A novel DNA-binding protein regulates sexual differentiation in malaria parasites and is essential for transmission.

R. Campelo Morillo,[1] X. Tong, [1] W. Xie ,[2] T. Lenz,[3] G. Batugedara,[3] N. Tabassum,[1] L. M. Orchard,[4] W. Daher,[5] D. J. Patel,[2] W. S. Noble,[6] M. Llinás, [4,7] K. G. Le Roch,[3] B. F.C. Kafsack [1,*]

Affiliations
[1] Department of Microbiology & Immunology, Weill Cornell Medicine, New York, NY 10065, USA.
[2] Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA.
[3] Department of Molecular, Cell and Systems Biology, University of California Riverside, Riverside, CA 92521, USA.
[4] Department of Biochemistry & Molecular Biology and Huck Center for Malaria Research, Pennsylvania State University, University Park, PA, 16802.
[5] Dynamique des Interactions Membranaires Normales et Pathologiques, UMR5235 CNRS, INSERM, Université de Montpellier, Montpellier.
[6] Department of Genome Sciences, University of Washington, Seattle, WA, USA.
[7] Department of Chemistry, Pennsylvania State University, University Park, PA, 16802, USA.
* Please send correspondence to bkafsack@med.cornell.edu

ABSTRACT

Transmission of *Plasmodium falciparum* and other malaria parasites requires their differentiation from asexual blood stages into gametocytes, the non-replicative sexual stage necessary for transmission to the mosquito vector. This transition involves changes in gene expression and chromatin reorganization mediating the activation and silencing of stage-specific genes. However, the genomes of malaria parasites have been noted for their dearth of transcriptional and chromatin regulators and the molecular mediators of these changes remain largely unknown. We recently identified HomeoDomain Protein 1 (HDP1) as a novel chromatin-associated DNA-binding protein that drives changes in chromatin structure and gene expression during early sexual differentiation. The discovery of a homeodomain-like DNA-binding protein marks a new class of transcriptional regulator in malaria parasites outside of the better-characterized ApiAP2 family. In this study, we demonstrate that HDP1 is required for gametocyte maturation and parasite transmission by facilitating the necessary upregulation of inner membrane complex components in early gametocytes.

INTRODUCTION

In order to complete its life cycle, *Plasmodium falciparum*, the most widespread and virulent of the human malaria parasites, must repeatedly differentiate into unique cell types that are able to access and exploit specialized niches within their human and mosquito hosts. One of these key developmental transitions occurs during the parasite's blood stage. Asexual blood-stages maintain a persistent infection through continuous lytic replication within erythrocytes but are not infectious to the mosquito vector, and therefore cannot mediate transmission to the next human host. This requires their differentiation into non-replicating, male and female gametocytes that are able to infect the mosquito once taken up during a blood meal. All differentiation requires the repression and activation of genes that underlie the specific phenotypes of the origin and destination cell types, respectively. To ensure complete commitment to one cell type or another, these transitions are often involve a bistable switch that controls the activity of a single master regulator at the top of the transcriptional cascade that underlies the differentiation program (*1-4*). Once activated, the master regulator then initiates the broader downstream changes in gene expression by altering the expression of additional transcriptional regulators and changing their access to cell type-specific promoters via chromatin re-organization.

Recent work has found that this paradigm also applies in malaria parasites, where the initiation of sexual differentiation is controlled by bistable expression of a master regulator, the transcription factor AP2-G (*5*, *6*). During asexual replication the *ap2-g* locus is silenced by heterochromatin (*5*, *7*, *8*) but, due to the presence of AP2-G binding sites within its own promoter region, incomplete repression of *ap2-g* in individual cells can trigger a transcriptional feedback loop that drives AP2-G expression to high levels, thereby locking cells into activating the sexual differentiation gene expression program (*5*, *9*, *10*). Under conditions that impair heterochromatin maintenance, this feedback loop is activated in a larger fraction of cells, thereby increasing the frequency of sexual differentiation (*8*, *11*, *12*).

Commitment to the sexual differentiation program also triggers a substantial re-distribution of heterochromatin during the early stages of gametocyte development (*13*, *14*). These changes involve substantial expansion of sub-telomeric heterochromatin domains on several chromosomes to silence asexual blood-stage specific genes. One such example is a cluster of genes on chromosome 2 that are required for trafficking of PfEMP1, a parasite-encoded adhesin on the erythrocyte surface that is critical for immune evasion in asexually-replicating parasites but not expressed in gametocytes (*15*). Conversely, heterochromatin is lost from several loci that

were previously silenced in asexual stages but upregulated during early gametocyte development (*13*, *14*). While AP2-G is critical for the initiation of sexual differentiation, it is expressed only during a small window that begins with sexually committed schizonts and ends after the first 48 hours of gametocyte development (*10*, *16*). This means that the many of changes in gene expression during the 12-14 days process of gametocyte development are regulated by additional waves of transcriptional regulators and is consistent with the our previous observations that AP2-G upregulates a number of putative transcription factors and chromatin remodeling enzymes (*9*, *10*).

While most species of malaria parasites form spherical gametocytes that mature in 2-6 days, sexual differentiation in *P. falciparum* takes substantially longer, 12-14 days, and produces gametocytes with the characteristic falciform morphology that give the parasite its name. Unsurprisingly, this prolonged gametocyte maturation is accompanied by a wide array of gene expression changes (*17-20*). However, relatively little is known about the transcriptional regulators downstream of AP2-G that mediate these changes. Compared to other single-celled eukaryotes, DNA-binding proteins are notable underrepresented in the genomes of malaria parasites (*21*). Like AP2-G, most belong to the ApiAP2 family (*21*), but only a small number have been shown to function specifically during gametocyte development of *P. falciparum* (*22*). In the rodent malaria parasite *P. berghei*, PbAP2-G2 functions as a transcriptional repressor of asexual-specific gene expression (*6*, *23*), while PbAP2-FG (PyAP2-G3 in *P. yoelii*) was shown to mediate upregulation of female-specific transcripts (*24*, *25*) but whether these functions are conserved in *P. falciparum* remains to be determined.

In this study, we identify HDP1, a previously uncharacterized DNA-binding protein that is first expressed during early sexual differentiation. Parasites lacking HDP1 are unable to upregulate genes required for the expansion of the inner membrane complex and fail to progress beyond the day 3 of gametocyte development. Additionally, loss of HDP1 expression results in an increase in chromatin accessibility, reduced telomere clustering, and leaky transcription of heterochromatin-associated genes in early gametocytes. Our findings indicate that HDP1 is an essential regulator of gene expression and chromatin during early gametocytogenesis.

102 **RESULTS**

103 ***hdp1* is essential for gametocyte development.**

104 Searching for possible regulators of gene expression during *P. falciparum* sexual differentiation,
105 we identified HomeoDomain-like Protein 1 *(hdp1, PF3D7_1466200)* encoding a 3078 amino acid
106 protein with a C-terminal homeodomain-like domain, containing a helix-turn-helix structural
107 motif commonly involved in DNA binding (fig. S1A) (*26*). Syntenic orthologs of *hdp1* could be
108 readily identified in other malaria parasites with homology to the homeodomain-like domain
109 extending to the coccidia but apparently absent from other apicomplexan clades (fig. S1B).
110 Analysis of *hdp1* expression in *P. falciparum* blood stages by qRT-PCR detected only minimal
111 expression in cultures of asexual blood stages (which always contain smalls number of
112 gametocytes) with substantial upregulation during the early stages of gametocytogenesis (fig.
113 S1C). AP2-G, the transcriptional master switch that controls the initiation of the sexual
114 differentiation gene expression program (*5*, *9*, *10*, *27*), binds at two sites located upstream of *hdp1*
115 in early gametocytes (*10*), consistent with our hypothesis that AP2-G activates additional
116 regulators of gene expression during early gametocytogenesis.

117

118 In an effort to determine HDP1's subcellular localization, we inserted an in-frame N-terminal
119 HaloTag at the endogenous *hdp1* locus (fig. S2A) in order to avoid possible interference with the
120 predicted DNA-binding domain (DBD) located at the very C-terminus. As expected based on
121 transcript abundance, no Halo-tagged protein was detected in asexual stages. However, when we
122 attempted to determine Halo-HDP1 levels in the sexual stages, we found that *halo-hdp1* cultures
123 were unable to produce the characteristic crescent shapes of maturing *P. falciparum* gametocytes
124 (Fig. 1A-B). Subsequent tagging at the HDP1 C-terminus with either GFP or a triple Ty1 epitope
125 tag (*hdp1-gfp* and *hdp1-Ty1*, fig. S2B-C), yielded parasite lines that produce gametocytes
126 indistinguishable in numbers and morphology from the wildtype parent (Fig. 1A-B), despite the
127 proximity to the putative DNA-binding domain. To test whether N-terminal tagging resulted in a
128 loss of HDP1 function, we generated a Δ*hdp1* knockout line for comparison by replacing 1.4 kb
129 at the 5' end of the *hdp1* locus with a selectable marker cassette (fig. S2D). The resulting Δ*hdp1*
130 parasites exhibited no discernible change in phenotype in asexual blood-stages, but like the *halo-*
131 *hdp1* parasites, were unable to form viable mature gametocytes (Fig. 1D-F). More detailed
132 analyses using synchronous induction of gametocytogenesis found that both *halo-hdp1* and
133 Δ*hdp1* have sexual commitment rates comparable to the NF54 parent (fig. S3) but are unable to
134 complete gametocyte development (Fig. 1A-C).

135

136 As the substantial length of the *hdp1* coding sequence made genetic complementation infeasible,
137 we generated inducible HDP1 knockdown parasites by inserting a triple Ty1 epitope tag followed
138 by the autocatalytic *glmS* ribozyme at the 3' end of the endogenous *hdp1* coding sequence through
139 genome editing (*hdp1-glmS,* fig. S2E) (*28*). In the absence of glucosamine, the resulting *hdp1-*
140 *glmS* gametocytes expressed HDP1 protein at levels comparable to *hdp1-Ty1* parasites lacking
141 the ribozyme and produced wild-type gametocytes in terms of number and morphology (fig. S4).
142 Supplementation of the culture medium with 5 mM glucosamine during the first 5 days of
143 gametocyte development had no discernible effect on *hdp1-Ty1* parasites but diminished HDP1
144 expression by 70% and reduced the number of falciform gametocytes by 80% in *hdp1-glmS*
145 parasites, recapitulating the phenotype of the *halo-hdp1* and Δ*hdp1* lines (Fig. 1G-H). Since *hdp1*
146 transcript levels remain relatively constant during gametocyte development (fig. S1C), we wanted
147 to test whether *hdp1* transcripts were required throughout gametocyte maturation. Knockdown of
148 *hdp1* levels with glucosamine reduced gametocyte maturation prior to day 5 but not thereafter,
149 indicating that sufficient HDP1 protein had been produced by that time to support gametocyte
150 maturation (fig. S4).
151
152 **HDP1 is a chromatin-associated nuclear protein in gametocytes that recognizes a GC-rich**
153 **DNA motif as a dimer.** An earlier study had reported a variety of subcellular localizations based
154 on antibodies raised against a low-complexity region of HDP1 (fig. S1A), including export to the
155 erythrocyte membrane of early gametocytes (*29*). This was surprising given the absence of a
156 signal peptide, the presence of two predicted nuclear localization signals (fig. S1A), and
157 localization of the *Toxoplasma gondii* ortholog to the nucleus of tachyzoites (fig. S5). To resolve
158 this apparent disagreement, we carried out live-cell and immunofluorescence microscopy of
159 *hdp1-gfp* and *hdp1-Ty1* gametocytes, respectively. In both lines, HDP1 localized exclusively to
160 the gametocyte nucleus (Fig. 2A-B) and we were unable to replicate the localization(s) described
161 in the earlier study. Unfortunately, according to the corresponding author the antisera used in that
162 study are no longer available, precluding a direct comparison. HDP1 protein was undetectable in
163 asexual blood stages by both microscopy and western blotting (fig. S6, Fig. 2C) but showed
164 increasing expression from day 2 of gametocytogenesis (Stage I-II) onward, reaching maximal
165 levels by day 5 (Stage III) before dropping again by day 8 (Stage IV) (Fig. 2C). Analysis of
166 subcellular fractions from day 5 *hdp1-Ty1* gametocytes found HDP1 almost exclusively in the
167 nuclear fraction, with about 70% resistant to solubilization by up to 600 mM NaCl, indicating a
168 tight association with chromatin (Fig. 2D) and validating HDP1 as a nuclear protein.

169 Since HDP1 is an integral component of chromatin and homeo-like domains typically mediate
170 protein-DNA interactions, we evaluated HDP1's ability to bind DNA using a protein-binding
171 microarray (PBM). Binding of recombinant HDP1 DBD was highly enriched on probes
172 containing the palindromic hexamer GTGCAC (Fig. 2E, fig. S7A-C). Since homeo-domains
173 generally recognize DNA as dimers (*26*, *30*), we carried out isothermal titration calorimetry to
174 measure the interaction of HDP1 with double-stranded DNA containing a tandem motif with a 5
175 bp spacer that places the motifs one helical turn apart. The results found that binding saturated at
176 a 2:1 protein to DNA molar ratio with a dissociation constant of 2.8 μM, indicating that DNA
177 recognition by the HDP1 DBD also occurs as a dimer (Fig. 2F). We subsequently confirmed
178 dimeric binding using DNA gel-shift assays with double-stranded DNA probes containing either
179 no match, a single binding-motif, or a tandem motif. When compared to the tandem-motif probe,
180 the gel-shift of the single motif probe was substantially weaker but identical in size (Fig. 2G),
181 again consistent with DNA recognition as a dimer even when only a single motif is present.
182 Intriguingly, a scan of the 3D7 reference genome for instances of GTGCAC tandem motifs with
183 a 5-6 bp spacer found them exclusively located at each chromosome end with the first tandem
184 located within either 7 or 28 bp of the ultimate *rep20* sub-telomeric repeats (*31*), with the
185 exception of those ends where this region was deleted as the result of telomere healing events
186 (*32*) (fig. S7D, Data Set 1A). Earlier work had predicted these conserved sequences to be subset
187 of motifs recognized by the ApiAP2-domain protein SIP2 in asexual blood stages based on its
188 binding preference *in vitro* (*33*), however, their unique spacing and palindromic nature clearly
189 distinguishes these symmetrical <u>GTGCACC</u>nnn<u>GGTGCAC</u> tandem motifs from the arrays of
190 asymetrical <u>GTGCA</u>nnnn<u>GGTGCG</u> SIP2-binding sites further downstream. While the functional
191 role of these motifs in asexual and sexual blood stages remains unclear, their exclusive
192 localization at chromosome ends is highly conserved across *P. falciparum* isolates of diverse
193 geographic origins (Data Set 1B).
194

195 **Loss of HDP1 leads to dysregulation of gene expression in early gametocytes.**

196 Based on HDP1's ability to bind DNA, we wanted to test whether it plays a role in the regulation
197 of gene expression during early gametocytogenesis. In order to identify changes in expression
198 that may account for the aberrant development of Δ*hdp1* gametocytes rather than looking at the
199 consequences of the loss in viability that occurs in Δhdp1 gametocytes, we decided to analyze the
200 transcriptome of Δ*hdp1* and parental NF54 on Day 2, when HDP1 is first detectable in wild-type
201 gametocytes but before any change in viability or morphology occurs in Δhdp1 gametocytes. (Fig.
202 1D, fig. S3B). Global comparison of transcript abundances found that the majority of genes were

