Homeodomain protein 1 is an essential regulator of gene expression during sexual differentiation of malaria parasites

Riward Campelo Morillo [1], Xinran Tong [1], Wei Xie [2], Todd Lenz [3], Gayani Batugedara [3], Nusrat Tabassum [1], Lindsey M. Orchard [4], Wassim Daher [5], Dinshaw J. Patel [2], William S. Noble [6], Manuel Llinás [4,7], Karine G. Le Roch [3], Björn F.C. Kafsack [1, *]

[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.
* Corresponding Author: bkafsack@med.cornell.edu

ABSTRACT

Transmission of *Plasmodium falciparum* and other malaria parasites requires differentiation from asexual blood stages into male and female gametocytes, a non-replicative sexual stage necessary for transmission to the mosquito vector. This particular transition relies on chromatin reorganization mechanisms that coordinate the silencing and activation of a large number of stage specific genes. However, the transcriptional and chromatin regulators that mediate the changes during sexual differentiation remain largely unknown. Here, we identify the previously uncharacterized Homeodomain Protein 1 (HDP1) as an essential regulator of this process. HDP1 binds DNA in a sequence-specific manner and is tightly associated with chromatin in early gametocytes and required for the critical expansion of the inner membrane complex. Loss of HDP1 leads to deregulation of heterochromatin-associated gene silencing, decreased telomere clustering and increased chromatin accessibility. We propose that HDP1 plays a crucial role in restructuring the parasite's chromatin landscape during early sexual differentiation.

INTRODUCTION

In order to complete its life cycle, *Plasmodium falciparum*, the most widespread and virulent of the protozoan parasites that cause malaria in humans, must differentiate repeatedly into unique cell types that are able access and exploit specialized niches within their human and mosquito hosts. One of these key developmental transitions occurs during the parasite's blood stage. To maintain a persistent infection within the human host, malaria parasites undergo continuous rounds of asexual replication within erythrocytes. However, these replicating asexual cells are not infectious to their mosquito vector and therefore cannot mediate transmission to the next human host. Transmission requires the formation of non-replicating male and female sexual forms, called gametocytes, that mate in the mosquito midgut after being taken up during a blood meal and develop into the motile ookinete stage that infects the mosquito. While most malaria parasite species form round gametocytes within 2-6 days, sexual differentiation in *P. falciparum* takes substantially longer, 12-14 days, and produces gametocytes with a characteristic falciform morphology that give the parasite its name.

All differentiation requires the repression and activation of genes that underlie the specific phenotypes of the origin and destination cell types, respectively. To ensure commitment to one cell type or another, these transitions are often controlled by a bistable switch that controls the activity of a key regulator at the top of the transcriptional cascade that underlies the differentiation program (Bhattacharya et al., 2010; Norman et al., 2015; Park et al., 2012; Satory et al., 2011). The broader downstream changes in gene expression are then accomplished by both changing the availability of specific transcription factors 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 (Kafsack et al., 2014; Sinha et al., 2014). During asexual replication the *ap2-g* locus is silenced by heterochromatin (Brancucci et al., 2014; Kafsack et al., 2014; Lopez-Rubio et al., 2009) but, due to the presence of AP2-G binding sites within its own promoter region, incomplete repression of *ap2-g* in individual cells can activate a transcriptional feedback loop that drives AP2-G expression to high levels,

thereby committing cells to the gene expression program underlying sexual differentiation (Josling et al., 2020; Kafsack et al., 2014; Poran et al., 2017). Under conditions that impair heterochromatin maintenance, this feedback loop is activated in a larger fraction of cells, thereby increasing the frequency of sexual differentiation (Brancucci et al., 2014; Coleman et al., 2014; Filarsky et al., 2018).

Commitment to the sexual differentiation program also triggers a substantial re-distribution of heterochromatin during the early stages gametocyte development (Bunnik et al., 2018; Fraschka et al., 2018). 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 (Tibúrcio et al., 2013). Conversely, heterochromatin is lost from several loci that were previously silenced in asexual stages but upregulated during early gametocyte development (Bunnik et al., 2018; Fraschka et al., 2018).

While AP2-G is critical for the initiation of sexual differentiation, it is expressed only during a small window beginning with sexually committed schizonts and is no longer required after the first 48 hours of gametocyte development (Bancells et al., 2019; Josling et al., 2020). This suggests that the changes in gene expression underlying gametocyte development are carried out by a second wave of transcriptional regulators, which is consistent with the observation that AP2-G upregulates a number of putative transcription factors and chromatin remodeling enzymes (Josling et al., 2020; Poran et al., 2017).

Unsurprisingly, gametocyte maturation over nearly two weeks and five distinct morphological stages is accompanied by a wide array of gene expression changes (Kent et al., 2018; Painter et al., 2017; van Biljon et al., 2019; Young et al., 2005). However, relatively little is known about the transcriptional regulators downstream of AP2-G that mediate these changes. Compared to other single-celled eukaryotes, the genomes of malaria parasites contain a small fraction of DNA binding proteins. Like AP2-G, most belong to the ApiAP2 family (Balaji et al., 2005), but only a small number have been shown to function specifically during gametocyte development (Modrzynska et al., 2017).

92 In the rodent malaria parasite *P. berghei* the DNA-binding protein PbAP2-G2 functions as
93 a transcriptional repressor of asexual-specific gene expression (Sinha et al., 2014; Yuda
94 et al., 2015), while PbAP2-FG (PyAP2-G3 in *P. yoelii*) was shown to mediate upregulation
95 of female-specific transcripts (Yuda et al., 2019; Zhang et al., 2017). Whether these
96 functions are conserved in *P. falciparum* remains to be determined.

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

105 **RESULTS**

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

107 Searching for possible regulators of gene expression during *P. falciparum* sexual
108 differentiation, we identified Homeodomain-like protein 1 *(hdp1, PF3D7_1466200)*
109 encoding a 3078 amino acid protein with a C-terminal homeodomain-like domain, a helix-
110 turn-helix structural motif commonly involved in DNA binding (Figure 1A) (Bürglin and
111 Affolter, 2016). Syntenic orthologs of *hdp1* could readily be identified in other malaria
112 parasites with homology to the homeodomain-like domain extending to the coccidia but
113 apparently absent from other apicomplexan clades (Figure 1B). Analysis of *hdp1*
114 expression in *P. falciparum* blood stages by qRT-PCR detected only minimal expression
115 in asexual blood stages with substantial upregulation during the early stages of
116 gametocytogenesis (Figure 1C). AP2-G, the transcriptional master switch that initiate the
117 sexual differentiation gene expression program (Kafsack et al., 2014), binds at two sites
118 located upstream of *hdp1* coding sequence early during the gametocytogenesis (Josling
119 et al., 2020), consistent with our hypothesis that AP2-G activates additional regulators of
120 gene expression during early gametocytogenesis.

121 In an effort to determine HDP1's subcellular localization, we inserted an in-frame N-
122 terminal HaloTag at the endogenous locus (Figure S1A) in order to avoid possible
123 interference with the putative DNA-binding activity of the C-terminal homeodomain-like
124 domain. As expected, no Halo-tagged protein could be detected in asexual stages (data
125 not shown). However, when we attempted to determine Halo-HDP1 levels, we found that
126 *halo-hdp1* cultures were unable to produce the characteristic crescent shapes of maturing
127 *P. falciparum* gametocytes (Figure 2A-B). Tagging the HDP1 C-terminus with either GFP
128 or a triple Ty1 epitope tag (*hdp1-gfp* and *hdp1-Ty1*, Figure S1B-C), yielded parasite lines
129 that produce gametocytes indistinguishable in numbers and morphology from the wildtype
130 parent (Figure 2A-B), despite the proximity to the putative DNA-binding domain.

131 To test whether insertion of the N-terminal tagging resulted in a loss of HDP1 function,
132 we generated a HDP1 knockout line for comparison. We used genome editing to replace
133 1.4 kb at the 5' end of the *hdp1* locus with a selectable marker cassette (Figure S1D).
134 The resulting Δ*hdp1* parasites exhibited no discernible change in phenotype in asexual
135 blood-stages, but like the *halo-hdp1* parasites, were unable to generate falciform
136 gametocytes (Figure 2D-E). More detailed analyses using synchronous induction of
137 gametocytogenesis found that both *halo-hdp1* and Δ*hdp1* have sexual commitment rates
138 comparable to the NF54 parent (Figure S2A) but are unable to complete gametocyte
139 development. We found that on day 2 of development Δ*hdp1* Stage I gametocytes had
140 normal levels of viability (Figure S2B) as indicated by robust mitochondrial membrane
141 potential but had lost the ability to maintain this potential by day 5 of development when
142 gametocytes from other lines had matured to stage III (Figure 2C, F).

143 As the substantial length of the *hdp1* coding sequence made genetic complementation
144 infeasible, we generated an inducible *hdp1* knockdown parasite line by inserting a triple
145 Ty1 epitope tag followed by with the autocatalytic glmS ribozyme at the 3' end of the
146 endogenous *hdp1* coding sequence through genome editing (Figure S1E) (Prommana et
147 al., 2013). In the absence of glucosamine, the resulting *hdp1-glmS* gametocytes
148 expressed HDP1 protein at levels comparable to the *hdp1-Ty1* parasites without the
149 ribozyme and produced normal gametocytes in term of number and morphology (Figure
150 S3). Supplementation of the culture medium with 5mM glucosamine during the first 5 days

151 of gametocyte development had no discernible effect on *hdp1-Ty1* parasites but
152 diminished HDP1 expression by 70% and reduced falciform gametocytes by 80% in *hdp1-*
153 *glmS* parasites, recapitulating the phenotype of the *halo-hdp1* and Δ*hdp1* lines (Figure
154 2G-H). Since *hdp1* transcript levels remain relatively constant during gametocyte
155 development (Figure 1C), we wanted to test whether continuous expression of HDP1 was
156 required for gametocyte maturation. Knockdown of *hdp1* levels with glucosamine reduced
157 gametocyte maturation when added prior to day 5 but not thereafter, indicating that
158 sufficient HDP1 protein had been produced by that time to support gametocyte maturation
159 (Figure S3).

160 **HDP1 localizes to the nucleus of early gametocytes.**

161 An earlier study had reported a variety of subcellular localizations for HDP1, including
162 export to the erythrocyte membrane of early gametocytes, based on antibodies raised
163 against the low-complexity repeat region (Nixon et al., 2018). This was surprising given
164 the absence of a predicted signal peptide or transmembrane domains and localization of
165 the *Toxoplasma gondii* ortholog to the nucleus of tachyzoites (Figure S4). To resolve this
166 apparent disagreement, we carried out live-cell and immunofluorescence microscopy in
167 gametocytes of *hdp1-gfp* and *hdp1-Ty1*, respectively. In each case, HDP1 clearly
168 localized to the gametocyte nucleus (Figure 3A-B). Analysis of HDP1 protein levels by
169 western blotting of *hdp1-Ty1* lysates with antibodies against the Ty1 epitope tag detected
170 no expression in asexual blood stages but showed increasing HDP1 levels from day 2 of
171 gametocytogenesis (Stage I-II) onward, reaching maximal levels by day 5 (Stage III)
172 before dropping again by day 8 (Stage IV) (Figure 3C). Analysis of subcellular fractions
173 from day 5 *hdp1-Ty1* gametocytes showed HDP1 almost exclusively in the parasite
174 nucleus, where about 70% is resistant to solubilization with up to 600mM NaCl, indicating
175 a tight association with chromatin (Figure 3D) and validating HDP1 as a nuclear protein.

176 **The HDP1 homeodomain-like domain dimerizes to efficiently bind DNA *in vitro.***

177 To evaluate HDP1's ability to bind DNA, we expressed a GST-fusion of the HDP1 homeo-
178 domain in *E. coli.* Following protein isolation, we found that much of the fusion protein
179 remained bound to bacterial DNA but could be released by the addition of the poly-cation
180 polyethylenimine during isolation (*data not shown*). Analysis of the domain's sequence

181 specificity by protein-binding microarray showed a preference for the palindromic
182 hexamer GTGCAC (Figure 4A, Figure S5). Since homeo-domains generally recognize
183 DNA as dimers (Bürglin and Affolter, 2016; Pradhan et al., 2012), we carried out
184 isothermal titration calorimetry to measure the interaction of HDP1 with double-stranded
185 DNA containing a tandem motif with a 5 bp spacer that places the motifs one helical turn
186 apart. The results indicated optimal binding at 2:1 protein to DNA molar ratio with a
187 dissociation constant of 2.8 μM supporting DNA recognition as a dimer (Figure 4B). Next,
188 we performed DNA gel-shift assays with double-stranded probes containing either no
189 match, a single binding-motif, or a tandem motif. When compared to the tandem-motif
190 probe, the gel-shift of the single motif probe was substantially weaker but identical in size
191 (Figure 4C), again consistent with DNA recognition as a dimer. Intriguingly, a scan of the
192 3D7 reference genome for instances of GTGCAC tandem motifs with a 5-6 bp spacer
193 found them exclusively located at each chromosome end with the first tandem located
194 within either 7 or 28bp of the ultimate *rep20* sub-telomeric repeat, with the exception of
195 those ends where this region was deleted as the result of telomere healing events
196 (Calhoun et al., 2017) (Figure 4D, Data Set 1). Earlier work had predicted these
197 conserved sequences to be recognized by the ApiAP2-domain protein SIP2 in asexual
198 blood stages based on its ability to bind them *in vitro* (Flueck et al., 2010), however, their
199 unique spacing and palindromic nature clearly distinguishes these tandem motifs from
200 the arrays of SIP2 binding sites just downstream. While the function of these motifs in
201 asexual and sexual blood stages remains unclear, their exclusive localization at
202 chromosome ends is highly conserved across *P. falciparum* isolates (*data not shown*).

203 **Loss of HDP1 leads to dysregulation of gene expression during early gametocyte**
204 **development.**

205 Based on HDP1's ability to bind DNA, we wanted to test whether it plays a role in the
206 regulation of gene expression during early gametocytogenesis. To minimize the chance
207 of looking at the consequences of aberrant gametocyte development we decided to
208 sequence total RNA from Δ*hdp1* and parental NF54 Stage I gametocytes on Day 2 of
209 development when HDP1 is first expressed but prior to any change in viability or
210 morphology in HDP1 deficient parasites (Figure 2D, Figure S2B).

211 Global comparison of transcript abundances by RNA-seq showed that the majority of
212 genes were expressed at similar levels including the markers of very early gametocytes
213 *pfs16* and *gexp5* (Figure 5A), supporting our earlier observation that HDP1-deficient
214 parasites initiate gametocyte development at similar rates. However, we observed
215 reduced transcript levels for 188 genes and increased levels for 150 genes in Δ*hdp1*
216 Stage I gametocytes compared to those of the NF54 parent line. Gene set enrichment
217 analysis (GSEA) found that transcripts encoding components of the inner membrane
218 complex were significantly down-regulated (Figure 5B) while transcripts from
219 heterochromatin-associated multi-copy gene families were significantly over-represented
220 among up-regulated genes (Figure 5C), including members of the *var*, *rifin*, *stevor*, and
221 *PHISTa/b/c* gene families (Figure 5D, Fig S6A-B). While fold-changes for these genes
222 could be substantial, the absolute increase in transcripts for these genes was often
223 relatively small but significantly above levels in WT cells (Data Set 2), possibly indicating
224 impaired heterochromatin-mediated silencing allowing for leaky background transcription
225 rather than full transcriptional activation. Comparison of transcript levels of *var* and *rifin*
226 family genes found similar levels of upregulation regardless of whether these were
227 located in subtelomeric or non-subtelomeric regions of heterochromatin (Figure S6C).
228 Analysis of *var* gene expression in asexual Δ*hdp1* ring-stages found expression of a
229 single major *var* gene as expected for a recently cloned parasite line and indicating that
230 mutually exclusive *var* gene expression remained unaffected (Figure S6D).

