier reports [35], [36], [37].

261 **Paralysed promastigotes are cleared from sand flies**

262 Whilst flagellar motility is generally believed to be required for development in sand flies,
263 enabling *Leishmania* migration from the midgut to the mouthparts [38-40], this has not been
264 directly tested. To interrogate the phenotypes of larger cohorts of *Leishmania* mutants in
265 parallel, we developed a multiplexed bar-seq strategy inspired by pioneering phenotyping

266 screens in yeast [41] and the malaria parasite *Plasmodium berghei* [42,43]. We pooled
267 mutant *L. mexicana* lines that were each tagged with a unique 17 bp barcode. This enabled
268 us to measure the relative abundance of each line at different time points after sand fly
269 infection (S11 Figure). Seventeen were flagellar mutants described above and five were
270 parental control cell lines tagged with unique barcodes in their small subunit (SSU) ribosomal
271 RNA locus. We also generated a barcoded Δ*LPG1* KO mutant, which is only defective in
272 LPG synthesis [22,44] and three barcoded mutants defective in the pathway leading to
273 mannose activation for synthesis of LPG and other glycoconjugates: KOs of
274 phosphomannose isomerase [45] (Δ*PMI*), phosphomannomutase [46] (ΔPMM) and GDP-
275 mannose pyrophosphorylase [47] (ΔGDP-MP). These mutants were included as control lines
276 expected to be outcompeted by the parental cell lines. The barcoded cell lines were pooled in
277 equal proportions and used to infect *L. longipalpis*. The relative abundance of each line was
278 determined by sequencing DNA isolated from the mixed promastigote pool and from flies at
279 two, six and nine days after infection. The results show progressively diminishing proportions
280 for the control mutants defective in LPG synthesis (Δ*LPG1*) or a broader range of
281 glycoconjugates including LPG (Δ*PMI*, Δ*PMM* and Δ*GDP-MP*) (Figure 6, S9 Table) indicating
282 that parasites lacking these molecules were at a competitive disadvantage in these
283 infections. This effect was apparent as early as two days after infection, consistent with a
284 protective role for PG-containing glycoconjugates in the digesting bloodmeal [48] and a role
285 for LPG in *L. mexicana* attachment to *L. longipalpis* [49].
286 Paralysed and uncoordinated mutants also became noticeably scarcer as the infection
287 progressed (Figure 6, S9 Table). The aflagellate Δ*IFT88* mutant showed the most severe
288 phenotype and a significant decrease over time was also measured for Δ*PF16,* Δ*CFAP43,*
289 Δ*CFAP44,* Δ*IC140,* Δ*dDC2* and Δ*RSP4/6.* By contrast, mutants with a mild swimming defect
290 (slower swimmers Δ*LmxM.21.1110,* Δ*FM458* and Δ*LmxM.18.1090* and faster Δ*LC4-like*)
291 (Figure 3D) remained as abundant as the normal swimmers throughout the infection (Figure
292 6, S9 Table). The exceptions were the slower swimmers Δ*Kharon1,* which is also defective in
293 the transport of a flagellar glucose transporter [50], and Δ*ARL-3A,* which has a short
294 flagellum (Figure 5). Both of these were rarer in the fly compared to the starting pool. To gain
295 anatomical resolution and determine whether an immotile mutant fails to migrate to anterior
296 portions of the fly gut, we infected separate batches of *L. longipalpis* with motile parasite
297 lines and complemented KO lines as controls, with the motile Δ*BBS2* mutant, which lacks a
298 component of the BBSome complex [51] which is expected to play a role in flagellar
299 membrane trafficking, and with the paralysed Δ*PF16* mutant (Figure 7). The Δ*PF16* mutants
300 are among the least motile cells that retain a long flagellum (Figure 5), while having only

301 moderate levels of flagellar curling (S9 Figure). The axonemal defect resulting in paralysis is
302 a well-characterised disruption of the central pair in kinetoplastids (Figure 3B and [22,52,53])
303 and is similar to the defect of the *pf16 Chlamydomonas reinhardtii* mutant [54] indicating it is
304 a well-conserved core axoneme component. Two days post blood-meal (PBM), the *L.*
305 *mexicana* wild type and *L. mex* Cas9 T7 [22] control cell lines and the Δ*BBS2* mutant
306 developed well, with infection rates above 70%; the Δ*PF16* mutant produced the lowest
307 infection rate (below 50%). The introduction of an add-back copy of *PF16* into the Δ*PF16* line
308 restored infection levels (Figure 7A). In all lines, promastigotes were localized in the
309 abdominal midgut, within the bloodmeal enclosed in the peritrophic matrix (Figure 7B). After
310 defecation (day 6 PBM), all control lines and the Δ*BBS2* mutant replicated well and
311 developed late-stage infections with colonisation of the whole mesenteron including the
312 stomodeal valve (Figure 7B) which is a prerequisite for successful transmission. Their
313 infection rates ranged from 56% to 83%. By contrast, Δ*PF16 Leishmania* failed to develop;
314 the infection rate was less than 2% (a single positive fly out of 62 dissected (Figure 7A), with
315 parasites restricted to the abdominal midgut (Figure 7B)), indicating that Δ*PF16* parasites
316 were lost during defecation and were unable to develop late stage infections in *L. longipalpis*.
317 Our data provide strong evidence that flagellar motility is an essential requirement for
318 successful development in sand flies and, by implication, transmission.

319 Discussion

320 This study demonstrates the power of high-throughput CRISPR-Cas9 knockout screens to
321 discover mutant phenotypes in *Leishmania*. We defined a high-confidence flagellar proteome
322 and used these data in conjunction with transcriptomics data and prior knowledge of
323 conserved axonemal proteins to demonstrate a role in motility for >50 genes from a set of
324 one hundred. We also show the importance of flagellar motility in the colonisation of sand
325 flies. The data from the pooled mutant population show a progressive loss of paralysed or
326 uncoordinated swimmers over nine days from infection. Because these data report total
327 abundance of each genotype in the whole fly without discriminating between regions of the
328 gut, we probed this question further in infections with the Δ*PF16* mutant, which is essentially
329 paralysed and incapable of sustained directional motility due to a defined defect in the central
330 pair complex of the axoneme [22]. The results show that Δ*PF16 Leishmania* were rapidly lost
331 from most of the dissected flies, consistent with the depletion of this mutant from the mixed
332 pool, and additionally shows that none of the few remaining parasites reached anterior parts
333 of the alimentary tract. These findings show that parasite motility is required for completion of

334 the *Leishmania* life cycle, in line with the essential role of motility in other vector-transmitted
335 protists. For example, Rotureau *et al.*, [55] showed that loss of forward motility, caused by
336 ablation of outer dynein arms though KO of *DNAI1*, rendered *T. brucei* unable to reach the
337 tsetse fly foregut. It seems likely that loss of motility also contributed to the inability of *L.*
338 *amazonensis* to progress beyond the abdominal midgut of *L. longipalpis* when the parasites
339 overexpressed GTP-locked ADP-ribosylation factor-like protein 3A (Arl-3A) and as a result
340 grew only short flagella [56].
341 Observations of attached *Leishmania* in dissected sand flies show adhesion specifically via
342 the flagellum but the precise molecular interactions between flagellum and the microvillar gut
343 lining remain to be clarified. The dominant cell surface glycoconjugate LPG which covers the
344 entire parasite surface including the flagellum is known to be important in *Leishmania*
345 attachment to sand fly guts [57] and our results support the view that LPG plays an important
346 role in *L. mexicana* infection of *L. longipalpis* [49]. The proportion of Δ*LPG1* mutants had
347 decreased by two days after infection and reduced further as infection progressed. The
348 observed loss of fitness of the Δ*PMM*, Δ*GDP-MP* and Δ*PMI* mutants is likely the cumulative
349 effect of the loss of LPG and a broader range of mannose-containing glycoconjugates which
350 were shown to protect *Leishmania* in the digesting bloodmeal [48]. The consistency of the
351 pooled mutant data with the reported phenotypes of individual glycoconjugate-deficient
352 mutants demonstrates the power of this new rapid method for mutant phenotyping in
353 *Leishmania*. However, whilst a role for LPG in *L. mexicana* attachment to the fly is well
354 established, the possible contribution of flagellum-specific surface molecules [58] has not yet
355 been conclusively resolved. Zauli et al., [36] reported isolation of *L. braziliensis* from a
356 patient's skin lesion which differentiated to promastigotes with an "atypical" morphology.
357 These cells had a short flagellum barely protruding from the flagellar pocket, with an
358 amorphous tip suggestive of a defect in flagellum elongation. In experimental infections of *L.*
359 *longipalpis*, these parasites persisted in the fly following defecation of the blood meal,
360 suggesting that they remained sufficiently anchored without a long flagellum. It would be
361 interesting to follow up the subsequent development of this mutant in the fly.
362 Several lines of evidence suggest a role for the trypanosomatid flagellum in environmental
363 sensing [40,59-61]. Evidence for specific signal transduction pathways aiding promastigote
364 navigation through the sand fly is however limited. Cyclic nucleotide signal transduction
365 pathways may have important roles in coupling environmental sensing with regulation of
366 flagellar beat patterns [62,63]. In our flagellar proteome we identified several adenylate
367 cyclases (ACs), cAMP-specific phosphodiesterases (PDEs) and PKA subunits and mapped
368 their localisations to distinct flagellar subdomains by protein tagging (S7 Figure). The motility

369 assays showed that deletion of PKA subunits (Δ*LmxM.34.4010* (partial KO only) and
370 Δ*FM458*) reduced swimming speed, whereas deletion of two different PDEs
371 (Δ*LmxM.18.1090* and Δ*LmxM.08_29.2440*) increased it, pointing to an activating role for
372 cAMP in *Leishmania* motility. Knockout of receptor-type adenylate cyclase a-like protein
373 LmxM.36.3180 had no effect on swimming speed in our motility assay but given the possible
374 redundancy with other flagellar ACs, this preliminary finding should be followed up by
375 examination of other AC mutants individually and in combinations. In our pooled KO screen
376 in sand flies, KOs of PDE LmxM.18.1090 and PKA RSU (FM458) remained as abundant as
377 the controls, indicating that the mild motility phenotypes measured *in vitro* did not significantly
378 impair colonisation of flies.
379 Perturbation of the flagellar membrane might be expected to interfere with sensory functions
380 mediated through the flagellum. Ablation of membrane proteins LmxM.17.0870 and
381 LmxM.23.1020 (S7 Figure) did not significantly enhance or reduce the relative abundance of
382 the respective mutants in sand flies over the nine-day observation period. BBS2 is an integral
383 part of the core BBSome complex which is highly conserved across ciliated eukaryotes [64]
384 and functions as a cargo adaptor for ciliary membrane protein trafficking in *Chlamydomonas*
385 flagella and metazoan cilia [65]. Our pooled mutant data and infections with the *BBS2*
386 deletion mutant alone found that loss of this gene had no discernible detrimental effect on
387 survival in sand flies and the parasites' ability to reach the anterior gut. By contrast, KO of
388 Kharon1, a protein shown to be required for trafficking of the glucose transporter LmGT1,
389 and perhaps other proteins, to the promastigote flagellum [50] led to slightly reduced fitness
390 in the flies from the earliest time point. The Δ*Arl-3A* mutants were also less abundant
391 compared to the controls. This is reminiscent of the previously published abortive phenotype
392 of *L. amazonensis* overexpressing the constitutively GTP bound *Ld*ARL-3A-Q70L [56]. This
393 mutant formed only a short flagellum, similar to the Δ*Arl-3A* mutant generated in the present
394 study. Failed attachment as a result of the shortened flagellum was thought to be a likely
395 cause for the rapid clearance of *Ld*ARL-3A-Q70L-expressing parasites but it was noted that
396 an inability to migrate at later stages of development would also lead to the disappearance of
397 the mutants [56]. In our study the phenotype of the Δ*Arl-3A* mutants was however mild
398 compared to the aflagellate (Δ*IFT88*) or paralysed mutants. Arl-3A acts as guanine
399 nucleotide exchange factor in the transport of lipidated proteins to the flagellar membrane
400 [66] and protein mis-targeting could contribute to the phenotype in addition to flagellar
401 shortening. Further insights into the contribution of flagellar membrane proteins to attachment
402 or directional swimming behaviour may be uncovered by further biochemical studies into