203 expressed at similar levels, including canonical markers of very early gametocytes such as *pfs16*
204 and *gexp5* (Fig. 3A). This confirms our earlier observation that HDP1-deficient parasites initiate
205 gametocyte development at similar rates and that changes in gene expression are not due to
206 change in viability at that point. However, we observed reduced transcript levels for 188 genes
207 and increased levels for 150 genes in Δ*hdp1* Stage I gametocytes compared to those of the NF54
208 parent line. Gene set enrichment analysis (GSEA) found that transcripts encoding components of
209 the inner membrane complex were significantly down-regulated (Fig. 3B) while transcripts from
210 heterochromatin-associated multi-copy gene families were significantly over-represented among
211 up-regulated genes (Fig. 3C), including members of the *var*, *rifin*, *stevor*, and *PHISTa/b/c* gene
212 families (Fig. 3D, fig. S8A-B). While fold-changes for these genes could be substantial, the
213 absolute increase in transcripts for these genes was generally small but nevertheless significantly
214 above their levels in wild-type cells (Data Set 2), suggesting a possible impairment of that
215 heterochromatin-mediated silencing leading to leaky background transcription but not full
216 transcriptional activation. Several of these heterochromatin-regulated families exhibit
217 transcriptional variation due to expression switching between members but this does not lead to
218 a broad upregulation across all family members as observed here. Comparison of transcript levels
219 of *var* and *rifin* family genes found similar levels of upregulation regardless of whether these
220 were located in subtelomeric or non-subtelomeric regions of heterochromatin (fig. S8C). Analysis
221 of *var* gene expression in asexual Δ*hdp1* ring-stages found expression of a single major *var* gene
222 as expected for a recently cloned parasite line and indicating that mutually exclusive *var* gene
223 expression remained unaffected in asexual blood stages (fig. S8D).
224

225 **HDP1 is required for extension of the inner membrane complex in early gametocytes.**

226 Among the genes with reduced transcript levels in HDP1 knockout parasites, those associated
227 with the inner membrane complex were substantially enriched (Fig. 3A-B, Fig. 4A). This was
228 notable since Δ*hdp1* gametocytes fail to elongate from spherical stage I gametocytes into the
229 oblong Stage II, a process that depends on the extension of the inner membrane complex. Six
230 separate inner membrane complex genes were confirmed by qRT-PCR to have greater than 2-
231 fold lower transcript levels in HDP1 knockout gametocytes (Fig. 4B). These included PhIL1 and
232 the PhIL1-interacting protein 1 (PIP1), both of which are highly upregulated in early *P.*
233 *falciparum* gametocytes and essential for assembly of the inner membrane complex and
234 gametocyte maturation (*34*). Strikingly, the spherical morphology of *hdp1*-deficient gametocytes
235 perfectly resembles the phenotype of gametocytes of PhIL1 knockdown parasites. Comparison of
236 PhIL1 protein levels showed a 75% percent reduction in Δ*hdp1* gametocytes compared levels in

237 the wild-type parent while other proteins remained unchanged (Fig. 4C). This concomitant
238 reduction in PhIL1 expression (Fig. 4D) and inner membrane complex extension (Fig. 4E) when
239 HDP1 is knocked down gametocytes confirmed the dependence of PhIL1 expression and inner
240 membrane complex extension on HDP1.
241
242 **Loss of HDP1 increases chromatin accessibility and decreases telomeric clustering.**
243 Since our experiments clearly showed that HDP1 can bind DNA and is tightly associated with
244 nuclear chromatin (Fig. 2), we set out to determine where in the genome HDP1 is binding by
245 chromatin immunoprecipitations sequencing. However, numerous attempts at HDP1 ChIP-seq
246 with both the GFP- and Ty1-tagged lines were unsuccessful in multiple labs (see Materials and
247 Methods for details). As an alternative approach to detect altercation in chromatin organization,
248 we used high-throughput chromatin-conformation capture (Hi-C) to look for differences in
249 chromatin organization of day 2 gametocytes between the NF54 and Δ*hdp1* lines. Correlation
250 analysis of Hi-C replicates indicated a clear difference between these lines (Fig. 5A), genome-
251 wide intra-chromosomal interactions appeared largely unchanged (fig. S9). However, further
252 statistical analysis revealed an overall reduction in long distance interactions in Δ*hdp1* (Fig. 5B).
253 Genome-wide mapping of interaction changes revealed a substantial reduction in interaction
254 frequency between nearly all chromosomes ends (Fig. 5C), consistent with reduced but not
255 abrogated telomere clustering. When combined with our observation of leaky transcription of
256 heterochromatin associated genes located at these sites, this reduction in telomeric clustering led
257 us to ask whether chromatin compaction is impaired in HDP1-deficient gametocytes, leading to
258 an increase in chromatin accessibility. Indeed, brief treatment with micrococcal nuclease (MNase)
259 revealed accelerated digestion of chromatin from Δ*hdp1* gametocytes (Fig. 5D), consistent with
260 an increase in chromatin accessibility.
261
262 **DISCUSSION**
263 Our understanding of the regulatory machinery that underlies sexual differentiation in *P.*
264 *falciparum* has improved substantially in recent years. Much of this work has focused on the
265 regulation of AP2-G, the master switch of this developmental decision. However, it is becoming
266 clear that the role of AP2-G is constrained to the initiation of the transcriptional program that
267 drives the nearly two week-long process of gametocyte development (*16*). This suggests that a
268 second wave of hitherto unknown transcriptional regulators is required to drive gametocyte-
269 specific gene expression during early gametocyte development.
270

271 In this study, we showed that HDP1 is a nuclear DNA-binding protein that functions as regulator
272 of gene expression in early gametocytes and is essential for their development. In particular,
273 HDP1 regulates the expression of genes, like *phil1* and *pip1/2*, that drive the expansion of the
274 inner membrane complex to give *P. falciparum* gametocytes their characteristic crescent shape.
275 While its expression is gametocyte-specific in *P. falciparum,* the *P. berghei* ortholog
276 (PBANKA_1329600) may also have a role during asexual blood-stage replication as its
277 disruption leads to significant reduction in growth (*35*). This suggests that the gametocyte-specific
278 function of HDP1 evolved more recently in the Laveranian clade of malaria parasites which
279 uniquely produce falciform gametocytes. While essential during early gametocyte development,
280 whether HDP1 is required in other parasite stages remains to be determined. Since the knockdown
281 system used in this study regulates expression at the transcript level, we can infer that additional
282 translation of HDP1 is not required during the later stages of gametocyte maturation but HDP1
283 protein expressed during the earlier stages of gametocytogenesis may still be required during the
284 later stages. Previous transcriptomic studies indicate that HDP1 is also expressed in ookinetes,
285 suggesting a role even after mating in the mosquito midgut. Intriguingly, homeodomain-like
286 proteins have been implicated in mating processes of other haploid protozoa, such as
287 Dictyostelium (*36*).
288
289 Our experiments clearly show that HDP1 has the ability to bind DNA and is an integral
290 component of chromatin in gametocytes (Fig. 2), but chromatin immuno-precipitation remained
291 unsuccessful, despite exhaustive attempts (see Materials and Methods for details). One likely
292 explanation is that the C-terminal tags are not accessible to antibodies when HDP1 is component
293 of chromatin. This explains why chromatin immunoprecipitation of HDP1 with has been
294 unsuccessful, as well as the diffuse nuclear localization we observed (Fig. 2A-B). Unfortunately,
295 N-terminal tagging of HDP1 is not a viable alternative as it results in loss of function (Fig. 1).
296 Another possibility is that HDP1 resides in a fraction of chromatin that is especially difficult to
297 solubilize, which has been shown to lead to incomplete recovery of chromatin under standard
298 ChIP conditions (*37*).
299
300 Parasites lacking HDP1 exhibit defects in efficient heterochromatin-mediated gene silencing
301 during early sexual differentiation that manifests as a low but significant up-regulation across a
302 wide array of heterochromatin-associated gene families (Fig. 3), reduced clustering of the sub-
303 telomeric heterochromatin regions (Fig. 5C) in the nuclear periphery, and greater chromatin
304 accessibility (Fig. 5D). One explanation is that HDP1 promotes chromatin packaging in early

305 gametocytes, leading to the overall increase in chromatin accessibility observed in HDP1-
306 deficient cells. This would also explain the leaky expression of heterochromatin-associated genes
307 as efficient silencing depends on limiting promoter accessibility while the promoters of
308 euchromatin-associated genes are already generally accessible and therefore a reduction in
309 packaging would therefore be predicted to have little impact on their transcript levels.
310
311 Additional studies will focus on identifying interaction partners of HDP1 and elucidating its
312 structure-function relationship. Only 2% of its protein sequence is comprised of the
313 homeodomain-like DNA-binding domain while the remainder contains no other identifiable
314 domains and conservation is generally weak outside this region. The fact that insertion of a
315 large tag at the N-terminus results in loss of function indicates that critical interactions occur in
316 this region, but which other regions are essential for HDP1's function remains unclear.
317 Similarly, identifying interactions with other nuclear proteins and genomic locations will offer
318 important insights into its function. Of particular interest will be whether HDP1 regulates gene
319 expression by altering chromatin structure or is more directly involved in the recruitment of
320 RNA polymerase II.

321 **MATERIALS & METHODS**

322

323 **Parasite culture**

324 Unless otherwise noted, *P. falciparum* parasites were grown in 0.5% AlbuMAX II supplemented
325 malaria complete medium using stablished cell culture techniques (*38*) at 3% hematocrit and
326 below 3% parasitemia. Strains expressing selectable markers were maintained under constant
327 drug-selection. Toxoplasma tachyzoites were cultured as described in (*39*).

328

329 **Gametocyte induction and isolation**

330 Gametocytes were induced synchronously as previously described in (*9*). Gametocyte maturation
331 was monitored by Giemsa-stained thin blood smears and gametocytemia was counted on the fifth
332 day of development. The sexual commitment rate was calculated by dividing the day 5
333 gametocytemia by the day 1 parasitemia, counted before addition of N-acetyl-D-glucosamine.
334 For knockdown experiments in the HDP1-glmS line, at the gametoring stage either 5 mM
335 glucosamine or solvent control was added. Gametocytes were purified from culture at the required
336 development stage using magnetic columns (LS columns, Miltenyi Biotec).

337

338 **Generation of transgenic strains**

339 Transfection of ring-stage parasites were performed as previously described (*40*). Genome
340 editing was performed by CRISR/Cas9 technology using the system described by (*41*). The pL6-
341 DX2 plasmid flanking homology region were PCR amplified using Advantage Genomic LA
342 polymerase (Takara) from NF54 genomic DNA (for list of primers, see Supplementary Table)
343 and cloned into the *AflII* and *SpeI* sites by Gibson assembly. *Plasmodium* codon optimized
344 sequences for HALO-tag and triple Ty1 epitope tag were synthetized as gene-Blocks (Genewiz).
345 The absence of undesired mutations in the homology regions and the sgRNA was confirmed by
346 Sanger sequencing. Genomic DNA from transfectant parasites was isolated with QIAamp DNA
347 blood Kit (Qiagen, Cat. N° 51106) and diagnostic PCR's were set using Taq phusion DNA
348 polymerase (Invitrogen). The TGME49_233160-HA parasite line was generated as part of an
349 earlier study by tagging of the endogenous locus in the *T. gondii* RH-ku80ko strain as described
350 in (*39*).

351

352 **Flow cytometric analysis of gametocyte viability**

353 Gametocytes were stained with 16 µM Hoechst33342 and 50 nM DilC1 for 30 minutes at 37˚C.
354 Using a Cytek DxP12 flow cytometer, gametocytemia was determined by gating for DNA-

355 positive cells and gametocyte viability was inferred based on mitochondrial membrane potential
356 dependent accumulation of DilC1(5) for 1000 gametocytes (*42*). Analysis was carried out using
357 FlowJo 10. The gating strategy is shown in fig. S10.
358

359 **Nuclear extract preparation and chromatin high salt fractionation**

360 Nuclear isolation and extraction was carried out based on (*43*), with minor modifications. Briefly,
361 parasites released from RBC's by saponin treatment (0.01%) were lysed with ice-chilled CLB (20
362 mM HEPES, pH 7.9; 10 mM KCl; 1mM EDTA, pH 8.0; 1 mM EGTA, pH 8.0; 0.65% NP-40; 1
363 mM DTT, 1x Roche Complete protease inhibitors cocktail). Nuclei were pelleted at 3,000 x *g* for
364 20min at 4ºC and cytoplasmic fraction was removed. Nuclei were resuspended in digestion buffer
365 (20 mM Tris- HCl, pH 7.5, 15 mM NaCl, 60 mM KCl, 1 mM $CaCl_2$, 5 mM $MgCl_2$, 300 mM
366 sucrose, 0.4% NP-40, 1 mM DTT, 1x Roche Complete protease inhibitors cocktail EDTA-free)
367 and treated with 5U of micrococcal nuclease (ThermoFisher, Cat. Nº 88216) for 30 min in a water
368 bath at 37ºC. Soluble and insoluble nuclear fractions were recovered by centrifugation at 3,000 x
369 *g* for 10 min at 4º C. Insoluble nuclear fraction were treated with salt fractionation buffer (10 mM
370 Tris-HCl, pH 7.4; 2 mM $MgCl_2$; 2 mM EGTA pH 8.0; 0.1% Triton X-100; 0.1 mM PMSF; 1x
371 Roche Complete protease inhibitors cocktail) supplemented with increasing NaCl concentrations
372 (80-600 mM) while rotating at 4ºC for 30 min. All supernatants were recovered by centrifugation
373 at 700 x *g* for 4 min at 4ºC and last remaining pellet was resuspended in 1x PBS supplemented
374 with protease inhibitors cocktail. 5% of each fraction was prepared for Western blotting in order
375 to check quality of fractionation.
376

377 **Micrococcal nuclease digestion**

378 Day 2 gametocytes were isolated magnetically and crosslinked with 1% formaldehyde then lysed
379 with hypotonic lysis buffer (10 mM Tris-HCl, pH 8.0; 3 mM $MgCl_2$; 0.2% NP-40; 1 mM DTT
380 and 1x Roche Complete protease inhibitors cocktail). Mechanical lysis was enhanced by passing
381 the suspension ten times through a 26 G 1/2 needle. Nuclei were pelleted at 3,000 x *g* for 20 min
382 at 4ºC and cytoplasmic fraction was removed. Nuclei were resuspended in MNase buffer (50 mM
383 Tris-HCl, pH 8.0; 5 mM $CaCl_2$; 0.4% NP-40) and 1U of MNase (ThermoFisher) was added.
384 Digestion was carried for 5 min at 37ºC and was stopped by adding 20mM EGTA pH 8.0. Soluble
385 and insoluble fractions were recovered by centrifugation at 3,000 x *g* for 10 min at 4ºC. Samples
386 were decrosslinked and deproteinized by adding 0.1% SDS and 100ug/ml of Proteinase K
387 (Roche,) and incubated for 1 hour at 50ºC. DNA was isolated through standard phenol-
388 chloroform-isoamyl alcohol extraction.