231 **Loss of HDP1 increases chromatin accessibility and decreases telomeric**
232 **clustering in early gametocytes.**

233 Despite numerous attempts with the GFP- and Ty1-tagged lines, we were unable to
234 determine genome-wide binding of HDP1 by chromatin immunoprecipitation sequencing.
235 Instead, we used high-throughput chromatin-conformation capture (Hi-C) to look for
236 differences in chromatin organization of day 2 gametocytes between the NF54 and Δ*hdp1*
237 lines. While correlation analysis of Hi-C replicates indicated a clear difference between
238 these lines (Figure 6A), overall intra-chromosomal interactions appeared largely
239 unchanged (Figure S7). However, further statistical analysis revealed a genome-wide
240 reduction in long distance interactions in Δ*hdp1* (Figure 6B). Genome-wide mapping of

241 interaction changes revealed a substantial reduction in interaction frequency between
242 virtually all chromosome ends (Figure 6C), consistent with reduced but not abrogated
243 telomere clustering. Based on the small but significant upregulation of heterochromatin
244 associated genes and a reduction in telomeric clustering, we hypothesized that chromatin
245 compaction might be impaired in HDP1-deficient parasites resulting in increased
246 chromatin accessibility. Indeed, brief treatment with micrococcal nuclease (MNase)
247 revealed accelerated digestion of chromatin from Δ*hdp1* gametocytes (Figure 6D),
248 consistent with an increase in chromatin accessibility.

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

251 Among the genes with reduced transcript levels in HDP1 knockout parasites, those
252 associated with the inner membrane complex were substantially enriched (Figure 5A-B,
253 Figure 7A). This was notable since Δ*hdp1* gametocytes fail to elongate from spherical
254 stage I gametocytes into the oblong Stage II, a process that depends on the extension of
255 the inner membrane complex. Six separate inner membrane complex genes were
256 confirmed by qRT-PCR to have greater than 2-fold lower transcript levels in HDP1
257 knockout gametocytes (Figure 7B). These included PhIL1 and the PhIL1-interacting
258 protein 1 (PIP1), both of which are highly upregulated in early *P. falciparum* gametocytes
259 and essential for inner membrane complex assembly and gametocyte maturation (Parkyn
260 Schneider et al., 2017). Strikingly, the spherical morphology of *hdp1*-deficient
261 gametocytes perfectly resembles the phenotype of gametocytes when PhIL1 is knocked
262 down. Comparison of PhIL1 protein levels showed a 75% percent reduction in Δ*hdp1*
263 gametocytes compared to wild-type parent (Figure 7C). The concomitant reduction in
264 PhIL1 expression (Figure 7D) and inner membrane complex extension (Figure 7E) when
265 HDP1 is knocked down with glucosamine in *hdp1-glmS* gametocytes confirmed the
266 dependence of PhIL1 expression on HDP1.

267

DISCUSSION

Our understanding of the regulatory mechanisms that control sexual differentiation in *P. falciparum* has improved substantially in recent years. Much of this work has focused on the regulation of AP2-G, the master regulator of this developmental decision. However, it is becoming clear that the function of AP2-G is constrained to the initiation of the transcriptional program that underlies the nearly two week-long process of forming gametocytes (Bancells et al., 2019). This suggests that a second wave of hitherto unknown transcriptional regulators is required to drive gametocyte-specific gene expression during early gametocyte development.

In this study, we demonstrate that HDP1 functions as an essential regulator during sexual differentiation. A nuclear DNA-binding protein, HDP1 is specifically up-regulated during early gametocyte development and essential for their maturation. In particular, HDP1 regulates the expression of genes required for the expansion of the inner membrane complex that gives *P. falciparum* gametocytes their characteristic crescent shape. While HDP1 expression in *P. falciparum* is gametocyte specific, its *P. berghei* ortholog (PBANKA_1329600) may also function during asexual blood-stage replication as its disruption leads to significant reduction in growth. This suggests that the gametocyte-specific function of HDP1 may have evolved more recently in the Laveranian clade of malaria parasites which produce crescent gametocytes. While absent from the asexual blood-stages and essential during early gametocyte development, whether HDP1 is required in other parasite stages remains to be determined. Since the knockdown system used in this study regulates expression at the transcript level, we can infer that additional translation of HDP1 is not required during the later stages of gametocyte maturation but HDP1 protein expressed during the earlier stages may well be required throughout gametocyte development. Previous transcriptomic studies indicate that HDP1 is also expressed in ookinetes, suggesting a role even following mating. Intriguingly, homeodomain-like proteins have been implicated in mating processes of other haploid protozoa, such as Dictyostelium (Hedgethorne et al., 2017).

We also found that HDP1 is required for efficient heterochromatin-mediated gene silencing during sexual differentiation. In gametocytes lacking HDP1, we observed low

298 but significant up-regulation across a wide array of heterochromatin-associated gene
299 families. Chromosome conformation capture analysis showed broad alterations to
300 chromatin organization in HDP1-deficient parasites with an overall reduction in long-
301 distance interactions that was most pronounced at the end of chromosomes, consistent
302 with reduced, but not ablated, clustering of the sub-telomeric heterochromatin regions in
303 the nuclear periphery. While its cognate DNA motifs are exclusively found at the boundary
304 of the subtelomeric repeats, localization of HDP1 appeared evenly distributed throughout
305 the gametocyte nucleus rather than in foci at the nuclear periphery, suggesting that the
306 HDP1 does not localize to these sites exclusively. One explanation is that HDP1 promotes
307 chromatin packaging, leading to an overall increase in chromatin accessibility observed
308 in HDP1 deficient cells. This would also explain the leaky expression of heterochromatin-
309 associated genes as efficient silencing depends on limiting promoter accessibility while
310 the promoters of euchromatin-associated genes are already generally accessible and a
311 reduction in packaging would therefore be predicted to have little impact on transcript
312 levels.

313 Additional studies will focus on identifying interaction partners of HDP1 and elucidating
314 its the structure-function relationship. Only 2% of its protein sequence is comprised of the
315 homeodomain-like DNA-binding domain while the remainder contains no other
316 identifiable domains and conservation is generally weak outside this region. The fact that
317 insertion of a large tag at the N-terminus results in loss of function indicates that critical
318 interactions occur in this region but which other regions are essential for HDP1's function
319 remains unclear. Similarly, identifying interactions with other nuclear proteins and
320 genomic locations will offer important insights into its function. Of particular interest will
321 be whether HDP1 regulates gene expression by altering chromatin structure or is more
322 directly involved in the recruitment of RNA polymerase.

323

MATERIALS & METHODS

Parasite culture

Unless otherwise noted, parasite strains were grown in 0.5% AlbuMAX II supplemented malaria complete medium using stablished cell culture techniques (Moll et al., 2008) at 3% hematocrit and below 3% parasitemia. Strains expressing selectable markers were maintained under constant drug-selection. Toxoplasma tachyzoites were cultured as described in (Morlon-Guyot et al., 2018).

Gametocyte induction

Gametocytes were induced synchronously as previously described in (Poran et al., 2017). Gametocyte maturation was monitored by Giemsa-stained thin blood smears and gametocytemia was counted on the fifth day of development. The gametocyte commitment rate was calculated by dividing the day 5 gametocytemia by the day 1 parasitemia, counted before addition of N-acetyl-D-glucosamine. Gametocytes were purified from culture at the required development stage using magnetic columns (LS columns, Miltenyi Biotec). For knockdown experiments in the HDP1-glmS line, at the gametoring stage and either 5 mM glucosamine or solvent control was added.

Generation of transgenic strains

Transfection of ring-stage parasites were performed as previously described (Rug and Maier, 2012) and selected on 4nM WR99210 and 1.5 μM DSM1 for two weeks. To obtain Δ*hdp1* parasites, selection with WR99210 was kept all the time. Genome editing was performed by CRISR/Cas9 technology using the system described by (Ghorbal et al., 2014). The pL6-DX2 plasmid, containing the sgRNA-expression cassette and the positive selectable marker cassette (*cam* promoter-*hdhfr-hrp2*/ts 3'*utr*) was used as backbone. Homology boxes were PCR amplified from NF54 genomic DNA, using specific primers flanked by appropriate overlapping region from pL6-DX2 plasmid (list of primers, see Supplementary Table) that allowed the cloning by Gibson assembly. All PCR reactions were set using Advantage Genomic LA polymerase (Takara, Cat. N° 639152). *Plasmodium* codon optimized sequences for HALO-tag and triple Ty1 epitope tag were synthetized as gene-Blocks (Genewiz). PCR products were purified from agarose gel using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Cat. N° D4008). Homology boxes were cloned through *AflII* and *SpeI* cutting sites, while annealed oligonucleotides paired encoding sgRNA with targeting sequence (Table) were inserted through the *XhoI* site, placed upstream of the sgRNA scaffold sequence. Sanger sequencing confirmed the absence of undesired mutations in the homology boxes and the sgRNA. Genomic DNA from transfectant parasites was isolated

356 with QIAamp DNA blood Kit (Qiagen, Cat. N° 51106) and diagnostic PCR's were set using Taq
357 phusion DNA polymerase (Invitrogen). The TGME49_233160-HA parasite line was generated as
358 part of an earlier study by tagging of the endogenous locus in the *T. gondii* RH-ku80ko strain as
359 described in (Morlon-Guyot et al., 2018).

360 **Flow cytometric analysis of gametocyte viability**

361 Gametocytes were stained with 16 µM Hoechst33342 and 50 nM DilC1 for 30min at 37˚C. Using
362 a Cytek DxP12 flow cytometer, gametocytemia was determined by gating for DNA-positive,
363 hemozoin-high cells and gametocyte viability was inferred based on mitochondrial membrane
364 potential dependent accumulation of DilC1(5) for 1000 gametocytes (Tanaka and Williamson,
365 2011).

366 **Nuclear extract preparation and high salt fractionation**

367 Nuclear enriched fraction was prepared following the protocol previously described in (Flueck et
368 al., 2009), with some modifications. Briefly, parasites released from RBC's by saponin treatment
369 (0.01%) were lysed with ice-chilled CLB (20mM HEPES pH 7.9; 10mM KCl; 1mM EDTA pH 8.0;
370 1mM EGTA pH 8.0; 0.65% NP-40; 1mM DTT, 1x Roche Complete protease inhibitors cocktail).
371 Nuclei were pelleted at 3,000 x *g* for 20min at 4°C and cytoplasmic fraction was removed. Nuclei
372 were resuspended in digestion buffer (20mM Tris- HCl, pH 7.5, 15mM NaCl, 60mM KCl, 1mM
373 $CaCl_2$, 5mM $MgCl_2$, 300mM sucrose, 0.4% NP-40, 1mM DTT, 1x Roche Complete protease
374 inhibitors cocktail EDTA-free) and treated with 5U of micrococcal nuclease (ThermoFisher, Cat.
375 N° 88216) for 30 min in a water bath at 37°C. Soluble and insoluble nuclear fractions were
376 recovered by centrifugation at 3,000 x *g* for 10 min at 4° C. Insoluble nuclear fraction were treated
377 with salt fractionation buffer (10mM Tris HCl pH 7.4; 2mM $MgCl_2$; 2mM EGTA pH 8.0; 0.1% Triton
378 X-100; 0.1mM PMSF; 1x Roche Complete protease inhibitors cocktail) supplemented with
379 increasing NaCl concentrations (80-600mM) while rotating at 4°C for 30 min. All supernatants
380 were recovered by centrifugation at 700 x *g* for 4 min at 4°C and last remaining pellet was
381 resuspended in 1x PBS supplemented with protease inhibitors cocktail. 5% of each fraction was
382 prepared for Western blotting in order to check quality of fractionation.

383 **Immunoblotting**

384 For SDS-PAGE, total protein lysates were prepared using saponin-lysed parasites resuspended
385 with 1x Laemmli loading buffer diluted in 1x PBS supplemented with 1x Roche Complete protease
386 inhibitors cocktail. Protein samples were separated in 4-15% polyacrylamide gels and transferred
387 to 0.2 µm Immobilion-P^{SQ} transfer membrane (Millipore, Cat. No ISEQ00010) using a Bio-Rad

388 transfer system. Membranes were blocked in 5% skim milk/1x TBS-Tween20 for 1 hour at RT.
389 Primary and secondary antibodies were prepared in 3% skim milk/1x TBS-Tween20 and
390 incubated for 1 hour at RT. Membranes were washed four times with 1x TBS-Tween20 for 10
391 min, after primary and secondary antibody incubations. The following primary antibodies were
392 used in this study: Anti-Ty1 BB2 mouse (1:2,500; Invitrogen Cat. Nº MA5-23513), anti-PhIL1 rabbit
393 (1:5,000 (Saini et al., 2017)), anti-PfHsp70 rabbit (1:5,000; StreesMarq Biosciences Cat. Nº SPC-
394 186D), anti-Histone 4 rabbit (1:2,000; Diagenode Cat. Nº C15410156-50). HRP-conjugated anti-
395 mouse and anti-rabbit antibodies were used (1:5,000, Millipore). Immunoblots were incubated
396 with the chemiluminescent substrate SuperSignal West Pico PLUS (ThermoFisher, Cat. Nº
397 34578) following manufacturer directions. Images were obtained using Azure c300 digital imaging
398 system (Azure Biosystems).

399 **Live-cell and Immunofluorescence microscopy**

400 For live-cell microscopy of *hdp1-gfp* and NF54 gametocytes, infected red blood cells were stained
401 with 16 µM Hoechst33342 in incomplete media for 15 min at 37º C and imaged at 1000×
402 magnification using a Leica DMI6000 microscope with differential interference contrast bright field
403 optics, DAPI, and GFP filter cubes with identical exposure times.

404 For immunofluorescence microscopy of *hdp1-Ty1* and *hdp1-glmS* gametocytes, cells were
405 seeded onto slides with Concanavilin A (5mg/ml; Sigma) as described in (Mehnert et al., 2019),
406 then fixed with a solution of 4% paraformaldehyde/0.0075% glutaraldehyde for 20 min at 37º C.
407 Parasites were permeabilized with 0.1% Triton X-100 for 15 min at RT followed by blocking with
408 3% BSA. Primary antibodies (anti-Ty1 BB2 mouse 1:1,000; anti-PhiL1 rabbit 1:400) were allowed
409 to bind for 1 hour in 3% BSA/PBS followed by three washes with blocking buffer for 5 min.
410 Secondary antibodies diluted 1:500 (anti-mouse-Alexa546 and anti-rabbit-Alexa488, Invitrogen)
411 plus 16 µM Hoechst33342 were added in fresh blocking buffer and incubated for 1 hour. Z-stacks
412 of stained specimens were collected at 1000× magnification using a Leica DMI6000 microscope
413 with differential interference contrast bright field optics, DAPI, and RFP filter cubes with identical
414 exposure times. Fluorescent channel z-stacks were deconvolved using the ImageJ
415 DeconvolutionLab2 plugin (NLLS algorithm) followed by maximum intensity z-projection and
416 background adjustment. Immunofluorescence microscopy of Toxoplasma tachyzoites was carries
417 out as previously described (Morlon-Guyot et al., 2018).

