**A unique GCN5 histone acetyltransferase complex controls erythrocyte invasion and virulence in the malaria parasite *Plasmodium falciparum***

Jun Miao[1,*], Chengqi Wang[2], Amuza Lucky[1], Xiaoying Liang[1], Hui Min[1], Swamy Rakesh Adapa[2], Rays Jiang[2], Kami Kim[1], Liwang Cui[1,*]

[1] Department of Internal Medicine, Morsani College of Medicine, University of South Florida, 3720 Spectrum Blvd, Tampa, FL 33612, USA

[2] Center for Global Health and Infectious Diseases Research, College of Public Health, University of South Florida, 3720 Spectrum Blvd, Tampa, Florida 33612, USA

**Short title:** GCN5 complex in malaria parasite

*** Correspondence:**

Jun Miao

Email: jmiao1@usf.edu

Liwang Cui

Email: liwangcui@usf.edu

## Abstract

The histone acetyltransferase GCN5-associated SAGA complex is evolutionarily conserved from yeast to human and functions as a general transcription co-activator in global gene regulation. In this study, we identified a divergent GCN5 complex in *Plasmodium falciparum*, which contains two plant homeodomain (PHD) proteins (PfPHD1 and PfPHD2) and a plant apetela2 (AP2)-domain transcription factor (PfAP2-LT). To dissect the functions of the PfGCN5 complex, we generated parasites with the bromodomain deletion in PfGCN5 and the PHD domain deletion in PfPHD1. The two deletion mutants closely phenocopied each other, exhibiting significantly reduced merozoite invasion of erythrocytes and elevated sexual conversion. These domain deletions caused dramatic decreases not only in histone H3K9 acetylation but also in H3K4 trimethylation, indicating synergistic crosstalk between the two euchromatin marks. Domain deletion in either PfGCN5 or PfPHD1 profoundly disturbed the global transcription pattern, causing altered expression of more than 60% of the genes. At the schizont stage, these domain deletions were linked to specific downregulation of merozoite genes involved in erythrocyte invasion, many of which harbor the DNA-binding motifs for AP2-LT and/or AP2-I, suggesting targeted recruitment of the PfGCN5 complex to the invasion genes by these specific transcription factors. Conversely, at the ring stage, PfGCN5 or PfPHD1 domain deletions disrupted the mutually exclusive expression pattern of the entire *var* gene family, which encodes the virulent factor PfEMP1. Correlation analysis between the chromatin state and alteration of gene expression demonstrated that up- and down-regulated genes in these mutants are highly correlated with the silenct and active chromatin states in the wild-type parasite, respectively. Collectively, the PfGCN5 complex represents a novel HAT complex with a unique subunit composition including the AP2 transcription factor, which signifies a new paradigm for targeting the co-activator complex to regulate general and parasite-specific cellular processes in this low-branching parasitic protist.

## Author Summary

Epigenetic regulation of gene expression plays essential roles in orchestrating the general and parasite-specific cellular pathways in the malaria parasite *Plasmodium falciparum*. Using tandem affinity purification and proteomic characterization, we identified a divergent transcription co-activator – the histone acetyltransferase GCN5-associated complex in *P. falciparum*, which contains nine core components, including two PHD domain proteins (PfPHD1 and PfPHD2) and a plant apetela2-domain transcription factor. To understand the functions of the PfGCN5 complex, we performed gene disruption in two subunits of this complex, PfGCN5 and PfPHD1. We found that the two deletion mutants displayed very similar growth phenotypes, including significantly reduced merozoite invasion rates and elevated sexual conversion. These two mutants were associated with dramatic decreases in histone H3K9 acetylation and H3K4 trimethylation, which led to global changes in chromatin states and gene expression. Genes significantly affected by the PfGCN5 and PfPHD1 gene disruption include those participating in parasite-specific pathways such as invasion, virulence, and sexual development. In conclusion, this study presents a new model of the PfGCN5 complex for targeting the co-activator complex to regulate general and parasite-specific cellular processes in this low-branching parasitic protist.

## Introduction

Packaging of the eukaryotic genomes with nucleosomes into chromatin affects all essential cellular processes such as transcription, DNA replication, and repair. A key mechanism for regulating chromatin structure involves post-translational modifications (PTMs) of nucleosomal histones, which can alter the accessibility of DNA and recruit distinct PTM readers and other effector proteins [1]. A multitude of histone PTMs such as acetylation, methylation, phosphorylation, ubiquitination, and sumoylation, act sequentially or combinatorially to create a "histone code" to facilitate or repress chromatin-mediated transcription [2-6]. Histone acetylation is a major PTM catalyzed by the histone acetyltransferase (HAT) enzymes and is correlated with active transcription [7]. HAT enzymes exist in large multimeric protein complexes such as the best-studied SAGA (Spt-Ada-Gcn5 acetyltransferase) complex [8] that is evolutionarily conserved from yeast to humans. The SAGA complex comprises 18–20 subunits, which are organized into functional modules including the HAT catalytic core, a histone deubiquitinase module, the TATA-binding protein (TBP) regulatory module, and the structural module [9]. In other SAGA-like complexes, the HAT catalytic core, consisting of the GCN5 acetyltransferase, ADA2, ADA3, and Sgf29, is conserved [9, 10]. Earlier studies suggested that SAGA regulates about 10% of genes in yeast and plants [11, 12], but a recent revisit of this issue in yeast revealed ubiquitous localization of SAGA at all gene promoters and reduced transcription of nearly all genes upon the disruption of SAGA [13, 14]. From these studies, SAGA appears to act as a general co-activator for all RNA polymerase II transcription, and its methyl reader (Sgf29) and acetyl reader (GCN5) subunits build synergistic crosstalk to coordinate transcription. As a co-activator complex that functions in the recruitment of the preinitiation complex, SAGA plays essential roles in metazoan development [15].

The human malaria parasite *Plasmodium falciparum* is an early-branching eukaryote, causing nearly half a million deaths in 2018 alone [16]. Its intricate lifecycle involving a vertebrate host and a mosquito vector requires precise regulation of transcription to cope with the comprehensive developmental program and environmental changes during host transitions [17-19]. Accumulating evidence indicates that the malaria parasite harbors unique properties of transcriptional regulation that are divergent from other eukaryotes even for the conserved transcription factors (TFs) [20-24]. Although the *Plasmodium* genome encodes the major