389 **Immunoblotting**

390 For SDS-PAGE, total protein lysates were prepared using saponin-lysed parasites resuspended
391 with 1x Laemmli loading buffer diluted in 1x PBS supplemented with 1x Roche Complete
392 protease inhibitors cocktail. Protein samples were separated in 4-15% polyacrylamide gels and
393 transferred to 0.2 µm Immobilion-P^{SQ} transfer membrane (Millipore, Cat. No ISEQ00010)
394 using a Bio-Rad transfer system. Membranes were blocked in 5% skim milk/1x TBS-Tween20
395 for 1 hour at RT. Primary and secondary antibodies were prepared in 3% skim milk/1x TBS-
396 Tween20 and incubated for 1 hour at RT. Membranes were washed four times with 1x TBS-
397 Tween20 for 10 min, after primary and secondary antibody incubations. The following primary
398 antibodies were used in this study: Anti-Ty1 BB2 mouse (1:2,500; Invitrogen Cat. N^{o} MA5-
399 23513), anti-PhIL1 rabbit (1:5,000 (*44*)), anti-PfHsp70 rabbit (1:5,000; StreesMarq Biosciences
400 Cat. N^{o} SPC-186D), anti-Histone 4 rabbit (1:2,000; Diagenode Cat. N^{o} C15410156-50). HRP-
401 conjugated anti-mouse and anti-rabbit antibodies were used (1:5,000, Millipore). Immunoblots
402 were incubated with the chemiluminescent substrate SuperSignal West Pico PLUS
403 (ThermoFisher, Cat. N° 34578) following manufacturer directions. Chemiluminescent images
404 were obtained using an Azure c300 digital imaging system (Azure Biosystems).

405

406 **Live-cell and Immunofluorescence microscopy**

407 For live-cell microscopy of *hdp1-gfp* and NF54 blood stages, infected red blood cells were stained
408 with 16 µM Hoechst33342 in incomplete media for 15 min at 37° C and imaged with identical
409 exposure settings at 1000× magnification using a Leica DMI6000 microscope with differential
410 interference contrast bright field optics, DAPI, and GFP filter cubes. For immunofluorescence
411 microscopy of *hdp1-Ty1* and *hdp1-glmS* gametocytes, cells were immobilized on glass slides with
412 Concanavalin A (5 mg/ml; Sigma) as described in (*45*), then fixed with a solution of 4%
413 paraformaldehyde/0.0075% glutaraldehyde for 20 min at 37° C. Parasites were permeabilized
414 with 0.1% Triton X-100 for 15 min at RT followed by blocking with 3% BSA. Primary antibodies
415 (anti-Ty1 BB2 mouse 1:1,000; anti-PhiL1 rabbit 1:400) were allowed to bind for 1 hour in 3%
416 BSA/PBS followed by three washes with blocking buffer for 5 min. Secondary antibodies were
417 diluted at 1:500 (anti-mouse-Alexa546 and anti-rabbit-Alexa488, Invitrogen) in fresh blocking
418 buffer containing 16 µM Hoechst 33342 and incubated for 1 hour. Z-stacks of stained specimens
419 were collected at 1000× magnification using a Leica DMI6000 microscope with differential
420 interference contrast bright field optics, DAPI, and RFP filter cubes with identical exposure times.
421 Fluorescent channel z-stacks were deconvolved using the ImageJ DeconvolutionLab2 plugin
422 (NLLS algorithm) followed by maximum intensity z-projection and background adjustment.

423 Immunofluorescence microscopy of Toxoplasma tachyzoites was carried out as previously
424 described (*39*).
425
426 **Protein Expression and Purification**
427 Expression of recombinant HDP1-DBD motif was done using the Glutathione S-transferase
428 (GST) gene fusion system (GE Healthcare). The pGEX-4T-1 plasmid was used as backbone for
429 cloning the codon optimized sequence comprising the last 87aa of HDP1. Plasmid pGEX-GST-
430 HDP1-DBD was transformed into BL21 (DE3) competent *E. coli* strain (NEB) and protein
431 expression was done following the manufacturer directions with some modifications. Briefly, an
432 overnight culture was inoculated with one bacterial colony in 2x YT media supplemented with
433 the corresponding antibiotic. Next day, culture was diluted 1:100 with fresh media and kept at
434 30º C with vigorous agitation. Bacterial growth was monitored until culture reach exponential
435 phase. At this point, IPTG (1 mM final concentration) was added and the culture was kept for
436 another 2 hours at 30º C with vigorous agitation. Cells were harvested and resuspended in lysis
437 buffer (50 mM Tris-HCl, pH 7.5; 100 mM NaCl; 1 mM DTT; 5% Glycerol; 1 mM PMSF; 1 mM
438 EDTA; 1x protease inhibitors cocktail) supplemented with Lysozyme (1mg/ml, Sigma). In order
439 to remove bacterial DNA from our putative DNA binding protein, lysate was treated with
440 polyethyleneimine (PEI) (0.1% v/v) (*46*). Lysate was sonicated, cleared by centrifugation at
441 14,000 x *g* for 30 min at 4º C. Protein extract was recovered and GST-HDP1-DBD protein was
442 purified using Pierce GST Spin purification kit (Cat. Nº 16106) following manufacturer
443 directions. Protein of interest was dialyzed using Slide-A-Lyzer Dialysis Cassette 10,000 MWCO
444 (ThermoScientific, Cat. Nº 66810) and concentrated using Amicon Ultra Centrifugal filters
445 10,000 MWCO (Millipore, Cat. Nº UFC901024). Purity was verified by Coomassie staining after
446 SDS-PAGE and concentration was measured by Bradford assay.
447
448 **Protein Binding Microarray**
449 GST-HDP1-DBD binding was analyzed twice on Protein Binding Microarrays (PBMs) as
450 previously described (*47*, *48*). In this study two different universal PBM arrays (Agilent
451 AMADIDs 016060 v9 and AMADID 015681 v11) were used covering all contiguous 8-mers, as
452 well as gapped 8-mers spanning up to 10 positions. Binding of purified GST-HDP1-DBD fusion
453 proteins was visualized on the PBMs using Alexa-488 conjugated anti-GST antibody. Data
454 analysis was carried out using the PBM analysis software suite downloaded from
455 http://thebrain.bwh.harvard.edu/PBMAnalysisSuite/index.html. Following normalization,
456 enrichment scores were calculated and the "Seed-and-Wobble" algorithm was applied to the

457 combined data to generate position weight matrices (PWMs). An enrichment score cut-off of 0.45
458 was used to separate high affinity binding from non-specific and low affinity binding. Secondary
459 motifs were identified by running the "rerank" program until E-scores below 0.45 were obtained.
460

461 **Isothermal Titration Calorimetry**

462 Sequence encoding HDP1 aa2991-3078 were cloned into the MCS1 of the pRSFDuet-1 vector
463 (Novagen) engineered with an N-terminal His-SUMO tag. The proteins were expressed in *E. coli*
464 strain BL21 CodonPlus (DE3)-RIL (Stratagene). Bacteria were grown in Luria-Bertani medium
465 at 37°C to OD600=0.8 and induced with 0.4 mM IPTG at 18°C overnight. Cells were collected
466 via centrifugation at 5000×g and lysed via sonication in Lysis Buffer (20 mM Tris-HCl, pH 8.0;
467 500 mM NaCl; 20 mM imidazole, and 5% Glycerol) supplemented with 1 mM
468 phenylmethylsulfonyl fluoride and 0.5% Triton X-100. Cellular debris was removed by
469 centrifugation at 20,000×g, and the supernatant was loaded onto 5 ml HisTrap FF column (GE
470 Healthcare) and eluted using the lysis buffer supplemented with 500 mM imidazole. The elution
471 was dialyzed at 4°C overnight against the buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20
472 mM imidazole, and 5 mM β-mercaptoethanol) with ULP1 protease added (lab stock). The sample
473 was reloaded on the HisTrap FF column to remove the His-SUMO tag. The flow-through was
474 loaded on the Heparin column (GE Healthcare) and eluted with a gradient of NaCl from 300 mM
475 to 1 M. The target protein was further purified by size exclusion chromatography (Superdex 200
476 [16/60], GE Healthcare) in the buffer (20 mM Tris-HCl, pH 7.5; 200 mM NaCl; 1 mM $MgCl_2$;
477 and 1 mM DTT). The high purity eluting fractions were detected by SDS-PAGE and concentrated
478 to around 20 mg/ml. The protein was flash-frozen in liquid nitrogen and stored at -80°C.
479 All the binding experiments were performed on a Microcal ITC 200 calorimeter. Purified HDP1-
480 DBD proteins were dialyzed overnight against the buffer (20 mM HEPES, pH 7.5, 150 mM NaCl,
481 1 mM DTT) at 4°C. DNA oligos were synthesized by Integrated DNA Technologies (IDT) and
482 dissolved in the same buffer. The assays perform with 1 mM DNA duplexes containing the
483 tandem motif (TA<u>GTGCAC</u>CTATG<u>GTGCAC</u>TT) with 0.1 mM HDP1-DBD proteins. Each
484 reaction's exothermic heat was measured by sequential injection of the 2 μL DNA duplexes into
485 proteins solution, spaced at intervals of 180 seconds. The titration was according to standard
486 protocol at 20°C and the data were fitted using the program Origin 7.0.

487 **Gel-shift Assays**

488 Electrophoretic mobility shift assays were performed using Light Shift EMSA kits (Thermo
489 Scientific) using 24 pg of protein and 40 fmol of probe, as previously described (*49*). Biotinylated
490 double-stranded were synthesized probes (ThermoFisher) with the indicated sequence.

491

492 **RNA Extraction, cDNA synthesis, and quantitative RT-PCR**

493 Total RNA from saponin-lysed parasites was extracted using Trizol (Invitgrogen) and Direct-Zol
494 RNA MiniPrep Plus kit (Zymo Research). The cDNA was prepared from 100-500ng total RNA
495 (pre-treated with 2U DNase I, amplification grade) using SuperScript III Reverse Transcriptase
496 kit (Invitrogen) and random hexamers. Quantitative PCR was performed on the Quant Studio 6
497 Flex (Thermo Fisher) using iTaq Sybr Green (Bio-Rad) with specific primers for selected target
498 genes (Table). Quantities were normalized to seryl-tRNA synthetase (PF3D7_0717700). Analysis
499 of expression of the *var* gene family was performed by using the primer set described in Salanti
500 et al. 2003 (*50*).

501

502 **RNA sequencing**

503 Following gametocyte induction, highly synchronous cultures containing committed schizonts
504 were added to fresh RBCs and allowed to reinvade for 12 hours prior to the addition of 50 mM
505 N-acetyl glucosamine to block the development of hemozoin-containing asexual trophozoites. On
506 day 2 of gametocyte development, stage I gametocytes were purified magnetically, and total RNA
507 was extracted as described above. Following RNA isolation, total RNA integrity was checked
508 using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). RNA concentrations were
509 measured using the NanoDrop system (Thermo Fisher Scientific, Inc., Waltham, MA).
510 Preparation of RNA sample library and RNA-seq were performed by the Genomics Core
511 Laboratory at Weill Cornell Medicine. rRNA was removed from Total RNA using Illumina Ribo
512 Zero Gold for human/mouse/rat kit. Using Illumina TruSeq RNA Sample Library Preparation v2
513 kit (Illumina, San Diego, CA), Messenger RNA was fragmented into small pieces using divalent
514 cations under elevated temperature. The cleaved RNA fragments were copied into first strand
515 cDNA using reverse transcriptase and random primers. Second strand cDNA synthesis followed,
516 using DNA Polymerase I and RNase H. The cDNA fragments then went through an end repair
517 process, the addition of a single 'A' base, and then ligation of the adapters. The products were
518 then purified and enriched with PCR to create the final cDNA library. Libraries were pooled and
519 sequenced on Illumina HiSeq4000 sequencer with single-end 50 cycles. Read files were checked
520 for quality by using FASTQC v0.11.5 (https://github.com/s-andrews/FastQC). Reads were

521 trimmed to remove low-quality positions and adapter sequences using cutadapt (version 1.16)
522 (*51*). Reads were mapped against the *P. falciparum* 3D7 reference genome v40 (*52*) using STAR
523 aligner (version 2.61) (*53*) and nuclear-encoded genes were analyzed for differential gene
524 expression using cufflinks (version 2.2.1) (*54*). Genes with a false discovery rate of <=0.05 with
525 a mean FPKM >5 in at least one strain were called significant. For genes with FPKM > 5 in one
526 strain and no detectable expression in the other, FKPM values were set to 0.1 for purposes of
527 fold-change calculation. Gene Set Enrichment Analysis was carried out with the FGSEA v1.16.0
528 Bioconductor v 3.12 package (*55*) with an FDR cutoff of <=0.05.
529
530 **Hi-C Sequencing**
531 Parasites were crosslinked in warm PBS containing 1.25% formaldehyde for 25 min on a rocking
532 platform and quenched by adding glycine to 150 mM followed by 15 min incubations at 37°C
533 and 4°C. The parasites were centrifuged at 660 x *g* for 20 min at 4°C, resuspended in 5 volumes
534 of ice-cold PBS, and incubated for 10 min at 4°C. Parasites were centrifuged at 660 x *g* for 15
535 min at 4°C, washed once in ice-cold PBS, and stored as a pellet at −80°C. Parasite pellets were
536 then thawed on ice in Hi-C lysis buffer (10 mM Tris-HCl, pH8.0; 10 mM NaCl, 2 mM AEBSF,
537 Roche Complete Mini EDTA-free protease inhibitor cocktail, 0.25% Igepal CA-630). Parasite
538 membranes were disrupted by passing the lysate through a 26.5- gauge needle 15 times with a
539 syringe. Samples were spun and pellets were washed once with ice-cold Hi-C lysis buffer,
540 resuspended in 0.5% SDS and incubated at 62°C for 10 min to solubilize the chromatin. The SDS
541 was neutralized by the addition of nuclease-free water and 1% Triton X-100, for 15-minutes at
542 37°C. DNA was then digested using *MboI* restriction overnight at 37°C. After digestion, the
543 enzyme was heat inactivated at 62°C for 20 min, and then cooled to RT. 5' overhangs were filled
544 in by Biotin-14-dCTP (Invitrogen) using DNA Polymerase I and Large (Klenow) Fragment
545 (NEB) for 45 min at 37°C. Blunt-end ligation was performed using T4 DNA ligase at 20°C for 4
546 hours. The nuclei were spun down to remove random ligation products and to reduce the overall
547 reaction volume and resuspend in decrosslinking buffer (50 mM Tris-HCl, pH 8.0, 1% SDS, 1
548 mM EDTA, 500 mM NaCl) before adding RNaseA (20 mg/ml) for 45 min at 37°C and proteinase
549 K (NEB) treatment overnight at 45°C. DNA was extracted using Agencourt AMPure XP beads
550 (Beckman Coulter) and eluted in 10 mM Tris-HCl at pH 8.0. The purified DNA was sheared on
551 a Covaris S220 and biotinylated DNA fragments were pull-down using MyOne Streptavidin T1
552 magnetic beads (Invitrogen) at 55°C for 2 min, washed and resuspended in T4 DNA Ligase Buffer
553 to perform end-repair on DNA fragments bound to the beads containing DNA Polymerase I,
554 Large (Klenow) Fragment (NEB), and T4 DNA Polymerase (NEB) and incubated for 30 min at