418

419 **Protein Expression and Purification**

420 Expression of recombinant HDP1-DBD motif was done using the Glutathione S-transferase (GST)
421 gene fusion system (GE, Healthcare). The pGEX-4T-1 (Addgene Cat. Nº 27458001) plasmid was
422 used as backbone for cloning the codon optimized sequence comprising the last 87aa of HDP1.
423 Plasmid pGEX-GST-HDP1-DBD was transformed into BL21 (DE3) competent *E. coli* strain (NEB,
424 Cat. Nº C2527) and protein expression was done following the manufacturer directions with some
425 modifications. Briefly, an overnight culture was inoculated with one bacterial colony in 2x YT
426 media supplemented with the corresponding antibiotic. Next day, culture was diluted 1:100 with
427 fresh media and kept at 30º C with vigorous agitation. Bacterial growth was monitored until culture
428 reach exponential phase. At this point, IPTG (1mM final concentration) was added and the culture
429 was kept for another 2 hours at 30º C with vigorous agitation. Cells were harvested and
430 resuspended in lysis buffer (50mM Tris HCl pH 7.5; 100mM NaCl; 1mM DTT; 5% Glycerol; 1mM
431 PMSF; 1mM EDTA; 1x protease inhibitors cocktail) supplemented with Lysozyme (1mg/ml,
432 Sigma). In order to remove bacterial DNA from our putative DNA binding protein, lysate was
433 treated with polyethyleneimine (PEI) (0.1% v/v) (RR, 1991). Lysate was sonicated, cleared by
434 centrifugation at 14,000 x *g* for 30 min at 4º C. Protein extract was recovered and GST-HDP1-
435 DBD protein was purified using Pierce GST Spin purification kit (Cat. Nº 16106) following
436 manufacturer directions. Protein of interest was dialyzed using Slide-A-Lyzer Dialysis Cassette
437 10,000 MWCO (ThermoScientific, Cat. Nº 66810) and concentrated using Amicon Ultra
438 Centrifugal filters 10,000 MWCO (Millipore, Cat. Nº UFC901024). Purity was verified by
439 Coomassie staining after SDS-PAGE and concentration was measured by Bradford assay.

440 Sequence encoding HDP1 aa2991-3078 were cloned into the MCS1 of the pRSFDuet-1 vector
441 (Novagen) engineered with an N-terminal His-SUMO tag. The proteins were expressed in *E. coli*
442 strain BL21 CodonPlus(DE3)-RIL (Stratagene). Bacteria were grown in Luria-Bertani medium at
443 37°C to OD600=0.8 and induced with 0.4 mM IPTG at 18°C overnight. Cells were collected via
444 centrifugation at 5000×g and lysed via sonication in Lysis Buffer (20 mM Tris-HCl, pH 8.0, 500
445 mM NaCl, 20 mM imidazole, and 5% Glycerol) supplemented with 1 mM phenylmethylsulfonyl
446 fluoride and 0.5% Triton X-100. Cellular debris was removed by centrifugation at 20,000×g, and
447 the supernatant was loaded onto 5 ml HisTrap FF column (GE Healthcare) and eluted using the
448 lysis buffer supplemented with 500 mM imidazole. The elution was dialyzed at 4°C overnight
449 against the buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole, and 5 mM β-
450 mercaptoethanol) with ULP1 protease added (lab stock). The sample was reloaded on the
451 HisTrap FF column to remove the His-SUMO tag. The flow-through was loaded on the Heparin

452 column (GE Healthcare) and eluted with a gradient of NaCl from 300 mM to 1 M. The target
453 protein was further purified by size exclusion chromatography (Superdex 200 [16/60], GE
454 Healthcare) in the buffer (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM MgCl2, and 1 mM DTT).
455 The high purity eluting fractions were detected by SDS-PAGE and concentrated to around 20
456 mg/ml. The protein was flash-frozen in liquid nitrogen and stored at -80°C.

457 **Protein Binding Microarray**

458 All GST-HDP1-DBD binding was analyzed twice in the PBMs. The PBM experiments were
459 performed as previously described (Berger and Bulyk, 2009; Berger et al., 2006). Briefly, custom
460 designed oligonucleotide arrays are double-stranded using a universal primer, incubated with
461 GST-HDP1-DBD fusion proteins, visualized with Alexa-488 conjugated anti-GST antibody, and
462 scanned using an Axon 4200A scanner. Proteins were used at the maximum concentration
463 obtained from purification and represent one-fifth of the total reaction volume used on the PBM.
464 In this study three different universal platforms were used covering all contiguous 8-mers as well
465 as gapped 8-mers spanning up to 10 positions. After data normalization and calculation of
466 enrichment scores the “Seed-and-Wobble” algorithm was applied to combine the data from two
467 separate experiments and create position weight matrices (PWMs). An enrichment score cut-off
468 of 0.45 was used to distinguish high affinity binding data from low affinity and non-specific binding.
469 The score for each 8-mer reflects the affinity of a DNA binding domain for that sequence, with
470 higher scores representing tighter interactions. Secondary motifs were identified by running the
471 “rerank” program until E-scores below 0.45 were obtained. The PBM analysis suite was
472 downloaded from the_brain.bwh.harvard.edu/PBMAnalysisSuite/index.html.

473 **Isothermal Titration Calorimetry**

474 All the binding experiments were performed on a Microcal ITC 200 calorimeter. Purified HLH
475 proteins were dialyzed overnight against the buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM
476 DTT) at 4°C. DNA oligos were synthesized by Integrated DNA Technologies (IDT) and dissolved
477 in the same buffer. The assays perform with 1 mM DNA duplexes containing the tandem motif
478 (TA<u>GTGCAC</u>CTATG<u>GTGCAC</u>TT) with 0.1 mM HLH90 proteins. Each reaction's exothermic heat
479 was measured by sequential injection of the 2 μL DNA duplexes into proteins solution, spaced at
480 intervals of 180 seconds. The titration was according to standard protocol at 20°C and the data
481 were fitted using the program Origin 7.0.

482

483 **Gel-shift Assays**

484 Electrophoretic mobility shift assays were performed using Light Shift EMSA kits (Thermo
485 Scientific) using 24 pg of protein and 40 fmol of probe, as previously described (Campbell et al.,
486 2010). Biotinylated double-stranded probes were generated with the indicated sequence.

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

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

496 **RNA sequencing**

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

515 positions and adapter sequences using cutadapt (version 1.16) (Martin, 2011). Reads were
516 mapped against the *P. falciparum* 3D7 reference genome v.40 (Warrenfeltz et al., 2018) using
517 STAR aligner (version 2.61) (Dobin et al., 2012) and nuclear-encoded genes were analyzed for
518 differential gene expression using cufflinks (version 2.2.1) (Trapnell et al., 2013). Genes with a
519 false discovery rate of <=0.05 with a mean FPKM >5 in at least one strain were called significant.
520 For genes with FPKM > 5 in one strain and no detectable expression in the other, FKPM values
521 were set to 0.1 for purposes of fold-change calculation. Gene Set Enrichment Analysis was
522 carried out with the FGSEA Bioconductor package (Korotkevich et al., 2019) with a false discovery
523 cutoff of <=0.05.

524 **Hi-C Sequencing**

525 Parasites were crosslinked in 1.25% formaldehyde in warm PBS for 25 min on a rocking platform.
526 Glycine was added to a final concentration of 150 mM, followed by 15 min of incubation at 37°C
527 and 15 min of incubation at 4°C. The parasites were centrifuged at 660g for 20 min at 4 °C,
528 resuspended in 5 volumes of ice-cold PBS, and incubated for 10 min at 4°C. Parasites were
529 centrifuged at 660g for 15 min at 4°C, washed once in ice-cold PBS, and stored as a pellet at
530 −80°C. Parasite pellets were then thawed on ice in Hi-C lysis buffer (10 mM Tris-HCl at pH8.0,
531 10 mM NaCl, 2 mM AEBSF, Roche Complete Mini EDTA-free protease inhibitor cocktail,0.25%
532 Igepal CA-630). Parasite membranes were disrupted by passing the lysate through a 26.5- gauge
533 needle 15 times with a syringe. Samples were spun and pellets were washed once with ice-cold
534 Hi-C lysis buffer, resuspended in 0.5% SDS and incubated at 62°C for 10 min to solubilize the
535 chromatin. The SDS was neutralized by the addition of nuclease-free water and 1% Triton X-100,
536 for 15-minutes at 37°C. DNA was then digested using MboI restriction overnight at 37°C. After
537 digestion, the enzyme was heat inactivated at 62°C for 20 min, and then cooled to RT. 5'
538 overhangs were filled in by Biotin-14-dCTP (Invitrogen) using DNA Polymerase I and Large
539 (Klenow) Fragment (NEB) for 45 min at 37°C. Blunt-end ligation was performed using T4 DNA
540 Ligase at 20°C for 4 hours. The nuclei were spun down to remove random ligation products and
541 to reduce the overall reaction volume and resuspend in decrosslinking buffer (50 mM Tris-HCl,
542 pH 8.0, 1% SDS, 1 mM EDTA, 500 mM NaCl) before adding RNaseA (20 mg/ml) for 45 min at
543 37°C and proteinase K (NEB) treatment overnight at 45°C. DNA was extracted using Agencourt
544 AMPure XP beads (Beckman Coulter) and eluted in 10 mM Tris-HCl at pH 8.0. The purified DNA
545 was sheared on a Covaris S220 and biotinylated DNA fragments were pull-down using MyOne
546 Streptavidin T1 magnetic beads (Invitrogen) at 55°C for 2 min, washed and resuspended in T4
547 DNA Ligase Buffer to perform end-repair on DNA fragments bound to the beads containing DNA

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

561 **Analysis of Hi-C Data**

562 Bed files of virtual MboI (GATC) digest of the *P. falciparum* genome were generated with
563 digest_genome.py from the HiC-Pro suite (Servant et al., 2015). The paired-end reads were
564 processed (i.e. mapped, paired, and de-duplicated), binned at 10kb resolution, and normalized
565 using the ICE algorithm using HiC-Pro (Imakaev et al., 2012). Reproducibility and strain-
566 dependent differences (Figure 6A) between Hi-C samples was evaluated by correlation analysis
567 using HiCSpector (Yan et al., 2017) and GenomeDISCO (same correlation pattern as HiCSpector,
568 not shown) (Ursu et al., 2018). Interaction maps (Figure S7) were generated by plotting intra-
569 chromosomal ICE-normalized interactions. Summary statistics (Figure 6B) of intra-chromosomal
570 interaction distances were calculated for each strain and chromosome from normalized ICE
571 interaction counts using R. Differential interaction analysis was carried out at 10 kb resolution for
572 each replicate using selfish (Ardakany et al., 2019) based on ICE-normalized interaction counts
573 from the Hi-C Pro all-valid-pairs matrix and reformatted into the ".hic" file-format using Juicebox
574 (Robinson et al., 2018). All genome-wide bin pairs with significant changes in interaction
575 frequency between the *hdp1*-ty1 and Δ*hdp1* strains with a FDR <=0.1 in both replicate pairs are
576 shown in the top left of Figure 6C. For the analysis of chromosome-end interactions, interactions
577 fold-change were calculated for all interactions in bins between the chromosome end and the first
578 protein coding gene. As the final bin on each chromosome varies in size, an additional bin was
579 added for ends truncated by telomere healing (5R, 11R, and 14R) to ensure interaction counts
580 were calculated for at least 10kb.

581

ACKNOWLEDGEMENTS

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. Llinas 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. This work was supported by startup funds from Weill Cornell Medicine (BK), NIAID 1R01AI141965 (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.;

DECLARATIONS OF INTEREST

The authors declare no competing interests.

600 **REFERENCES**

601 Ardakany, A.R., Ay, F., Lonardi, S., 2019. Selfish: discovery of differential chromatin interactions via a
602 self-similarity measure. Bioinformatics 35, i145–i153. doi:10.1093/bioinformatics/btz362
603
604 Balaji, S., Babu, M.M., Iyer, L.M., Aravind, L., 2005. Discovery of the principal specific transcription factors
605 of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains.
606 Nucleic Acids Res 33, 3994–4006. doi:10.1093/nar/gki709
607
608 Bancells, C., Llorà-Batlle, O., Poran, A., Nötzel, C., Rovira-Graells, N., Elemento, O., Kafsack, B.F.C.,
609 Cortés, A., 2019. Revisiting the initial steps of sexual development in the malaria parasite Plasmodium
610 falciparum. Nat. Microbiol 4, 144–154. doi:10.1038/s41564-018-0291-7
611
612 Berger, M.F., Bulyk, M.L., 2009. Universal protein-binding microarrays for the comprehensive
613 characterization of the DNA-binding specificities of transcription factors. Nature Protocols 4, 393–411.
614 doi:10.1038/nprot.2008.195
615
616 Berger, M.F., Philippakis, A.A., Qureshi, A.M., He, F.S., Estep, P.W., Bulyk, M.L., 2006. Compact,
617 universal DNA microarrays to comprehensively determine transcription-factor binding site specificities.
618 Nat Biotechnol 24, 1429–1435. doi:10.1038/nbt1246
619
620 Bhattacharya, S., Conolly, R.B., Kaminski, N.E., Thomas, R.S., Andersen, M.E., Zhang, Q., 2010. A
621 Bistable Switch Underlying B-Cell Differentiation and Its Disruption by the Environmental Contaminant
622 2,3,7,8-Tetrachlorodibenzo-p-dioxin. Toxicol Sci 115, 51–65. doi:10.1093/toxsci/kfq035
623
624 Brancucci, N.M.B., Bertschi, N.L., Zhu, L., Niederwieser, I., Chin, W.H., Wampfler, R., Freymond, C.,
625 Rottmann, M., Felger, I., Bozdech, Z., Voss, T.S., 2014. Heterochromatin protein 1 secures survival and
626 transmission of malaria parasites. Cell Host Microbe 16, 165–176. doi:10.1016/j.chom.2014.07.004
627
628 Bunnik, E.M., Cook, K.B., Varoquaux, N., Batugedara, G., Prudhomme, J., Cort, A., Shi, L., Andolina, C.,
629 Ross, L.S., Brady, D., Fidock, D.A., Nosten, F., Tewari, R., Sinnis, P., Ay, F., Vert, J.-P., Noble, W.S.,
630 Le Roch, K.G., 2018. Changes in genome organization of parasite-specific gene families during the
631 Plasmodium transmission stages. Nature Communications 9, 1910. doi:10.1038/s41467-018-04295-5
632
633 Burgess, R.R., 1991. Use of polyethyleneimine in purification of DNA-binding proteins. Meth. Enzymol.
634 208, 3–10. doi:10.1016/0076-6879(91)08003-z
635
636 Bürglin, T.R., Affolter, M., 2016. Homeodomain proteins: an update. Chromosoma 125, 497–521.
637 doi:10.1007/s00412-015-0543-8
638
639 Calhoun, S.F., Reed, J., Alexander, N., Mason, C.E., Deitsch, K.W., Kirkman, L.A., Boyle, J.P., 2017.
640 Chromosome End Repair and Genome Stability in Plasmodium falciparum. mBio 8, e00547–17.
641 doi:10.1128/mBio.00547-17
642
643 Campbell, T.L., De Silva, E.K., Olszewski, K.L., Elemento, O., Llinás, M., 2010. Identification and
644 genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the
645 malaria parasite. PLoS Pathog 6, e1001165. doi:10.1371/journal.ppat.1001165
646
647 Coleman, B.I., Skillman, K.M., Jiang, R.H.Y., Childs, L.M., Altenhofen, L.M., Ganter, M., Leung, Y.,
648 Goldowitz, I., Kafsack, B.F.C., Marti, M., Llinás, M., Buckee, C.O., Duraisingh, M.T., 2014. A
649 Plasmodium falciparum Histone Deacetylase Regulates Antigenic Variation and Gametocyte
650 Conversion. Cell Host Microbe 16, 177–186. doi:10.1016/j.chom.2014.06.014
651