403 flagellar membrane composition and subjecting different mutants (with or without overt
404 motility phenotypes in culture) to chemotaxis assays and fly infections.
405 In contrast to the absolute requirement of motility for movement through the sand fly vector,
406 flagellar motility is dispensable for promastigote proliferation in culture. Promastigotes are
407 viable and able to divide even if they fail to assemble a flagellum at all, as demonstrated
408 originally by the deletion of cytoplasmic dynein-2 heavy chain gene *LmxDHC2.2* [35] and
409 IFT140 [37] and the phenotypes of knockouts of anterograde and retrograde IFT components
410 in the present study. The ensuing prediction that most gene deletions affecting flagellar
411 function are expected to yield viable promastigotes in the laboratory is borne out by our high
412 success rate of obtaining 96% of attempted knockouts. Thus, in *Leishmania*, flagellar mutant
413 phenotypes can be observed in replication-competent cells over many cell cycles and our
414 mutant library enables detailed systematic studies of KO phenotypes to probe protein
415 functions in flagellum assembly, motility and signal transduction.
416 A fruitful area for further studies will be dissection of PFR function and assembly
417 mechanisms. This extra-axonemal structure is required for motility as demonstrated through
418 deletion of the major structural PFR components, PFR1 and PFR2 in *Leishmania* [34,67] and
419 ablation of PFR2 by RNAi in *T. brucei* procyclic forms [68] but its precise role remains
420 unclear. The PFR comprises more than 40 proteins, some with structural roles, others with
421 roles in adenine nucleotide homeostasis, cAMP signalling, calcium signalling and many
422 uncharacterised components [69,70] and it may anchor metabolic and regulatory proteins as
423 well as influencing the mechanical properties of the flagellum. Our results showed that
424 fragmentation of the PFR caused by loss of LmxM.21.1110 reduced swimming speed to
425 levels similar to the structurally more severe PFR2 KO. Whether LmxM.21.1110 is required
426 for correct PFR assembly or stabilisation is currently unknown.
427 Motility mutants analysed in our screen also included deletions of genes with human
428 orthologs linked to ciliopathies (such as *hydin*) or male infertility (*CFAP43* and *CFAP44*) [71].
429 *Leishmania* offers a genetically tractable system to gain further mechanistic insight into their
430 functions. The *hydin* mutant has been extensively characterised in other species: in
431 mammals, mutations in the *hydin* gene cause early-onset hydrocephalus [72] and
432 subsequent studies on *C. reinhardtii*, *T. brucei* and mice showed that hydin localises to the
433 C2 projection of the central pair complex [73], and that loss of hydin function causes
434 mispositioning and loss of the CP [74] and motility defects [73-75]. The motility phenotype in
435 the *L. mexicana* Δ*hydin* mutant was consistent with these existing data and we made the
436 new observation that the mutant flagella show extensive curling (Figure 4, S9 Figure).
437 Interestingly, *hydin* knockdown in *C. reinhardtii* caused flagella to arrest at the switch point

438 between effective and recovery stroke, leaving cells with one flagellum pointing up and the
439 other down, prompting speculation that this may indicate a role for hydin in signal
440 transmission to dynein arms [73]. Consistent with this hypothesis, cilia of *hy3/hy3* mouse
441 mutants frequently stalled at the transition point between the effective and recovery stroke
442 [75]. Curling may represent the failure of flagellum bending to reverse during progression of
443 the normal flagellum waveform down the flagellum, leaving the flagellum locked at one
444 extreme of bending, analogous to the ciliary beat *hydin* phenotype. In *L. mexicana*, the
445 Δ*hydin* mutant presented the most severe manifestation of the curling phenotype, which was
446 also observed in a lower proportion of other mutants (S9 Figure). This phenotype may be a
447 consequence of mis-regulated dyneins and the set of mutants exhibiting curling will facilitate
448 further experiments to establish the mechanistic basis for flagellar curling.
449 Genetic, biochemical and structural studies have provided elegant and detailed models for
450 the mechanisms of flagellar motility [76,77]. Phylogenetic profiling and comparative
451 proteomics studies have yielded insights into the evolutionary history, core conserved
452 structures and lineage-specific adaptations of eukaryotic flagella. Our CRISPR-Cas9 KO
453 method enables rapid targeted gene deletion and characterisation of loss-of-function
454 phenotypes for large cohorts of *Leishmania* genes *in vitro* and *in vivo* and hence new
455 opportunities to interrogate the functions of hitherto poorly characterised flagellar proteins in
456 motility regulation, environmental sensing and axoneme remodelling from 9+2 to 9+0. The
457 bar-seq strategy for phenotyping of mutants can also be used to probe parasite-host
458 interactions in mammals.

459 Materials and Methods

460 ***Cell culture***

461 Promastigote-form *L. mexicana* (WHO strain MNYC/BZ/62/M379) were grown at 28°C in
462 M199 medium (Life Technologies) supplemented with 2.2 g/L $NaHCO_3$, 0.005% haemin, 40
463 mM 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7.4 and 10% FCS. The
464 modified cell line *L. mexicana SMP1:TYGFPTY* [17] was cultured in supplemented M199
465 with the addition of 40 µg/ml G-418 Disulfate. L. *mex* Cas9 T7 [22] was cultured in
466 supplemented M199 with the addition of 50 µg/ml Nourseothricin Sulphate and 32 µg/ml
467 Hygromycin B.

468 ***Deflagellation protocol***

469 To avoid proteolytic degradation, all procedures were performed on ice. $2{\cdot}10^9$ *L. mexicana*
470 *SMP1:TYGFPTY* cells were collected at 800*g* for 15 min at 4°C, washed once in 20 ml

471 phosphate buffered saline (PBS) and resuspended in 5 ml 10 mM PIPES [10 mM NaCl, 10
472 mM piperazine-N,N′-bis(2-ethanesulfonic acid, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 0.32 M sucrose,
473 adjusted to pH 7.2]. 0.375 ml of 1 M Ca^{2+} solution (final conc. 0.075 M) and a protease
474 inhibitor cocktail [final concentration, 5 µM E-64, 50 µM Leupeptin hydrochloride, 7.5 µM
475 Pepstatin A and 500 µM Phenylmethylsulfonyl fluoride (PMSF)] were added to the cell
476 suspension. Cells were deflagellated by passing them 100 times through a 200 µl gel loading
477 pipette tip (Starlab) attached to a 10 ml syringe. Flagella and cell bodies were separated
478 through density gradient centrifugation, using a modified version of the protocol in [78]. The
479 sample was loaded on top of the first sucrose-bed containing three layers of 10 mM PIPES
480 with 33% (upper), 53% (middle) and 63% (bottom) w/v sucrose [10 mM NaCl, 10 mM
481 piperazine-N,N′-bis(2-ethanesulfonic acid, 1 mM $CaCl_2$, 1 mM $MgCl_2$, adjusted to pH 7.2 with
482 either 0.96M, 1.55M or 1.84M sucrose] and centrifuged at 800*g* for 15 min at 4°C. The pellet
483 in the 63% sucrose layer was diluted with 10 ml 10 mM PIPES and centrifuged at 800*g* for 15
484 min at 4°C. The supernatant was discarded and the pellet resuspended in 40 µl 10 mM
485 PIPES. This was the cell body fraction. The top layer of the first sucrose-bed, containing
486 flagella, was collected and sucrose sedimentation was repeated with a second sucrose-bed
487 containing only one layer of 10 mM PIPES with 33% w/v sucrose. The resulting top layer of
488 the second sucrose bed was transferred to an ultra-centrifugation tube (Beckmann tubes)
489 and collected by ultra-centrifugation at 100,000*g* for 1 h at 4°C (Beckman Coulter). The
490 supernatant was discarded and the pellet resuspended in 40 µl 10 mM PIPES. This was the
491 flagellar fraction. All other sucrose layers contained a mixture of flagella and cell bodies and
492 were discarded. 1 µl of flagellar and cell body fractions was used for counting and imaging
493 and 36 µl of each fraction were used for proteomic analysis.

494 ***Protein Mass spectrometry***

495 Cell body and flagellar fractions were supplemented with 4 µl protease inhibitor cocktail (see
496 above) and 10 µl octylglycoside (1% (w/v) final conc.), incubated for 20 min on ice and
497 centrifuged at 18,500*g* for 1 h at 4°C to separate soluble (supernatant) from insoluble (pellet)
498 proteins. 50 µl ice cold reducing 2x Laemmli buffer was added to the resulting supernatant.
499 Pellets were dissolved in 100 µl 1x Laemmli buffer. To avoid aggregation of hydrophobic
500 proteins, fractions were not boiled prior to SDS-PAGE [79]. 20 µl of flagella fractions and 10
501 µl of cell body fractions (~5 – 20 µg protein) were pre-fractionated on a 10% polyacrylamide
502 gel, stained overnight with SYPRO Ruby Protein Gel Stain (Molecular Probes) and destained
503 for 30 min in 10% (v/v) Methanol / 7% (v/v) acetic acid. Sample preparation in the following
504 was carried out as described in [80]. Briefly, gel pieces were destained with 50% acetonitrile,
505 reduced with 10mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride) for 30 minutes at

506 RT, followed by alkylation with 55 mM Iodoacetamide for 60 minutes in the dark at RT.
507 Samples were deglycosylated with PNGase F over two days at RT and digested overnight at
508 37°C with 100 ng trypsin. Samples were acidified to pH 3.0 using 0.1% trifluoroacetic acid
509 and desalted by reversed phase liquid chromatography. Samples were analysed on an
510 Ultimate 3000 RSLCnano HPLC (Dionex, Camberley, UK) system run in direct injection
511 mode coupled to a QExactive Orbitrap mass spectrometer (Thermo Electron, Hemel
512 Hempstead, UK).

513 ***MS data analysis and SINQ enrichment plot***

514 MS-data were converted from .RAW to .MGF file using ProteoWizard (S6 Table) and
515 uploaded to the Central Proteomics Facilities Pipeline (CPFP [81]). Protein lists were
516 generated by using CPFP meta-searches (S6 Table) against the predicted *L. mexicana*
517 proteome (gene models based on [28], followed by label-free SINQ quantification (S1 and S6
518 Table). For SINQ enrichment plots detected geneIDs were filtered ($p \geq 0.95$, ≥ 2 peptides)
519 and plotted using normalized spectral indices. For missing indices pseudo spectral indices of
520 10^{-10} were inserted. The mass spectrometry proteomics data have been deposited to the
521 ProteomeXchange Consortium via the PRIDE [82] partner repository with the dataset
522 identifier PXD011057.

523 ***Gene tagging***

524 Tagging was achieved by insertion of a drug-selectable marker cassette and fluorescent
525 protein gene into the endogenous gene to produce an in-frame gene fusion. The fusion PCR
526 method described in Dean *et al.*, [83] was used for tagging with eYFP, using pJ1170 (pLENT-
527 YB) as the template plasmid for PCR and selection with 5 µg/ml Blasticidin-S deaminase.
528 The CRISPR-Cas9 method described in Beneke et al., [22] was used for tagging with
529 mNeonGreen. The online primer design tool www.LeishGEdit.net was used to design primers
530 for amplification of the 5' or 3' sgRNA template and primers for amplification of donor DNA
531 from pPLOTv1 blast-mNeonGreen-blast or pPLOTv1 puro-mNeonGreen-puro. Transfectants
532 were selected with either 5 µg/ml Blasticidin-S deaminase or 20 µg/ml Puromycin
533 Dihydrochloride.