555 20°C. The beads were washed twice before performing A-tailing using Klenow Fragment (3′→5′
556 exo–nuclease) (NEB)) for 30 min at 37°C. The beads were then washed twice and resuspended
557 in T4 DNA Ligase Buffer to perform adapter ligation using the NEBNext Illumina Adapter for
558 15 min at 20°C in a thermomixer. The beads were then resuspended in 100 µl of 10 mM Tris-HCl
559 at pH 8.0 and transferred to a new tube to amplify the library using the HiFi HotStart ReadyMix
560 (KAPA Biosystems) as well as the universal forward primer and barcoded reverse primer before
561 being incubated with the following PCR program: 45 sec at 98°C, 12 cycles of 15 sec at 98°C, 30
562 sec at 55°C, 30 sec at 62°C and a final extension of 5 min at 62°C. The library was then purified
563 using double-SPRI size selection, with 0.5Å~ right-side selection (25 µl AMPure XP beads) and
564 1.0Å~ left-side selection (25 µl AMPure XP beads). Libraries were quantified by NanoDrop
565 (Thermo Scientific) and Bioanalyzer (Agilent), prior to multiplexing and sequencing in a 75-bp
566 paired-end run on a NextSeq500 (Illumina).
567

568 **Analysis of Hi-C Data**

569 Bed files of virtual complete *MboI* (cutting at GATC) digest of the *P. falciparum* genome were
570 generated with digest_genome.py from the HiC-Pro suite (*56*). The paired-end reads were
571 processed (i.e. mapped, paired, and de-duplicated), binned at 10kb resolution, and normalized
572 using the ICE algorithm using HiC-Pro v2.11.4 (*57*). Reproducibility and strain-dependent
573 differences (Fig. 5A) between Hi-C samples was evaluated by correlation analysis using
574 HiCSpector v1.0.1 (*58*) and GenomeDISCO (same correlation pattern as HiCSpector, not shown)
575 (*59*). Interaction maps (Fig. S9) were generated by plotting intra-chromosomal ICE-normalized
576 interactions. Summary statistics (Fig. 5B) of intra-chromosomal interaction distances were
577 calculated for each strain and chromosome from normalized ICE interaction counts using R
578 v4.0.0. Differential interaction analysis was carried out at 10 kb resolution for each replicate using
579 selfish (*60*) based on ICE-normalized interaction counts from the Hi-C Pro all-valid-pairs matrix
580 and reformatted into the ".hic" file-format using Juicebox v1.9.9 (*61*). All genome-wide bin pairs
581 with significant changes in interaction frequency between the *hdp1*-ty1 and Δ*hdp1* strains with
582 an FDR <=0.1 in both replicate pairs are shown in the top left of Fig. 5C. For the analysis of
583 chromosome-end interactions, interactions fold-change were calculated for all interactions in bins
584 between the chromosome end and the first protein coding gene. As the final bin on each
585 chromosome varies in size, an additional bin was added for ends truncated by telomere healing
586 (5R, 11R, and 14R) to ensure interaction counts were calculated for at least 10kb.

587 **Description of Chromatin Immunoprecipitation Sequencing attempts.**

588 Since our experiments clearly that HDP1 has the ability to bind DNA and is tightly associated
589 with nuclear chromatin (Fig. 2), we have spent the last 2.5 years attempting determine where in
590 the genome HDP1 is bound by chromatin immunoprecipitation sequencing (ChIP-seq). Repeated
591 independent efforts using crosslinked chromatin from the Day 5 *hdp1-Ty1* gametocytes were
592 unsuccessful in the Kafsack (Weill Cornell), and Bartfai (Radboud University) labs while control
593 ChIP-seq for known chromatin-associated proteins was successful. We then generated the *hdp1-*
594 *gfp* line in case the Ty1 epitope tags were no accessible. But independent ChIP-seq attempts with
595 crosslinked chromatin in the Kafsack and Llinás labs with *hdp1-gfp* gametocytes were also
596 unsuccessfully while controls worked. Finally, we adapted the CUT&RUN (a method for native,
597 *in situ* ChIP) (*62*, *63*) for use in malaria parasites (manuscript forthcoming). Again, CUT&RUN
598 worked great for other chromatin-associated proteins but failed repeatedly for HDP1.

599 **REFERENCES AND NOTES**
600

601 1. B. O. Park, R. Ahrends, M. N. Teruel, Consecutive Positive Feedback Loops Create a Bistable
602 Switch that Controls Preadipocyte-to-Adipocyte Conversion. *Cell Reports*. **2**, 976–990 (2012).

603 2. S. Bhattacharya *et al*., A Bistable Switch Underlying B-Cell Differentiation and Its Disruption by
604 the Environmental Contaminant 2,3,7,8-Tetrachlorodibenzo-p-dioxin. *Toxicol Sci*. **115**, 51–65
605 (2010).

606 3. T. M. Norman, N. D. Lord, J. Paulsson, R. Losick, Stochastic Switching of Cell Fate in Microbes.
607 *Annu Rev Microbiol*. **69**, 381–403 (2015).

608 4. D. Satory, A. J. Gordon, J. A. Halliday, C. Herman, Epigenetic switches: can infidelity govern
609 fate in microbes? *Curr Opin Microbiol*. **14**, 212–217 (2011).

610 5. B. F. C. Kafsack *et al*., A transcriptional switch underlies commitment to sexual development in
611 malaria parasites. *Nature*. **507**, 248–252 (2014).

612 6. A. Sinha *et al*., A cascade of DNA-binding proteins for sexual commitment and development in
613 Plasmodium. *Nature*. **507**, 253–257 (2014).

614 7. J. J. Lopez-Rubio, L. Mancio-Silva, A. Scherf, Genome-wide analysis of heterochromatin
615 associates clonally variant gene regulation with perinuclear repressive centers in malaria
616 parasites. *Cell Host Microbe*. **5**, 179–190 (2009).

617 8. N. M. B. Brancucci *et al*., Heterochromatin protein 1 secures survival and transmission of malaria
618 parasites. *Cell Host Microbe*. **16**, 165–176 (2014).

619 9. A. Poran *et al*., Single-cell RNA sequencing reveals a signature of sexual commitment in malaria
620 parasites. *Nature*. **551**, 95–99 (2017).

621 10. G. A. Josling *et al*., Dissecting the role of PfAP2-G in malaria gametocytogenesis. *Nature*
622 *Communications*. **11**, 1–13 (2020).

623 11. B. I. Coleman *et al*., A Plasmodium falciparum Histone Deacetylase Regulates Antigenic
624 Variation and Gametocyte Conversion. *Cell Host Microbe*. **16**, 177–186 (2014).

625 12. M. Filarsky *et al*., GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-
626 dependent gene silencing. *Science*. **359**, 1259–1263 (2018).

627 13. E. M. Bunnik *et al*., Changes in genome organization of parasite-specific gene families during the
628 Plasmodium transmission stages. *Nature Communications*. **9**, 1910 (2018).

629 14. S. A. Fraschka *et al*., Comparative Heterochromatin Profiling Reveals Conserved and Unique
630 Epigenome Signatures Linked to Adaptation and Development of Malaria Parasites. *Cell Host*
631 *Microbe*. **23**, 407–420.e8 (2018).

632 15. M. Tibúrcio *et al*., Early gametocytes of the malaria parasite Plasmodium falciparum specifically
633 remodel the adhesive properties of infected erythrocyte surface. *Cell Microbiol*. **15**, 647–659
634 (2013).

635 16. C. Bancells *et al*., Revisiting the initial steps of sexual development in the malaria parasite
636 Plasmodium falciparum. *Nat. Microbiol*. **4**, 144–154 (2019).

637 17. J. A. Young *et al*., The Plasmodium falciparum sexual development transcriptome: a microarray
638 analysis using ontology-based pattern identification. *Mol Biochem Parasitol*. **143**, 67–79 (2005).

639 18. H. J. Painter, M. Carrasquilla, M. Llinás, Capturing in vivo RNA transcriptional dynamics from
640 the malaria parasite Plasmodium falciparum. *Genome Res*. **27**, 1074–1086 (2017).

641 19. R. van Biljon *et al*., Hierarchical transcriptional control regulates Plasmodium falciparum sexual
642 differentiation. *BMC Genomics*. **20**, 920 (2019).

643 20. R. S. Kent *et al*., Inducible developmental reprogramming redefines commitment to sexual
644 development in the malaria parasite Plasmodium berghei. *Nat. Microbiol*. **3**, 1206–1213 (2018).

645 21. S. Balaji, M. M. Babu, L. M. Iyer, L. Aravind, Discovery of the principal specific transcription
646 factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding
647 domains. *Nucleic Acids Res*. **33**, 3994–4006 (2005).

648 22. K. Modrzynska *et al*., A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting
649 Transcriptional Regulators Controlling the Plasmodium Life Cycle. *Cell Host Microbe*. **21**, 11–
650 22 (2017).

651 23. M. Yuda, S. Iwanaga, I. Kaneko, T. Kato, Global transcriptional repression: An initial and
652 essential step for Plasmodium sexual development. *Proceedings of the National Academy of*
653 *Sciences* (2015), doi:10.1073/pnas.1504389112.

654 24. C. Zhang *et al*., Systematic CRISPR-Cas9-Mediated Modifications of Plasmodium yoelii
655 ApiAP2 Genes Reveal Functional Insights into Parasite Development. *mBio*. **8**, 610 (2017).

656 25. M. Yuda, I. Kaneko, S. Iwanaga, Y. Mura, T. Kato, Female-specific gene regulation in malaria
657 parasites by an AP2-family transcription factor. *Mol Microbiol*. **6**, 1–51 (2019).

658 26. T. R. Bürglin, M. Affolter, Homeodomain proteins: an update. *Chromosoma*. **125**, 497–521
659 (2016).

660 27. O. Llorà-Batlle *et al*., Conditional expression of PfAP2-G for controlled massive sexual
661 conversion in Plasmodium falciparum. *Science advances*. **6**, eaaz5057 (2020).

662 28. P. Prommana *et al*., Inducible Knockdown of Plasmodium Gene Expression Using the glmS
663 Ribozyme. *PLoS ONE*. **8**, e73783 (2013).

664 29. C. P. Nixon *et al*., Antibodies to PfsEGXP, an Early Gametocyte-Enriched Phosphoprotein,
665 Predict Decreased Plasmodium falciparum Gametocyte Density in Humans. *J Infect Dis*. **218**,
666 1792–1801 (2018).

667 30. L. Pradhan *et al*., Crystal Structure of the Human NKX2.5 Homeodomain in Complex with DNA
668 Target. *Biochemistry*. **51**, 6312–6319 (2012).

669 31. P. Oquendo *et al*., Characterisation of a repetitive DNA sequence from the malaria parasite,
670 Plasmodium falciparum. *Mol Biochem Parasitol*. **18**, 89–101 (1986).

671 32. S. F. Calhoun *et al*., Chromosome End Repair and Genome Stability in Plasmodium falciparum.
672 *mBio*. **8**, e00547–17 (2017).

673 33. C. Flueck *et al*., A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in
674 chromosome end biology. *PLoS Pathog*. **6**, e1000784 (2010).

675 34. M. Parkyn Schneider *et al*., Disrupting assembly of the inner membrane complex blocks
676 Plasmodium falciparum sexual stage development. *PLoS Pathog*. **13**, e1006659 (2017).

677 35. E. Bushell *et al*., Functional Profiling of a Plasmodium Genome Reveals an Abundance of
678 Essential Genes. *Cell*. **170**, 260–272.e8 (2017).

679 36. K. Hedgethorne *et al*., Homeodomain-like DNA binding proteins control the haploid-to-diploid
680 transition in Dictyostelium. *Science advances*. **3**, e1602937 (2017).

681 37. S. Henikoff, J. G. Henikoff, A. Sakai, G. B. Loeb, K. Ahmad, Genome-wide profiling of salt
682 fractions maps physical properties of chromatin. *Genome Res*. **19**, 460–469 (2009).

683 38. K. Moll, I. Ljungström, H. Perlmann, A. Scherf, Methods in malaria research. *Manassas* (2008).

684 39. J. Morlon-Guyot *et al*., A proteomic analysis unravels novel CORVET and HOPS proteins
685 involved in Toxoplasma gondii secretory organelles biogenesis. *Cell Microbiol*. **20**, e12870
686 (2018).

687 40. M. Rug, A. G. Maier, in *Malaria* (Humana Press, Totowa, NJ, Totowa, NJ, 2012), vol. 923 of
688 *Methods in Molecular Biology*, pp. 75–98.

689 41. M. Ghorbal *et al*., Genome editing in the human malaria parasite Plasmodium falciparum using
690 the CRISPR-Cas9 system. *Nat Biotechnol*. **32**, 819–821 (2014).

691 42. T. Q. Tanaka, K. C. Williamson, A malaria gametocytocidal assay using oxidoreduction
692 indicator, alamarBlue. *Mol Biochem Parasitol*. **177**, 160–163 (2011).

693 43. C. Flueck *et al*., Plasmodium falciparum Heterochromatin Protein 1 Marks Genomic Loci Linked
694 to Phenotypic Variation of Exported Virulence Factors. *PLoS Pathog*. **5**, e1000569 (2009).

695 44. E. Saini *et al*., Ph otosensitized I NA- L abelled protein 1 (PhIL1) is novel component of the
696 inner membrane complex and is required for Plasmodium parasite development. *Sci. Rep*. **7**,
697 15577 (2017).

698 45. A.-K. Mehnert, C. S. Simon, J. Guizetti, Immunofluorescence staining protocol for STED
699 nanoscopy of Plasmodium-infected red blood cells. *Mol Biochem Parasitol*. **229**, 47–52 (2019).

700 46. B. RR, Use of polyethyleneimine in purification of DNA-binding proteins. *Meth. Enzymol*. **208**,
701 3–10 (1991).

702 47. M. F. Berger *et al*., Compact, universal DNA microarrays to comprehensively determine
703 transcription-factor binding site specificities. *Nat Biotechnol*. **24**, 1429–1435 (2006).

704 48. M. F. Berger, M. L. Bulyk, Universal protein-binding microarrays for the comprehensive
705 characterization of the DNA-binding specificities of transcription factors. *Nature Protocols*. **4**,
706 393–411 (2009).

707 49. T. L. Campbell, E. K. De Silva, K. L. Olszewski, O. Elemento, M. Llinás, Identification and
708 genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from
709 the malaria parasite. *PLoS Pathog*. **6**, e1001165 (2010).

710 50. A. Salanti *et al*., Selective upregulation of a single distinctly structured var gene in chondroitin
711 sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Mol*
712 *Microbiol*. **49**, 179–191 (2003).

713 51. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads.
714 *EMBnet.journal*. **17**, 10–12 (2011).

715 52. S. Warrenfeltz *et al*., EuPathDB: The Eukaryotic Pathogen Genomics Database Resource.
716 *Methods Mol Biol*. **1757**, 69–113 (2018).

717 53. A. Dobin *et al*., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. **29**, 15–21 (2012).

718 54. C. Trapnell *et al*., Differential analysis of gene regulation at transcript resolution with RNA-seq.
719 *Nat Biotechnol*. **31**, 46–53 (2013).

720 55. G. Korotkevich, V. Sukhov, A. Sergushichev, Fast gene set enrichment analysis. *bioRxiv*. **10**,
721 060012 (2019).

722 56. N. Servant *et al*., HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome*
723 *Biol*. **16**, 1–11 (2015).

724 57. M. Imakaev *et al*., Iterative correction of Hi-C data reveals hallmarks of chromosome
725 organization. *Nat Meth*. **9**, 999–1003 (2012).

726 58. K.-K. Yan, G. G. Yardımcı, C. Yan, W. S. Noble, M. Gerstein, HiC-spector: a matrix library for
727 spectral and reproducibility analysis of Hi-C contact maps. *Bioinformatics*. **33**, 2199–2201
728 (2017).