652 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M.,
653 Gingeras, T.R., 2012. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.
654 doi:10.1093/bioinformatics/bts635
655
656 Filarsky, M., Fraschka, S.A., Niederwieser, I., Brancucci, N.M.B., Carrington, E., Carrió, E., Moes, S.,
657 Jenoe, P., Bartfai, R., Voss, T.S., 2018. GDV1 induces sexual commitment of malaria parasites by
658 antagonizing HP1-dependent gene silencing. Science 359, 1259–1263. doi:10.1126/science.aan6042
659
660 Flueck, C., Bartfai, R., Niederwieser, I., Witmer, K., Alako, B.T.F., Moes, S., Bozdech, Z., Jenoe, P.,
661 Stunnenberg, H.G., Voss, T.S., 2010. A major role for the Plasmodium falciparum ApiAP2 protein
662 PfSIP2 in chromosome end biology. PLoS Pathog 6, e1000784. doi:10.1371/journal.ppat.1000784
663
664 Flueck, C., Bartfai, R., Volz, J., Niederwieser, I., Salcedo-Amaya, A.M., Alako, B.T.F., Ehlgen, F., Ralph,
665 S.A., Cowman, A.F., Bozdech, Z., Stunnenberg, H.G., Voss, T.S., 2009. Plasmodium falciparum
666 Heterochromatin Protein 1 Marks Genomic Loci Linked to Phenotypic Variation of Exported Virulence
667 Factors. PLoS Pathog 5, e1000569. doi:10.1371/journal.ppat.1000569
668
669 Fraschka, S.A., Filarsky, M., Hoo, R., Niederwieser, I., Yam, X.Y., Brancucci, N.M.B., Mohring, F.,
670 Mushunje, A.T., Huang, X., Christensen, P.R., Nosten, F., Bozdech, Z., Russell, B., Moon, R.W., Marti,
671 M., Preiser, P.R., Bartfai, R., Voss, T.S., 2018. Comparative Heterochromatin Profiling Reveals
672 Conserved and Unique Epigenome Signatures Linked to Adaptation and Development of Malaria
673 Parasites. Cell Host Microbe 23, 407–420.e8. doi:10.1016/j.chom.2018.01.008
674
675 Ghorbal, M., Gorman, M., Macpherson, C.R., Martins, R.M., Scherf, A., Lopez-Rubio, J.J., 2014. Genome
676 editing in the human malaria parasite Plasmodium falciparum using the CRISPR-Cas9 system. Nat
677 Biotechnol 32, 819–821. doi:10.1038/nbt.2925
678
679 Hedgethorne, K., Eustermann, S., Yang, J.-C., Ogden, T.E.H., Neuhaus, D., Bloomfield, G., 2017.
680 Homeodomain-like DNA binding proteins control the haploid-to-diploid transition in Dictyostelium.
681 Science advances 3, e1602937. doi:10.1126/sciadv.1602937
682
683 Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J.,
684 Mirny, L.A., 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat
685 Meth 9, 999–1003. doi:10.1038/nmeth.2148
686
687 Josling, G.A., Russell, T.J., Venezia, J., Orchard, L., van Biljon, R., Painter, H.J., Llinás, M., 2020.
688 Dissecting the role of PfAP2-G in malaria gametocytogenesis. Nature Communications 11, 1–13.
689 doi:10.1038/s41467-020-15026-0
690
691 Kafsack, B.F.C., Rovira-Graells, N., Clark, T.G., Bancells, C., Crowley, V.M., Campino, S.G., Williams,
692 A.E., Drought, L.G., Kwiatkowski, D.P., Baker, D.A., Cortés, A., Llinás, M., 2014. A transcriptional
693 switch underlies commitment to sexual development in malaria parasites. Nature 507, 248–252.
694 doi:10.1038/nature12920
695
696 Kent, R.S., Modrzynska, K.K., Cameron, R., Philip, N., Billker, O., Waters, A.P., 2018. Inducible
697 developmental reprogramming redefines commitment to sexual development in the malaria parasite
698 Plasmodium berghei. Nat. Microbiol 3, 1206–1213. doi:10.1038/s41564-018-0223-6
699
700 Korotkevich, G., Sukhov, V., Sergushichev, A., 2019. Fast gene set enrichment analysis. bioRxiv 10,
701 060012. doi:10.1101/060012
702
703 Lopez-Rubio, J.J., Mancio-Silva, L., Scherf, A., 2009. Genome-wide analysis of heterochromatin
704 associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. Cell
705 Host Microbe 5, 179–190. doi:10.1016/j.chom.2008.12.012
706

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12.

Mehnert, A.-K., Simon, C.S., Guizetti, J., 2019. Immunofluorescence staining protocol for STED nanoscopy of Plasmodium-infected red blood cells. Mol Biochem Parasitol 229, 47–52. doi:10.1016/j.molbiopara.2019.02.007

Modrzynska, K., Pfander, C., Chappell, L., Yu, L., Suarez, C., Dundas, K., Gomes, A.R., Goulding, D., Rayner, J.C., Choudhary, J., Billker, O., 2017. A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle. Cell Host Microbe 21, 11–22. doi:10.1016/j.chom.2016.12.003

Moll, K., Ljungström, I., Perlmann, H., Scherf, A., 2008. Methods in malaria research. Manassas.

Morlon-Guyot, J., Hajj, El, H., Martin, K., Fois, A., Carrillo, A., Berry, L., Burchmore, R., Meissner, M., Lebrun, M., Daher, W., 2018. A proteomic analysis unravels novel CORVET and HOPS proteins involved in Toxoplasma gondii secretory organelles biogenesis. Cell Microbiol 20, e12870. doi:10.1111/cmi.12870

Nixon, C.P., Nixon, C.E., Michelow, I.C., Silva-Viera, R.A., Colantuono, B., Obeidallah, A.S., Jha, A., Dockery, D., Raj, D., Park, S., Duffy, P.E., Kurtis, J.D., 2018. Antibodies to PfsEGXP, an Early Gametocyte-Enriched Phosphoprotein, Predict Decreased Plasmodium falciparum Gametocyte Density in Humans. J INFECT DIS 218, 1792–1801. doi:10.1093/infdis/jiy416

Norman, T.M., Lord, N.D., Paulsson, J., Losick, R., 2015. Stochastic Switching of Cell Fate in Microbes. Annu Rev Microbiol 69, 381–403. doi:10.1146/annurev-micro-091213-112852

Painter, H.J., Carrasquilla, M., Llinás, M., 2017. Capturing in vivo RNA transcriptional dynamics from the malaria parasite Plasmodium falciparum. Genome Res 27, 1074–1086. doi:10.1101/gr.217356.116

Park, B.O., Ahrends, R., Teruel, M.N., 2012. Consecutive Positive Feedback Loops Create a Bistable Switch that Controls Preadipocyte-to-Adipocyte Conversion. Cell Reports 2, 976–990. doi:10.1016/j.celrep.2012.08.038

Parkyn Schneider, M., Liu, B., Glock, P., Suttie, A., McHugh, E., Andrew, D., Batinovic, S., Williamson, N., Hanssen, E., McMillan, P., Hliscs, M., Tilley, L., Dixon, M.W.A., 2017. Disrupting assembly of the inner membrane complex blocks Plasmodium falciparum sexual stage development. PLoS Pathog 13, e1006659. doi:10.1371/journal.ppat.1006659

Poran, A., Nötzel, C., Aly, O., Mencia-Trinchant, N., Harris, C.T., Guzman, M.L., Hassane, D.C., Elemento, O., Kafsack, B.F.C., 2017. Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. Nature 551, 95–99. doi:10.1038/nature24280

Pradhan, L., Genis, C., Scone, P., Weinberg, E.O., Kasahara, H., Nam, H.-J., 2012. Crystal Structure of the Human NKX2.5 Homeodomain in Complex with DNA Target. Biochemistry 51, 6312–6319. doi:10.1021/bi300849c

Prommana, P., Uthaipibull, C., Wongsombat, C., Kamchonwongpaisan, S., Yuthavong, Y., Knuepfer, E., Holder, A.A., Shaw, P.J., 2013. Inducible Knockdown of Plasmodium Gene Expression Using the glmS Ribozyme. PLoS ONE 8, e73783. doi:10.1371/journal.pone.0073783

Robinson, J.T., Turner, D., Durand, N.C., Thorvaldsdóttir, H., Mesirov, J.P., Aiden, E.L., 2018. Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. Cels 6, 256–258.e1. doi:10.1016/j.cels.2018.01.001

Rug, M., Maier, A.G., 2012. Transfection of Plasmodium falciparum, in: Malaria, Methods in Molecular Biology. Humana Press, Totowa, NJ, Totowa, NJ, pp. 75–98. Doi:10.1007/978-1-62703-026-7_6

Saini, E., Zeeshan, M., Brady, D., Pandey, R., Kaiser, G., Koreny, L., Kumar, P., Thakur, V., Tatiya, S., Katris, N.J., Limenitakis, R.S., Kaur, I., Green, J.L., Bottrill, A.R., Guttery, D.S., Waller, R.F., Heussler, V., Holder, A.A., Mohmmed, A., Malhotra, P., Tewari, R., 2017. Ph otosensitized I NA- L abelled protein 1 (PhIL1) is novel component of the inner membrane complex and is required for Plasmodium parasite development. Sci. Rep. 7, 15577. Doi:10.1038/s41598-017-15781-z

Salanti, A., Staalsoe, T., Lavstsen, T., Jensen, A.T.R., Sowa, M.P.K., Arnot, D.E., Hviid, L., Theander, T.G., 2003. Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. Mol Microbiol 49, 179–191. Doi:10.1046/j.1365-2958.2003.03570.x

Satory, D., Gordon, A.J., Halliday, J.A., Herman, C., 2011. Epigenetic switches: can infidelity govern fate in microbes? Curr Opin Microbiol 14, 212–217. Doi:10.1016/j.mib.2010.12.004

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., Barillot, E., 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 16, 1–11. Doi:10.1186/s13059-015-0831-x

Sinha, A., Hughes, K.R., Modrzynska, K.K., Otto, T.D., Pfander, C., Dickens, N.J., Religa, A.A., Bushell, E., Graham, A.L., Cameron, R., Kafsack, B.F.C., Williams, A.E., Llinás, M., Berriman, M., Billker, O., Waters, A.P., 2014. A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. Nature 507, 253–257. Doi:10.1038/nature12970

Tanaka, T.Q., Williamson, K.C., 2011. A malaria gametocytocidal assay using oxidoreduction indicator, alamarBlue. Mol Biochem Parasitol 177, 160–163. Doi:10.1016/j.molbiopara.2011.02.005

Tibúrcio, M., Silvestrini, F., Bertuccini, L., Sander, A.F., Turner, L., Lavstsen, T., Alano, P., 2013. Early gametocytes of the malaria parasite Plasmodium falciparum specifically remodel the adhesive properties of infected erythrocyte surface. Cell Microbiol 15, 647–659. Doi:10.1111/cmi.12062

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., Pachter, L., 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol 31, 46–53. Doi:10.1038/nbt.2450

Ursu, O., Boley, N., Taranova, M., Wang, Y.X.R., Yardimci, G.G., Stafford Noble, W., Kundaje, A., 2018. GenomeDISCO: a concordance score for chromosome conformation capture experiments using random walks on contact map graphs. Bioinformatics 34, 2701–2707. Doi:10.1093/bioinformatics/bty164

van Biljon, R., van Wyk, R., Painter, H.J., Orchard, L., Reader, J., Niemand, J., Llinás, M., Birkholtz, L.-M., 2019. Hierarchical transcriptional control regulates Plasmodium falciparum sexual differentiation. BMC Genomics 20, 920. Doi:10.1186/s12864-019-6322-9

Warrenfeltz, S., Basenko, E.Y., Crouch, K., Harb, O.S., Kissinger, J.C., Roos, D.S., Shanmugasundram, A., Silva-Franco, F., 2018. EuPathDB: The Eukaryotic Pathogen Genomics Database Resource. Methods Mol Biol 1757, 69–113. Doi:10.1007/978-1-4939-7737-6_5

Yan, K.-K., Yardımcı, G.G., Yan, C., Noble, W.S., Gerstein, M., 2017. HiC-spector: a matrix library for spectral and reproducibility analysis of Hi-C contact maps. Bioinformatics 33, 2199–2201. doi:10.1093/bioinformatics/btx152

817 Young, J.A., Fivelman, Q.L., Blair, P.L., la Vega, de, P., Le Roch, K.G., Zhou, Y., Carucci, D.J., Baker,
818 D.A., Winzeler, E.A., 2005. The Plasmodium falciparum sexual development transcriptome: a
819 microarray analysis using ontology-based pattern identification. Mol Biochem Parasitol 143, 67–79.
820 doi:10.1016/j.molbiopara.2005.05.007
821
822 Yuda, M., Iwanaga, S., Kaneko, I., Kato, T., 2015. Global transcriptional repression: An initial and
823 essential step for Plasmodium sexual development. Proceedings of the National Academy of Sciences.
824 doi:10.1073/pnas.1504389112
825
826 Yuda, M., Kaneko, I., Iwanaga, S., Mura, Y., Kato, T., 2019. Female-specific gene regulation in malaria
827 parasites by an AP2-family transcription factor. Mol Microbiol 6, 1–51. doi:10.1111/mmi.14334
828
829 Zhang, C., Li, Z., Cui, H., Jiang, Y., Yang, Z., Wang, X., Gao, H., Liu, C., Zhang, S., Su, X.-Z., Yuan, J.,
830 2017. Systematic CRISPR-Cas9-Mediated Modifications of Plasmodium yoelii ApiAP2 Genes Reveal
831 Functional Insights into Parasite Development. mBio 8, 610. doi:10.1128/mBio.01986-17
832