components of the general transcription machinery, there is a general deficiency of specific *Plasmodium* TFs [25, 26]. Compared to the yeast *Saccharomyces cerevisiae* genome with ~170 specific TFs [27], the similarly-sized genome of *P. falciparum* only has ~30 TFs, including the 27 apicomplexan-specific AP2-domain TFs [26, 28, 29]. In contrast, the presence of almost a full complement of proteins involved in chromatin biology underlines the significance of epigenetic regulation in malaria parasites [23, 30]. One distinct feature of the *P. falciparum* epigenome is that it mainly consists of euchromatin with restricted heterochromatin regions at subtelomeres and a few internal loci [31-37]. The heterochromatin clusters are localized at the nuclear periphery and demarcated by high levels of H3K9me3 and binding of heterochromatin protein 1 (HP1), and they control antigenic variation, drug sensitivity, and gametocyte production [31, 32, 37-42]. In comparison, the *Plasmodium* euchromatin is characterized by low or no nucleosome occupancy at the transcription start sites (TSSs) and core promoters of highly expressed genes, which exhibits cyclic changes during the intraerythrocytic development cycle (IDC) [43-48]. Euchromatin is marked with the active histone PTMs such as H3K9ac, H4K8ac, and H3K4me3 [32, 49, 50], presumably deposited by the HAT enzymes PfGCN5 and PfMYST, and the methyltransferase PfSET1, respectively [51-54]. Of these euchromatic marks, H3K9ac at the promoter regions correlates well with the transcriptional status of the genes, whereas H3K4me3 shows stage-specific regulation and does not exhibit general correlation with transcription [50]. Despite the importance of the euchromatin structures as witnessed by the essence of the writers of these histone marks [53, 55], the mechanisms by which these active histone marks are deposited, maintained, and dynamically regulated during development are unknown. More intriguingly, since most of the genes encoding the diverse cellular pathways reside in the euchromatic regions, it is not clear how the cascade-like gene expression pattern is achieved.

The *P. falciparum* genome encodes a single GCN5 protein, PfGCN5, with a long, unique N-terminal extension lacking similarity to known protein domains, and a conserved C-terminal HAT enzyme domain that can acetylate histone H3 at K9 and K14 *in vitro* [51]. During the IDC, PfGCN5 is present as a full-length form, which also undergoes proteolytic processing by a cysteine protease-like enzyme [56]. PfGCN5 is essential for the IDC of the parasites; thus its function has been probed by chemical inhibition of its activity, which caused overall disturbance of transcription and gross reduction of H3K9ac, establishing a potential link between PfGCN5 and H3K9ac in the parasite [52, 57]. Recent efforts aiming to identify readers of the PTMs in *P.*

*falciparum* led to the identification of putative PfGCN5-associated protein complex(es), which is highly divergent from the evolutionarily conserved SAGA complex [58]. Here, we used a tandem affinity purification (TAP) procedure to define a unique PfGCN5 complex and performed functional analyses of its key subunits. This work established the essential functions of this PfGCN5 complex in regulating cellular and metabolic pathways that are critical for parasite-specific processes such as antigenic variation, erythrocyte invasion, and sexual development.

## Results

### PfGCN5 Forms A Unique Complex Highly Divergent From The SAGA Complex

The evolutionarily conserved SAGA complex in eukaryotes comprises 18–20 subunits, which are organized into several modules including the HAT catalytic core consisting of GCN5, ADA2, ADA3, and Sgf29 [9, 10]. However, bioinformatic analysis of the *Plasmodium* genomes using conserved modular components of the SAGA complexes only identified two ubiquitous subunits GCN5 and ADA2 [51, 59], and a potential Tra1 homolog (PF3D7_1303800) [60], suggesting that the GCN5 complex(es) in these early-branching, parasitic protists might be highly divergent from the SAGA complex. Our recent work aiming to identify "readers" of modified histones with the H3K4me3 peptide surprisingly pulled down a putative PfGCN5 complex containing the PfGCN5, PfADA2, and two large proteins containing multiple plant homeodomains (PHDs) named PfPHD1 and PfPHD2 [58]. To precisely define the GCN5 complex(es) in *P. falciparum*, we tagged the C-terminus of the endogenous *PfGCN5* gene in the 3D7 strain with a PTP tag consisting of a protein C epitope, a tobacco etch virus (TEV) protease cleavage site, and two protein A domains (**Figure S1A, B**), which allows efficient TAP of protein complexes under native conditions with extremely low backgrounds [61, 62]. Integration-specific PCR confirmed the correct integration of the *PTP* tag at the *PfGCN5* locus (**Figure S1C)**. The transgenic parasites showed no noticeable *in vitro* growth defects (not shown). Western blot analysis using the anti-protein C antibody revealed that PfGCN5::PTP was expressed in all developmental stages of the IDC with the peak protein level in early trophozoites (**Figure S1D**). Six protein bands were detected and the band pattern agreed with that detected with an antibody against the PfGCN5 C-terminal fragment [56], confirming proteolytic processing of PfGCN5 (**Figure S1D**). Also, the processing of PfGCN5 was partially blocked

with the cysteine proteinase inhibitor E64 (**Figure S1E**). Live-cell imaging of the green fluorescent protein (GFP)-tagged PfGCN5 parasite line [63] showed that the PfGCN5::GFP protein was expressed throughout the IDC and localized in the nucleus (**Figure S2**). Thus, we performed the TAP procedure using nuclear extracts from $10^9$ synchronized trophozoites of the PfGCN5::PTP parasite, which was followed by liquid chromatography and tandem mass spectrometry (LC-MS/MS) for accurate protein identification. The MS data were subjected to Significance Analysis of INTeractome (SAINT) using a threshold of probability above 94% and false discovery rate (FDR) below 1% [64].

Three independent experiments of TAP and LC-MS/MS consistently identified nine proteins (**Figure 1A**, **Table S1A, B**), presumably representing the core subunits of this PfGCN5 complex (**Figure 1A**). This is in sharp contrast to the detection of only some abundant cellular proteins in the control pulldown experiments (**Table S1A**). Seven proteins identified using the TAP procedure were also present in the PfGCN5-associated proteins identified by a single-step pulldown procedure [58]. In agreement with our earlier work showing interactions between PfGCN5 and PfADA2 [59], these two proteins were among the most enriched proteins in the PfGCN5::PTP pulldown, demonstrating the high efficiency of the TAP procedure and preserved integrity of the complex. Consistent with the recent PfGCN5::GFP pulldown [58], the PfGCN5 core complex includes two large proteins PfPHD1 (PF3D7_1008100) and PfPHD2 (PF3D7_1433400), each containing four PHD zinc fingers (**Figure 1A, B, S3A, B**). PfPHD1 also contains two AT hooks, which are DNA-binding domains with a preference for AT-rich regions [65]. Sequence analysis indicated that these PHDs belong to the PHD superfamily with some containing additional cysteine and histidine residues (called extended PHD, ePHD). The ePHD has been found to bind dsDNA, methylated H3K4, or other TFs [66-69]. Only the fourth PHD in PfPHD1 conforms to the canonical PHDs that bind to H3K4me3/2 [70] (**Figure S3C**). Our recent study confirmed that this domain indeed preferentially binds H3K4me3/2 [58]. Furthermore, these two proteins were found to harbor large numbers of acetylation sites in our acetylome study [71], indicating that they are substrates of protein lysine acetyltransferases. An AP2-domain family TF (PF3D7_0802100), named PfAP2-LT, which is highly expressed at the late stages of the IDC [29], was consistently identified in all experimental replicates of the pulldown studies. Of note, the interaction between the PfGCN5 N-terminal fragment and AP2-LT has been identified in a genome-wide yeast two-hybrid screen [72]. Further, a histone

assembly protein PfNAPS (PF3D7_0919000) was also identified in the PfGCN5 interactome. Finally, in the PfGCN5 core complex are three hypothetical proteins with unknown functions (PF3D7_1019700, PF3D7_1364400, and PF3D7_1402800), which are conserved in all *Plasmodium* species (**Figure 1A, B**).