534 ***CRISPR-Cas9 gene knockouts***

535 Gene deletions were essentially done as described in Beneke et al., [22]. The online primer
536 design tool www.LeishGEdit.net was used to design primers for amplification of the 5' and 3'
537 sgRNA templates and for amplification of donor DNA from pTBlast and either pTPuro or
538 pTNeo. Primers for deletion of *PFR2A-C* were designed with the EuPaGDT primer design

539 tool [84] using the *PFR2* array sequence from *L. mex* Cas9 T7. For amplification of the
540 sgRNA template DNA, primers:
541 Nsg: 5'-gaaattaatacgactcactataggTGCGTGCGGAGGTTTGCACGgttttagagctagaaatagc-3' /
542 Csg: 5'- gaaattaatacgactcactataggAAGGGTGGACGCCATCTCCGgttttagagctagaaatagc-3'.
543 For amplification of a pTNeo donor DNA fragment with 80 bp homology arms, primers:
544 F: 5'GCCCACCCCTTTCACTCTTTCGCTGCTCTCTCACCTCACCGACCCCCGCCTCTTT
545 CCATCTCTCACTGTGTGCTCCACCTgtataatgcagacctgctgc-3' /
546 R: 5'-AGCAGCCTTGAGACGACACCTGTAACAAAACCATCACCACAAGCTCCAAGGCGA
547 CAACATCGCGGGAAGACTTCGCCCCAccaatttgagagacctgtgc-3'.
548 For transfections on 96-well plates the protocol was modified as follows (similar to
549 descriptions in [85]): 52 x 10^7 late log phase *L. mex* Cas9 T7 cells (1 x 10^7 cells per reaction)
550 were collected by centrifugation at 800*g* for 15 min. A transfection buffer was prepared by
551 mixing 2 ml 1.5 mM $CaCl_2$, 6.5 ml modified 3x Tb-BSF buffer (22.3 mM Na_2HPO_4, 7.67 mM
552 NaH_2PO_4, 45 mM KCl, 450 mM sucrose, 75 mM HEPES pH 7.4) and 1.5 ml ddH_2O. The cells
553 were re-suspended in 3 ml of this transfection buffer and centrifuged again as above. The
554 heat-sterilised sgRNA and donor DNA PCR products were placed into 48 wells of a 96-well
555 disposable electroporation plate (4 mm gap, 250 µl, BTX) such that a given well combined all
556 of the donor DNAs and the sgRNA templates for a given target gene. After centrifugation,
557 cells were re-suspended in 5.2 ml transfection buffer and 100 µl of the cell suspension
558 dispensed into each of the 48 wells containing the PCR products. Plates were sealed with foil
559 and placed into the HT-200 plate handler of a BTX ECM 830 Electroporation System.
560 Transfection used the following settings: 1500 V, 24 pulses, 2 counted pulses, 500 ms
561 interval, unipolar, 100 µs. After transfection cells were recovered in 1 ml supplemented M199
562 in 24-well plates and incubated for 8-16 h at 28°C / 5% CO_2 before addition of the relevant
563 selection drugs by adding 1 ml of supplemented M199 with double the concentration of the
564 desired drugs. Mutants were selected with 5 µg/ml Blasticidin-S deaminase in combination
565 with either 20 µg/ml Puromycin Dihydrochloride or 40 µg/ml G-418 Disulfate and further
566 incubated. Drug resistant populations typically emerged after one week.

567 ***Diagnostic PCR for knockout validation***

568 Drug-selected populations were passaged at least twice (one with at least a 1:100 dilution)
569 before extraction of genomic DNA using the protocol described in [86]. A diagnostic PCR was
570 done to test for the presence of the target gene ORF in putative KO lines and the parental
571 cell line (S8 Figure). Primer3 [87] was used to design, for the entire *L. mexicana* genome
572 (gene models based on [28]), primers to amplify a short 100 – 300 bp unique fragment of the
573 target gene ORF (S7 Table). In a second reaction, primers 518F: 5'-

574 CACCCTCATTGAAAGAGCAAC-3' and 518R: 5'-CACTATCGCTTTGATCCCAGGA-3' were
575 used to amplify the blasticidin resistance gene from the same genomic DNA samples. For
576 Δ*PFR2* additional primers were used to confirm deletions (S10 Figure;
577 1F: 5'-GCAGAAGGAGAAGAGCGAGC-3'; 1R: 5'-CCAGGAACTGCTGGTACTCC-3';
578 2F: 5'-CGCAGAAGGAGAAGAGCGAG-3'; 2R: 5'-GTTGTACACGGACAGCTCCA-3';
579 3F: 5'-ACCCCTTTCACTCTTTCGCTG-3'; 4R: 5'-ACCAACGACGTACACAGCAG-3').

580 ***Light and electron microscopy***

581 *Leishmania* cells expressing fluorescent fusion proteins were imaged live. Samples were
582 prepared as described in [17]. Cells were immediately imaged with a Zeiss Axioimager.Z2
583 microscope with a 63× numerical aperture (NA) 1.40 oil immersion objective and a
584 Hamamatsu ORCA-Flash4.0 camera or a 63× NA 1.4 objective lens on a DM5500 B
585 microscope (Leica Microsystems) with a Neo sCMOS camera (Andor Technology) at the
586 ambient temperature of 25–28°C.
587 For transmission electron microscopy, deflagellated cell bodies and isolated flagella were
588 prepared with a modified version of the chemical fixation protocol described in Höög et al.,
589 [88]. Pellets of cell fractions were fixed with 2.5% glutaraldehyde in 10 mM PIPES (buffer as
590 described above) overnight at 4°C. Pellets were washed four times for 15 min in 10 mM
591 PIPES and overlaid with 10 mM PIPES, containing 1% osmium tetraoxide and incubated at
592 4°C for 1 h in darkness, then washed five times with ddH_2O for 5 min each time and stained
593 with 500 µl of 0.5% uranyl acetate in darkness at 4°C overnight. Samples were dehydrated,
594 embedded in resin, sectioned and on-section stained as described previously [88]. Electron
595 micrographs were captured on a Tecnai 12 TEM (FEI) with an Ultrascan 1000 CCD camera
596 (Gatan).

597 ***Image processing and analysis***

598 Micrographs were processed using Fiji [89]. To enable comparisons between the parental
599 and tagged cell lines, the same display settings for the green fluorescence channel were
600 used. PFR length (defined by reporter signal) and flagellar length (distance between stained
601 kinetoplast DNA and flagellar tip) was measured with the ROI manager in Fiji [89].

602 ***Motility assays***

603 Mutant and parental cell lines were tracked using the previously described method in [32]
604 with three modifications. Firstly, the scripts were modified for batch analysis of multiple files.
605 Secondly, prior to calculation of swimming statistics any 'drift' due to bulk fluid flow in the
606 sample was subtracted. As swimming direction of each cell in the population is in a random
607 direction any drift is visible as a mean population movement between successive frames. We

608 treated drift as an apparent translation and scaling of cell positions between successive
609 video frames, which was then subtracted. Finally, the primary statistics we considered were
610 mean speed (using the path at 200 ms resolution) and directionality (mean velocity as a
611 fraction of mean speed). Swimming behaviour was measured in triplicates at approximately
612 $6{\cdot}10^6$ cells/ml and analysed from 5 µl of cell culture on a glass slide in a 250-µm deep
613 chamber covered with a 1.5 mm cover slip using darkfield illumination with a 10× NA 0.3
614 objective and a Hamamatsu ORCA-Flash4.0 camera on a Zeiss Axioimager.Z2 microscope
615 at the ambient temperature of 25–28°C. The sample was stored inverted prior to use, then
616 turned upright immediately prior to imaging to ensure consistent motion of immotile cells
617 through sedimentation between samples. A sample of the parental cell line killed with a final
618 concentration of 1% formaldehyde was used as a reference for motion of completely
619 paralysed cells through sedimentation and Brownian motion alone.

620 ***Lutzomyia longipalpis* infections**

621 Sand flies were either infected with pooled mutant populations of *L. mexicana* or individually
622 with *L. mexicana* MNYC/BZ/62/M379 wild type (WT), parental line *L. mex* Cas9 T7, knockout
623 cell line Δ*PF16*, its add-back (*PF16*AB) [22], knockout cell line Δ*BBS2* and its add-back
624 (*BBS2*AB). All parasites were cultivated at 23°C in M199 medium supplemented with 20%
625 foetal calf serum (Gibco), 1% BME vitamins (Sigma), 2% sterile urine and 250 µg/ml amikin
626 (Amikin, Bristol-Myers Squibb). Before experimental infections, logarithmic growing parasites
627 were washed three times in PBS and resuspended in defibrinated heat-inactivated rabbit
628 blood at a concentration of 10^6 promastigotes/ml. *Lutzomyia longipalpis* (from Jacobina,
629 Brazil) were maintained at 26°C and high humidity on 50% sucrose solution and a 12h
630 light/12h dark photoperiod. Sand fly females, 3-5 days old, were fed through a chick skin
631 membrane as described previously [90]. Fully-engorged females were separated and
632 maintained at 26° C with free access to 50% sucrose solution. They were dissected on days
633 2 or 6 post blood-meal (PBM) and the guts were checked for localisation and intensity of
634 infection by light microscopy. Parasite load was graded as described previously by [91] into
635 four categories: zero, weak (<100 parasites/gut), moderate (100–1000 parasites/gut) and
636 heavy (>1000 parasites/gut).

637 ***DNA extraction for bar-seq screen***

638 Mutant *Leishmania* lines were grown separately to late log phase and mixed in equal
639 proportions ($1{\cdot}10^7$ cells per cell line). This pool was divided equally into three aliquots. DNA
640 was extracted using the Roche High Pure Nucleic Acid Kit according to the manufacturers
641 instructions. Each aliquot was then used to infect three separate batches of 150 sand flies.

642 The three batches were kept separate and DNA was extracted from 50 whole sand flies two,
643 six and nine days post blood meal, using the same kit as follows: Sand flies were placed in
644 two 1.5 ml microcentrifuge tubes (25 flies per tube), overlaid with 100 µl tissue lysis buffer
645 and frozen at -20°C. The dead flies were defrosted and after addition of 100 µl tissue lysis
646 buffer and 40 µl proteinase K, flies were disrupted with a disposable plastic pestle (Bel-Art)
647 and DNA purified according to the kit manufacturer's instructions. DNA was eluted with 50 µl
648 elution buffer and eluates from the same timepoint were combined, yielding 100 µl in total.

649 ***Bar-seq library preparation and sequencing***

650 For each sample, 600 ng isolated DNA was incubated with exonuclease VII (NEB) overnight
651 at 37°C, purified using SPRI magnetic beads and amplified using PAGE purified custom
652 designed p5 and p7 primers (Life Technologies), containing indexes for multiplexing and
653 adapters for Illumina sequencing. Amplicons were again bead purified and multiplexed by
654 pooling them in equal proportions. The final sequencing pool was once again bead purified
655 and quantified by qPCR using NEB Library Quant Kit. Library size was determined using
656 Agilent High Sensitivity DNA Kit on a 2100 Bioanalyzer instrument. The pool was diluted to 4
657 nM and spiked with 30% single indexed *Leishmania* genomic DNA, prepared using Illumina
658 TruSeq Nano DNA Library kit according to the manufacturers instructions. The final library
659 was spiked with 1% PhiX DNA and the Illumina sequencer was loaded with 8 pM to allow low
660 cluster density (~800 K/mm²). Sequencing was performed using MiSeq v3 150 cycles kit
661 following the manufactures instructions with paired-end sequencing (2x75 cycles, 6 and 8
662 cycles index read).
663 MiSeq raw files were de-multiplexed using bcl2fastq (Illumina). De-multiplexed samples were
664 subjected to barcode counting using a bash script. Each gene in the *Leishmania* genome
665 was assigned a unique barcode – the number of times each of these barcodes appeared in
666 the sequencing output was recorded. The criteria for barcode counting was a 100% match to
667 the 17 nt total length. Counts for each barcode were normalized for each sample by
668 calculating their abundance relative to all 25 barcodes. One of the mutants selected for the
669 pooled screen was excluded from the analysis because sequencing of the cell line showed
670 eight nucleotide mismatches in the p5 sequence (likely introduced during oligonucleotide
671 synthesis) which precluded amplification of the barcode region. To calculate "fitness"
672 normalized barcode counts in the pooled population before feeding were divided by
673 normalized counts at the relevant time point post blood meal.

674 ***Orthofinder***

675 Orthofinder 1.1.2 was used to generate orthogroups for predicted protein coding genes from
676 48 eukaryotic genomes (32 ciliated species and 16 non-ciliated species): The 45 previously
677 used by Hodges *et al*. [64] (with *Leishmania major* replaced with *Leishmania mexicana*) and
678 supplemented with *Aspergillus nidulans*, *Plasmodium berghei* and *Schistosoma mansoni*.