729 59. O. Ursu *et al*., GenomeDISCO: a concordance score for chromosome conformation capture
730 experiments using random walks on contact map graphs. *Bioinformatics*. **34**, 2701–2707 (2018).

731 60. A. R. Ardakany, F. Ay, S. Lonardi, Selfish: discovery of differential chromatin interactions via a
732 self-similarity measure. *Bioinformatics*. **35**, i145–i153 (2019).

733 61. J. T. Robinson *et al*., Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data.
734 *cels*. **6**, 256–258.e1 (2018).

735 62. P. J. Skene, J. G. Henikoff, S. Henikoff, Targeted in situ genome-wide profiling with high
736 efficiency for low cell numbers. *Nature Protocols*. **13**, 1006–1019 (2018).

737 63. M. P. Meers, T. D. Bryson, J. G. Henikoff, S. Henikoff, Improved CUT&RUN chromatin
738 profiling tools. *elife*. **8**, 245 (2019).

ACKNOWLEDGMENTS

General: We wish to thank Dr. Pawan Malhotra for the generous gift of anti-PhIL1 antibodies, V. Carruthers for the generous gift of anti-Ty1 antibodies, the Weill Cornell Medicine genomics core for technical support, M. Llinás and R. Bartfai for separate HDP1 ChIP-seq attempts with *hdp1*-gfp and *hdp1*-ty1 gametocytes, respectively, as well as K. Deitsch and J. King for valuable feedback on the manuscript.

Funding: This work was supported by startup funds from Weill Cornell Medicine (BK), 1R01AI141965 (BK), 1R01AI138499 (BK), 1R01 AI125565 (ML), 1R01 AI136511 (KLR), R21 AI142506-01 (KLR), the University of California, Riverside (NIFA-Hatch-225935-KLR), and support from the Mathers Foundation (DJP).

Author contributions: Conceptualization: B.F.C.K.; Methodology: B.F.C.K., R.C.M., W.X., W.N., K.G.L.; Investigation: R.C.M., X.T., W.X., T.L., G.B., L.O., W.D.; Software, Formal Analysis, Data Curation: B.F.C.K., T.L.; Writing – Original Draft: R.C.M.; Writing – Review & Editing: B.F.C.K., R.C.M.; Visualization: R.C.M., B.F.C.K.; Supervision: B.F.C.K., K.G.L., M.L., W.N, D.P., Project Administration: B.F.C.K; Funding Acquisition: B.F.C.K.

Competing interests: The authors declare no competing interests.

Data and materials availability:

Raw high throughput sequencing data have been deposited in the NCBI Sequence Read Archive under accession number SRPXXXX.

Analysis code is available at https://github.com/KafsackLab/HDP1.

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

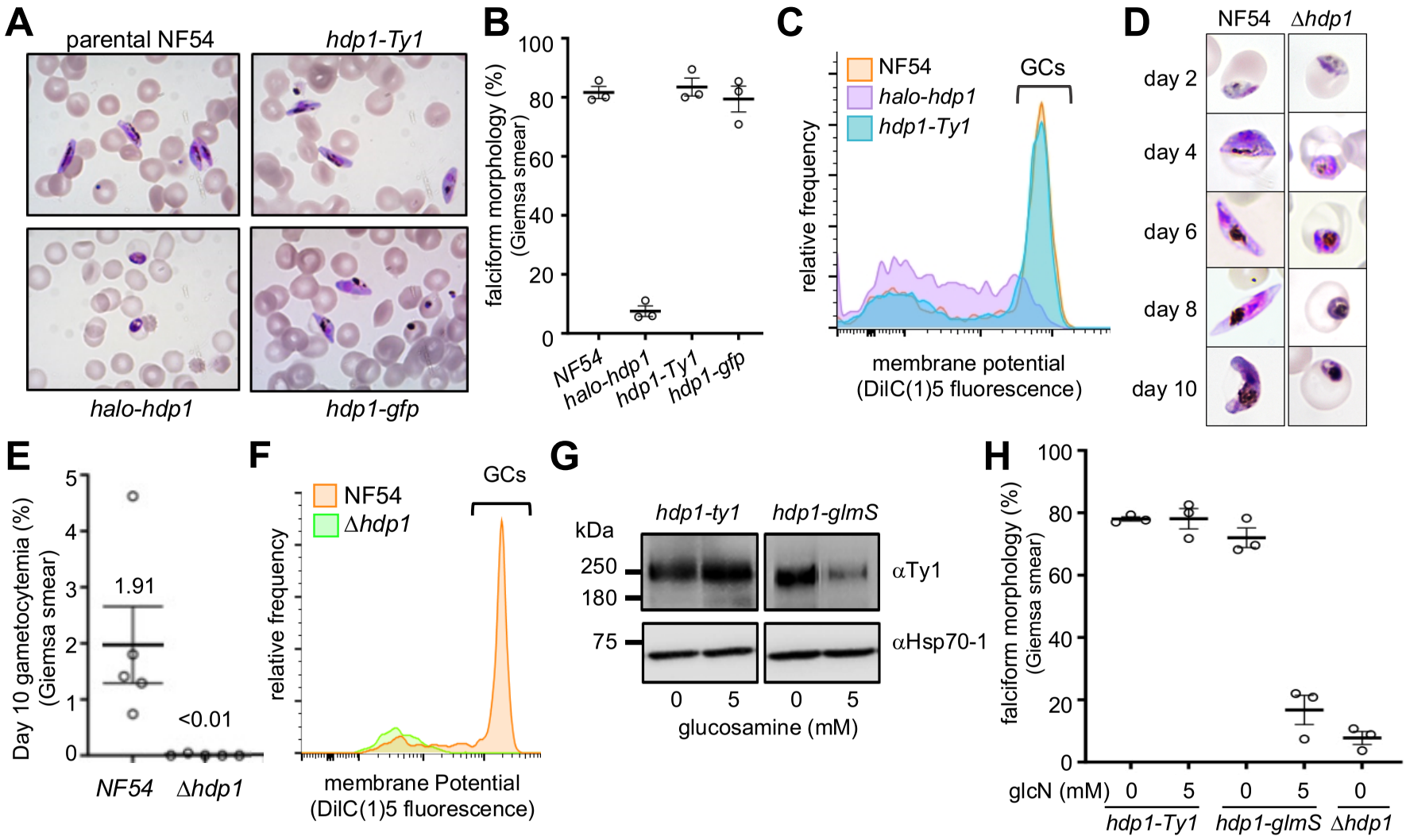


Fig. 1: Loss of HDP1 function disrupts gametocyte maturation. (A-C) N-terminal tagging of the endogenously encoded HDP1 (*halo-hdp1*) blocked maturation of gametocytes (GCs) while C-terminal tagging of the endogenous locus with either GFP (*hdp1-gfp*) or a triple Ty1 (*hdp1-Ty1*) epitope had no effect on gametocyte morphology or viability, as determined by membrane potential staining with DilC(1)-5 on day 5 of gametocyte maturation. **(D-F)** Targeted disruption of the *hdp1* locus (Δ*hdp1*) blocked formation of late gametocytes. Graphs show mean ± s.e.m of n=3-5 **(G)** Glucosamine-inducible knockdown of HDP1 in *hdp1-Ty1* and *hdp1-glmS* day 5 gametocytes. Representative of n=3. **(H)** Percentage of falciform day 5 gametocytes in response to 5 mM glucosamine (GlcN). All Images and flow cytometry plots are representative of n=3.

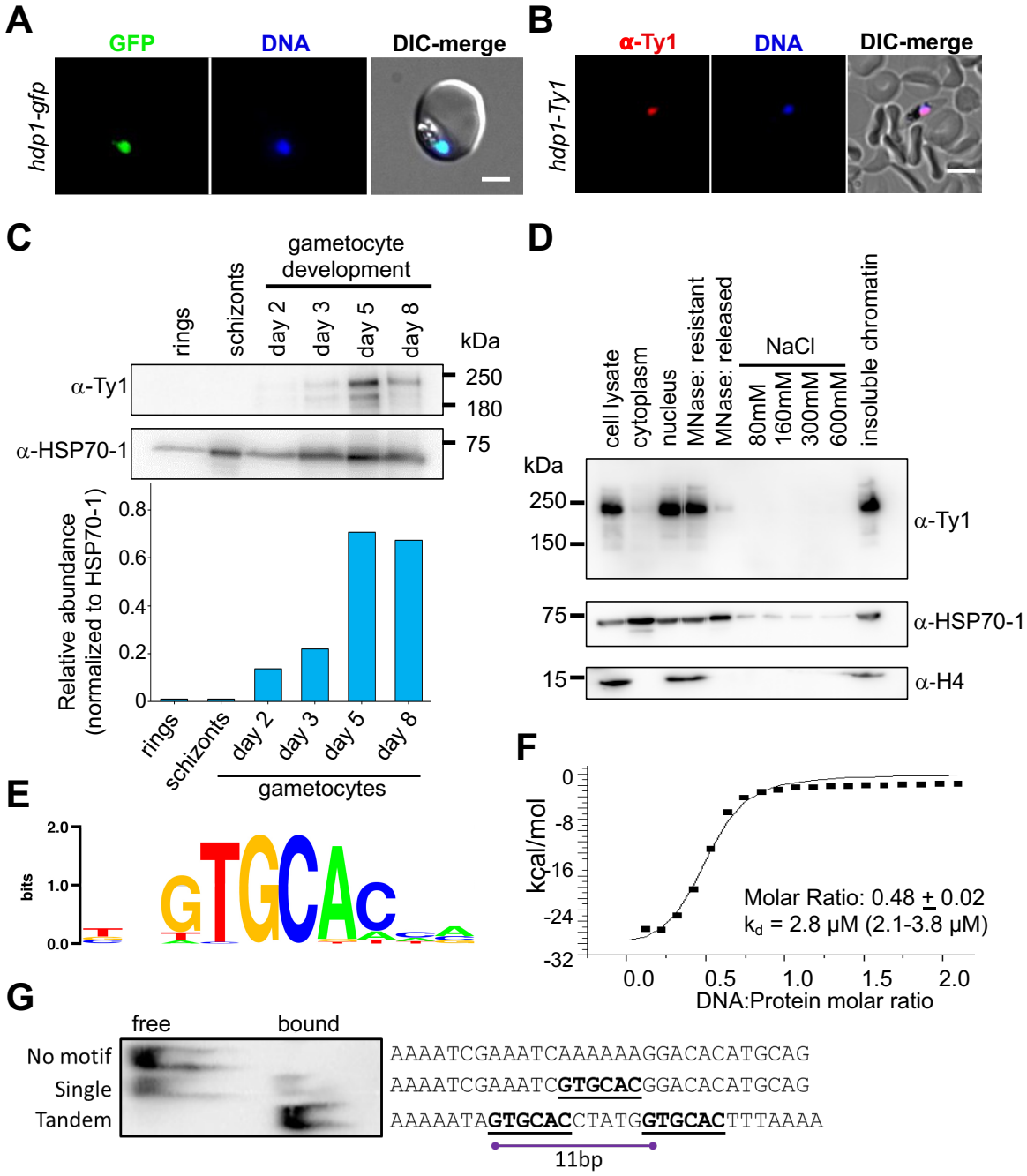
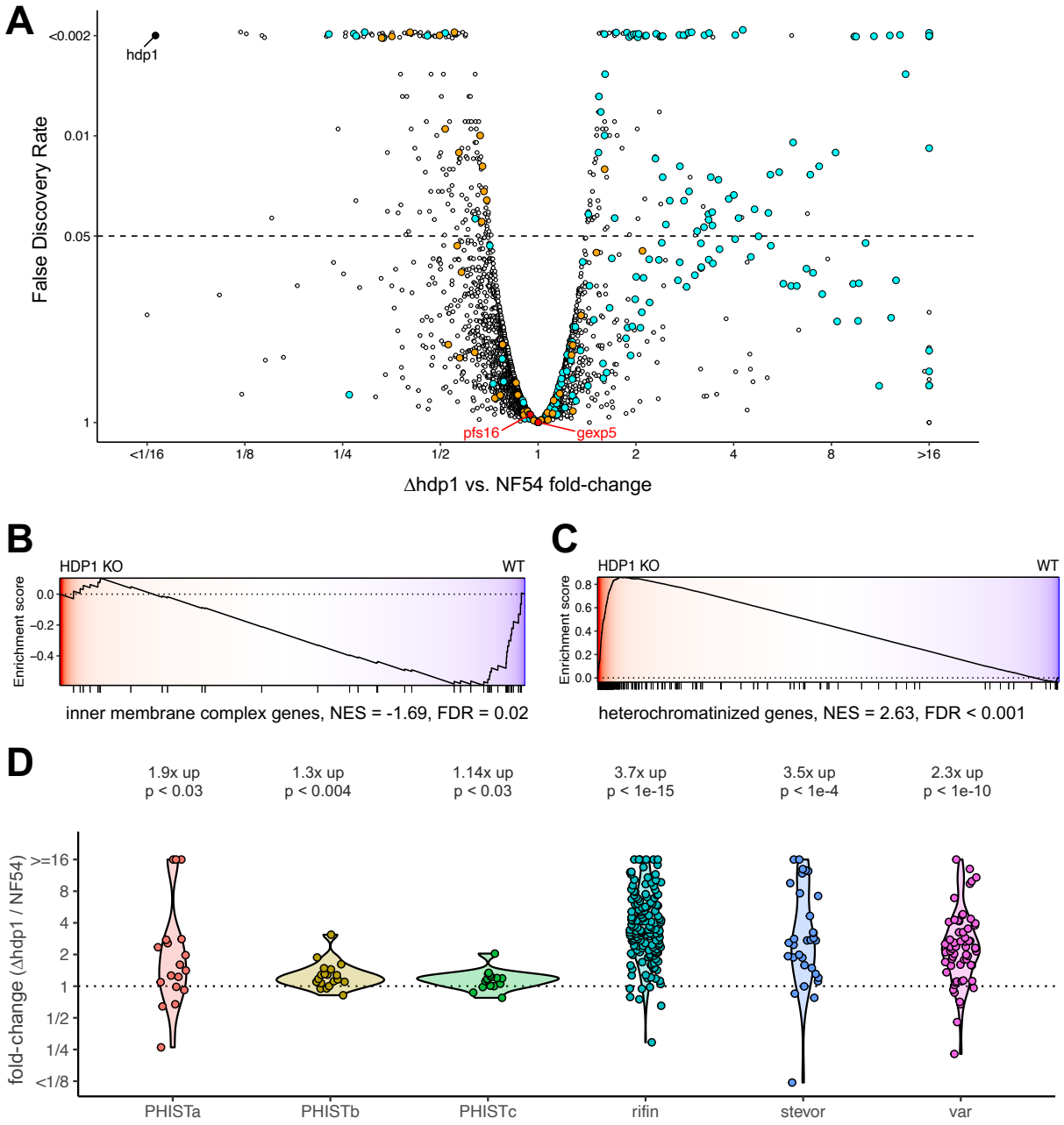


Fig. 2: HDP1 is chromatin-associated nuclear protein in gametocytes that recognizes a GC-rich DNA motif as a dimer. (A) Live-cell fluorescence microscopy of *hdp1-gfp* on day 5 of maturation stained with the DNA dye Hoechst33342 (blue). Scale Bar: 2mm. Representative of n=2. **(B)** Immunofluorescence microscopy of *hdp1-Ty1* and parental NF54 gametocytes on day 5 of maturation co-stained with anti-Ty1 antibodies (red) and Hoechst33342 (blue). Scale Bar: 5 µm. Representative of n=2. **(C)** Western blotting of parasite lysates in asexual stages and during gametocyte maturation shows HDP1 is expressed during the stages of gametocytogenesis. Representative of n=3. **(D)** Western blot of cytoplasmic and nuclear extracts of *hdp1-Ty1* gametocytes on day 5 of maturation stained with antibodies against the Ty1 epitope tag, the histone H4, and HSP70-1. Representative of n=3. **(E)** Maximum enrichment DNA motif for the GST-HDP1 DBD from a PBM. **(F)** Isothermal calorimetry indicates the HDP1-DBD domain recognizes DNA as a dimer. n=2. **(G)** Optimal gel-shift was observed for probes containing a tandem motif with a 5bp spacer compared to probes with either a single or no motif. Representative of n=3.