To confirm that PfPHD1 and PfPHD2 are the core constituents of the PfGCN5 complex, rabbit polyclonal antibodies were generated against peptides of PfPHD1 (aa 3685-3702) and PfPHD2 (aa 5738-5756). Indirect immunofluorescence assay (IFA) showed that the pre-immune sera did not react with parasitized red blood cells (RBCs) (data not shown), whereas the anti-PfPHD1 and -PfPHD2 antibodies detected fluorescent signals in the parasite nuclei (**Figure S4A**). Nuclear extracts of trophozoites from the PfGCN5::PTP line were subjected to immunoprecipitation (IP) with anti-PfPHD1 and anti-PfPHD2 antibodies, and the immunoprecipitated proteins were analyzed by immunoblotting with the anti-protein C antibodies. In contrast to the control IP with the pre-immune sera, PfGCN5-PTP was only detected in the IP with the anti-PfPHD1 or -PfPHD2 antibodies, confirming co-purification of PfGCN5 with PfPHD1 and PfPHD2 (**Figure S4B**).

IP and proteomic analysis with the tagged PfPHD1 and PfPHD2 subunits suggests they may be present in different versions of the PfGCN5 complex [58]. Whereas IP from PfPHD1::3×HA only purified PfPHD1, PfGCN5, ADA2, and PF3D7_1402800, IP from PfPHD2::GFP identified 12 putative subunits with marginal enrichment of PfPHD1 but no pulldown of PF3D7_1402800 [58]. To clarify this discrepancy with our TAP results, we separately tagged two subunits in these two putative versions of the PfGCN5 complexes for reciprocal IP: the PfPHD1 with a C-terminal c-Myc tag and PF3D7_1019700 with a C-terminal GFP tag (**Figure S5**). Correct integration of the *c-Myc* tag at the *PfPHD1* locus and the *GFP* tag at the *PF3D7_1019700* locus was confirmed by Southern blot analysis and integration-specific PCR, respectively (**Figure S5A, B, S6A, B**). Nuclear localization of the PfPHD1::Myc and PF3D7_1019700::GFP proteins was verified by cellular fractionation–Western blot (**Figure S5C**) and live-cell imaging analysis (**Figure S6C**), respectively. Affinity purification of the trophozoite nuclear extracts from the PfPHD1::Myc parasites with the Myc-trap beads followed by LC-MS/MS consistently identified all 9 subunits of the putative PfGCN5 complex, including PfPHD2 (**Figure 1A, Table S1C, D**). Similarly, two IP replicates from the

PF3D7_1019700::GFP identified 7 of the 9 core components of the PfGCN5 complex, including both PfPHD1 and PfPHD2 (**Table S1E, F**). These results are consistent with the presence of both PfPHD1 and PfPHD2 in the same PfGCN5 complex.

To estimate the size of the PfGCN5 complex, the purified native complex by TAP from the PfGCN5::PTP parasites was subjected to gel filtration, and fractions were assayed for HAT activity and Western blots using anti-PTP, -PfPHD1, and -PfPHD2 antibodies (**Figure 1C**). The results showed that PfGCN5 and its associated HAT activity, PfPHD1, and PfPHD2 all were detected in fractions 15 and 16, which is compatible with one major PfGCN5 complex in *P. falciparum* asexual stages. Based on the calibration using molecular mass standards, the size of the complex was approximately 2.3 MDa, which is comparable to the size (2.26 MDa) estimated based on the predicted molecular masses of the 9 core subunits (**Figure 1B)**.

**Domain Deletions in PfGCN5 and PfPHD1 Cause Severe Growth Defects in Parasites**

To characterize the function of the PfGCN5 complex in transcription regulation, we wanted to knock out the *PfGCN5, PfADA2, PfPHD1,* and *PfPHD2* genes by double-crossover homologous recombination but without success after multiple attempts (data not shown), indicating these genes are essential for parasite survival. This result is consistent with the mutagenesis scores in the genome-wide *piggyback* transposon mutagenesis study showing the essentiality of these genes [73]. Since PTM-binding domains are important for anchoring and withholding the respective proteins or complexes to the chromatin, we speculated that deleting these domains might disturb histone modifications without causing lethality to the parasite. Thus, we attempted to delete the bromodomain (BrD) and the PHD from the C-termini of PfGCN5 and PfPHD1, respectively, using a single-crossover gene disruption strategy, and meanwhile tag the C-termini of these truncated proteins with a GFP tag for sorting parasites with truncated PfGCN5 or PfPHD1 (**Figure S7A, C**). After transfection, the parasites were selected with WR99210, and GFP-positive parasites were cloned by sorting GFP-positive parasites using flow cytometry. Correct integration of the plasmids at the *PfGCN5* and the *PfPHD1* loci in the parasite genome was confirmed by Southern blots (**Figure S7B, D**). Phenotypic analyses of the parasites with the domain deletions in these two proteins revealed that the parasites with PfGCN5 BrD deletion (GCN5-ΔBrD) and parasites with PfPHD1 PHD domain deletion (PHD1-ΔPHD), to the greatest

extent, pheno-copied each other (**Figure 2**). Both the GCN5-ΔBrD and PHD1-ΔPHD parasites grew significantly more slowly than the wild-type (WT) parasites; they only reached ~1% parasitemia on day 7 as compared to ~10% in WT parasites (**Figure 2A**). A more detailed analysis of the growth defects in these domain deletion lines showed that mature schizonts in these mutant parasites produced similar numbers of merozoites as the WT parasites (**Figure 2B**), but these merozoites had substantially reduced efficiency (by almost 80%) in the invasion of RBCs ($P < 0.05$, paired Wilcoxon test; **Figure 2C**). In addition, these domain deletion mutants also had a 2–3 h longer IDC than the WT parasites ($P < 0.05$, paired Wilcoxon test; **Figure 2D**), and a more extended ring stage (**Figure 2E, F, S6E**). Further, these domain deletion parasites were inclined to produce more gametocytes than WT ($P < 0.05$, paired Wilcoxon test; **Figure 2G**).