679 **Acknowledgements**

680 We like to thank Svenja Hester, Philip Charles and Benjamin Thomas Central Proteomics
681 Facility at the Sir William Dunn School of Pathology for help with protein mass spectrometry,
682 Errin Johnson (Dunn School Bioimaging Facility) for assistance with electron microscopy,
683 Amanda Williams (University of Oxford) for help with Illumina sequencing, Maxime Cesca
684 (Université de Paris Sud) and Laura Makin (University of Oxford) for assistance with mutant
685 characterisation, Oliver Billker (Wellcome Sanger Institute) and members of O.B.'s lab for
686 helpful discussions about the bar-seq method, Sébastien Pomel (Université de Paris Sud) for
687 drawing our attention to the PMI, PMM and GDP-MP mutants, Diane McMahon-Pratt (Yale
688 University) for antibody 2E10 (F-4) and Keith Gull (University of Oxford) for advice and
689 helpful comments on the manuscript, antibodies L8C4 and L13D6 and access to equipment,
690 and members of K.G.'s lab for helpful discussions, particularly Samuel Dean and Jack Sunter
691 for help with development of high-throughput transfection protocols.

692 **Funding statement**

693 This research was jointly funded by the UK Medical Research Council (MRC) and the UK
694 Department for International Development (DFID) under the MRC/DFID Concordat
695 agreement; grant no. MR/R000859/1.

696 EG was supported by a Royal Society University Research Fellowship (UF100435 and
697 UF160661).

698 RW was supported by the Wellcome Trust, grant nos. [211075/Z/18/Z, 103261/Z/13/Z,
699 104627/Z/14/Z]

700 PV, JSad and TL were supported by European Regional Development (ERD) Funds, project
701 CePaViP (CZ.02.1.01/0.0/0.0/ 16_019/0000759).

702 This work was supported by MRC PhD studentships to TB (15/16_MSD_836338) and to JV
703 (13/14_MSD_OSS_363238), by BBSRC Interdisciplinary Biosciences DTP studentships to
704 HJ and SS, by an Oxford Radcliffe Scholarship to HJ and by Erasmus grants to TB and FD.

NA was funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Moran J, McKean PG, Ginger ML (2014) Eukaryotic Flagella: Variations in Form, Function, and Composition during Evolution. BioScience 64: 1103-1114.
2. Ginger ML, Portman N, McKean PG (2008) Swimming with protists: perception, motility and flagellum assembly. Nat Rev Microbiol 6: 838-850.
3. Vincensini L, Blisnick T, Bastin P (2011) 1001 model organisms to study cilia and flagella. Biol Cell 103: 109-130.
4. Badano JL, Mitsuma N, Beales PL, Katsanis N (2006) The ciliopathies: an emerging class of human genetic disorders. Annu Rev Genomics Hum Genet 7: 125-148.
5. van Dam TJ, Wheway G, Slaats GG, Huynen MA, Giles RH (2013) The SYSCILIA gold standard (SCGSv1) of known ciliary components and its applications within a systems biology consortium. Cilia 2: 7.
6. Broadhead R, Dawe HR, Farr H, Griffiths S, Hart SR, et al. (2006) Flagellar motility is required for the viability of the bloodstream trypanosome. Nature 440: 224-227.
7. Pazour GJ, Agrin N, Leszyk J, Witman GB (2005) Proteomic analysis of a eukaryotic cilium. J Cell Biol 170: 103-113.
8. Mayer U, Kuller A, Daiber PC, Neudorf I, Warnken U, et al. (2009) The proteome of rat olfactory sensory cilia. Proteomics 9: 322-334.
9. Nakachi M, Nakajima A, Nomura M, Yonezawa K, Ueno K, et al. (2011) Proteomic profiling reveals compartment-specific, novel functions of ascidian sperm proteins. Mol Reprod Dev 78: 529-549.
10. Nevers Y, Prasad MK, Poidevin L, Chennen K, Allot A, et al. (2017) Insights into Ciliary Genes and Evolution from Multi-Level Phylogenetic Profiling. Mol Biol Evol 34: 2016-2034.
11. Krüger T, Engstler M (2015) Flagellar motility in eukaryotic human parasites. Semin Cell Dev Biol 46: 113-127.
12. Hochstetter A, Pfohl T (2016) Motility, Force Generation, and Energy Consumption of Unicellular Parasites. Trends Parasitol 32: 531-541.
13. Langousis G, Hill KL (2014) Motility and more: the flagellum of *Trypanosoma brucei*. Nat Rev Microbiol 12: 505-518.
14. Engstler M, Pfohl T, Herminghaus S, Boshart M, Wiegertjes G, et al. (2007) Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. Cell 131: 505-515.
15. Vaughan S (2010) Assembly of the flagellum and its role in cell morphogenesis in *Trypanosoma brucei*. Curr Opin Microbiol 13: 453-458.
16. Alsford S, Turner DJ, Obado SO, Sanchez-Flores A, Glover L, et al. (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Res 21: 915-924.
17. Wheeler RJ, Gluenz E, Gull K (2015) Routes to a 9+0 flagellum: Basal body multipotency and axonemal plasticity. Nature Communications 6: 8964.
18. Gluenz E, Höög JL, Smith AE, Dawe HR, Shaw MK, et al. (2010) Beyond 9+0: noncanonical axoneme structures characterize sensory cilia from protists to humans. FASEB J 24: 3117-3121.

752 19. Gadelha AP, Cunha-e-Silva NL, de Souza W (2013) Assembly of the *Leishmania*
753 amazonensis flagellum during cell differentiation. J Struct Biol 184: 280-292.
754 20. Killick-Kendrick R, Molyneux DH, Ashford RW (1974) *Leishmania* in phlebotomid
755 sandflies. I. Modifications of the flagellum associated with attachment to the mid-gut
756 and oesophageal valve of the sandfly. Proc R Soc Lond B Biol Sci 187: 409-419.
757 21. Lawyer PG, Ngumbi PM, Anjili CO, Odongo SO, Mebrahtu YB, et al. (1990) Development
758 of *Leishmania major* in *Phlebotomus duboscqi* and *Sergentomyia schwetzi* (Diptera:
759 Psychodidae). Am J Trop Med Hyg 43: 31-43.
760 22. Beneke T, Madden R, Makin L, Valli J, Sunter J, et al. (2017) A CRISPR Cas9 high-
761 throughput genome editing toolkit for kinetoplastids. R Soc Open Sci 4: 170095.
762 23. Tull D, Vince JE, Callaghan JM, Naderer T, Spurck T, et al. (2004) SMP-1, a member of
763 a new family of small myristoylated proteins in kinetoplastid parasites, is targeted to
764 the flagellum membrane in *Leishmania*. Mol Biol Cell 15: 4775-4786.
765 24. Trudgian DC, Ridlova G, Fischer R, Mackeen MM, Ternette N, et al. (2011) Comparative
766 evaluation of label-free SINQ normalized spectral index quantitation in the central
767 proteomics facilities pipeline. Proteomics 11: 2790-2797.
768 25. Lubec G, Afjehi-Sadat L (2007) Limitations and pitfalls in protein identification by mass
769 spectrometry. Chem Rev 107: 3568-3584.
770 26. Varga V, Moreira-Leite F, Portman N, Gull K (2017) Protein diversity in discrete
771 structures at the distal tip of the trypanosome flagellum. Proc Natl Acad Sci U S A
772 114: E6546-E6555.
773 27. Dean S, Moreira-Leite F, Varga V, Gull K (2016) Cilium transition zone proteome reveals
774 compartmentalization and differential dynamics of ciliopathy complexes. Proc Natl
775 Acad Sci U S A 113: E5135-5143.
776 28. Fiebig M, Kelly S, Gluenz E (2015) Comparative lifecycle transcriptomics revises
777 *Leishmania mexicana* genome annotation and links a chromosome duplication with
778 parasitism of vertebrates. PLoS Pathog 11: e1005186.
779 29. Lahav T, Sivam D, Volpin H, Ronen M, Tsigankov P, et al. (2011) Multiple levels of gene
780 regulation mediate differentiation of the intracellular pathogen *Leishmania*. FASEB J
781 25: 515-525.
782 30. Emms D, Kelly S (2015) OrthoFinder: solving fundamental biases in whole genome
783 comparisons dramatically improves orthologous gene group inference accuracy. .
784 Genome Biol. 16:157.
785 31. Aslett M, Aurrecoechea C, Berriman M, Brestelli J, Brunk BP, et al. (2010) TriTrypDB: a
786 functional genomic resource for the Trypanosomatidae. Nucleic Acids Res 38: D457-
787 462.
788 32. Wheeler RJ (2017) Use of chiral cell shape to ensure highly directional swimming in
789 trypanosomes. PLoS Comput Biol 13: e1005353.
790 33. Edwards BFL, Wheeler RJ, Barker AR, Moreira-Leite FF, Gull K, et al. (2018) Direction of
791 flagellum beat propagation is controlled by proximal/distal outer dynein arm
792 asymmetry. Proc Natl Acad Sci U S A 115: E7341-E7350.
793 34. Maga JA, Sherwin T, Francis S, Gull K, LeBowitz JH (1999) Genetic dissection of the
794 *Leishmania* paraflagellar rod, a unique flagellar cytoskeleton structure. J Cell Sci 112
795 (Pt 16): 2753-2763.
796 35. Adhiambo C, Forney JD, Asai DJ, LeBowitz JH (2005) The two cytoplasmic dynein-2
797 isoforms in *Leishmania mexicana* perform separate functions. Molecular and
798 Biochemical Parasitology 143: 216-225.
799 36. Zauli RC, Yokoyama-Yasunaka JK, Miguel DC, Moura AS, Pereira L, et al. (2012) A
800 dysflagellar mutant of *Leishmania* (*Viannia*) *braziliensis* isolated from a cutaneous
801 leishmaniasis patient. Parasit Vectors 5: 11.
802 37. Fowlkes-Comninellis T, Beverley SM (2015) *Leishmania* IFT140 mutants show normal
803 viability but lack external flagella: a tool for the study of flagellar function through the
804 infectious cycle. Cilia 4 (Suppl 1): P49.