783



784

785 **Fig. 3: Disruption of HDP1 results in leaky expression of heterochromatin-associated genes and**
786 **reduced expression of inner membrane complex genes in early gametocytes. (A)** Genome-wide
787 comparison of differential gene expression in Δ*hdp1* and parental NF54 gametocytes on day 2 of
788 gametocytogenesis (stage I, n=2). *hdp1* (solid black), heterochromatin-associated genes (cyan), IMC genes
789 (orange) and the two canonical early gametocyte markers, *pfs16* and *gexp5 (red)* are highlighted. Gene set
790 enrichment analysis (GSEA) indicated significant downregulation of IMC genes **(B)** and global
791 upregulation of heterochromatin associated genes **(C)**. **(D)** Heterochromatin associated gene families with
792 significant up regulation in Δ*hdp1*. Geometric mean fold-changes and p-values (two-sided, one sample t-
793 test) are indicated.

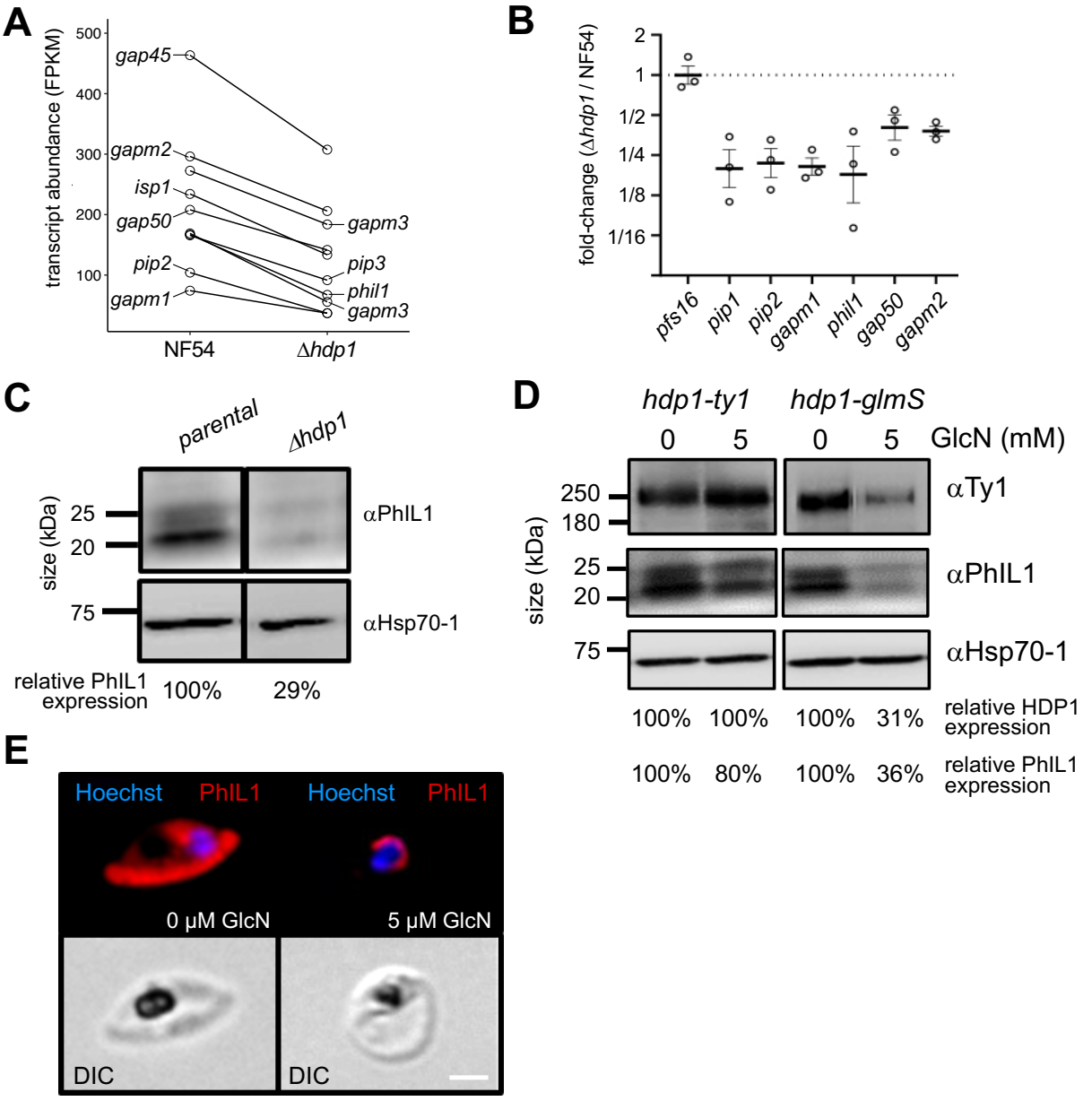


Fig. 4: HDP1 regulates PhIL1-dependent expansion of the inner membrane complex in early gametocytes. (A) Transcripts encoding inner membrane complex proteins are down-regulated in Δ*hdp1* day 2 gametocytes. Normalized transcript levels (FPKM) in NF54 and Δ*hdp1* day 2 gametocytes. **(B)** Validation of down-regulation by qRT-PCR. Levels of the canonical early gametocyte marker *pfs16* remained unchanged (n=3). **(C)** PhiL1 protein levels in parental and Δ*hdp1* day 5 gametocytes. Hsp70-1 abundance shown as the loading control. Representative result of n=2. **(D)** Day 5 morphology of *hdp1-ty1* and *hdp1- glmS* gametocytes under 0 and 5 mM glucosamine. HDP1 and PhiL1 protein levels in *hdp1-ty1* and Δ*hdp1* day 5 gametocytes. Hsp70-1 abundance shown as the loading control. Representative result of n=3. **(E)** Immunofluorescence microscopy of PhiL1 distribution in day 5 gametocytes of *hdp1-glmS* under 0 and 5 mM glucosamine. Scale Bar: 3 µm

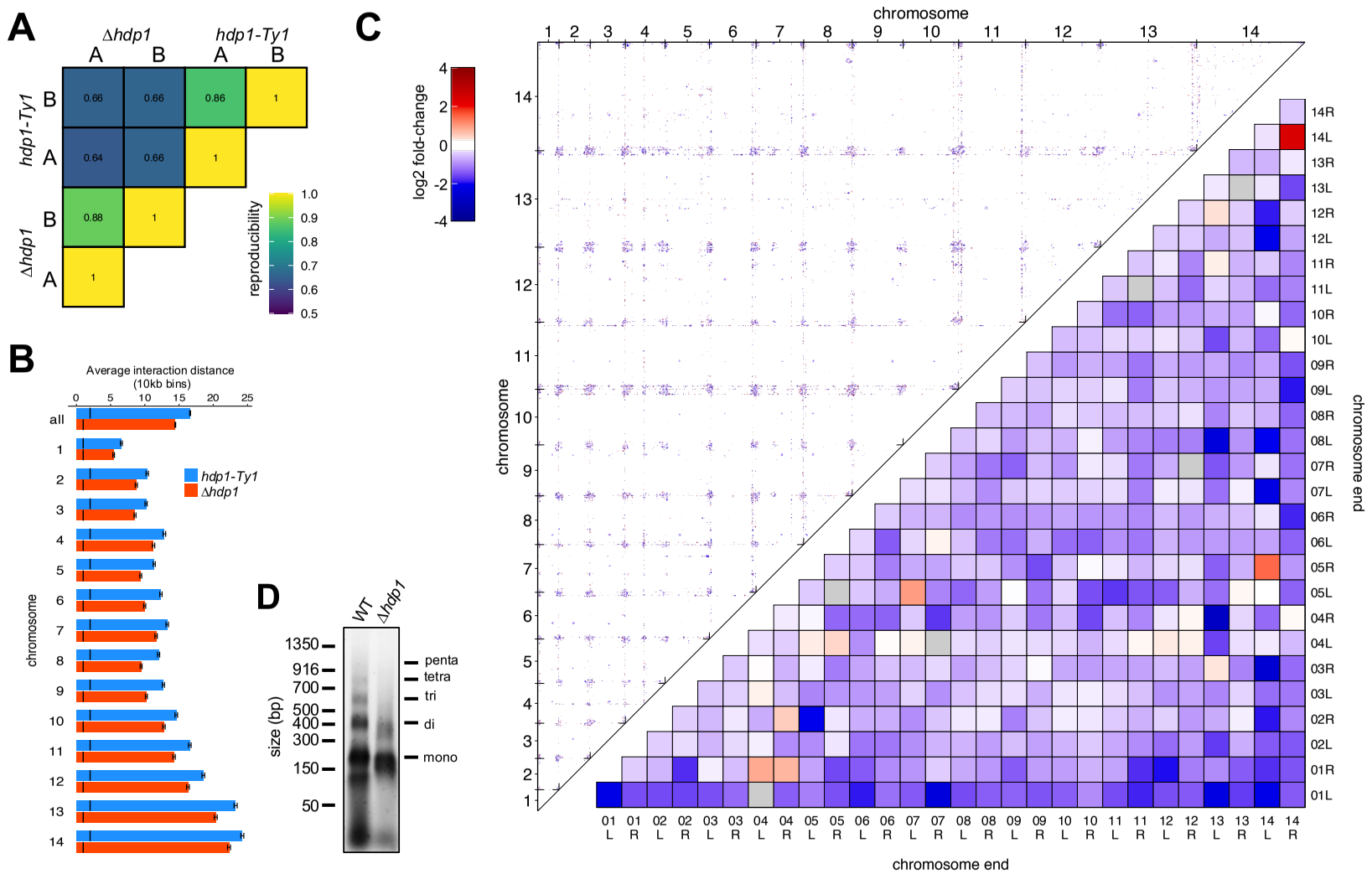


Fig. 5: Loss of HDP1 results in reduced telomere clustering. (A) Analysis of Hi-C interaction maps for each sample by HiCSpector highlights reproducibility of replicates and illustrate systematic differences between the strains. Scale ranges from 0-1.0 but only 0.6-1.0 is shown for improved visibility. **(B)** Mean HiC interaction distance in *hdp1-ty1* (blue) and Δ*hdp1* (red) stage I gametocytes. Error-bars indicate the 95% confidence interval and the black line indicates the median interaction distance. **(C)** Top triangle shows significant genome-wide changes in the Δ*hdp1* compared to *hdp1-Ty1* with inside ticks indicating chromosome boundaries. FDR < 0.1. The average interaction fold-changes for all pair-wise interactions of chromosome ends are shown in the bottom triangle. **(D)** Accelerated MNase digestion of Δ*hdp1* chromatin from stage I gametocytes indicates reduced chromatin compaction.

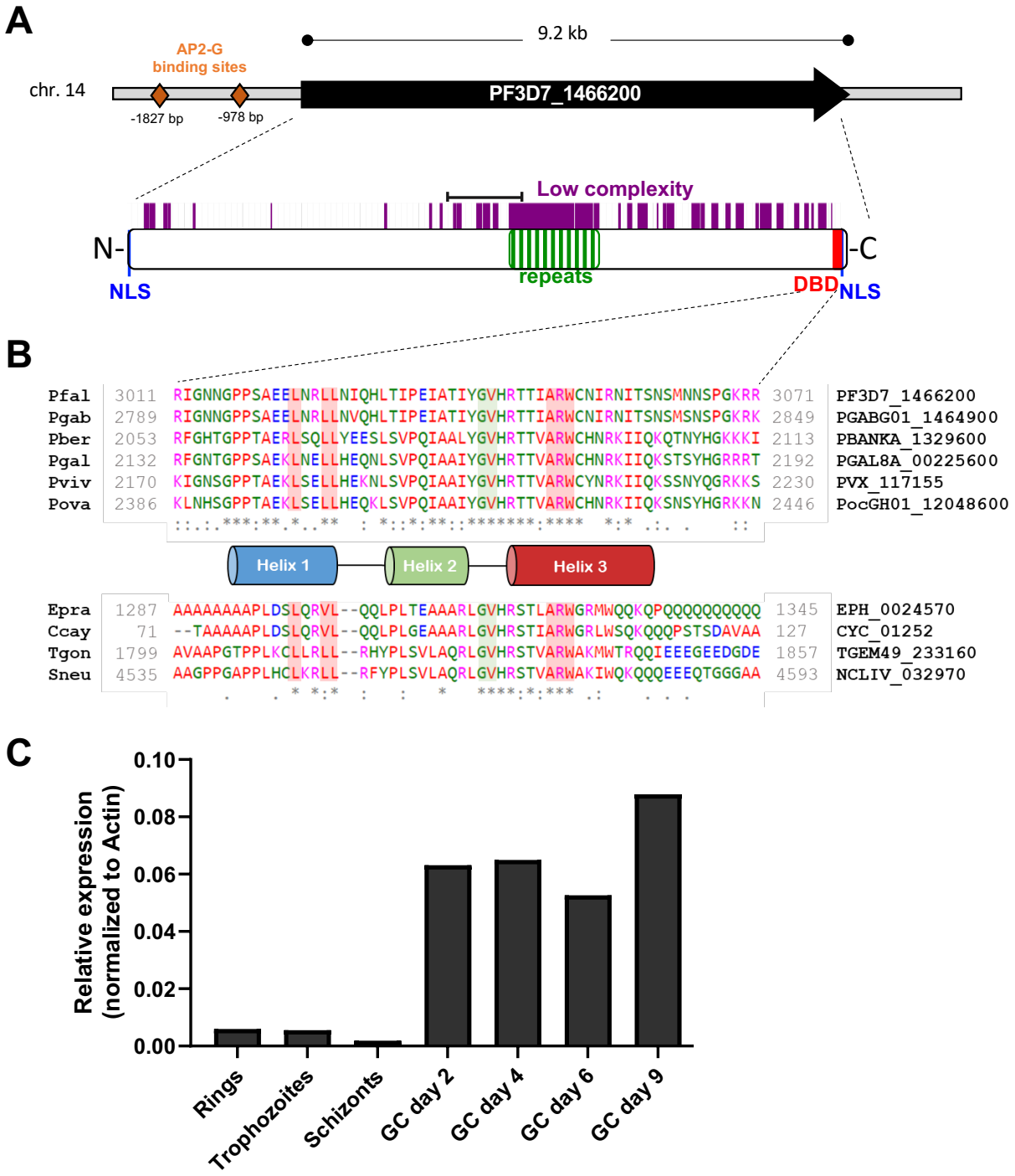


fig. S1: The predicted DNA-binding protein HDP1 is expressed in gametocytes. (A) The single exon locus Pf3D7_1466200 encodes a large 3078aa protein with a predicted C-terminal Helix-Turn-Helix DNA-binding domain (DBD) and two nuclear localization signals (NLS). Multiple AP2-G binding sites are located in the 2kb promoter region. Black bracket indicates the antigen used for generating antisera used in (29). **(B)** Alignment of the helix-turn-helix domain for homologs from other apicomplexan parasites. **(C)** Quantitative RT-PCR of *hdp1* transcripts found minimal expression in asexual blood stages with upregulation during gametocyte development. (mean of n=2).