**Domain Deletions in PfGCN5 and PfPHD1 Are Associated with Globally Reduced H3K9ac and H3K4me3**

*P. falciparum* has an extensively euchromatic epigenome with a preponderance of the histone marks H3K9ac and H3K4me3 [32, 50]. The presence of BrD and PHD in the PfGCN5 complex, which bind to acetylated H3K9/14 and H3K4me3/2 marks, respectively, strongly suggests that both domains may be required for anchoring the PfGCN5 complex to chromosomal regions to reinforce the euchromatic state. To determine the impacts of BrD and PHD deletions on the overall euchromatic histone marks, histones were purified from parasites of different developmental stages, and several histone marks were analyzed by Western blots. Consistent with PfGCN5 being the major HAT mediating H3K9 and H3K14 acetylation, deletion of BrD in PfGCN5 led to a significant reduction of the H3K9ac and H3K14ac levels, but the effect was more pronounced in the trophozoite stage, corresponding to the time of peak PfGCN5 expression (**Figure 3A**). Similarly, deletion of the PHD in PfPHD1 also resulted in the reduction of H3K9 and H3K14 acetylation. In comparison, domain deletions in PfGCN5 and PfPHD1 did not cause noticeable changes in H4 tetra-acetylation (at positions H4K5, 8, 14, and 20), which is mediated by another HAT protein, PfMYST [53]. Interestingly, domain deletions in these two subunits of the PfGCN5 complex also resulted in significantly reduced levels of H3K4me3 in trophozoites (**Figure 3A**), another major euchromatin mark conferred by the PfSET1 histone

methyltransferase, highlighting the presence of extensive crosstalk between the two euchromatin marks. This result echoes well with findings from studies of the SAGA complexes in model organisms, where GCN5 deletion or the Sgf29 Tudor domain deletion reduced the levels of both H3K9ac and H3K4me3 [13, 14, 74]. Taken together, these results indicate that both BrD in PfGCN5 and PHD in PfPHD1 are important for anchoring the PfGCN5 complex to maintain the euchromatic histone marks.

Spatial compartmentalization of chromatins in the nucleus is critical for gene regulation in malaria parasites. The active chromatin marks H3K9ac and the H3K14ac overlap extensively with the DAPI staining (often used to define the euchromatin domains), whereas the heterochromatin mark H3K9me3 mostly occupies the nuclear periphery outside of DAPI [31, 75]. To analyze whether these changes in histone modifications were associated with altered spatial organization of the chromatins, we used IFA and live microscopy to observe the nuclear locations of these histone modifications as well as the truncated PfGCN5 and PfPHD1. In the PfGCN5-ΔBrD::GFP parasites, Western blot with the anti-GFP antibodies detected reduced expression of the truncated PfGCN5 protein, while the pattern of the PfGCN5 fragments remained similar (**Figure S8A**). Live microscopy of the GFP-tagged truncated PfGCN5-ΔBrD::GFP and PfPHD1-ΔPHD::GFP parasites showed that the GFP signals, while overlapping largely with the parasite nuclei from DAPI staining (**Figure S8B, C**), were more diffused than those in the WT parasites (**Figure S2**). Consistent results were obtained by IFA with anti-GFP and H3K9ac antibodies in WT and PfGCN5-ΔBrD::GFP parasites (**Figure 3B, C**). In the PfGCN5::GFP parasites, there were high levels of overlaps between the DAPI and PfGCN5 ($r^2$=0.83–0.91), and between PfGCN5 and H3K9ac ($r^2$=0.89–1.0), indicating that PfGCN5 is tightly associated with H3K9ac in the active euchromatin area demarcated by DAPI staining (**Figure 3B**). However, the levels of the co-localization substantially decreased in the PfGCN5-ΔBrD::GFP parasites ($r^2$=0.37–0.74) (**Figure 3C**). Some signals of the truncated PfGCN5 were localized beyond the DAPI area, suggesting that PfGCN5 might have spread to the perinuclear heterochromatic area. Altogether, these results indicated that domain deletions in PfGCN5 and PfPHD1 altered the nuclear distribution of the truncated proteins and reduced the levels of euchromatin marks in the parasites.

## PHD Deletion in PfPHD1 Affects the Integrity of the PfGCN5 Core Complex

To determine whether BrD deletion in PfGCN5 and PHD deletion in PfPHD1 affected the integrity of the PfGCN5 complex, we performed IP by using the GFP-Trap antibodies with nuclear extracts from the individual domain deletion lines and analyzed the affinity-purified proteins by LC-MS/MS. Four replicates of IP using the PfGCN5::GFP parasites as the positive control consistently detected the core components of the PfGCN5 complex (**Figure 3D, Table S2A, B**). IP with the PfGCN5-ΔBrD::GFP parasites also confidently purified the 9 core subunits of the PfGCN5 complex, suggesting the deletion of the BrD from PfGCN5 did not affect the integrity of the complex (**Figure 3D, Table S2C, D**). However, only four major components of the PfGCN5 complex (PfGCN5, PfADA2, PfPHD1, and PF3D7_1402800) were detected from the PfPHD1-ΔPHD::GFP parasites (**Figure 3D, Table S2E, F**).

## BrD and PHD Deletions Profoundly Affect Global Transcription

Since GCN5-associated complexes facilitate transcription of target genes by bridging transcriptional activator and the preinitiation complex, deletion of domains from subunits of the complex that interact with histone tails weakens the anchoring and retention of the complex, leading to reduced transcriptional activation of the target genes [74, 76, 77]. To gain a mechanistic understanding of the PfGCN5 complex in transcriptional regulation, we compared the transcriptomes of the WT parasites and parasites with domain deletions. Parasites were highly synchronized from purified schizonts with a 3 h window, and RNA-seq analysis was performed during the IDC at 10, 20, 30 and 40 h after RBC invasion in the WT parasites and at 10, 23, 33 and 43 h in the parasites with BrD and PHD deletions to more closely match the developmental stages of the WT and domain-deletion parasites based on comparison of their IDC (**Figure 2E, F**). Compared to the phaseogram of WT parasites displaying a clear cascade-like gene expression pattern [18], PfGCN5 BrD deletion profoundly disturbed the global transcription pattern, causing 3533 (62.6%) genes to be differentially expressed in at least one of the four IDC time points analyzed (**Figure 4A**, **Table S3, S4**). Specifically, BrD deletion resulted in down-regulation of 997, 799, 861 and 902 genes, and up-regulation of 1127, 780, 846, and 368 genes at the ring, early trophozoite, late trophozoite, and schizont stage, respectively (**Figure 4B-D, S9A, B**). Noticeably, the numbers of up- and down-regulated

transcripts were comparable at all stages except at the schizont stage, where 2.5-fold more transcripts were down-regulated than up-regulated (**Figure 4D**). In comparison, PHD deletion in PfPHD1 caused a similar but more profound disturbance of gene expression during the IDC, with 3870 (68.6%) transcripts being differentially expressed in at least one of the four stages analyzed (**Figure 4A**), which is congruent with the more substantial disruption of the PfGCN5 complex upon PHD deletion (**Figure 4**). The PfPHD1 PHD deletion resulted in the down-regulation of 872, 1021, 557, and 787 genes, and up-regulation of 1481, 1266, 648, and 1028 genes at the ring, early trophozoite, late trophozoite and schizont stage, respectively (**Figure 4D-F, S9C, D)**. Of note, only in late trophozoites did PfPHD1 PHD deletion disturb the expression of fewer genes than PfGCN5 Brd deletion (1205 vs. 1707 genes) (**Figure 4D**).

With both PfGCN5 and PfPHD1 being integral members of the PfGCN5 complex, the global transcription changes resulted from the domain deletions of these two genes were remarkably similar; their transcriptomes showed a significant correlation between the respective stages with correlation coefficients ranging from 0.67 to 0.82 (**Figure S9E**). In addition, 44.1–63.8% of the down-regulated genes in each stage were shared between the two domain deletion strains (**Figure 4G**). In comparison, 36.1–49.2% of the up-regulated genes were shared between the two domain deletion mutants in the ring, early and late trophozoites, but this up-regulated gene repertoire in the two domain deletion strains only shared 18.8% during the schizont stage (**Figure 4H**).