805 38. Sunter J, Gull K (2017) Shape, form, function and *Leishmania* pathogenicity: from
806 textbook descriptions to biological understanding. Open Biol 7.
807 39. Bates PA (2008) *Leishmania* sand fly interaction: progress and challenges. Curr Opin
808 Microbiol 11: 340-344.
809 40. Leslie G, Barrett M, Burchmore R (2002) *Leishmania mexicana*: promastigotes migrate
810 through osmotic gradients. Exp Parasitol 102: 117-120.
811 41. Smith AM, Heisler LE, Mellor J, Kaper F, Thompson MJ, et al. (2009) Quantitative
812 phenotyping via deep barcode sequencing. Genome Res 19: 1836-1842.
813 42. Gomes AR, Bushell E, Schwach F, Girling G, Anar B, et al. (2015) A genome-scale
814 vector resource enables high-throughput reverse genetic screening in a malaria
815 parasite. Cell Host Microbe 17: 404-413.
816 43. Bushell E, Gomes AR, Sanderson T, Anar B, Girling G, et al. (2017) Functional Profiling
817 of a *Plasmodium* Genome Reveals an Abundance of Essential Genes. Cell 170: 260-
818 272 e268.
819 44. Ryan KA, Garraway LA, Descoteaux A, Turco SJ, Beverley SM (1993) Isolation of
820 virulence genes directing surface glycosyl-phosphatidylinositol synthesis by functional
821 complementation of *Leishmania*. Proc Natl Acad Sci U S A 90: 8609-8613.
822 45. Garami A, Ilg T (2001) The role of phosphomannose isomerase in *Leishmania mexicana*
823 glycoconjugate synthesis and virulence. J Biol Chem 276: 6566-6575.
824 46. Garami A, Mehlert A, Ilg T (2001) Glycosylation defects and virulence phenotypes of
825 *Leishmania mexicana* phosphomannomutase and dolicholphosphate-mannose
826 synthase gene deletion mutants. Mol Cell Biol 21: 8168-8183.
827 47. Garami A, Ilg T (2001) Disruption of mannose activation in *Leishmania mexicana*: GDP-
828 mannose pyrophosphorylase is required for virulence, but not for viability. EMBO J
829 20: 3657-3666.
830 48. Sacks DL, Modi G, Rowton E, Späth G, Epstein L, et al. (2000) The role of
831 phosphoglycans in Leishmania-sand fly interactions. Proc Natl Acad Sci U S A 97:
832 406-411.
833 49. Jecna L, Dostalova A, Wilson R, Seblova V, Chang KP, et al. (2013) The role of surface
834 glycoconjugates in *Leishmania* midgut attachment examined by competitive binding
835 assays and experimental development in sand flies. Parasitology 140: 1026-1032.
836 50. Tran KD, Rodriguez-Contreras D, Vieira DP, Yates PA, David L, et al. (2013) KHARON1
837 mediates flagellar targeting of a glucose transporter in *Leishmania mexicana* and is
838 critical for viability of infectious intracellular amastigotes. J Biol Chem 288: 22721-
839 22733.
840 51. Nachury MV, Loktev AV, Zhang Q, Westlake CJ, Peranen J, et al. (2007) A core complex
841 of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane
842 biogenesis. Cell 129: 1201-1213.
843 52. Branche C, Kohl L, Toutirais G, Buisson J, Cosson J, et al. (2006) Conserved and
844 specific functions of axoneme components in trypanosome motility. J Cell Sci 119:
845 3443-3455.
846 53. Ralston KS, Lerner AG, Diener DR, Hill KL (2006) Flagellar motility contributes to
847 cytokinesis in *Trypanosoma brucei* and is modulated by an evolutionarily conserved
848 dynein regulatory system. Eukaryot Cell 5: 696-711.
849 54. Dutcher SK, Huang B, Luck DJ (1984) Genetic dissection of the central pair microtubules
850 of the flagella of *Chlamydomonas reinhardtii*. J Cell Biol 98: 229-236.
851 55. Rotureau B, Ooi CP, Huet D, Perrot S, Bastin P (2014) Forward motility is essential for
852 trypanosome infection in the tsetse fly. Cell Microbiol 16: 425-433.
853 56. Cuvillier A, Miranda JC, Ambit A, Barral A, Merlin G (2003) Abortive infection of
854 *Lutzomyia longipalpis* insect vectors by aflagellated LdARL-3A-Q70L overexpressing
855 *Leishmania amazonensis* parasites. Cell Microbiol 5: 717-728.
856 57. Pimenta PF, Saraiva EM, Rowton E, Modi GB, Garraway LA, et al. (1994) Evidence that
857 the vectorial competence of phlebotomine sand flies for different species of

858 *Leishmania* is controlled by structural polymorphisms in the surface
859 lipophosphoglycan. Proc Natl Acad Sci U S A 91: 9155-9159.
860 58. Warburg A, Tesh RB, McMahon-Pratt D (1989) Studies on the attachment of *Leishmania*
861 flagella to sand fly midgut epithelium. J Protozool 36: 613-617.
862 59. Sharma AI, Olson CL, Engman DM (2017) The Lipid Raft Proteome of African
863 Trypanosomes Contains Many Flagellar Proteins. Pathogens 6: 39.
864 60. Rodriguez-Contreras D, Aslan H, Feng X, Tran K, Yates PA, et al. (2015) Regulation and
865 biological function of a flagellar glucose transporter in *Leishmania mexicana*: a
866 potential glucose sensor. FASEB J 29: 11-24.
867 61. Pozzo LY, Fontes A, de Thomaz AA, Santos BS, Farias PM, et al. (2009) Studying taxis
868 in real time using optical tweezers: applications for *Leishmania amazonensis*
869 parasites. Micron 40: 617-620.
870 62. Porter ME, Sale WS (2000) The 9 + 2 axoneme anchors multiple inner arm dyneins and
871 a network of kinases and phosphatases that control motility. J Cell Biol 151: F37-42.
872 63. Mukhopadhyay AG, Dey CS (2016) Reactivation of flagellar motility in demembranated
873 *Leishmania* reveals role of cAMP in flagellar wave reversal to ciliary waveform. Sci
874 Rep 6: 37308.
875 64. Hodges ME, Scheumann N, Wickstead B, Langdale JA, Gull K (2010) Reconstructing the
876 evolutionary history of the centriole from protein components. J Cell Sci 123: 1407-
877 1413.
878 65. Jin H, White SR, Shida T, Schulz S, Aguiar M, et al. (2010) The conserved Bardet-Biedl
879 syndrome proteins assemble a coat that traffics membrane proteins to cilia. Cell 141:
880 1208-1219.
881 66. Ismail SA, Chen YX, Miertzschke M, Vetter IR, Koerner C, et al. (2012) Structural basis
882 for Arl3-specific release of myristoylated ciliary cargo from UNC119. EMBO J 31:
883 4085-4094.
884 67. Santrich C, Moore L, Sherwin T, Bastin P, Brokaw C, et al. (1997) A motility function for
885 the paraflagellar rod of *Leishmania* parasites revealed by *PFR-2* gene knockouts. Mol
886 Biochem Parasitol 90: 95-109.
887 68. Bastin P, Sherwin T, Gull K (1998) Paraflagellar rod is vital for trypanosome motility.
888 Nature 391: 548.
889 69. Portman N, Lacomble S, Thomas B, McKean PG, Gull K (2009) Combining RNA
890 interference mutants and comparative proteomics to identify protein components and
891 dependences in a eukaryotic flagellum. J Biol Chem 284: 5610-5619.
892 70. Portman N, Gull K (2010) The paraflagellar rod of kinetoplastid parasites: from structure
893 to components and function. Int J Parasitol 40: 135-148.
894 71. Coutton C, Vargas AS, Amiri-Yekta A, Kherraf ZE, Ben Mustapha SF, et al. (2018)
895 Mutations in CFAP43 and CFAP44 cause male infertility and flagellum defects in
896 *Trypanosoma* and human. Nat Commun 9: 686.
897 72. Davy BE, Robinson ML (2003) Congenital hydrocephalus in hy3 mice is caused by a
898 frameshift mutation in Hydin, a large novel gene. Hum Mol Genet 12: 1163-1170.
899 73. Lechtreck KF, Witman GB (2007) *Chlamydomonas reinhardtii* hydin is a central pair
900 protein required for flagellar motility. J Cell Biol 176: 473-482.
901 74. Dawe HR, Shaw MK, Farr H, Gull K (2007) The hydrocephalus inducing gene product,
902 Hydin, positions axonemal central pair microtubules. BMC Biol 5: 33.
903 75. Lechtreck KF, Delmotte P, Robinson ML, Sanderson MJ, Witman GB (2008) Mutations in
904 Hydin impair ciliary motility in mice. J Cell Biol 180: 633-643.
905 76. Lindemann CB, Lesich KA (2010) Flagellar and ciliary beating: the proven and the
906 possible. J Cell Sci 123: 519-528.
907 77. Lin J, Nicastro D (2018) Asymmetric distribution and spatial switching of dynein activity
908 generates ciliary motility. Science 360.
909 78. Oberholzer M, Langousis G, Nguyen HT, Saada EA, Shimogawa MM, et al. (2011)
910 Independent analysis of the flagellum surface and matrix proteomes provides insight

911 into flagellum signaling in mammalian-infectious *Trypanosoma brucei*. Mol Cell
912 Proteomics 10: M111 010538.
913 79. Sagne C, Isambert MF, Henry JP, Gasnier B (1996) SDS-resistant aggregation of
914 membrane proteins: application to the purification of the vesicular monoamine
915 transporter. Biochem J 316 (Pt 3): 825-831.
916 80. Gundry RL, White MY, Murray CI, Kane LA, Fu Q, et al. (2009) Preparation of proteins
917 and peptides for mass spectrometry analysis in a bottom-up proteomics workflow.
918 Curr Protoc Mol Biol Chapter 10: Unit10 25.
919 81. Trudgian DC, Thomas B, McGowan SJ, Kessler BM, Salek M, et al. (2010) CPFP: a
920 central proteomics facilities pipeline. Bioinformatics 26: 1131-1132.
921 82. Vizcaino JA, Csordas A, Del-Toro N, Dianes JA, Griss J, et al. (2016) 2016 update of the
922 PRIDE database and its related tools. Nucleic Acids Res 44: 11033.
923 83. Dean S, Sunter J, Wheeler RJ, Hodkinson I, Gluenz E, et al. (2015) A toolkit enabling
924 efficient, scalable and reproducible gene tagging in trypanosomatids. Open Biol 5:
925 140197.
926 84. Peng D, Tarleton R (2015) EuPaGDT: a web tool tailored to design CRISPR guide RNAs
927 for eukaryotic pathogens. Microb Genom 1: e000033.
928 85. Dyer P, Dean S, Sunter J (2016) High-throughput Gene Tagging in *Trypanosoma brucei*.
929 J Vis Exp.
930 86. Rotureau B, Gego A, Carme B (2005) Trypanosomatid protozoa: a simplified DNA
931 isolation procedure. Exp Parasitol 111: 207-209.
932 87. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, et al. (2012) Primer3--new
933 capabilities and interfaces. Nucleic Acids Res 40: e115.
934 88. Höög JL, Gluenz E, Vaughan S, Gull K (2010) Ultrastructural investigation methods for
935 *Trypanosoma brucei*. Methods Cell Biol 96: 175-196.
936 89. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, et al. (2012) Fiji: an
937 open-source platform for biological-image analysis. Nat Methods 9: 676-682.
938 90. Volf P, Volfova V (2011) Establishment and maintenance of sand fly colonies. J Vector
939 Ecol 36 Suppl 1: S1-9.
940 91. Myskova J, Votypka J, Volf P (2008) *Leishmania* in sand flies: comparison of quantitative
941 polymerase chain reaction with other techniques to determine the intensity of
942 infection. J Med Entomol 45: 133-138.
943 92. Dupe A, Dumas C, Papadopoulou B (2015) Differential Subcellular Localization of
944 *Leishmania* Alba-Domain Proteins throughout the Parasite Development. PLoS One
945 10: e0137243.
946 93. Subota I, Julkowska D, Vincensini L, Reeg N, Buisson J, et al. (2014) Proteomic analysis
947 of intact flagella of procyclic *Trypanosoma brucei* cells identifies novel flagellar
948 proteins with unique sub-localization and dynamics. Mol Cell Proteomics 13: 1769-
949 1786.
950 94. Kozlowski LP (2016) IPC - Isoelectric Point Calculator. Biol Direct 11: 55.
951 95. Kohl L, Sherwin T, Gull K (1999) Assembly of the paraflagellar rod and the flagellum
952 attachment zone complex during the *Trypanosoma brucei* cell cycle. J Eukaryot
953 Microbiol 46: 105-109.
954 96. Ismach R, Cianci CM, Caulfield JP, Langer PJ, Hein A, et al. (1989) Flagellar membrane
955 and paraxial rod proteins of *Leishmania*: characterization employing monoclonal
956 antibodies. J Protozool 36: 617-624.
957

958 # Legends

959 **Figure 1. Isolation of flagella and deflagellated cell bodies**

960 **(A)** Overview of the deflagellation and differential centrifugation protocol. Percentage sucrose
961 concentration (w/v) is indicated. **(B)** Micrographs show merged phase and fluorescence
962 channel (SMP-1::GFP, green; Hoechst DNA stain, red) for each isolation stage (i-iv) depicted
963 in (A). **(i)** *L. mexicana* SMP1::GFP cells before deflagellation, **(ii)** cells after deflagellation,
964 **(iii)** isolated flagella (F) and **(iv)** deflagellated cell bodies (CB). Scale bars represent 20 µm.
965 **(C)** Protein gel stained with SYPRO RUBY. Numbers on the left indicate molecular weight in
966 kDa, numbers below indicate cell equivalent of protein loaded on each lane. Each sample
967 lane of F_S, F_I, C_S and C_I was cut into eight pieces and analysed by mass spectrometry (two
968 biological replicates). **(D)** Venn diagram shows total number of all detected proteins (≥ 2
969 peptides detected, p-value > 0.95) of both replicates.