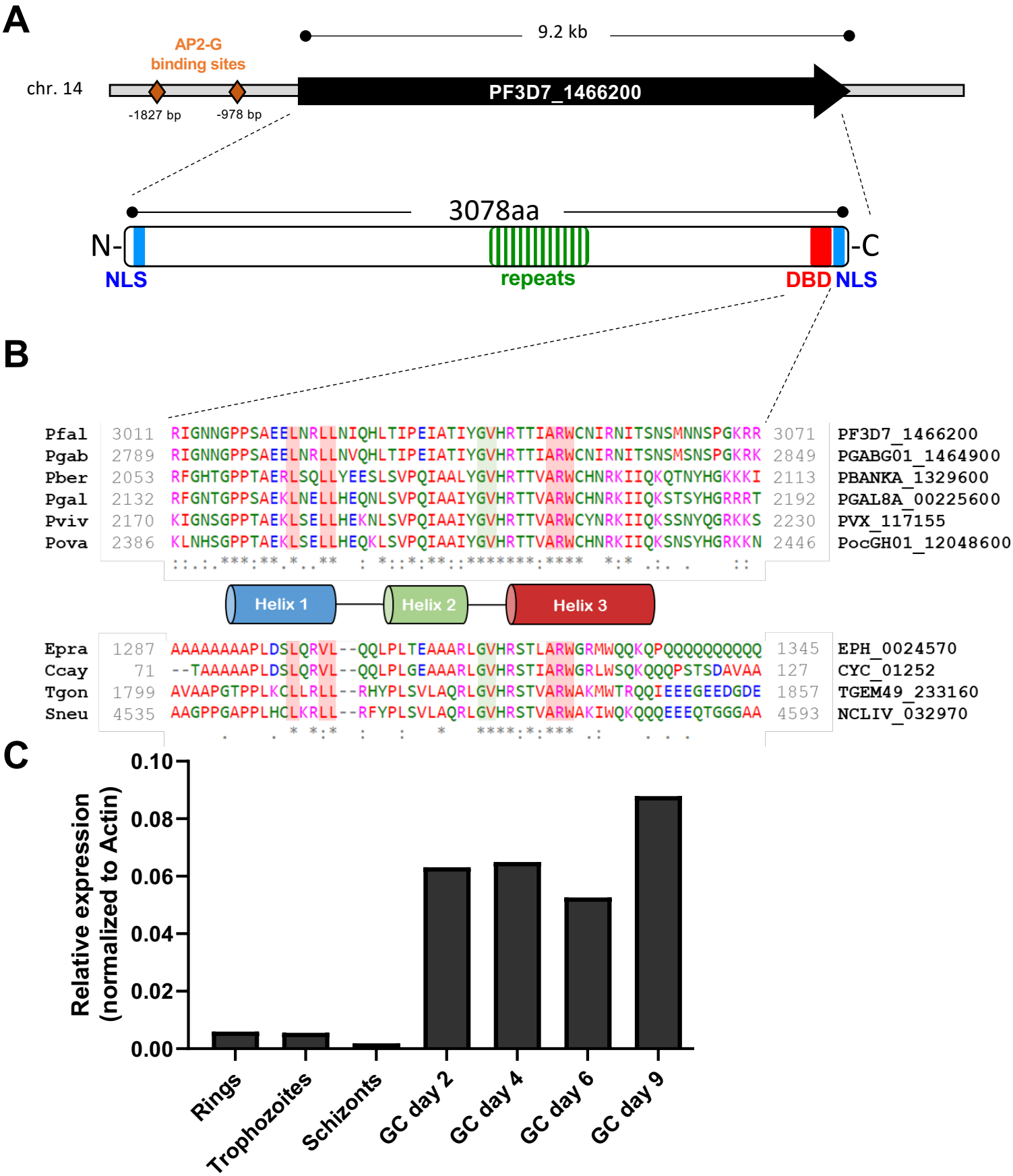


Figure 1: The predicted DNA-binding protein HDP1 is expressed in gametocytes. (A) The single exon locus Pf3D7_1466200 encodes a large 3078aa protein with a predicted C-terminal Helix-Turn-Helix DNA-binding domain (DBD) and two nuclear localization signals (NLS). Multiple AP2-G binding sites are located in the 2kb promoter region. **(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).

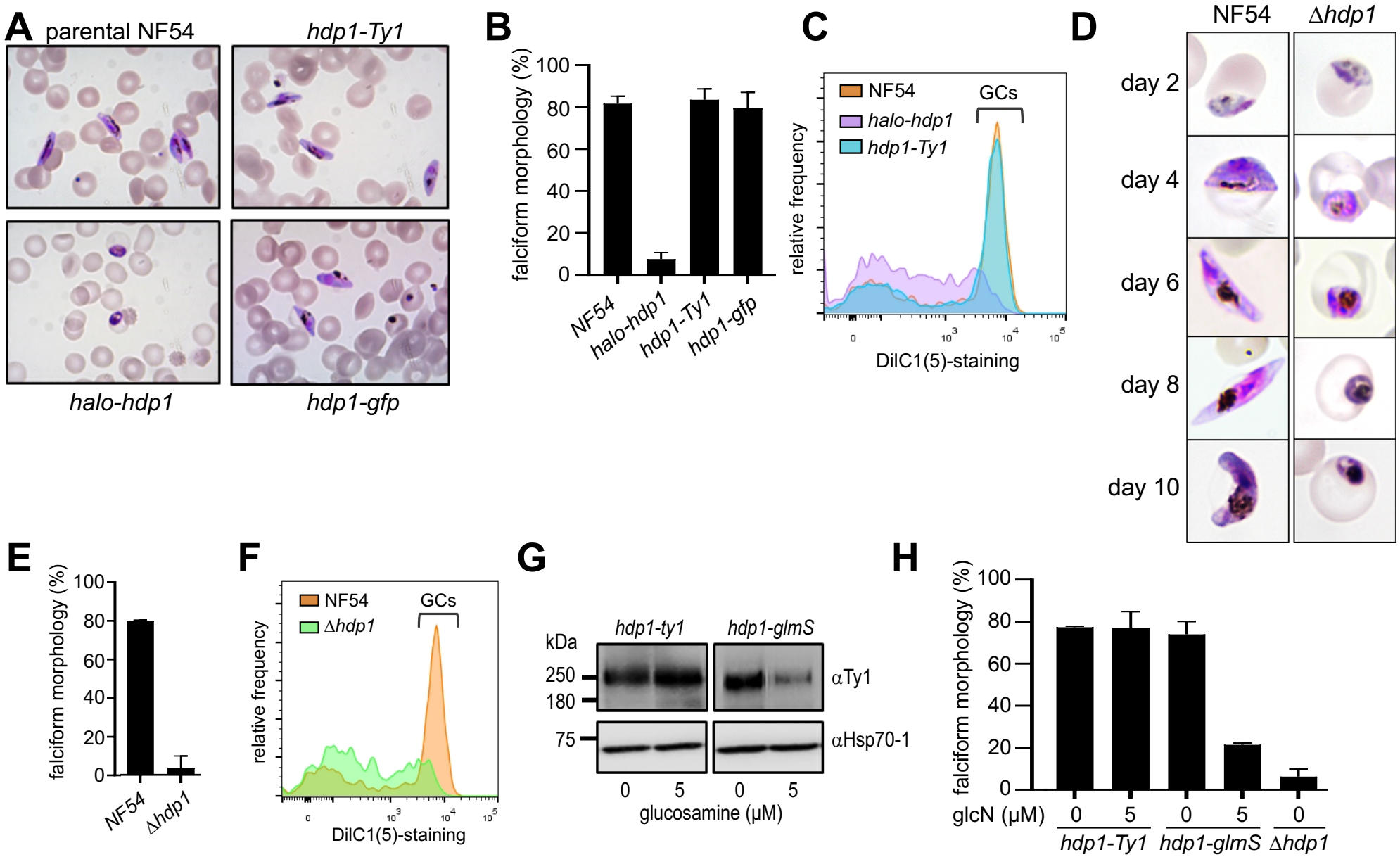


Figure 2: Loss of PfHDP1 function disrupts gametocyte maturation. (A-C) N-terminal tagging of the endogenously encoded HDP1 (*halo-hdp1*) blocked maturation of gametocytes (GCs) while C-terminal tagging of the endogenous locus with either GFP (*hdp1-gfp*) or a triple Ty1 (*hdp1-ty1x3*) epitope had no 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*) similarly blocked gametocyte maturation. Bar graphs show mean±s.e.m of n=3 **(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). Images and flow cytometry plots are are representative of n=3.

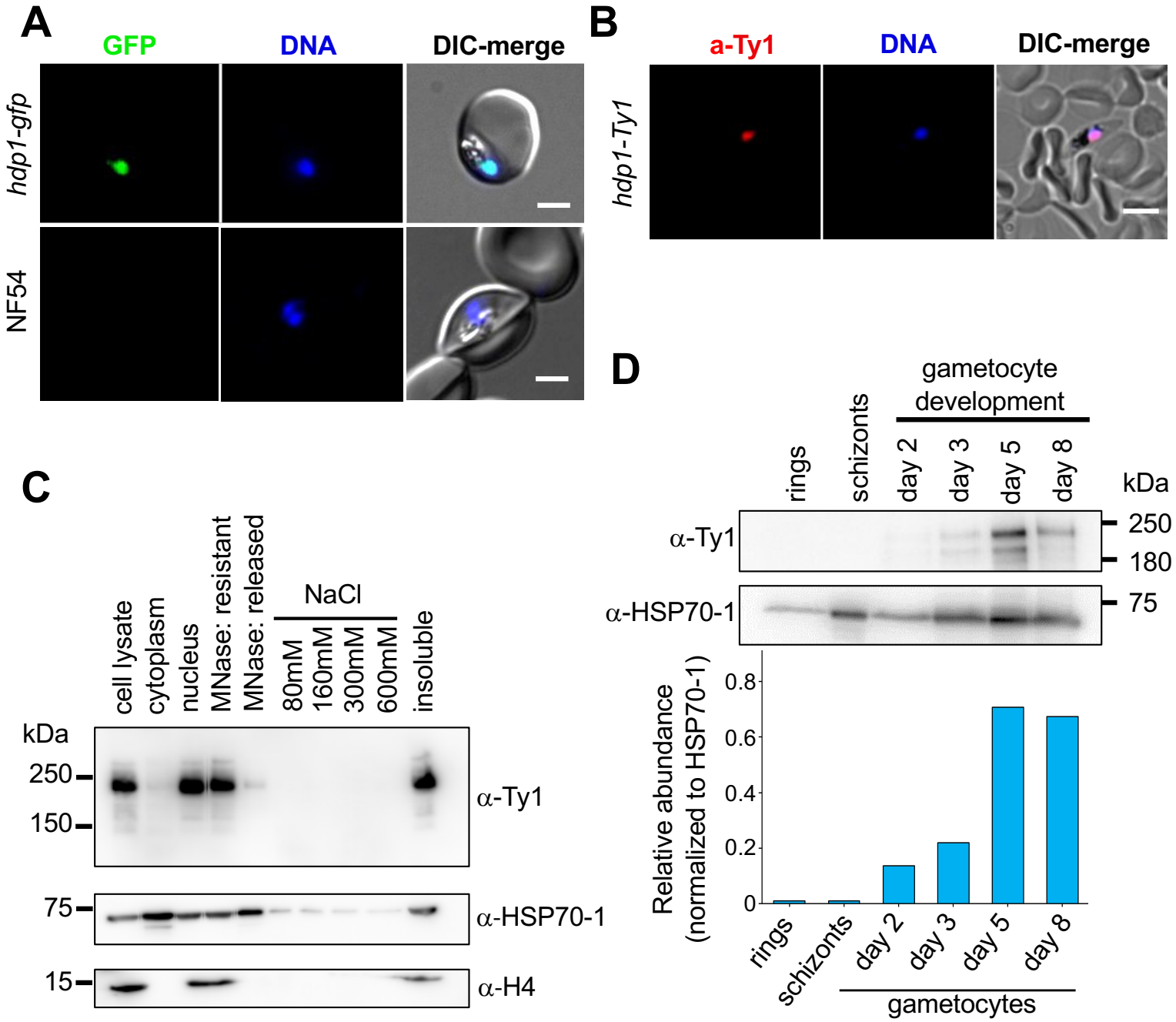


Figure 3: PfHDP1 localizes to the nucleus of gametocytes. (A) Live-cell fluorescence microscopy of *hdp1-gfp* and parental NF54 gametocytes on day 5 of maturation stained with the DNA dye Hoechst33342 (blue). Scale Bar: 2μm. Representative of n=2. **(B)** Immunofluorescence microscopy of hdp1-ty1x3 and parental NF54 gametocytes on day 5 of maturation co-stained with anti-ty1 antibodies (red) and Hoechst33342 (blue). Scale Bar: 5μm. Representative of n=2. **(C)** Western blot of cytoplasmic and nuclear extracts of hdp1-ty1x3 gametocytes on day 5 of maturation with stained with antibodies against the ty1 epitope tag, the histone H4, and HSP70-1. Representative of n=3. **(D)** 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.