**BrD and PHD Deletions Affect Parasite-specific Cellular Pathways**

To determine whether deletion of PfGCN5 BrD or PfPHD1 PHD affected specific biological processes, gene ontology (GO) enrichment analysis was performed on the genes with significantly altered expression (**Figure 5A, B, S10, Table S5**). During the early IDC stages, transcripts associated with cytoadherence, merozoite invasion, DNA replication, and organellar activities were up-regulated (**Figure 5A, S10A, Table S5**). Of particular interest, the 60 *var* genes, encoding the virulent factor PfEMP1 that mediates cytoadherence, were dramatically upregulated in the early stages, with their overall transcripts increased ~4-fold upon BrD deletion and ~25-fold upon PHD deletion (**Figure 5C, Table S6**). It is noteworthy that many *var* members were up-regulated, albeit *var2csa* was a major *var* gene expressed in PfGCN5-ΔBrD,

suggesting activation of the overall *var* gene family. Western blot using antibodies against the conserved cytoplasmic ATS domain of the PfEMP1 proteins detected more complicated expression patterns and higher abundance of the PfEMP1 on the surface of RBC infected by trophozoite-stage parasites upon PfGNC5 BrD and PfPHD1 PHD deletions (**Figure 5D**). To further determine whether domain deletion in PfGCN5 and PfPHD1 activated multiple *var* members in a single infected RBC, we performed single-cell RNA-fluorescent in situ hybridization (FISH) using the type B *var* exon 2 as the probe [31], which is predicted to hybridize to 22 type B *var* genes. *Var* genes are clustered into 6-8 foci at the nuclear periphery and colocalized with the "telomere bouquets" [34], while the active *var* gene is localized to a *var*-specific expression site [78, 79]. Consistent with the mutually exclusive expression of *var* genes in single cells, the majority of the RNA-FISH positive cells contained one fluorescent spot (mean ± standard deviation, 1.04±0.21, n=67) indicating expression of a type B *var* (**Figure 5E**). In contrast, 43.9% and 58.2% positive rings had more than one hybridization signal in the PfGCN5-ΔBrD (1.55±0.66, n=66) and PfPHD1-ΔPHD parasites (1.91±0.98, n=79), respectively. In addition, 28.7% (45/157) of the WT 3D7 ring-stage parasites showed hybridization, which increased to 34.5% (61/177) and 41.3% (71/172) in the PfGCN5-ΔBrD and PfPHD1-ΔPHD rings, respectively. It is noteworthy that the hybridization signals were mostly localized in areas at the periphery of the DAPI staining. Thus, these results indicate the presence of multiple *var* expression sites in the PfGCN5-ΔBrD and PfPHD1-ΔPHD parasites.

Conversely, genes in the biological processes of translation and transcription were significantly enriched in the down-regulated genes upon BrD or PHD deletion during the early IDC, which was probably responsible for the slowing down of development (**Figure 5B, S10B, C, Table S6**). During the late IDC (late trophozoites and schizonts), genes involved in RBC invasion were greatly reduced, which is consistent with the phenotype of reduced RBC invasion rates of the PfGCN5-ΔBrD and PfPHD1-ΔPHD merozoites (**Figure 5B, 6A, S10B, Table S6**). Of the 86 putative invasion-related genes [80], 76 showed peak expression at the late stages of IDC in WT parasites (**Figure 6A**). Except for *MSRP1* that was up-regulated, 75 genes were significantly down-regulated at the late stages in the PfGCN5-ΔBrD and PfPHD1-ΔPHD parasites (**Figure 6A**). These data collectively indicate the involvement of the PfGCN5 complex in the regulation of the general cellular processes such as transcription, translation,

oxidoreduction, and organellar function, as well as parasite-specific processes of pathogenesis and host cell invasion.

**Potential Coordination of the PfGCN5 Complex by AP2 Transcription Factors**

The profound effects of PfGCN5 BrD/PfPHD1 PHD deletion on particular pathways such as invasion and cell adhesion suggest that these coordinated changes may involve the participation of specific TFs of the ApiAP2 domain family [81, 82]. The consistent pulldown with the PfGCN5 complex of the AP2-LT that is expressed abundantly in late stages of the IDC and occasionally other AP2 proteins (**Figure 1A**, **3D**) [58] suggests coordination of chromatin modification with transcription activators/repressors. To this end, we analyzed the transcriptional changes of all 27 ApiAP2 TFs after BrD or PHD deletion. Indeed, the cascade of AP2 transcriptions was substantially disturbed; some AP2-TFs such as AP2-SP2 and AP2-O3 were activated at stages when they are normally silenced in WT parasites, whereas some (e.g., AP2-O, AP2-I, and AP2-LT) were down-regulated at the stages when they are supposed to be active (**Figure 6B, Table S6**). PfSIP2 is associated with the chromosomal end clusters and is required for heterochromatin formation and genome integrity, including the silencing of subtelomeric *var* genes [83]. Its down-regulation upon BrD and PHD domain deletion may influence the organization of the subtelomeric heterochromatin, resulting in the overall de-repression of *var* genes (**Figure 6B**). The master regulator of gametocytogenesis AP2-G [84, 85] was consistently up-regulated upon Brd and PHD deletions (**Figure 6B**), which agrees with the increased gametocytogenesis detected in the mutant parasite lines (**Figure 2G**). The downregulation of AP2-I and AP2-LT might partially explain the downregulation of the invasion-related gene upon domain deletions. Of the 75 invasion-related genes down-regulated upon BrD or PHD deletion, 19 (mostly associated with the rhoptry) are targets of AP2-I, and 33 were predicted targets of AP2-LT (**Figure 6A, Table S6**) [29, 86]. The AP2-LT subunit of the PfGCN5 complex is predicted to bind to 986 genes with the motif sequence ACACA [29]. Analysis of the genes altered upon BrD deletion in PfGCN5 and PHD domain deletion in PfPHD1 revealed that genes down-regulated in late trophozoite and schizont stages were significantly enriched in those containing the AP2-LT binding motif (**Figure 6C, S10D**). This finding is consistent with the decreased recruitment of the PfGCN5 complex in the deletion mutants by AP2-LT to genes with

AP2-LT binding motifs, resulting in extensive down-regulation of the gene categories during the late-stage development (**Figure 5B**). The unique presence of AP2 TFs in the GCN5 complexes in apicomplexan parasites suggests that the GCN5 complexes are specifically recruited to regulate the expression of certain clusters of genes [87].