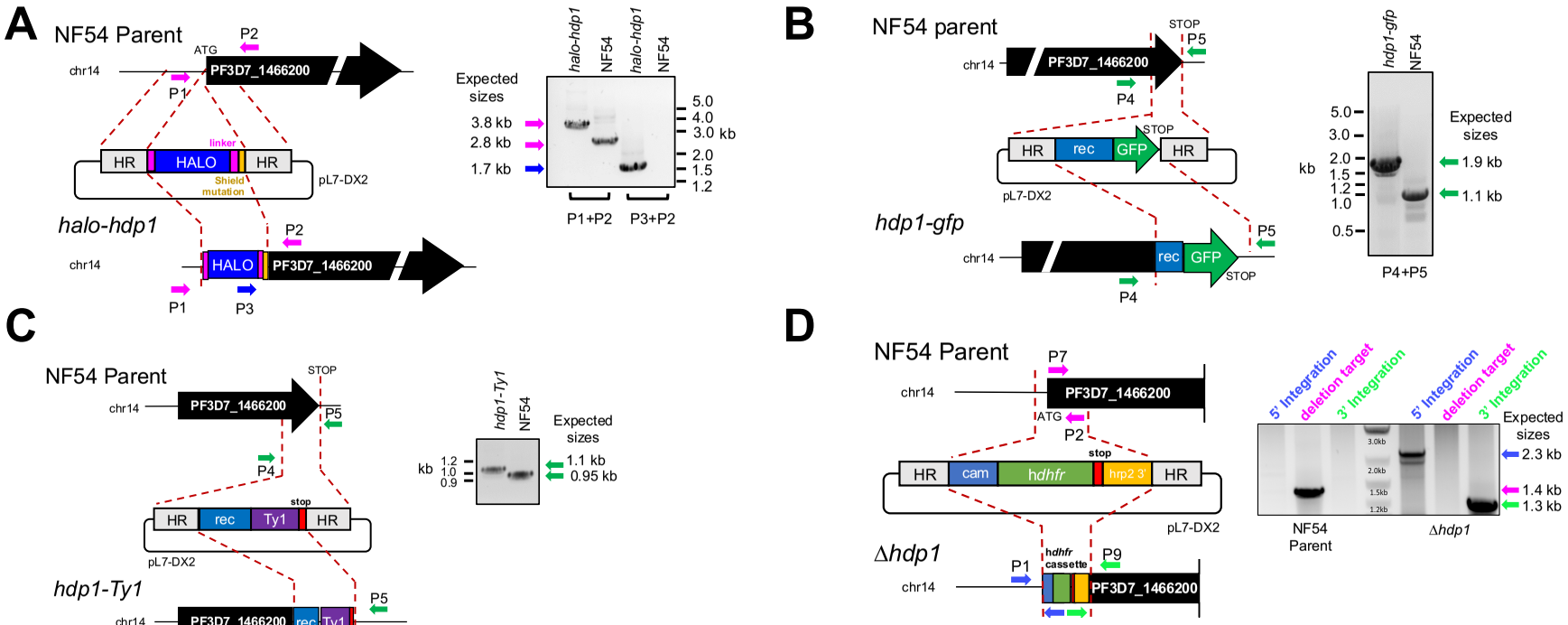
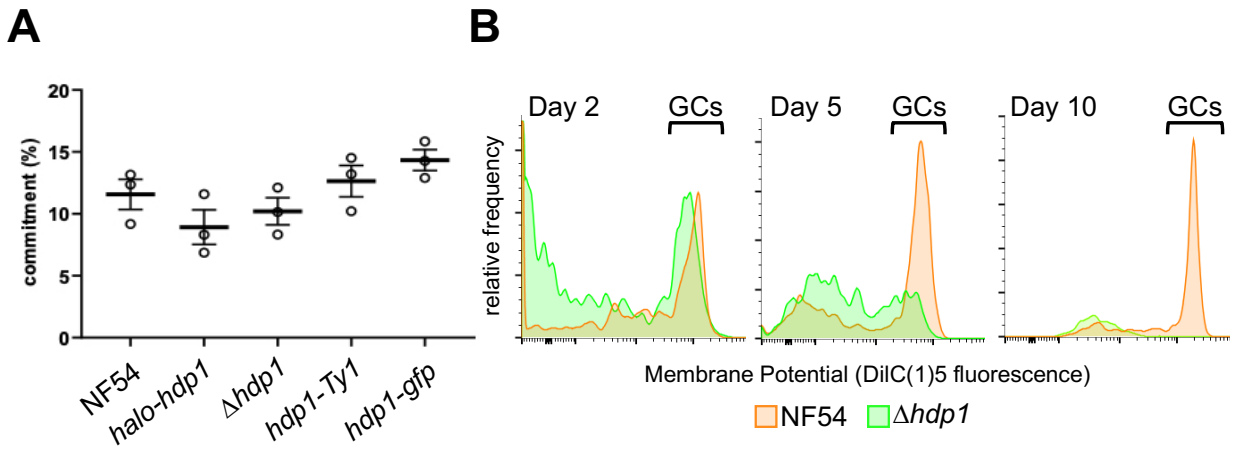


fig. S2: Validation of engineered parasite lines. (A) Generation of h*alo-hdp1* parasites by CAS9 genome editing. Insertion of the N-terminal HALO tag at the 5' end the *hdp1* coding sequence was confirmed by PCR and checked for mutations by Sanger sequencing of the 3.8kb PCR product (not shown). **(B)** Generation of *hdp1-gfp* parasites by CAS9 genome editing. Insertion of the C-terminal GFP tag at the 3' end the *hdp1* coding sequence was confirmed by PCR and checked for mutations by Sanger sequencing of the 1.9kb PCR product (not shown). **(C)** Generation of *hdp1-Ty1* parasites by CAS9 genome editing. Insertion of the C-terminal triple Ty1 epitope tag at the 3' end the *hdp1* coding sequence was confirmed by PCR and checked for mutations by Sanger sequencing of the 1.1kb PCR product (not shown). **(D)** Generation of Δ*hdp1* parasites by CAS9 genome editing. Replacement of 1.4 kb flanking the *hdp1* start codon by a hDHFR selectable marker cassette was confirmed by PCR. **(E)** Generation of *hdp1-glmS* parasites by CAS9 genome editing. Insertion of the C-terminal triple Ty1 epitope tag and the *glmS* ribozyme at the 3' end the *hdp1* coding sequence was confirmed by PCR and checked for mutations by Sanger sequencing of the 1.4 kb PCR product (not shown).

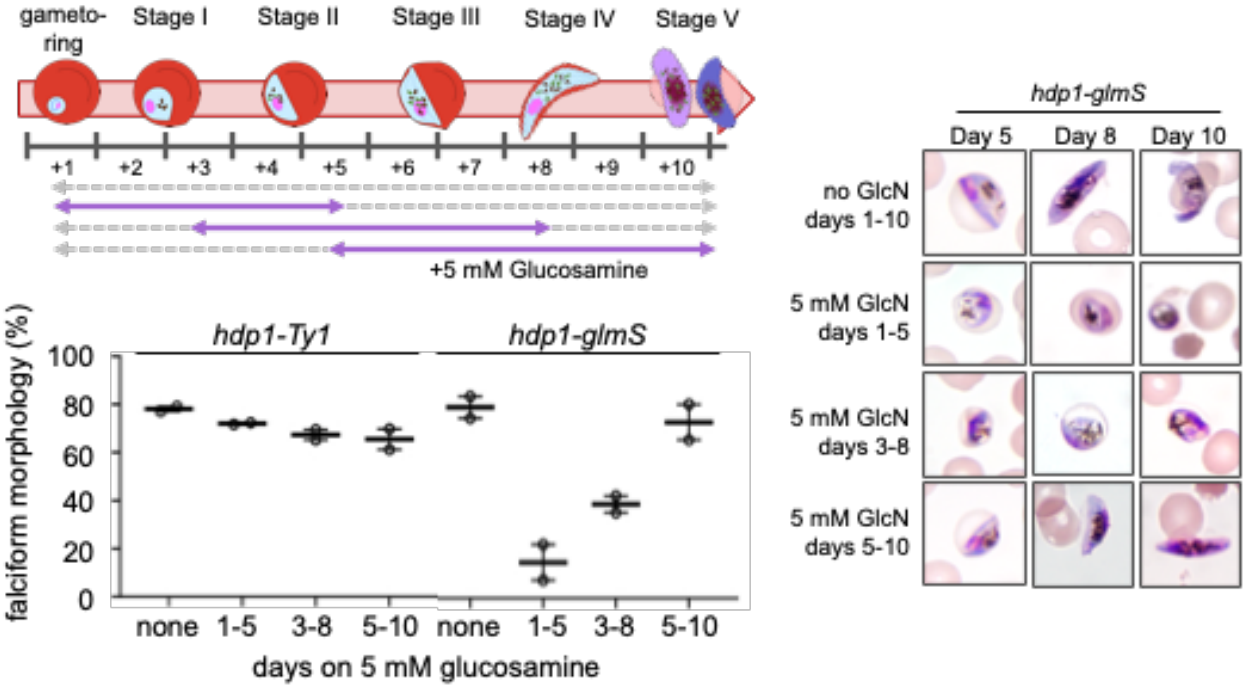
818



819

820 **Fig. S3: Loss of HDP1 does not alter the sexual commitment frequency or Stage I gametocyte**
821 **viability. (A)** The sexual commitment frequency (day 5 gametocytes per day 1 ring stages) is not
822 significantly affected in *halo-hdp1* and Δ*hdp1* parasites. n=3 **(B)** Mitochondrial membrane potential of
823 gametocytes (GCs), as measured by DilC(1)5 staining, indicates similar viability on day 2, but not days 5
824 or 10, for NF54 (orange) and Δ*hdp1* (green) gametocytes. Representative of n=2.

825



826

827 **fig. S4: Inducible knockdown of HDP1 reduces gametocyte maturation in early but not late**
828 **gametocytes.** Representative morphology (right) of *hdp1-glmS* gametocytes in response to 5 mM
829 glucosamine on days 1-5, 3-8, 5-10, or in the absence of glucosamine. Percentage of falciform gametocytes
830 on Day 10 in response to 5 mM glucosamine on days 1-5, 3-8, 5-10, or in the absence of glucosamine for
831 *hdp1-Ty1* or *hdp1-glmS* parasites (bottom). mean±s.e.m of n=2.

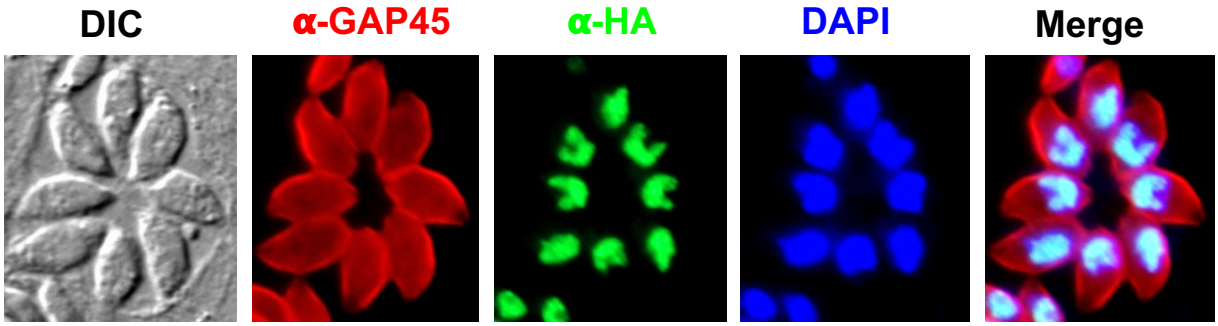


fig. S5: Immunofluorescence microscopy localizes the HA-tagged ortholog TGME49_233160 to the nucleus of *Toxoplasma gondii* tachyzoites.

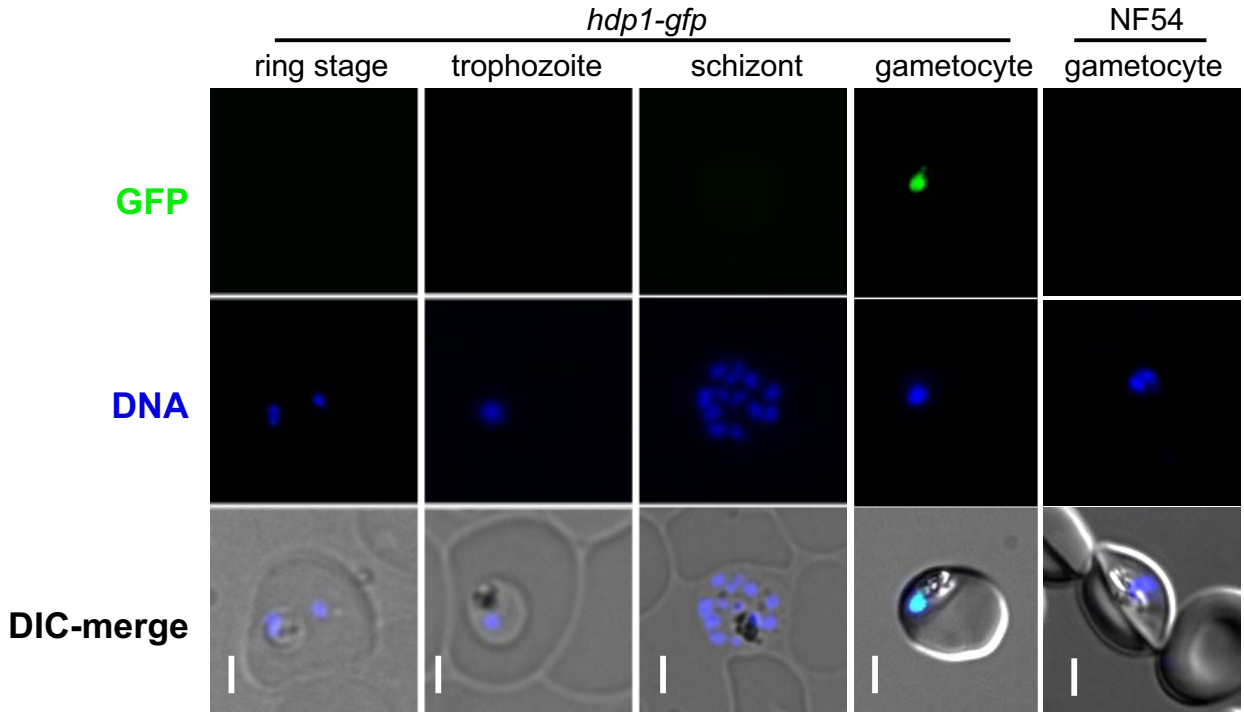


fig. S6: HDP1-GFP localized to the nucleus of *hdp1-gfp* gametocytes. No signal was observed in *hdp1-gfp* asexual blood stages or gametocytes of the untagged NF54 parent line. Scale bar is 3 microns. Exposure and brightness/contrast settings are uniform across the images shown. Representative of n=3.