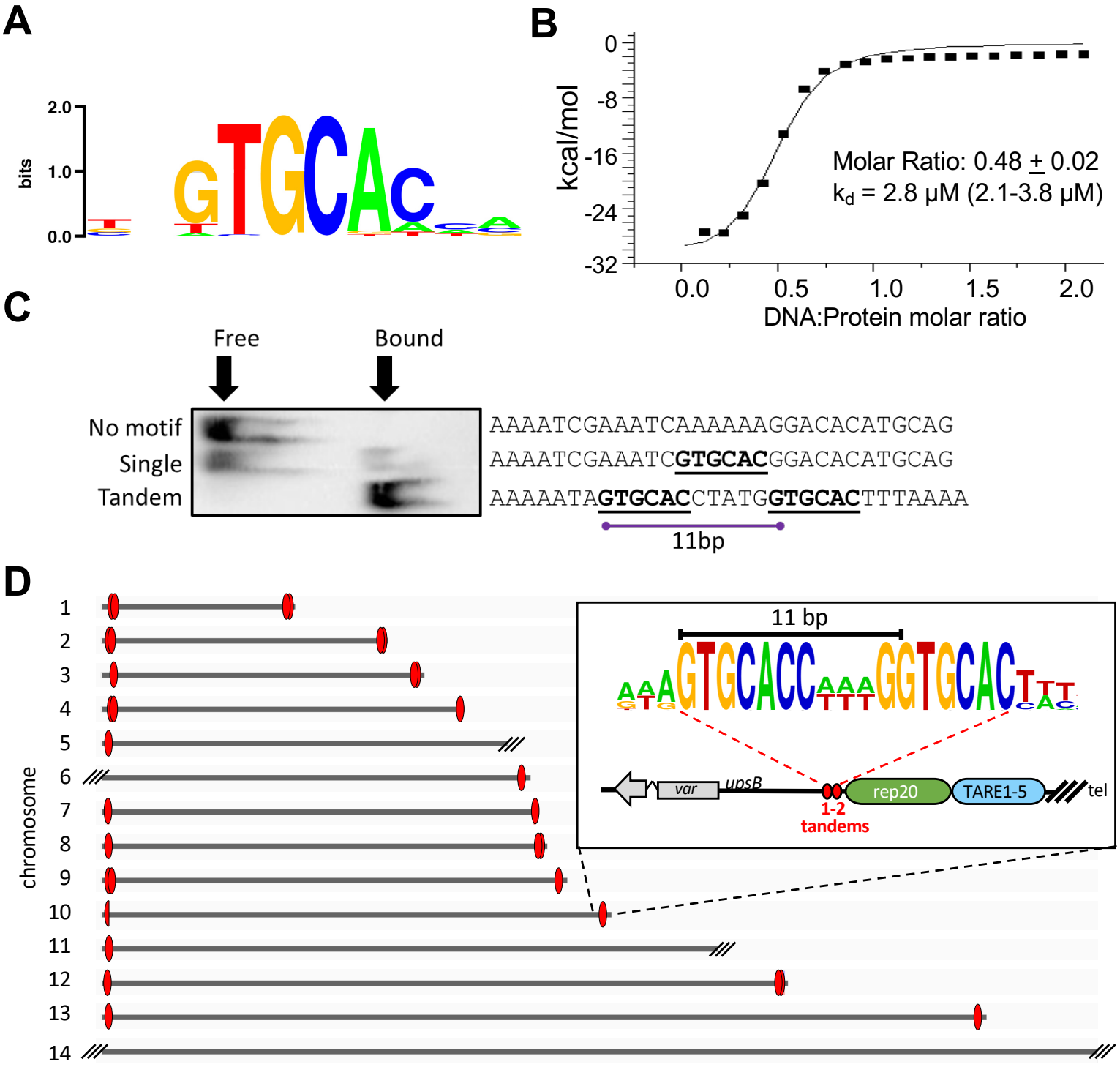


Figure 4: The HDP1 HTH domain dimers recognize a tandem GC-rich motif found at chromosome ends. (A) Maximum enrichment DNA motif for for the GST-HDP1 HTH domain from a protein-binding microarray. **(B)** Isothermal calorimetry indicates the HDP1-HTH domain recognizes DNA as a dimer. n=2. **(C)** Optimal 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. **(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.

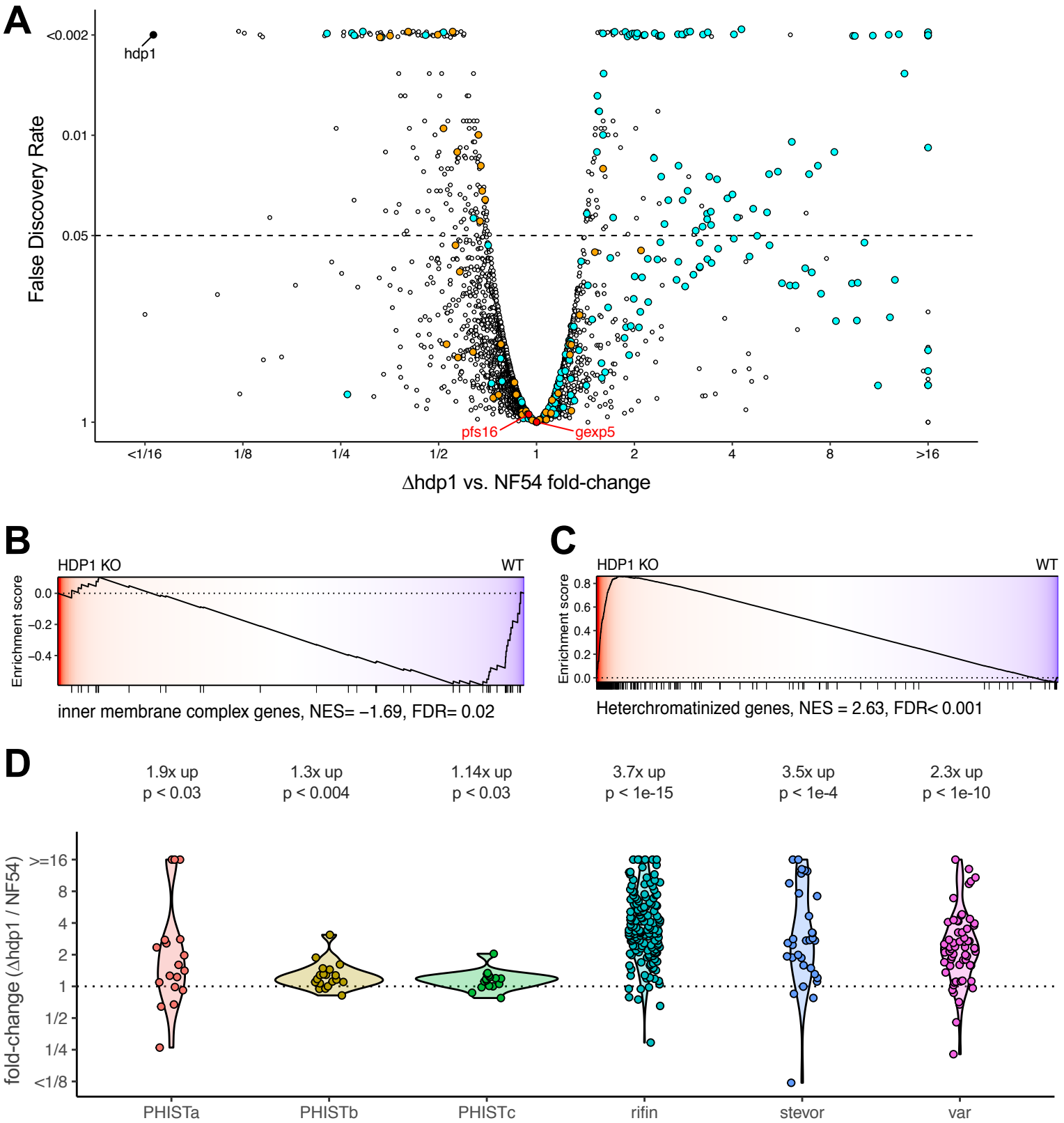


Figure 5: Disruption of HDP1 results in leaky expression of heterochromatin-associated genes and reduced expression of inner membrane complex genes in early gametocytes. **(A)** Genome-wide comparison of differential gene expression in Δ*hdp1* and parental NF54 gametocytes on day 2 of gametocytogenesis (stage I, n=2). *hdp1* (solid black), heterochromatin-associated genes (cyan), IMC genes (orange) and the two canonical early gametocyte markers, *pfs16* and *gexp5 (red)* are highlighted. Gene set enrichment analysis (GSEA) indicated significant downregulation of IMC genes **(B)** and global upregulation of heterochromatin associated genes **(C)**. **(D)** Heterochromatin associated gene families with significant up regulation in Δhdp1. Geometric mean fold-changes and p-values (two-sided, one sample t-test) are indicated.

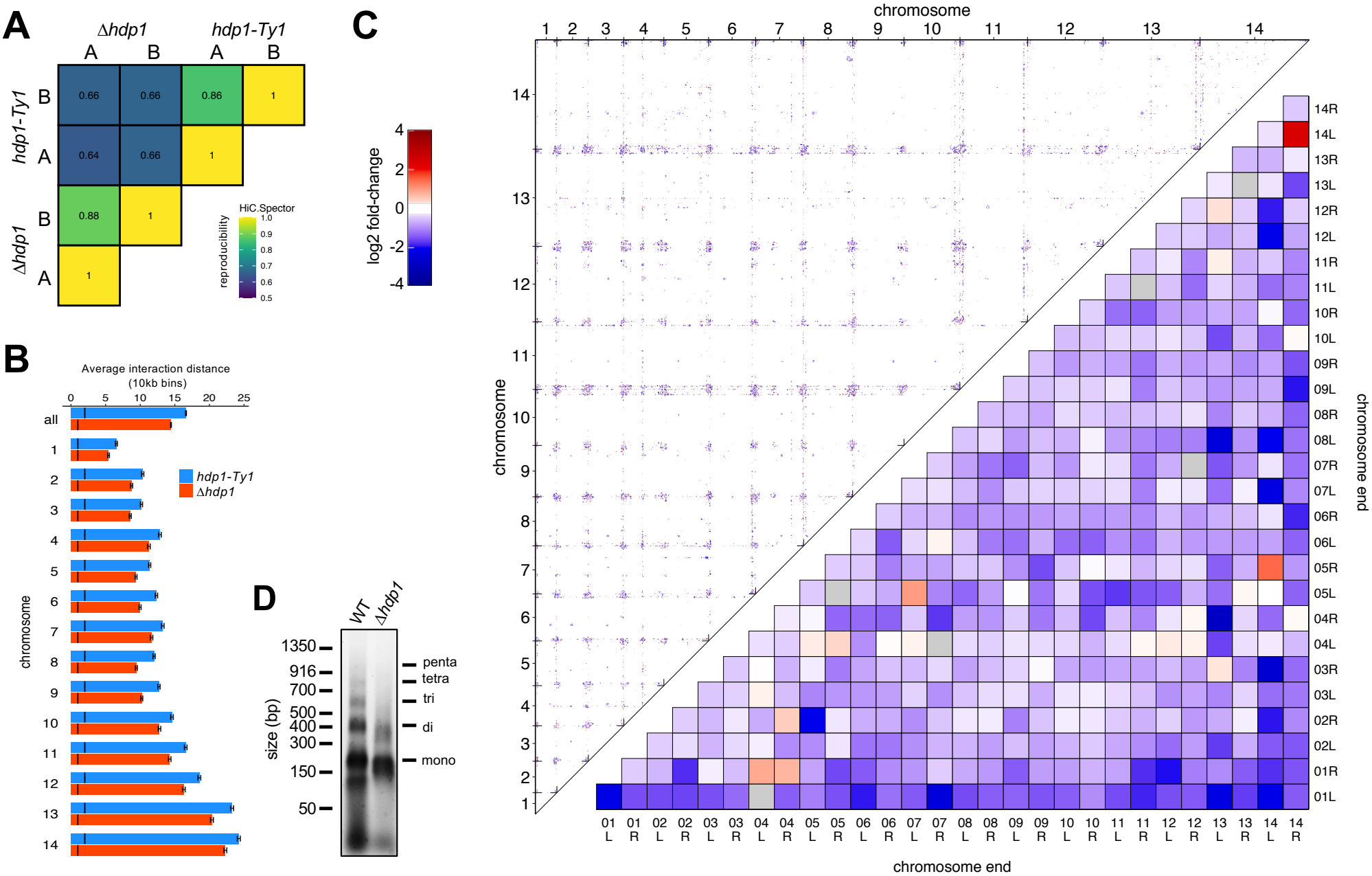


Figure 6: Loss of HDP1 results in reduced telomere clustering. (A) Analysis of Hi-C interaction maps for each sample by HiCSpector highlights reproducibility of replicates and illustrate systematic differences between the strains. Scale ranges from 0-1.0 but only 0.6-1.0 is shown for improved visibility. **(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 Δhdp1 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.