**BrD and PHD Deletions Broadly Alter Chromatin Structures**

The transcriptomic phaseograms showed that many genes were expressed out of the "phase" in the PfGNC5-ΔBrD and PfPHD1-ΔPHD parasites (**Figure 4A**); genes that are normally active were down-regulated, whereas genes supposed to be silent were active at the wrong time during the IDC. Since epigenetic regulation of gene expression in *P. falciparum* is most evident in heterochromatic regions [23], while gene expression from euchromatic regions correlates positively with the chromatin accessibility [48], we compared the chromatin status and accessibility of genes with altered expression upon PfGCN5 and PfPHD1 domain deletions. We first compared the up- and down-regulated genes with the accessibility of their promoters determined by the assay for transposase accessible chromatin sequencing (ATAC-seq) [48]. In all time points, the down-regulated genes upon BrD or PHD deletion are enriched in genes with more open chromatin structures at the promoters in the WT parasites, whereas up-regulated genes are significantly enriched in genes with less open promoters in the WT parasites (**Figure 7A**). Conversely, up-regulated genes upon BrD or PHD deletion are significantly more often associated with the heterochromatin loci that are normally enriched with HP1 and repressed during the IDC (**Figure 7B, Table S7**). This group of genes includes many variant gene families (*var*, *rifin* and *stevor*) and AP2-G (**Figure 5B, S10B, Table S7**). Among the genes that were up-regulated upon BrD or PHD deletion but have low accessibility in their promoters during the IDC in the WT parasites are the genes specific for sexual-stage development, which are normally silent during the IDC in the WT parasites. BrD deletion led to significant up-regulation of 353 gametocyte- and 401 ookinete-specific genes, respectively (**Figure 7C, D, Table S8**), and many were up-regulated at the ring stage. Similarly, PHD deletion caused up-regulation of 403 gametocyte- and 401 ookinete-specific genes, respectively (**Figure 7C, D, Table S8**). Among them, 151 gametocyte- and 199 ookinete-specific genes are shared between both deletional mutants. Taken together, both domain deletions similarly affected chromatin structures and led to

the activation of genes involved in sexual development.

To verify that the PfGCN5 complex directly regulated the accessibility and chromatin state of the promoters, we selected three genes to evaluate the recruitment of the complex by chromosome immunoprecipitation (ChIP) followed by qPCR analysis. Compared to the WT parasites, the truncated PfGCN5-ΔBrD and PfPHD1-ΔPHD were significantly depleted at the *PfMSP1* promoter in schizonts, whereas they were enriched at the ring stage (**Figure S11A**). Their dynamic associations with the *PfMSP1* promoter were correlated with the down- and up-regulation of *PfMSP1* in these two domain deletion lines. Consistently, the PfGCN5-ΔBrD and PfPHD1-ΔPHD proteins were significantly enriched at the promoter of the *var2csa* gene and the gametocyte-specific gene *Pfg27/25* at the ring stage (**Figure S11B, C**), which was correlated with significant up-regulation of these two genes in the domain deletion parasites.

## Discussion

An unusual aspect of the chromatin-mediated gene regulation in the malaria parasite *P. falciparum* is that the parasite epigenome is dominantly euchromatic marked extensively with H3K9ac and H3K4me3 [32, 49], whereas heterochromatin-associated histone modifications H3K9me3 and H3K36me3 are localized to genes undergoing variable expression [31, 88]. This is contrasted to most eukaryotes where these heterochromatin marks are found in genes throughout the genome. Here we identified a unique GCN5 complex in *P. falciparum* responsible for depositing the euchromatic marks H3K9ac and H3K14ac, which is drastically different from the canonical SAGA complex that is conserved from yeast to human. Functional characterization of two major subunits demonstrated the crucial functions of the PfGCN5 complex in regulating global gene expression and parasite biology during its development in the host RBCs.

The evolutionarily conserved SAGA complex has a modular structure that supports multiple activities including histone acetylation, deubiquitination, and interactions with TFs [9]. Using multiple approaches, we refined the subunits of the GCN5 complex in *P. falciparum*, which completely lacks the deubiquitinase, the TBP-associated proteins, and the core structural module with the histone folds. Even for the HAT catalytic core [10], the *P. falciparum* GCN5 complex is also distinctive with the conservation of only the GCN5 and ADA2 homologs.

Interestingly, the PfGCN5 complex also appears to differ substantially from the *Toxoplasma gondii* GCN5b complex [87, 89], as the latter contains additional proteins potentially involved in RNA binding and transcription elongation. In the PfGCN5 complex, the H3K4me3-binding activity, mediated by the tandem tudor domains of Sgf29 in the SAGA complexes [74, 90], is replaced by the PfPHD1 protein [58]. Moreover, PfPHD2 in the PfGCN5 complex contains four atypical PHDs, and may bind to other histone modifications. Importantly, since deletion of either BrD or PHD affected their localization and reduced the levels of both H3K9ac and H3K4me3, these domains are needed for anchoring and retention of the PfGCN5 complex on chromatin, similar to what was observed in the SAGA complex [13, 14, 74]. This result also implies the presence of synergistic crosstalk between the PfGCN5 complex and the histone H3K4me3 methyltransferase complex. As in model organisms, the binding of GCN5 BrD to H3K9Ac likely promotes H3 acetylation, which in return augments H3K4me3, since histone methyltransferases have a preference for the acetylated H3 tail [91-94]. The interactions between PfGCN5 and PfSET1 identified through yeast two-hybrid analysis further attests to the crosstalk between these euchromatic histone marks [72]. Given that the H3K4me3 levels are gradually increased toward the late stages of the IDC [50] and that most genes affected during PfSET1 knockdown are also expressed at the late stages [55], the intricate interplay between the writer complexes of these euchromatin marks needs to be further dissected.

Our earlier suggestion that PfPHD1 and PfPHD2 may represent different flavors of the PfGCN5 complex was based on the identification of only four proteins (PfGCN5, PfADA2, PfPHD1 and PF3D7_1402800 ) in the PfPHD1::3×HA pulldown and a complete lack of PF3D7_1402800 from the PfPHD2::GFP pulldown [58]. Here we provided evidence suggesting that both PfPHD1 and PfPHD2 are subunits of the same PfGCN5 complex. First, reciprocal pulldown with both the PfPHD1::Myc and PF3D7_1019700::GFP, which belong to the two putative PfGCN5 subcomplexes suggested earlier, identified the core components of the PfGCN5 complex with the abundant presence of both PfPHD1 and PfPHD2 (**Figure 1, Table S1**). Second, both PfPHD1 and PfPHD2 were co-eluted in the same fractions during gel filtration of the PfGCN5 complex. Third, the predicted size of the single PfGCN5 complex is compatible with the summation of the core subunits, while missing either of these two large PHD proteins would drastically reduce the size of the complex. Thus, this discrepancy may be due to the use of different tags for PfPHD1 (c-Myc vs 3×HA tag) and the different stringency of the analysis (1%

FDR used here vs 10% used earlier). Also, we found that PfPHD1 could not be tagged with a larger tag such as the GFP (not shown), suggesting that tagging PfPHD1 with 3HA may interfere with the integrity of the PfGCN5 complex. Interestingly, pulldowns with both PfPHD1::3×HA and PfPHD1-ΔPHD::GFP identified the same four subunits of the PfGCN5 complex (**Figure 3D**). Thus, studies employing biochemical and cryogenic electron microscopy will allow a further resolution of the PfGCN5 complex.