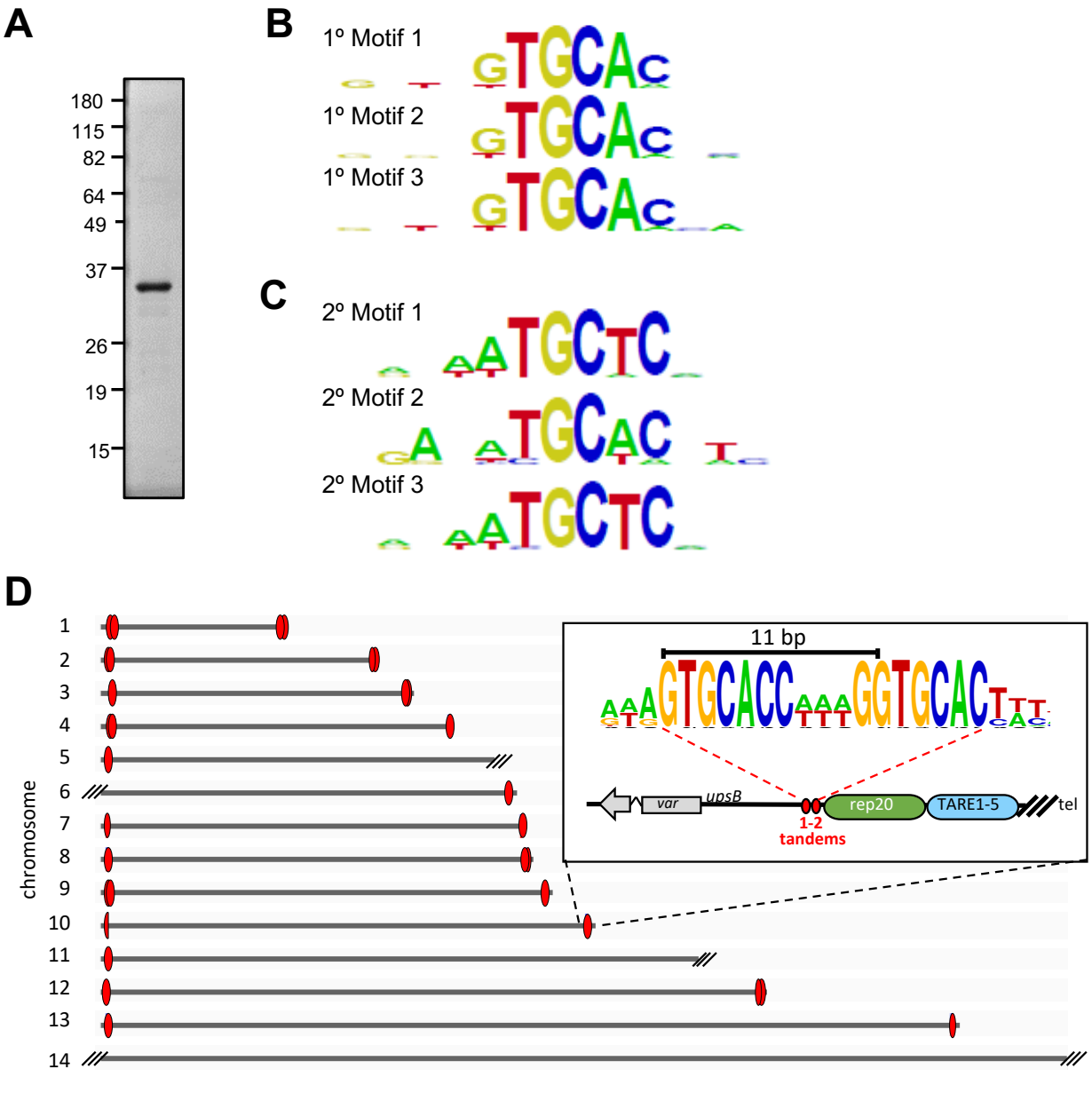


fig. S7: The HDP1 DNA-binding domain dimers recognize a tandem GC-rich motif found at chromosome ends. (A) Coomassie stain of GST-HDP1-DBD used for PBM analysis. **(B)** The three most highly enriched DNA motifs for the GST-HDP1 DBD domain on the protein-binding microarray. **(C)** Top three secondary motif hits after removal of primary hits. **(D)** Tandem HDP1 motifs with a 5-6bp spacer occur exclusively in arrays of 1-2 copies within either 7 bp or 28 bp of the last *rep20* sub-telomeric repeats at all full-length chromosome ends. The motif on the left end of chromosome 10 is degenerate but its positioning is conserved. Truncation by telomere healing resulted in the loss of rep20 adjacent region at five other chromosome ends.

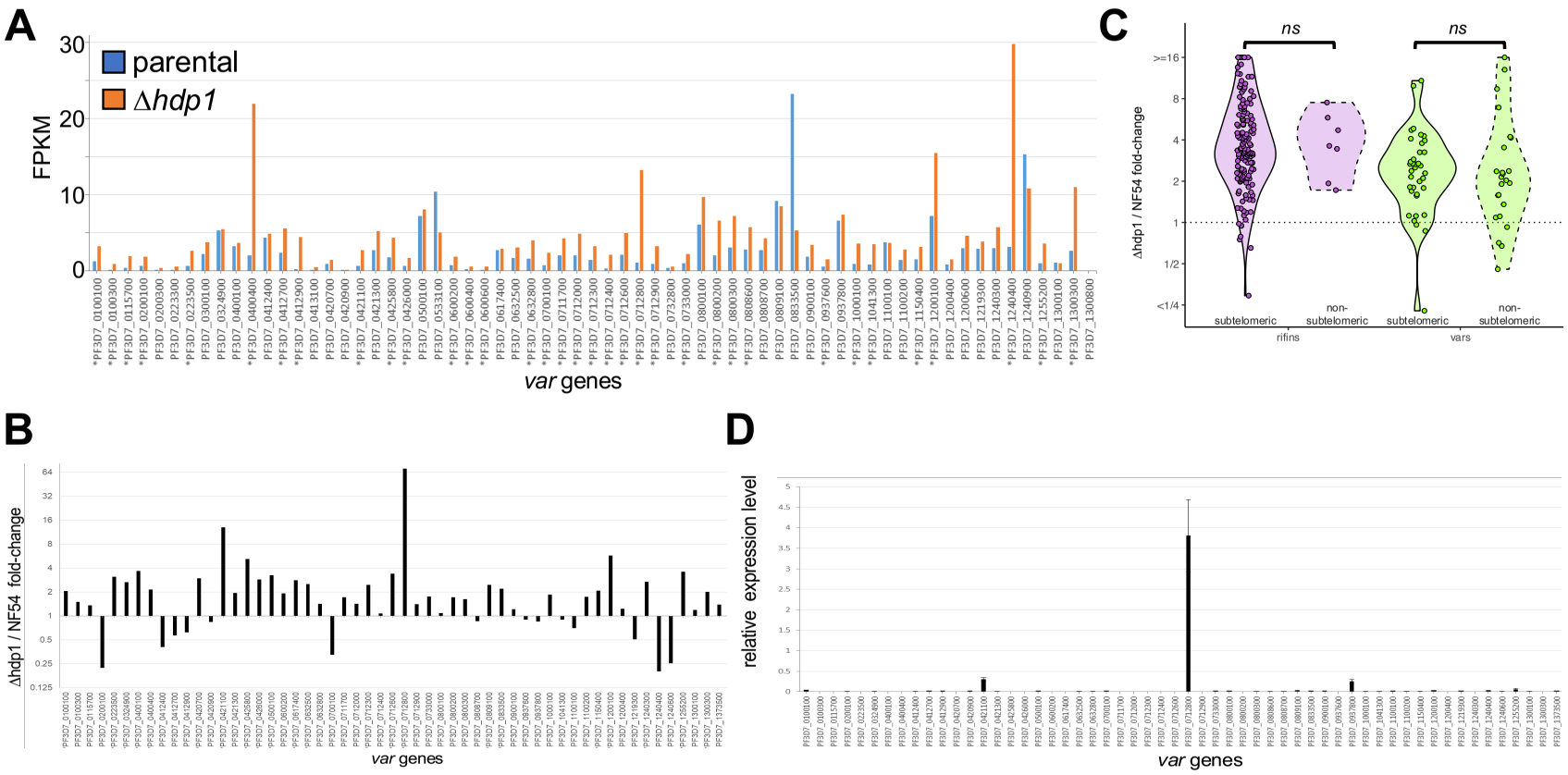
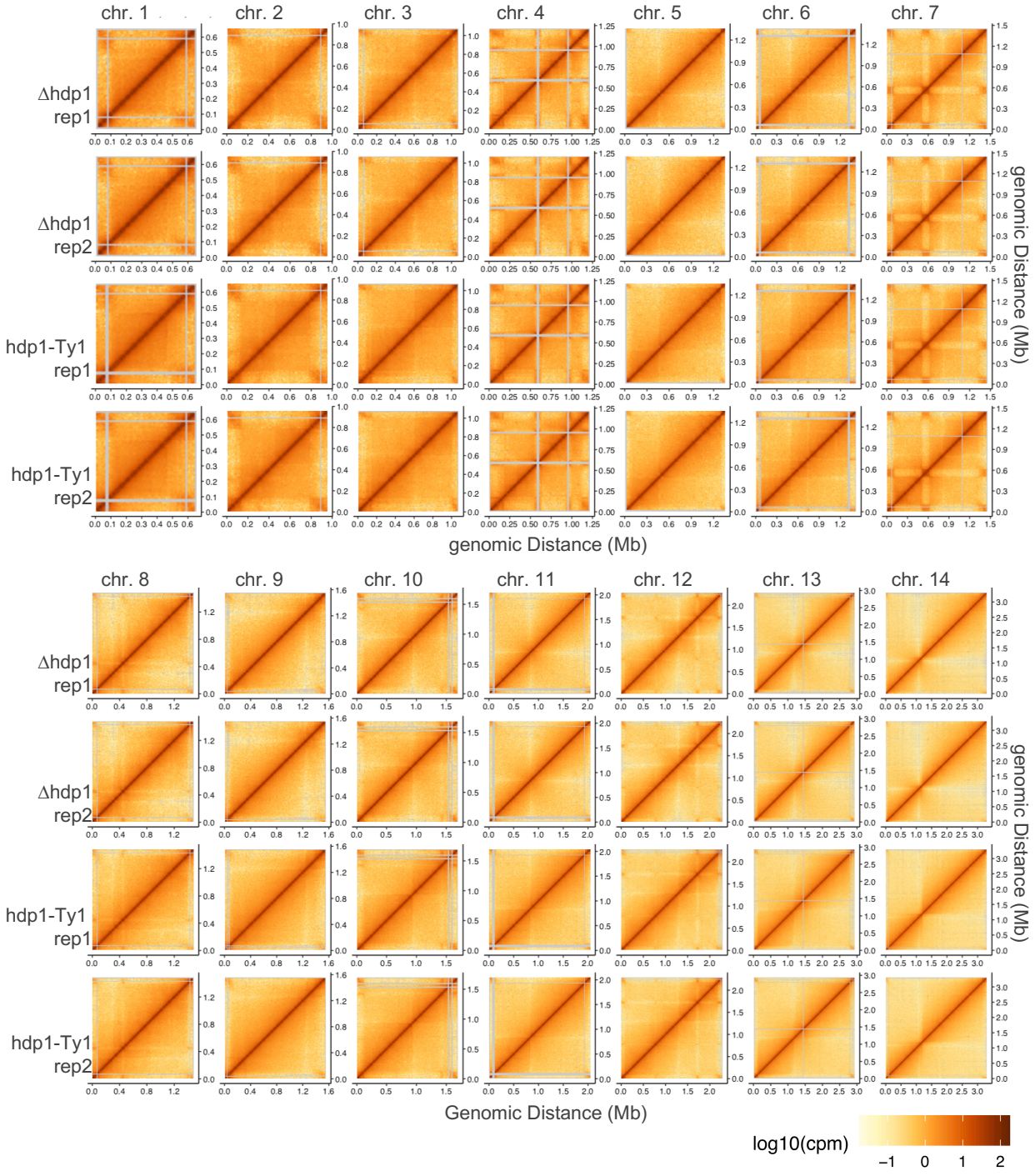


fig. S8: *var* gene expression is altered in early Δ*hdp1* gametocytes but not asexual blood stages. (A) Normalized abundance of reads uniquely mapping to *var* genes in day 2 gametocytes from Δ*hdp1* (orange) or parental NF54 parasites (blue). Significantly upregulated genes are marked with asterisks. (n=2) **(B)** qRT-PCR confirmation of *var* gene up-regulation in Δ*hdp1* vs NF54 day 2 gametocytes. **(C)** Upregulation of *rifin* (purple) and *var* genes (green) in Δ*hdp1* day 2 gametocytes is independent of chromosomal position in subtelomeric or non- subtelomeric heterochromatin clusters. **(D)** qRT-PCR analysis of *var* transcript abundance (normalized seryl tRNA synthetase expression) in Δ*hdp1* asexual ring stages found expression a single dominant var as expected indicating that mutually exclusive expression remains unchanged in asexual stages. (n=2)

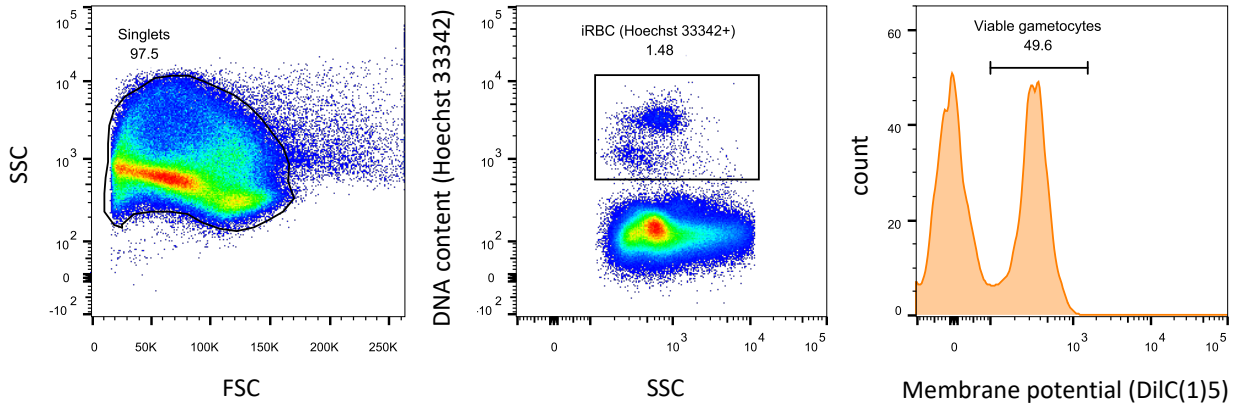
852



853

854 **fig. S9: Intra-chromosomal HiC maps from Δ*hdp1* and *hdp1-Ty1* of day 2 gametocytes.** Color scale
855 indicates the normalized log10 interaction frequency (counts per million). Gray bars indicate bins of high-
856 homology that could not be uniquely mapped with high confidence.

857



858
859 **fig. S10: Gating schema for viable gametocytes.** Populations were gated for single cells based on
860 forward (FSC) and side scatter (SSC). Viable gametocytes were identified based on DNA content and
861 mitochondrial membrane potential based on Hoechst33342 and DilC(1)5 staining.