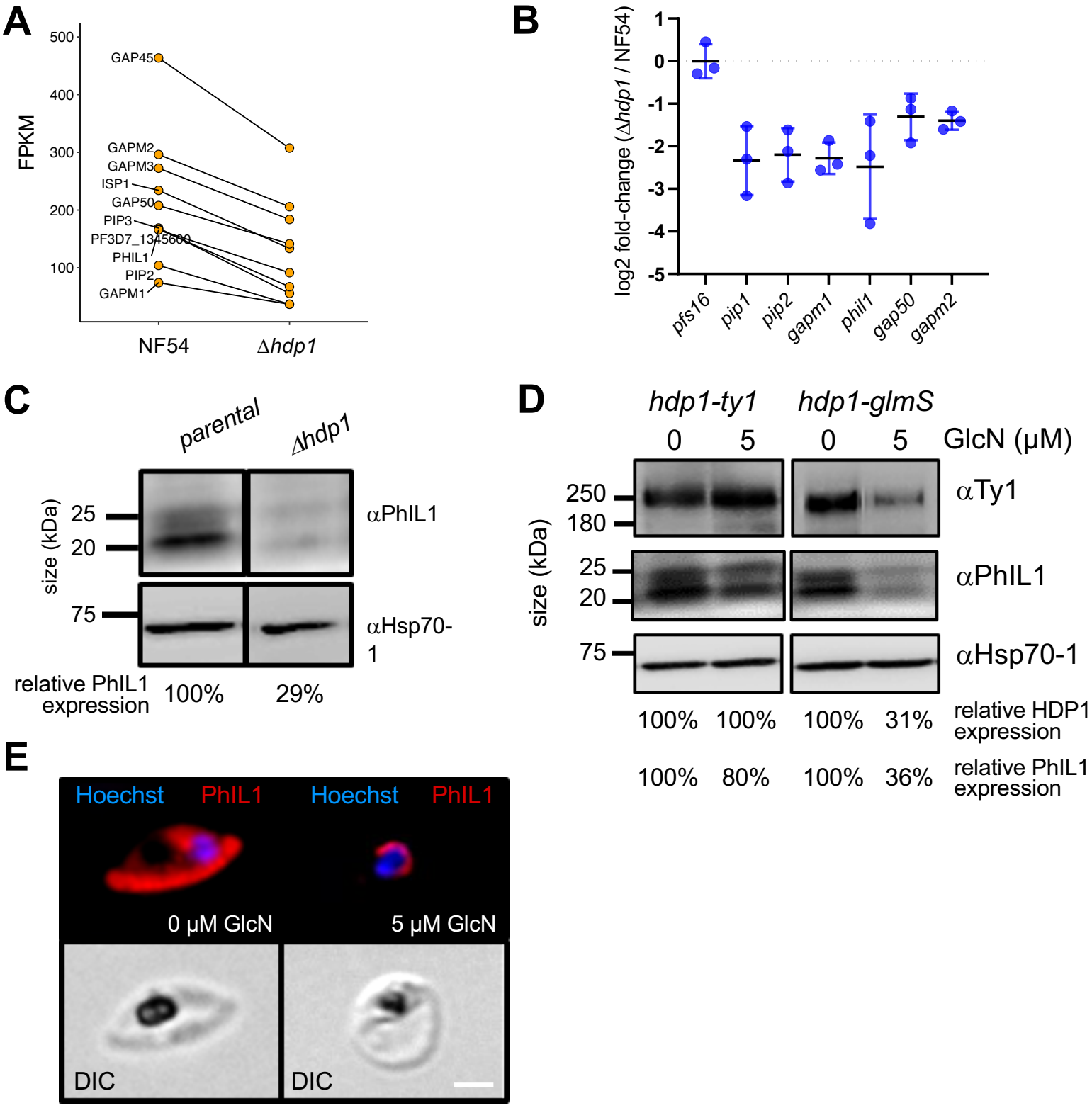
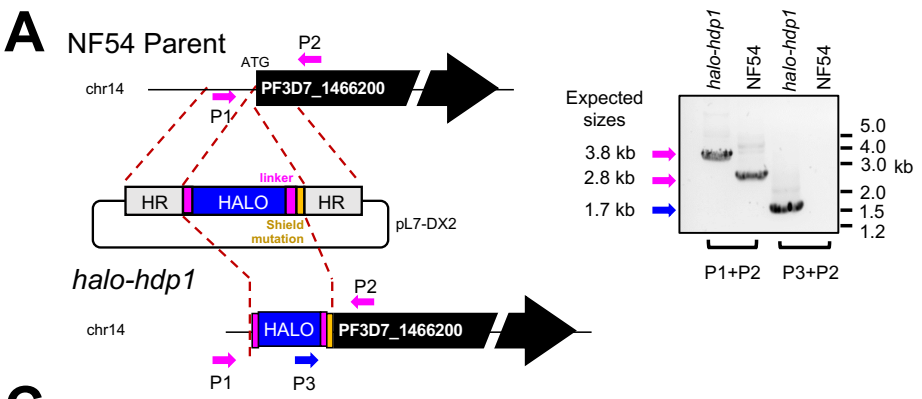
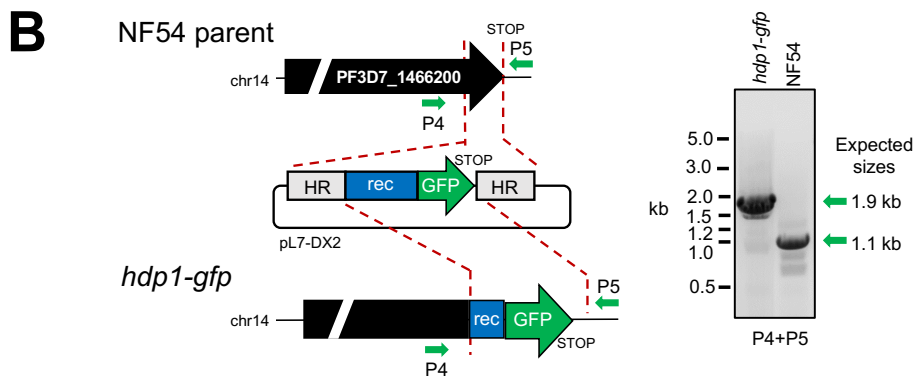
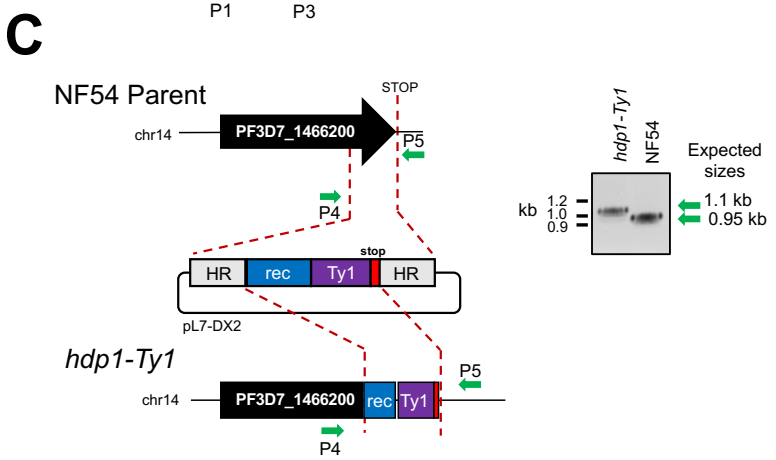
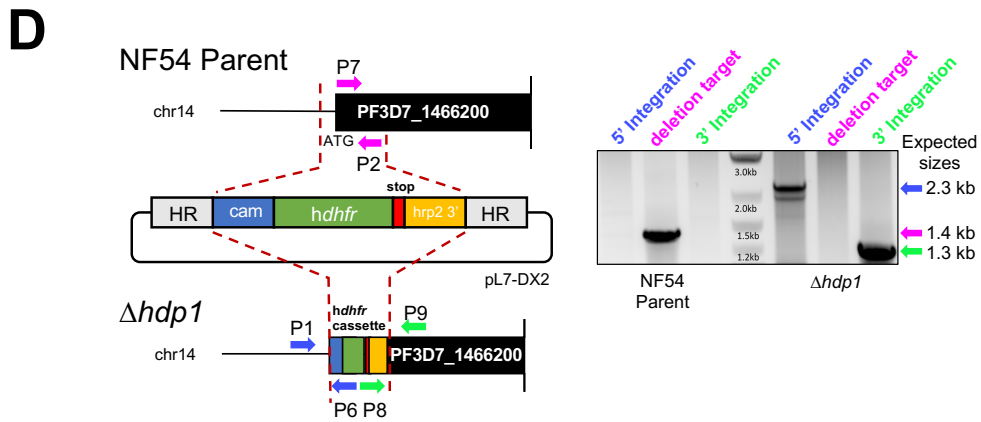


Figure 7: HDP1 regulates PhIL1-dependent expansion of the inner membrane complex in early gametocytes. (A) Transcripts encoding inner membrane complex proteins are down-regulated in Δ*hdp1* day 2 gametocytes. Normalized transcript levels (FPKM) in NF54 and Δ*hdp1* day 2 gametocytes. **(B)** Validation 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 μM 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 μM glucosamine. Scale Bar: 3μm









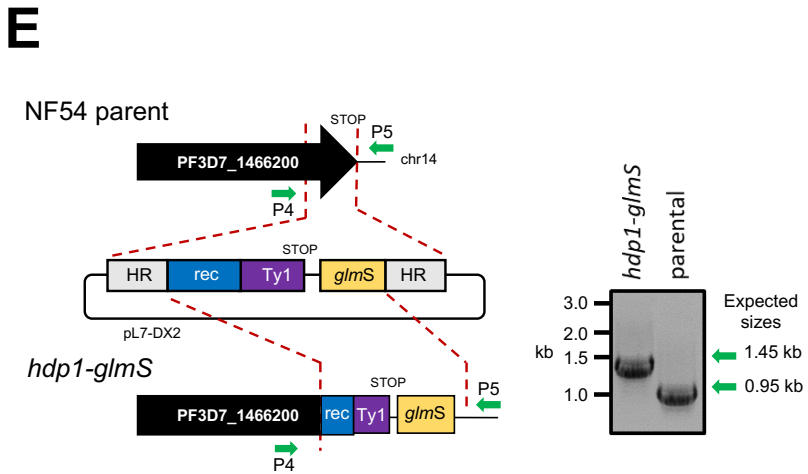


Figure S1: Validation of engineered parasite lines. (A) Generation of h*alo-hdp1* parasites by CAS9 genome editing. Insertion of the N-terminal HALO tag at the 5' end the *hdp1* coding sequence was confirmed by PCR and checked for mutations by Sanger sequencing of the 3.8kb PCR product (not shown). **(B)** 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).

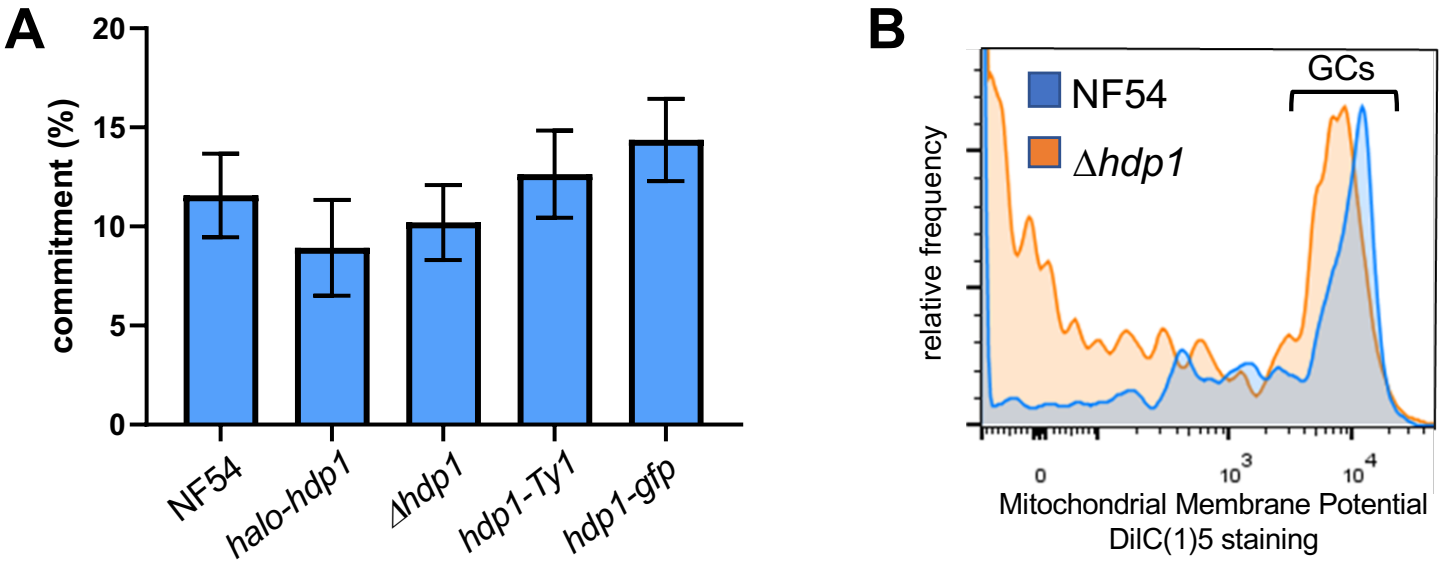


Figure S2: Loss of HDP1 does not alter the sexual commitment frequency or Stage I gametocyte viability. (A) The sexual commitment frequency (day 5 gametocytes per day 1 ring stages) is not significantly affected in *halo-hdp1* and Δ*hdp1* parasites. n=3 **(B)** Mitochondrial membrane potential as measured by DilC(1)5 staining indicates similar viability of day 2 gametocytes (GCs) from NF54 (blue) and Δ*hdp1* (orange). Representative of n=2.

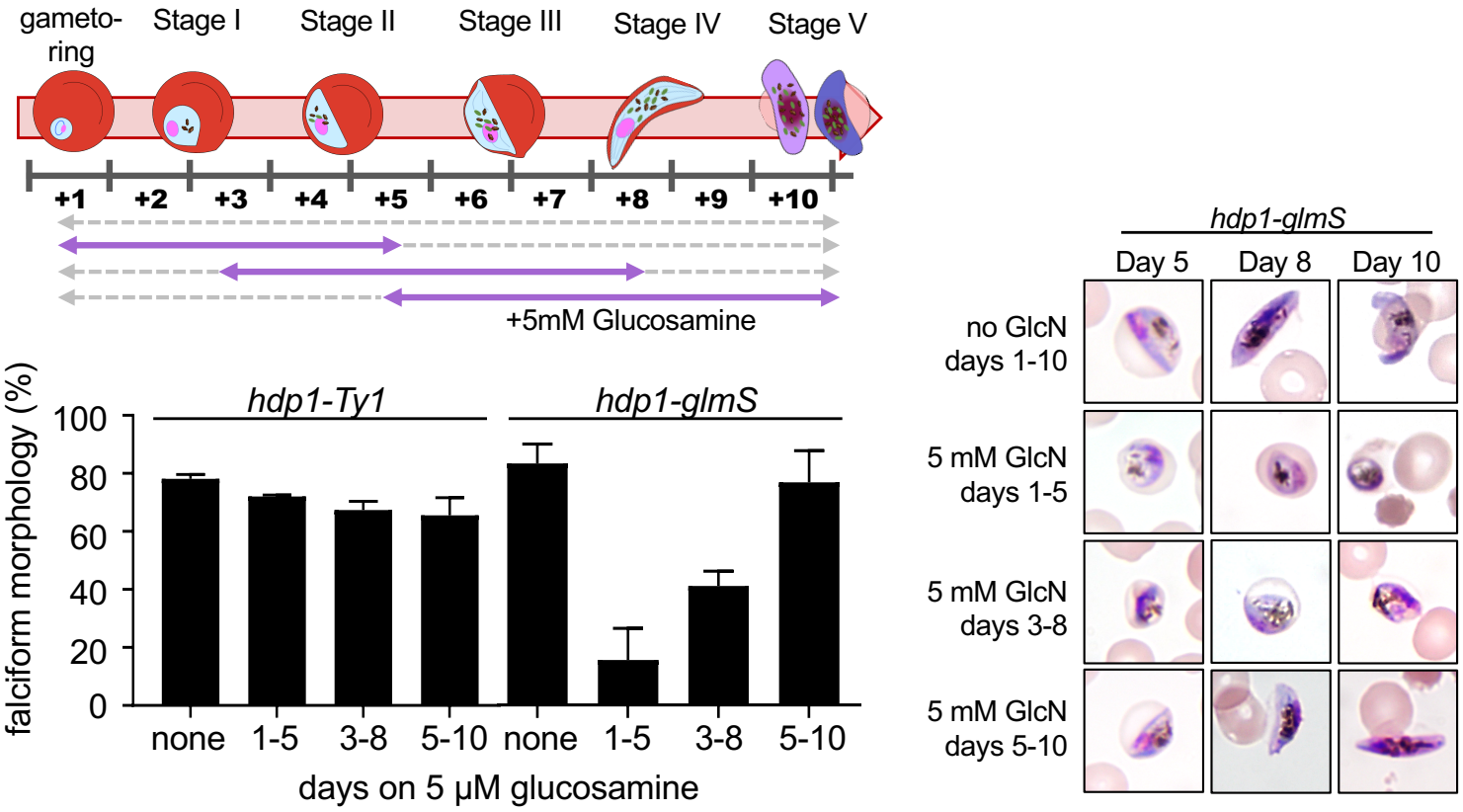


Figure S3: Inducible knockdown of HDP1 reduces gametocyte maturation in early but not late gametocytes. Representative morphology (right) of hdp1-glmS gametocytes in response to 5mM glucosamine on days 1-10, 3-8, 5-10, or in the absence of glucosamine. Percentage of falciform gametocytes on Day 10 in response to 5 mM glucosamine on days 1-10, 3-8, 5-10, or in the absence of glucosamine for hdp1-Ty1 or hdp1-glmS parasites (bottom). mean±s.e.m of n=3.