The SAGA co-activator complex plays a critical role in regulating global gene expression [13, 14]. In *P. falciparum* with an unusual, dominantly euchromatic epigenome, the significance of the PfGCN5 complex is demonstrated by the essence of several core subunits for asexual development. Although domain deletion partially relieved the problem of lethality, deletion of either the BrD in PfGCN5 or PHD in PfPHD1 caused considerable growth defects during the IDC and altered expression of >60% of genes with an approximately equal number of up- and down-regulated genes. This defective gene expression pattern may be due to reduced levels of H3K9ac and H3K4me3, as well as mislocalization of the PfGCN5 complex. This may have led to the opposite changes in chromatin state, which are correlated with the genome-wide "out of phase" expression pattern. These global transcriptional changes are reminiscent of those when parasites were treated with the histone deacetylase inhibitor apicidin, which caused a reduction of H3K9Ac and H3K4me3 [95]. Analysis of genes with altered expression upon domain deletions for their chromatin state (HP1 occupancy) and promoter openness (ATAC-seq) provided indirect evidence supporting the mislocalization of the PfGCN5 complex in the deletion mutants. It is also noteworthy that PHD deletion in PfPHD1 caused more severe effects on gene expression, which can be explained by the disturbance of the integrity of the PfGCN5 complex upon PHD deletion, suggesting that PfPHD1 may play a scaffolding role for the structural integrity of the complex. Interestingly, the transcriptomic changes after BrD and PHD domain deletions (presumably due to reduced expression and mislocalization effects) are distinct from those after deletion of the SAGA subunits in yeast and human cells, which showed global downregulation of transcription [13, 14].

SAGA is recruited to the promoters through the interactions between Tra1 and TFs. Although a Tra1 homolog is present in the *P. falciparum* genome, it was not identified in the PfGCN5 complex. Another distinguishing feature of the PfGCN5 complex is the presence of an AP2-domain TF (PfAP2-LT) as a consistent member of the complex. Of note, the core GCN5b

complex in the apicomplexan parasite *T. gondii* tachyzoites also contains multiple AP2 factors [87], suggesting a conserved characteristic in these lower-branching eukaryotes. These TFs would allow direct recruitment of the GCN5 complex to the target gene promoters, circumventing the need for a bridging factor such as Tra1. In addition to the identification of AP2-LT in the complex, two other AP2 factors (AP2-I and PF3D7_1239200) were also identified in the PfPHD1::Myc pulldown, suggesting that they might be either loosely associated with the GCN5 complex or represent additional flavors of the minor GCN5 complexes. In support of this notion, PfGCN5 was also identified in the AP2-I pulldown [86]. It is noteworthy that all pulldown experiments in this study were performed in the late trophozoite stage when AP2-LT is highly expressed, suggestive of the possibility that other AP2 factors may be associated with the PfGCN5 complex during different developmental stages. With the H3K9ac mark at the promoter regions dynamically following the pattern and level of transcription throughout the IDC [50], it is logical to propose that the dynamic recruitment of the PfGCN5 complex to different promoters is mediated by different AP2 factors. This hypothesis is compatible with the recruitment of the GCN5 complex by AP2-I to the promoters of invasion-related genes to acetylate the histones, which then recruit the BrD protein PfBDP1 [86]. In line with this, the genes predicted to be the targets of AP2-LT were mostly down-regulated upon PfGCN5 BrD and PfPHD1 PHD deletions, consistent with the AP2-LT-mediated recruitment of the PfGCN5 complex to the promoters of these genes. Of the 86 invasion-related genes expressed in merozoites, 33 were predicted to be the AP2-LT target genes, and 15 of these genes also are AP2-I targets (**Figure 6A**). Moreover, the majority of those AP2-LT target genes specifically affected by the domain deletions in PfGCN5 and PfPHD1 are expressed in late stages, coincidental with the peak expression of the AP2-LT (**Figure 6C**).

Antigenic variation in *P. falciparum* is mediated by the monoallelic expression of the ~60 members of the *var* gene family [96]. *Var* gene clusters are located in the heterochromatin regions of the nuclear periphery and their expression is associated with the relocation and change of the promoter to a euchromatic state [38, 78, 97]. Such a mutually exclusive, monoallelic expression pattern of *var* genes is completely disrupted when the PfGCN5 BrD or PfPHD1 PHD domain was deleted, evidenced by the simultaneous expression of multiple *var* genes in single infected RBCs. This phenotype is similar to what was observed when the histone deacetylate PfHda2, PfSir2A, B, and the heterochromatin marker PfHP1 were experimentally knocked out or

knocked down [98-101]. Given the presence of potential binding elements for AP2 factors in the *var* promoters, we cannot fully exclude the possibility that the activation of the *var* gene family upon PfGCN5 or PfPHD1 domain deletion may be due to altered recruitment of the PfGCN5 to *var* promoters through AP2 factors present in subsets of the GCN5 complex. However, the increased expression of most genes located in the heterochromatin regions marked by PfHP1 suggests that the expression of the whole *var* gene family may reflect a general loss of heterochromatin-based silencing instead of a specific effect on the mutually exclusive *var* gene expression. As the silent *var* genes cluster into 6-8 foci with the telomeres in the nuclear periphery, their activation upon PfGCN5 BrD or PfPHD1 PHD deletion did not seem to involve their "moving" to a single transcriptionally competent locus, but rather appearing as multiple active *var* loci in the nuclear periphery. This is consistent with the observed expansion of the truncated PfGCN5 or PfPHD1 to the outer nuclear compartments beyond the DAPI-stained central region. Moreover, the magnitude of all *var* transcripts compared with that in the WT parasite was significantly higher during PHD deletion (~25-fold) than during BrD deletion (~4-fold), which also agrees with the more severe effect of PHD deletion on the integrity of the PfGCN5 complex. This result emphasizes that maintaining the spatial organization of the different chromatin domains in *P. falciparum* is crucial for regulating antigenic variation.

This study has demonstrated the power of the TAP procedure for more precisely identifying protein complexes in malaria parasites. The nine subunits identified in this study may constitute the major PfGCN5 complex in the late stages of the IDC, while multiple variants of the PfGCN5 complex may exist to carry out different biological functions. Single-step IP identified large numbers of additional, less abundant proteins, which may represent those that are either associated less tightly with the PfGCN5 core complex or are the subunits of variant PfGCN5 complexes. In particular, a conserved histone modification reader protein Pf14-3-3I (PF3D7_0818200) was also identified in the pulldowns with both PfGCN5 and PfPHD1. Pf14-3-3I binds to purified parasite histones and H3 phosphopeptides [102] and its potential binding to H3Ser10p may further favor the recruitment of the PfGCN5 complex to acetylate H3K14 as identified in yeast [103], pointing to the presence of extensive crosstalk of epigenetic marks in *P. falciparum*. Other identified proteins with the GCN5 complex are mostly associated with the biology of chromatins including components of the SWI-SNF chromatin-remodeling complex and proteins with histone-interacting domains (e.g., BrD, WD40, PHD, CHD). Their potential

associations with the PfGCN5 complex highlight the complexity of epigenetic regulation of gene expression in *P. falciparum*.