970 **Figure 2. Enrichment plot of flagellar and cell body proteins**

971 Proteins detected by MS were quantified with SINQ. Plotted on the X-axis is the log2 fold-
972 change of the spectral index (C_S+C_I / F_S+F_I) and on the Y-axis the log2 fold change of the
973 spectral index (F_S+C_S / F_I+C_I). To allow plotting of proteins that were only detected in one
974 fraction, a value of 10^{-10} was inserted for missing spectral indices. The resulting diagonal
975 lines in each quadrant represent proteins uniquely detected in the respective fraction. **(A)**
976 Each data point represents one of 2414 proteins detected in MS run 1 and bubble size
977 reflects protein abundance. Coloured circles indicate representative proteins for different
978 flagellar sub-structures (GeneIDs in S2 Table). The plot can be interactively explored on
979 http://www.leishgedit.net/leishgedit_db/. **(B)** Correlation with RNA-seq data. All 2414 proteins
980 detected in MS run 1 were plotted as in (A) and colour-coded according to the log2 fold
981 change of differentially expressed transcripts between promastigotes (PRO) and amastigotes
982 (AMA) [28].

983 **Figure 3. Identification of motility phenotypes**

984 All deletion cell lines generated in this study and the Δ*PF16* mutant [22] were analysed for
985 defects in swimming speed or directionality (the ratio of velocity to speed). **(A)** Mean
986 swimming speeds. Speeds were measured three times and the mean of all three replicates
987 (●) and the individual replicates (○) are shown. Error bars represent the standard deviation.
988 Red dotted lines indicate two standard deviations above and below the parental cell line
989 (Cas9 T7) mean swimming speed. Cas9 T7 cells killed with 1% formaldehyde (Cas9 T7
990 fixed) were used as a completely immotile control. *** p<0.0005, ** p<0.005, * p<0.05

991 (Student's t-test compared to the parental cell line). For a sub-set of mutants, an addback
992 copy of the deleted gene was introduced and swimming speeds restored toward the wild-
993 type. **(B)** The swimming speeds of all knockout mutants (○), as in (A), plotted against mean
994 directionality. Error bars represent the standard deviation of the three replicates. Four main
995 mutant phenotype clusters are apparent: Paralysed (including mutants lacking a long
996 flagellum), uncoordinated (which move slowly, but with greatly reduced directionality), slow
997 (which move slowly, but with reduced directionality and speed) and fast (which move faster,
998 with a tendency for higher directionality). **(C)** Speed and directionality of all knockout mutants
999 in (B) with deletion of IFT components, mutants with a tendency for curling of the flagellum
1000 (S9 Figure) or disrupted PFR structure (Figure 5) highlighted. **(D)** Speed and directionality of
1001 knockout mutants in (B) with those passaged though sand flies (Figure 6) highlighted.

1002 **Figure 4. Phenotype categories**

1003 Categories of mutant phenotypes: among cells with normal morphology there were normal
1004 swimmers as well as slower or faster swimmers. Four categories of distinct morphological
1005 defects were observed, which all lead to impaired motility: no flagellum, short flagellum,
1006 curled flagellum and disrupted PFR. One representative mutant is shown for each category,
1007 the deleted gene is indicated. Green, PFR2::mNG signal; red, Hoechst-stained DNA. Scale
1008 bar, 2 µm.

1009 **Figure 5. Measurement of flagellar and PFR length in motility mutants**

1010 (**A**) Measurements of flagellar length (measured from kinetoplast DNA to flagellar tip; grey
1011 bar) and PFR2::mNG signal (green bar; data in S8 Table). Error bars show standard
1012 deviation. At least 70 measurements were recorded per cell line. The GeneIDs / gene names
1013 indicate the deleted gene. Numbers above bars show PFR : flagellar length ratio.
1014 Measurements were compared to the parental cell line expressing PFR2::mNG and p-values
1015 (Students t-test) for flagellar length (grey) and PFR length (green) are indicated: *** $p \leq 0.001$.
1016 (**B**) Fluorescence micrographs showing tagged mutant cell lines used for measurements in
1017 (A). Green, PFR2::mNG signal; red, Hoechst-stained DNA. Scale bar, 5 µm. White arrows
1018 indicate PFR defects.

1019 **Figure 6. Relative fitness of *Leishmania* mutants in *L. longipalpis* infections**

1020 The plots display abundance of barcodes at time points post bloodmeal (PBM) relative to the
1021 abundance of this barcode in the mixed parasite population used to infect *L. longipalpis*.
1022 Mutants are grouped according to the function of the deleted gene and severity of motility
1023 phenotype (**A**) flagellar mutants with severe motility defects (paralyzed, uncoordinated

1024 swimmers and aflagellate cells), (**B**) flagellar mutants with mild motility defects, (**C**) mutants
1025 lacking flagellar membrane proteins or proteins involved in protein trafficking to the flagellar
1026 membrane, (**D**) mutants lacking key enzymes for the synthesis of LPG and other
1027 glycoconjugates, (**E**) control mutants with wild type motility. Data points represent the
1028 average of three replicates. Error bars show the standard deviation of the mean of the three
1029 replicates. Dotted red lines indicate two standard deviations above and below the parental
1030 cell line (SBL1-5). Measurements were compared (two-sided t-test) to the average of all five
1031 parental controls and p-values are indicated: *≤0.05, **≤0.005, ***≤0.0005.

1032 **Figure 7. Development of *L. mexicana* ΔPF16 in *L. longipalpis***

1033 **(A)** Infection rates (% of infected females) and intensities of infections (four categories) of *L.*
1034 *mexicana* in *L. longipalpis*. Numbers above the bars indicate numbers of dissected females
1035 in the group. **(B)** Localization of *L. mexicana* infections in *L. longipalpis*. ESP: endoperitrophic
1036 space, AMG: abdominal midgut, TMG: thoracic midgut, CAR: cardia, SV: stomodeal valve.
1037 Numbers above the bars indicate numbers of infected females in the group.

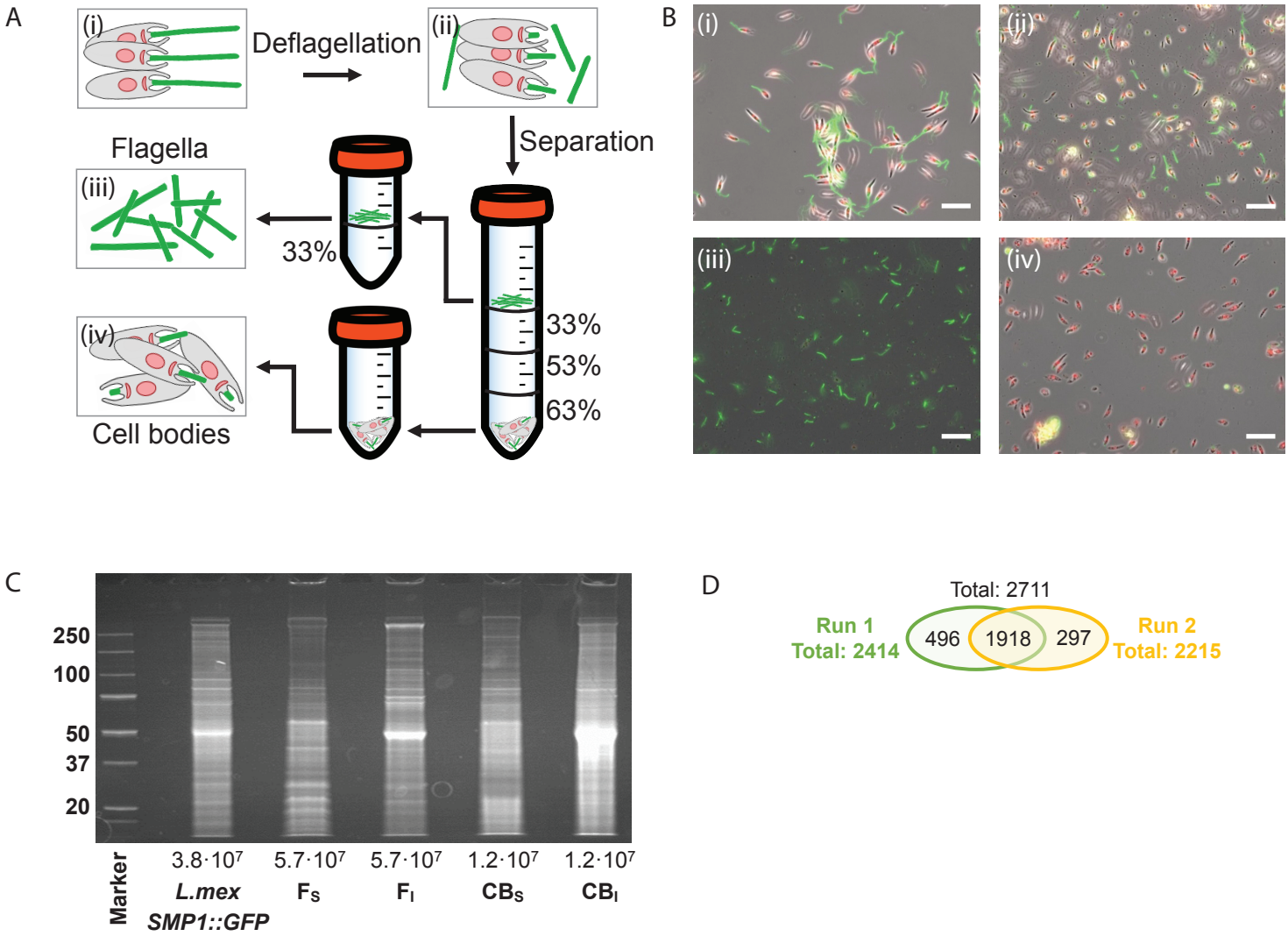

A
(i)
Deflagellation
(ii)
Separation
Flagella
(iii)
33%
33%
53%
63%
(iv)
Cell bodies
B
(i)
(ii)
(iii)
(iv)
C
250
100
50
37
20
Marker
3.8·10^7
L.mex SMP1::GFP
5.7·10^7
F_S
5.7·10^7
F_I
1.2·10^7
CB_S
1.2·10^7
CB_I
D
Total: 2711
Run 1
Total: 2414
496
1918
297
Run 2
Total: 2215