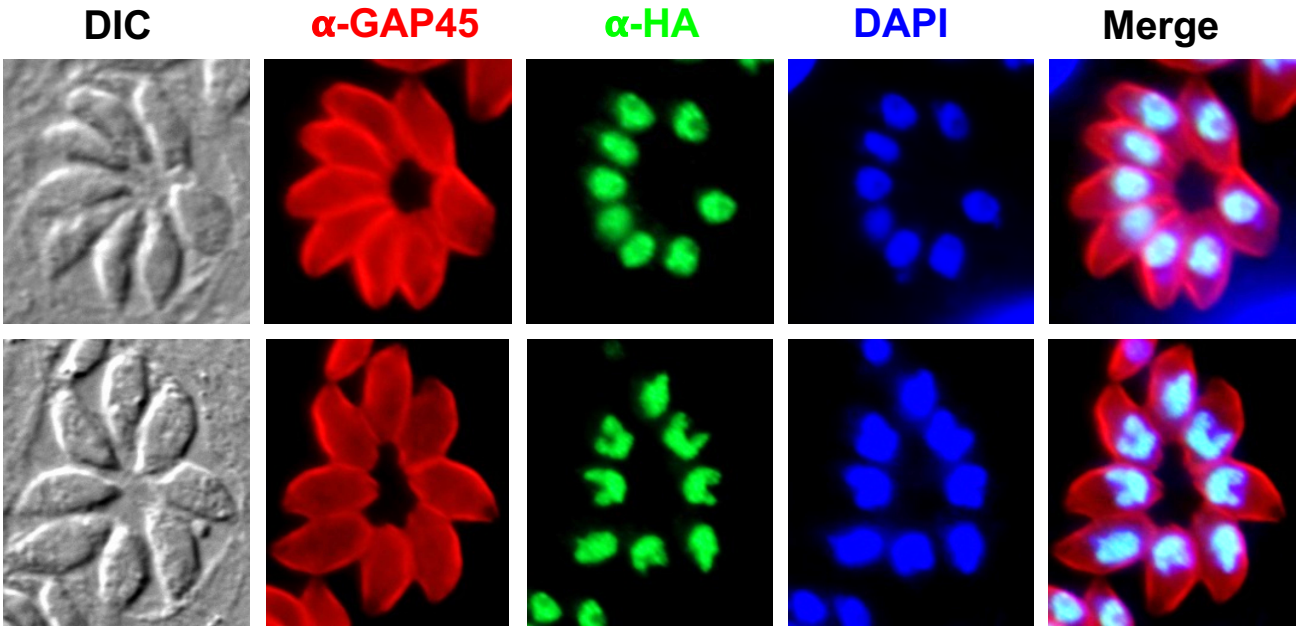


Figure S4: Immunofluorescence microscopy localizes the HA-tagged ortholog TGME49_233160 to the nucleus of *Toxoplasma gondii* tachyzoites.

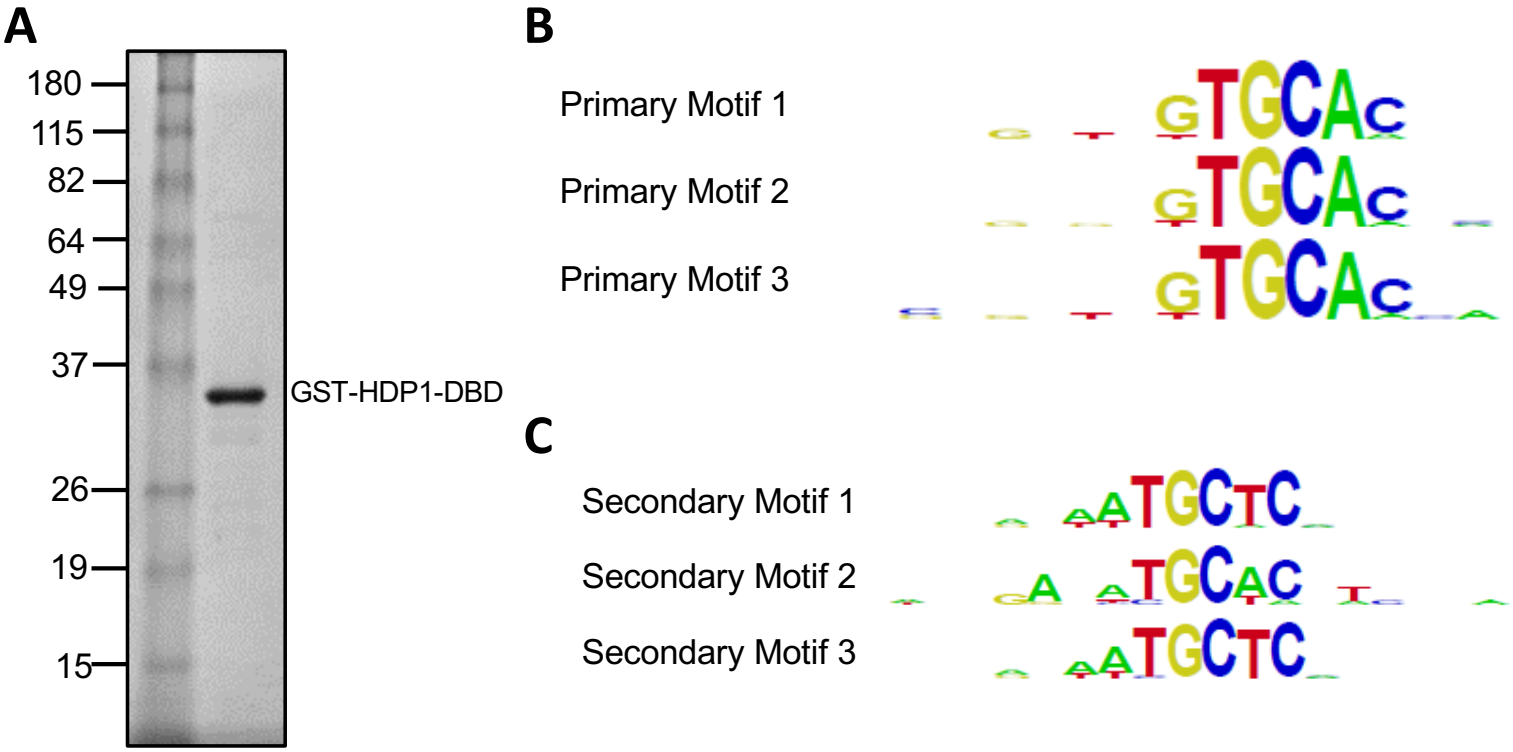


Figure S5: HDP1 binds to a GC-rich motif on Protein Binding Microarray (A) Coomassie stain of GST-HDP1-DBD used for PBM analysis. **(B)** Top three primary motif hits. **(C)** Top three secondary motif hits after removal of primary hits.

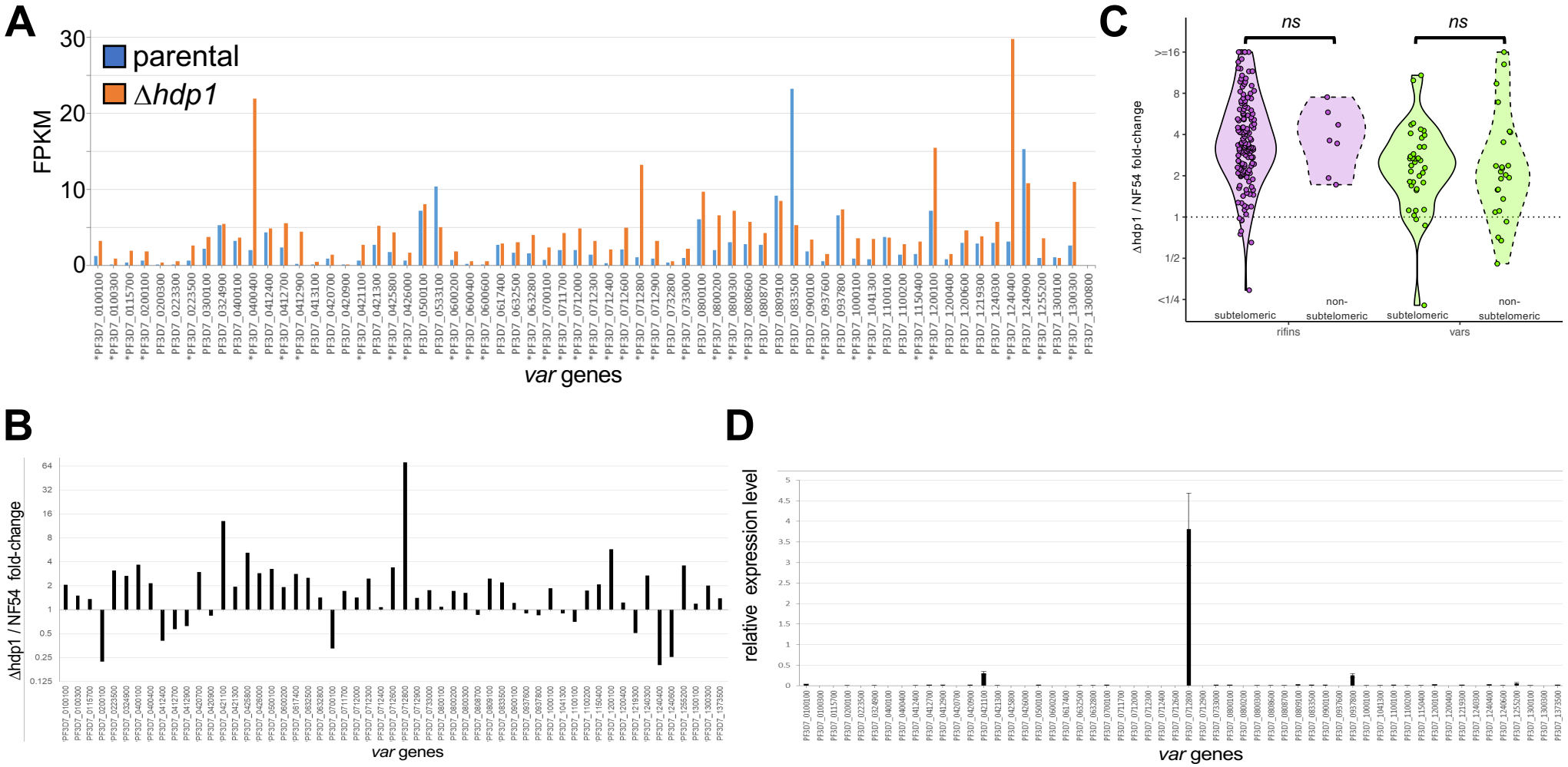
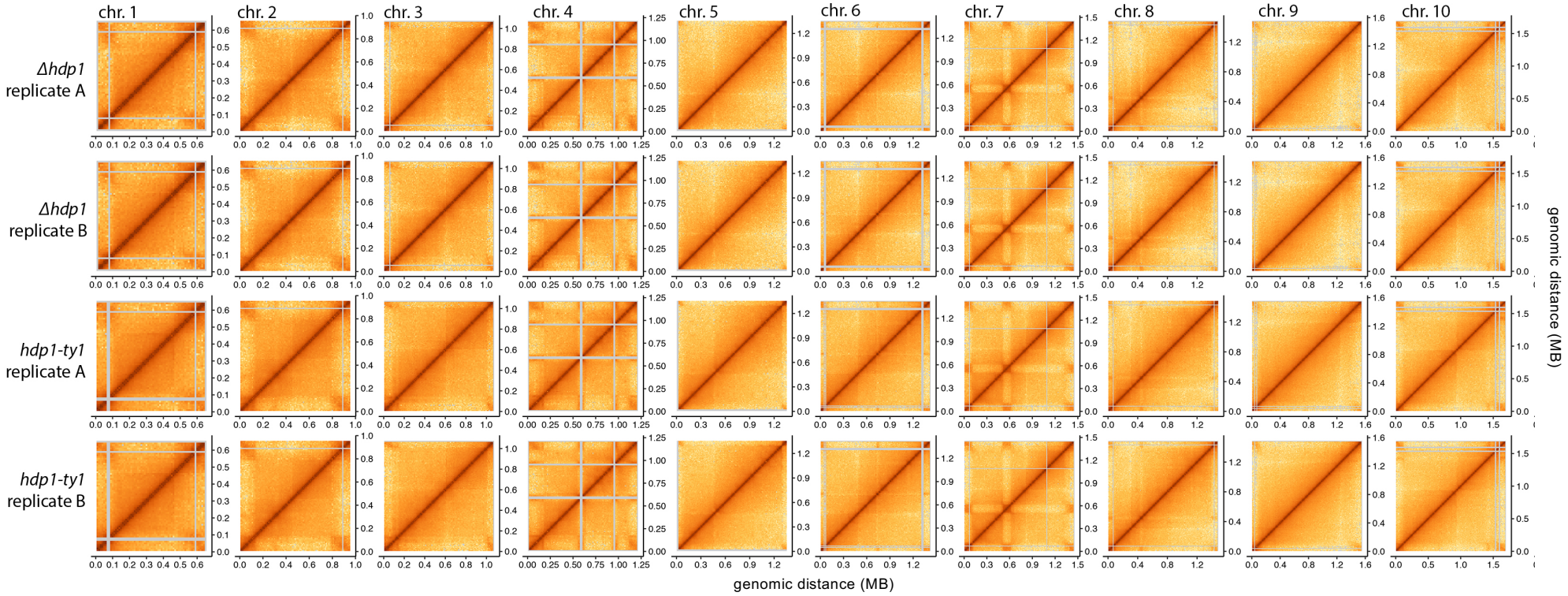


Figure S6: *var* gene expression is altered in early Δ*hdp1* gametocytes but not asexual blood stages. (A) Normalized abundance of reads uniquely mapping to *var* genes in day 2 gametocytes from Δ*hdp1* (orange) or parental NF54 parasites (blue). Significantly upregulated genes are marked with asterisks. (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)



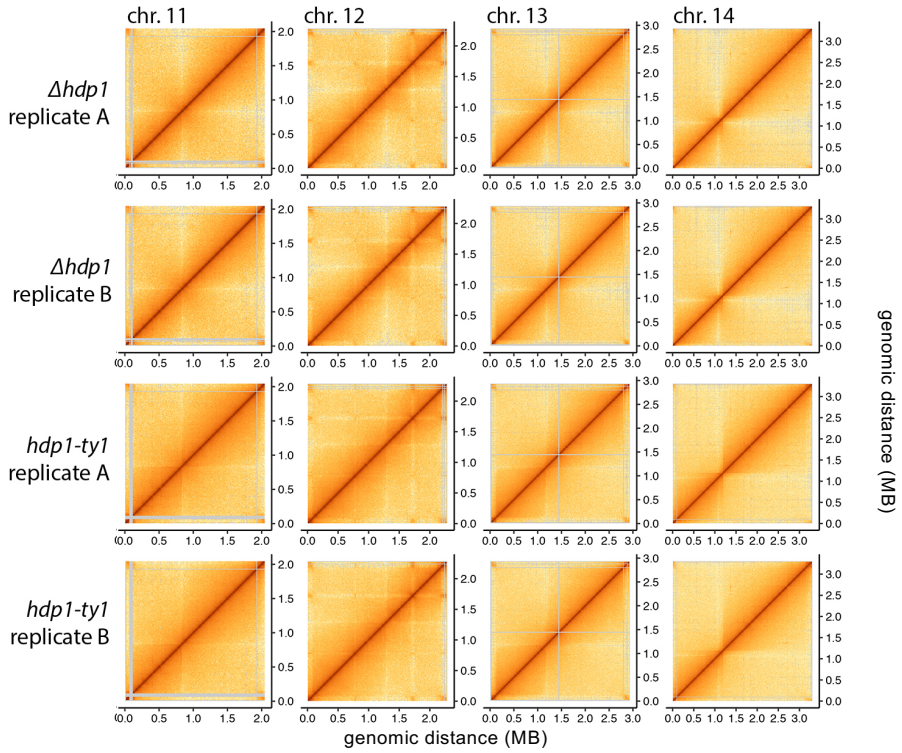


Figure S7 Intra-chromosomal HiC maps from Δ*hdp1* (top) and *hdp1-Ty1* (bottom) of day 2 gametocytes. Color scale indicates the normalized log10 interaction frequency (counts per million). Gray bars indicate bins of high-homology that could not be uniquely mapped with high confidence.