Figure 1

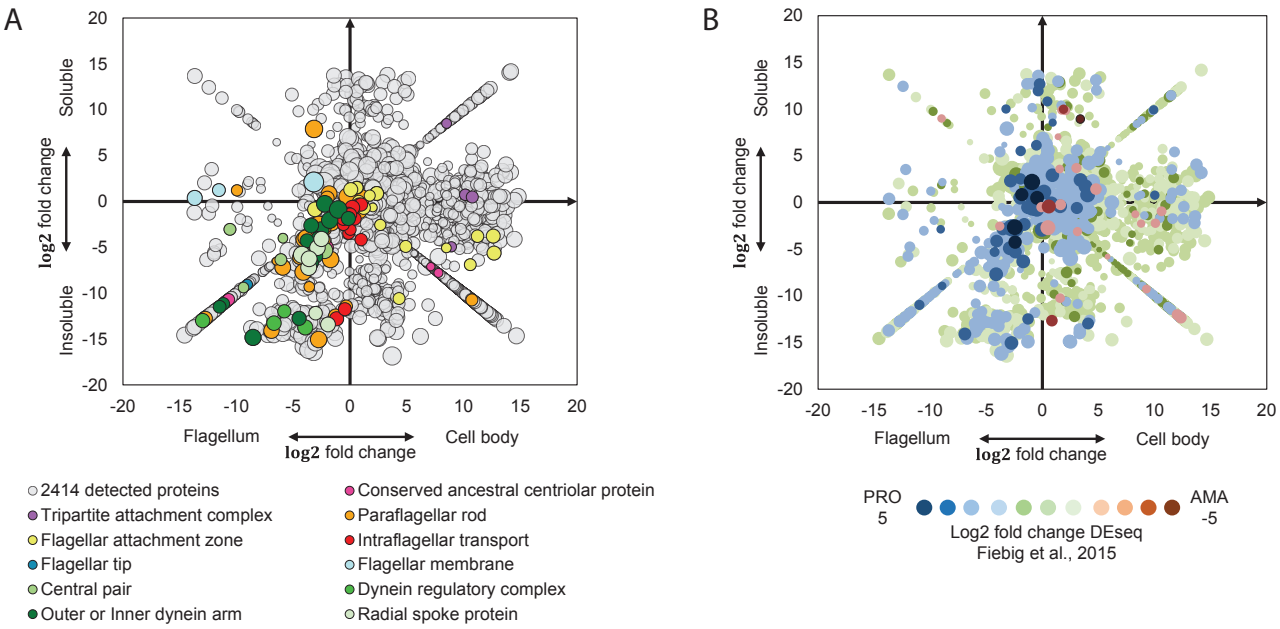
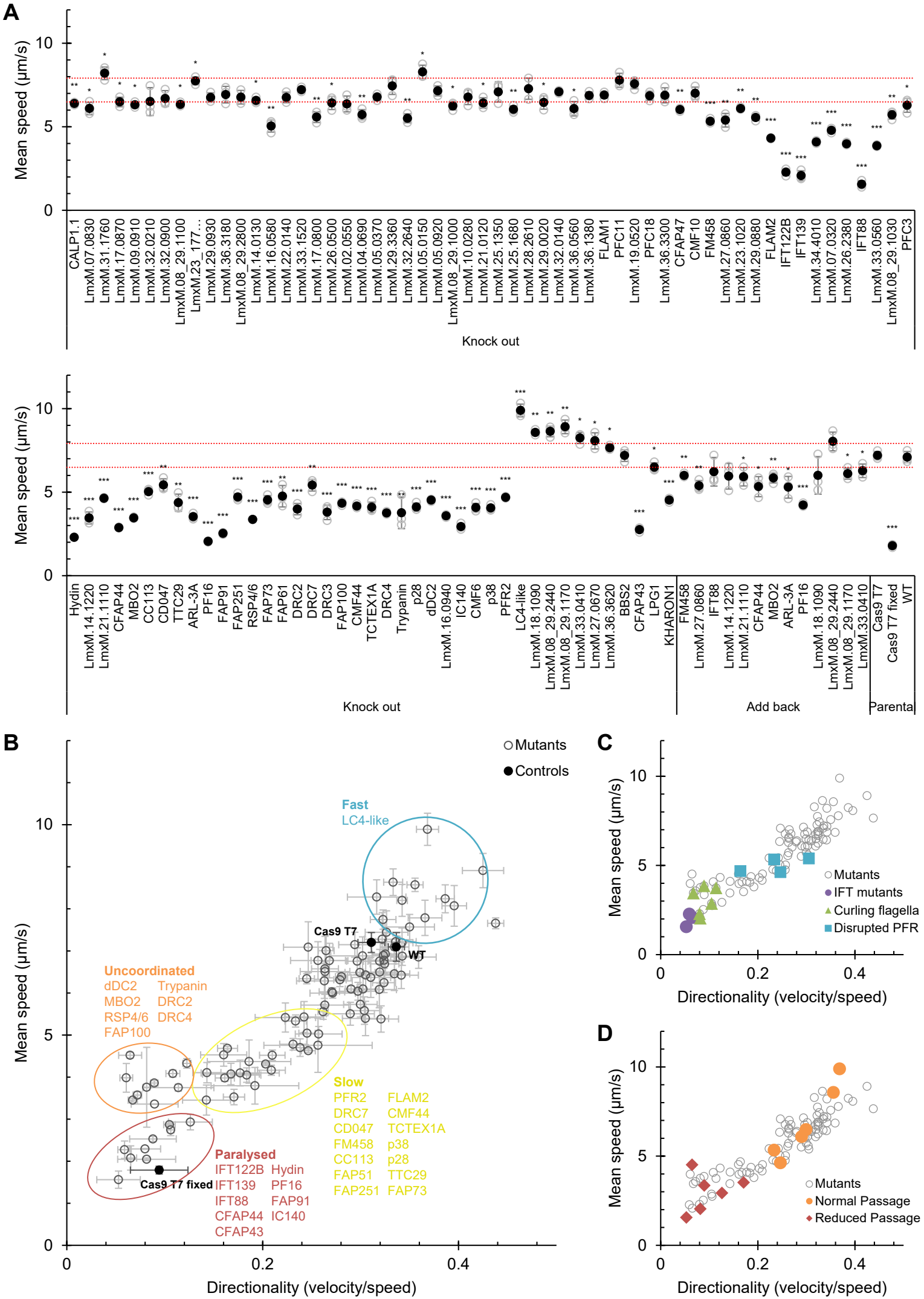

A
B
Soluble
Insoluble
log2 fold change
Flagellum
Cell body
log2 fold change
2414 detected proteins
Tripartite attachment complex
Flagellar attachment zone
Flagellar tip
Central pair
Outer or Inner dynein arm
Conserved ancestral centriolar protein
Paraflagellar rod
Intraflagellar transport
Flagellar membrane
Dynein regulatory complex
Radial spoke protein
PRO
5
AMA
-5
Log2 fold change DEseq
Fiebig et al., 2015

Figure 2


A
Mean speed (µm/s)
Knock out
Add back
Parental
B
Mutants
Controls
Fast
LC4-like
Cas9 T7
WT
Uncoordinated
dDC2 Trypanin
MBO2 DRC2
RSP4/6 DRC4
FAP100
Slow
PFR2 FLAM2
DRC7 CMF44
CD047 TCTEX1A
FM458 p38
CC113 p28
FAP51 TTC29
FAP251 FAP73
Paralysed
IFT122B Hydin
IFT139 PF16
IFT88 FAP91
CFAP44 IC140
CFAP43
Cas9 T7 fixed
Directionality (velocity/speed)
C
Mutants
IFT mutants
Curling flagella
Disrupted PFR
D
Mutants
Normal Passage
Reduced Passage

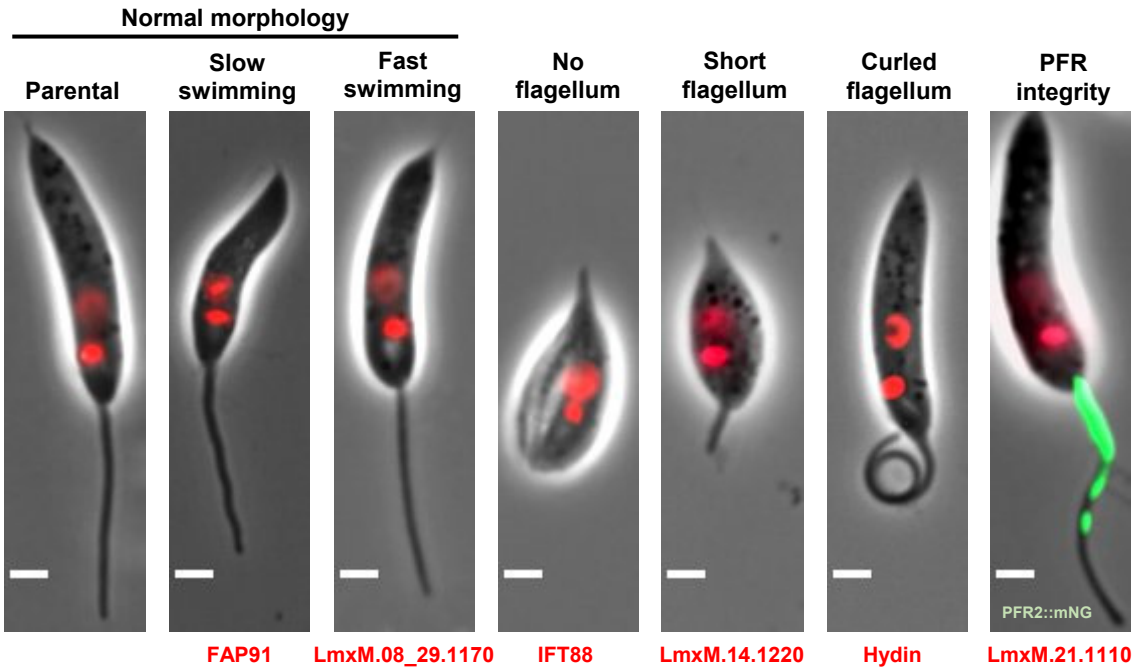
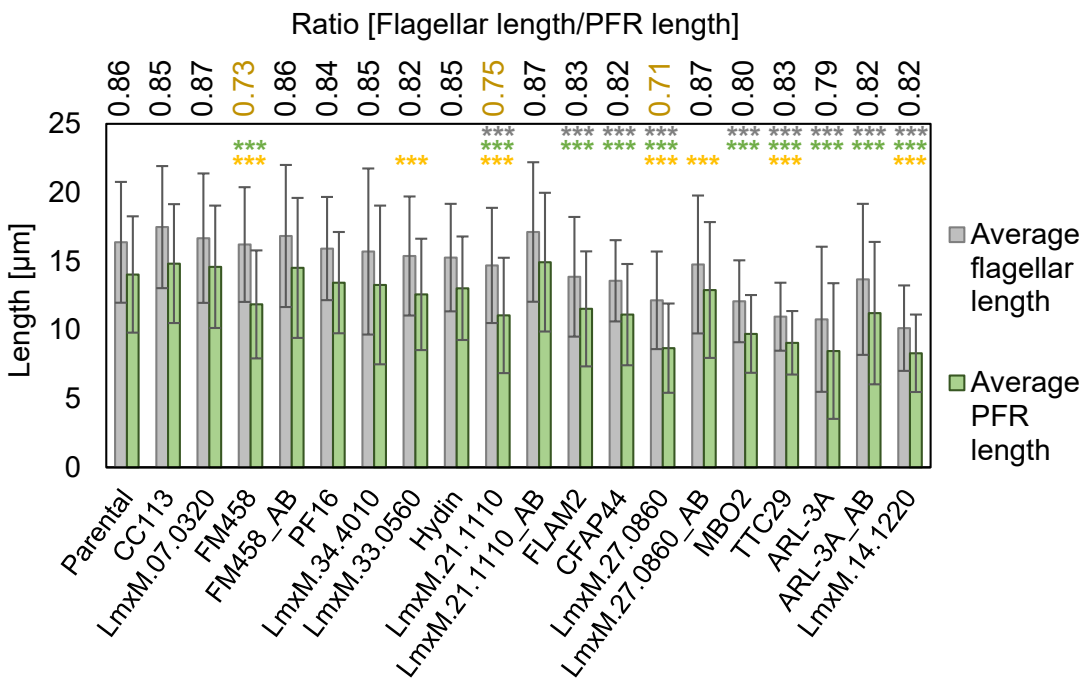

Normal morphology
Parental
Slow swimming
Fast swimming
No flagellum
Short flagellum
Curled flagellum
PFR integrity
PFR2::mNG
FAP91
LmxM.08_29.1170
IFT88
LmxM.14.1220
Hydin
LmxM.21.1110

Figure 4

A


Ratio [Flagellar length/PFR length]
0.86
0.85
0.87
0.73
0.86
0.84
0.85
0.82
0.85
0.75
0.87
0.83
0.82
0.71
0.87
0.80
0.83
0.79
0.82
0.82
Length [µm]
25
20
15
10
5
0
Average flagellar length
Average PFR length
Parental
CC113
LmxM.07.0320
FM458
FM458_AB
PF16
LmxM.34.4010
LmxM.33.0560
Hydin
LmxM.21.1110
LmxM.21.1110_AB
FLAM2
CFAP44
LmxM.27.0860
LmxM.27.0860_AB
MBO2
TTC29
ARL-3A
ARL-3A_AB
LmxM.14.1220

B

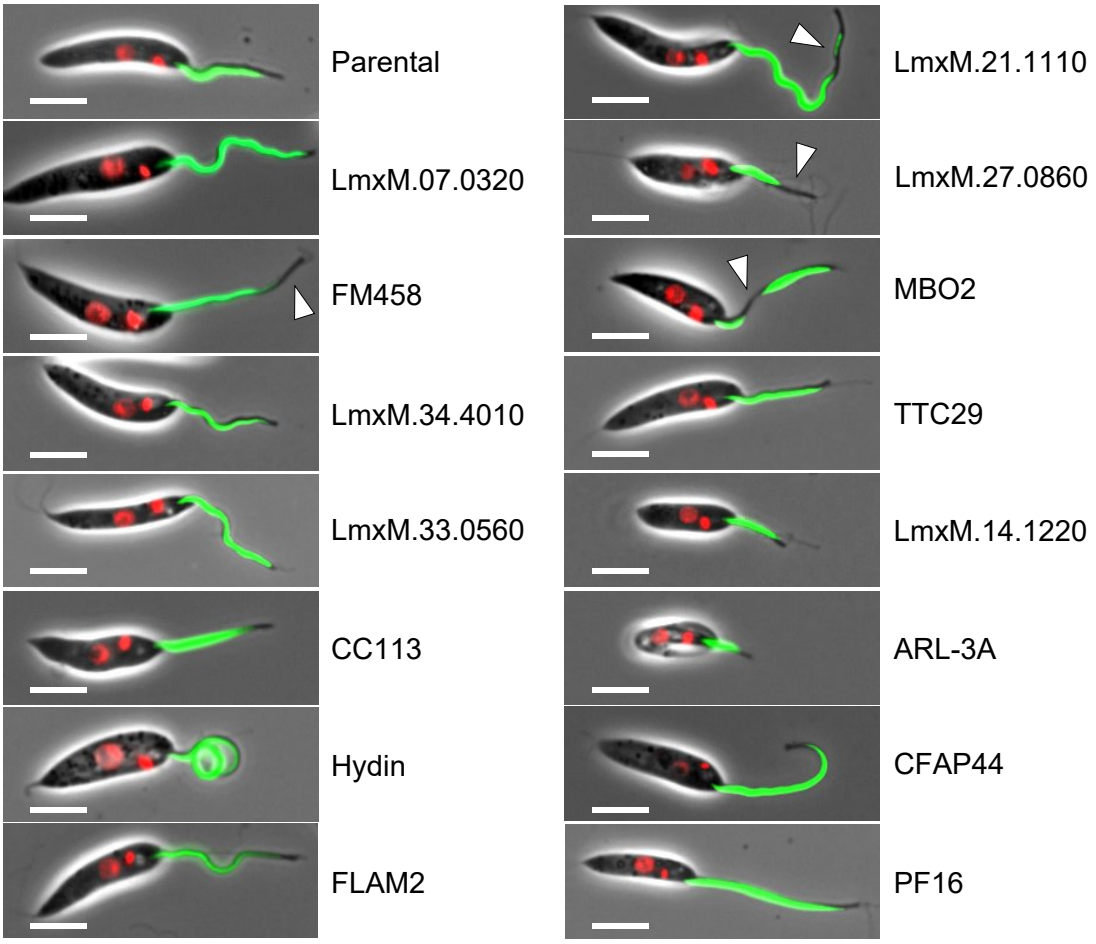

Parental
LmxM.07.0320
FM458
LmxM.34.4010
LmxM.33.0560
CC113
Hydin
FLAM2
LmxM.21.1110
LmxM.27.0860
MBO2
TTC29
LmxM.14.1220
ARL-3A
CFAP44
PF16

Figure 5

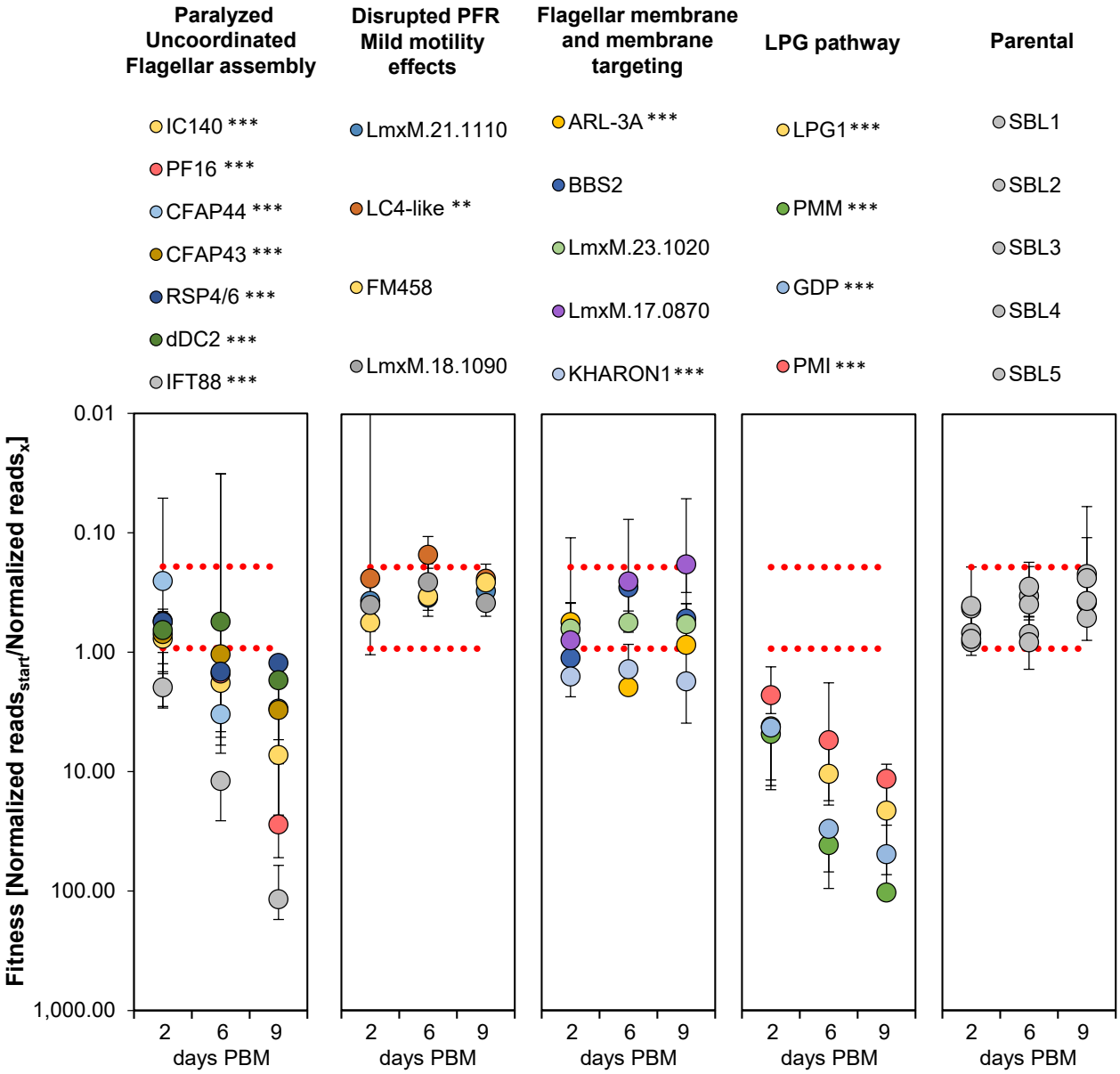

Paralyzed Uncoordinated Flagellar assembly
IC140 ***
PF16 ***
CFAP44 ***
CFAP43 ***
RSP4/6 ***
dDC2 ***
IFT88 ***
Disrupted PFR Mild motility effects
LmxM.21.1110
LC4-like **
FM458
LmxM.18.1090
Flagellar membrane and membrane targeting
ARL-3A ***
BBS2
LmxM.23.1020
LmxM.17.0870
KHARON1 ***
LPG pathway
LPG1 ***
PMM ***
GDP ***
PMI ***
Parental
SBL1
SBL2
SBL3
SBL4
SBL5
Fitness [Normalized reads$_{start}$/Normalized reads$_{x}$]
0.01
0.10
1.00
10.00
100.00
1,000.00
2
6
9
days PBM

Figure 6

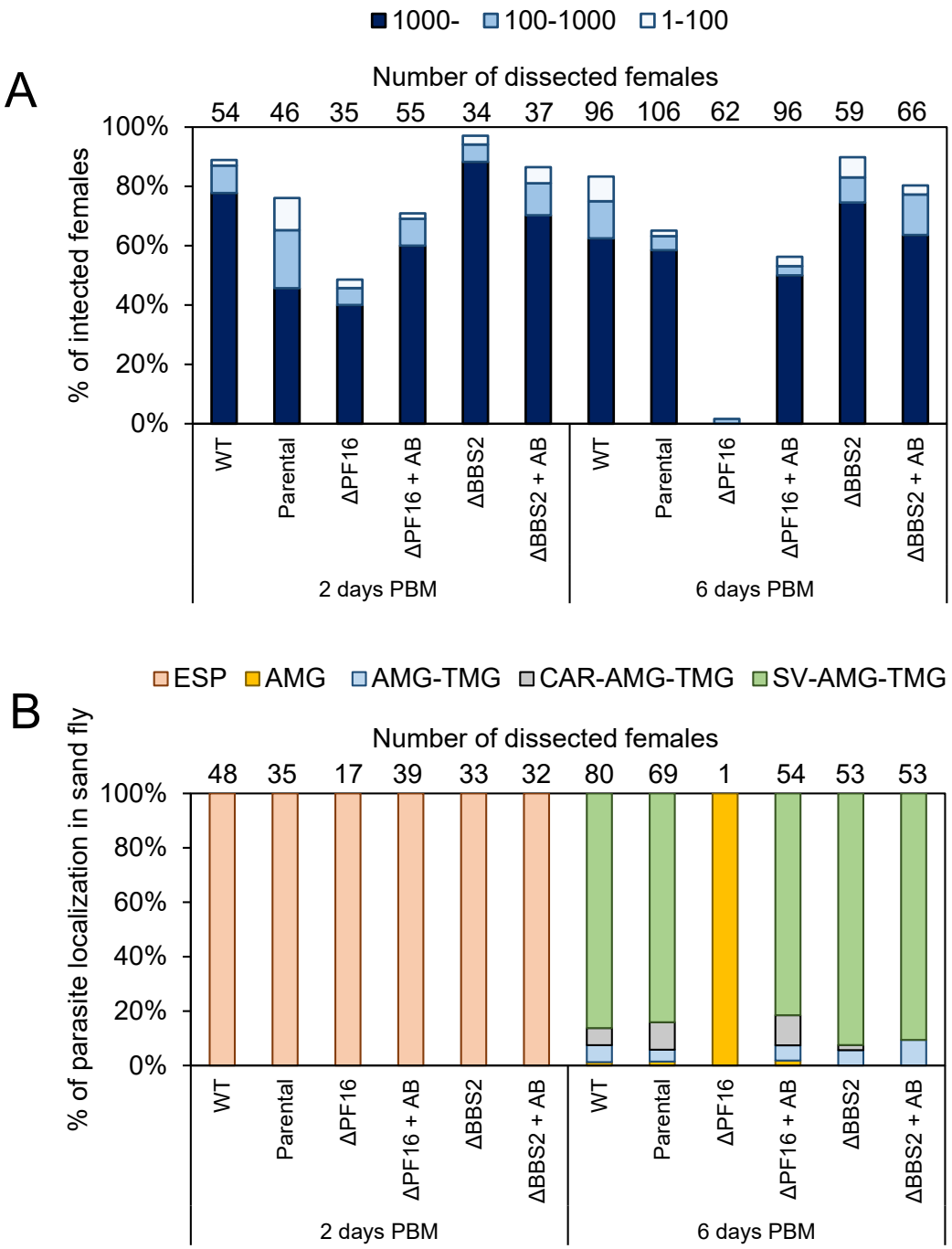

A
1000-
100-1000
1-100
Number of dissected females
54 46 35 55 34 37 96 106 62 96 59 66
% of intected females
100%
80%
60%
40%
20%
0%
WT
Parental
ΔPF16
ΔPF16 + AB
ΔBBS2
ΔBBS2 + AB
2 days PBM
6 days PBM
B
ESP
AMG
AMG-TMG
CAR-AMG-TMG
SV-AMG-TMG
Number of dissected females
48 35 17 39 33 32 80 69 1 54 53 53
% of parasite localization in sand fly

Figure 7