This study revealed a unique GCN5 complex in a lower eukaryotic parasite that is drastically distinct from the evolutionarily conserved SAGA complexes from yeast to human. The PfGCN5 complex, which is essential for regulating the stage-specific gene expression cascade, is also involved in the control of parasite-specific biologies such as RBC invasion and virulence. The critical role of the PfGCN5 complex in parasite biology and its significant divergence from the human host suggest that the PfGCN5 complex may be a vital target for chemotherapy against malaria parasites. Down this line, efforts are directed at identifying selective molecules inhibiting the GCN5 enzyme activity and selective inhibitors disrupting the interaction between the PfGCN5 BrD and acetylated histones [104, 105].

## Material and methods

### Parasite culture

The *P. falciparum* strain 3D7 and its genetically modified clones were cultured at 37°C in a gas mixture of 5% $CO_2$, 3% $O_2$ and 92% $N_2$ with type $O^+$ RBCs at 5% hematocrit in RPMI 1640 medium supplemented with 25 mM $NaHCO_3$, 25 mM HEPES, 50 mg/L hypoxanthine, 0.5% Albumax II and 40 mg/ml gentamicin sulfate [106]. Synchronization of asexual stages was performed by two rounds of sorbitol treatment at the rings stage or by incubation of schizonts with RBCs for 3 h to obtain highly synchronized ring-stage parasites [107].

### Genetic manipulation of *PfGCN5* and its associated genes

To tag the C-terminus of PfGCN5 with the PTP tag [61, 62], the C-terminal *PfGCN5* fragment [nucleotides (nt) 3778-4758] was amplified from *P. falciparum* genomic DNA using primers F1 × R1. For the deletion of the PfGCN5 BrD and the PfPHD1 PHD domain, the *PfGCN5* fragment (nt 3286-4044) and the *PfPHD1* fragment (nt 3286-4044) were amplified using primers F2 × R2 and F3 × R3, respectively. All amplified fragments were first cloned into a modified pBluescript SK plasmid to fuse with the PTP or GFP and *pDT* 3' UTR as described earlier [53, 108]. This cassette was then subcloned into pHD22Y at *Bam*HI and *Not*I sites to produce pHD22Y/PfGCN-PTP, pHD22Y/GCN5-ΔBrD-GFP, and pHD22Y/PHD1-ΔPHD-GFP, respectively. A similar

strategy was used to tag PfPHD1 with c-Myc and PF3D7_1019700 with GFP. All primers used for tagging, domain deletion, PCR verification of integration and probe for Southern blot are listed in **Table S9A**.

Parasite transfection was done using the RBC loading method [109]. Briefly, 100 µg of plasmid were introduced into fresh RBCs by electroporation. Purified schizonts were used to infect the RBCs pre-loaded with the plasmid and selection was done with 2.5 nM of WR99210 for approximately 4 weeks with weekly replenishment of fresh RBCs until resistant parasites appeared. Resistant parasites were subjected to three cycles of drug on-off selection and single clones of parasites with stable integration of the constructs were obtained by limiting dilution [107]. For the parasites transfected with constructs containing a GFP tag, fluorescence-activated cell sorting was employed to clone the GFP-positive parasites. Correct integrations of plasmids into the parasite genome were screened by integration-specific PCR or Southern blot with the digoxigenin (DIG)-labeled probes using an established protocol [110].

**Purification of protein complexes**

TAP was performed using the PTP-tagged PfGCN5 parasite line according to the published method [62]. Briefly, $10^9$ parasites were lysed in 5 volumes of the hypotonic buffer (10 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.5 mM EDTA) at 4°C for 10 min followed by centrifugation for 20 min at 500 × g. The resultant pellet (nucleus) was further lysed in 5 volumes of PA150 buffer (150 mM KCl, 20 mM Tris-HCl, pH 7.7, 3 mM $MgCl_2$, 0.5 mM DTT, and 0.1% Tween 20) containing a protease inhibitor cocktail (Roche). The lysate was centrifuged for 10 min at 16,000 × g and the supernatant was incubated with 100 µl (settled volume) of IgG agarose beads (GE Healthcare) at 4°C for 2 h. The beads were washed twice with PA150 and equilibrated twice with the TEV buffer (150 mM KCl, 20 mM Tris-HCl, pH 7.7, 3 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, and 0.1% Tween 20). To release the PfGCN5 and its associated proteins from IgG beads, the beads were incubated with 2 ml of TEV buffer containing 150 U of TEV protease and rotated overnight at 4°C. The supernatant was collected, and the beads were rinsed with another 4 ml of the PC150 buffer (150 mM KCl, 20 mM Tris-HCl, pH 7.7, 3 mM MgCl2, 1 mM $CaCl_2$, 0.1% Tween 20). Then, 7.5 µl of 1 M $CaCl_2$ were added to titrate the EDTA from the TEV buffer and the combined supernatant was incubated with the anti-protein C beads for 2 h at 4°C. The beads were washed four times with PC150 and

eluted with the buffer containing 10 nM EGTA/5 mM EDTA. For the single-step pulldown of GFP-tagged or Myc-tagged protein, GFP- or Myc-trap (Cat# gta-20, RRID:AB_2631357 or Cat# yta-20, RRID:AB_2631369, Chromotek) beads were used with lysates from $10^9$ parasites according to the manufacture's protocol.

**Mass spectrometry**

The proteins in the elution were concentrated by Amicon Ultra centrifugal filters (Millipore Sigma) and separated briefly in a 10% Bis-Tris SDS-PAGE gel for 10 min. Proteins in gel were excised, in-gel digested, and analyzed by nano-LC/MS/MS using a Waters NanoAcquity HPLC system interfaced to a Q Exactive™ Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific) [111]. Peptides were loaded on a trapping column and eluted over a 75 µm analytical column at 350 nL/min. MS and MS/MS were performed at 70,000 FWHM and 17,500 FWHM resolutions, respectively. The fifteen most abundant ions were selected for MS/MS. Parasite proteins were identified by searching the Uniprot *P. falciparum* protein database (v01/2014). Data were filtered at 1% protein and 0.2% peptide FDR, and at least two unique peptides per protein. Mascot DAT files were parsed into the Scaffold software for validation and filtering to create a non-redundant list per sample. The available mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [112] partner repository with the dataset identifier PXD023389 and 10.6019/PXD023389.

**Antibody generation**

To generate antibodies against PfPHD1 and PfPHD2, a PfPHD1 peptide (DNGKLQKVDGRKKRRYHK, aa 3685-3702) and a PfPHD2 peptide (DDNVKAEDYKDENNDNDGD, aa 5738-5756) were synthesized and rabbits were immunized with these peptides. After three times immunizations, the antibodies were purified by affinity purification with peptides conjugated to the beads (Proteintech Group).

**Gel filtration**

To access the size of the PfGCN5 complex, nuclear extract from the PfGCN5::PTP was incubated with IgG beads, eluted by TEV protease cleavage as described above, and applied to a Superose 6 gel filtration column (GE Healthcare). Molecular mass standards (Gel Filtration Calibration Kit HMW, GE Healthcare) were run under the same conditions to estimate the size

of the complex. The fractions were analyzed by Western blotting using antibodies against the PTP tag, PfPHD1 and PfPHD2, while HAT activity in the fractions was measured as described previously [53].

**Growth phenotype analysis**

The growth phenotypes of GCN5-ΔBrD::GFP and PHD1-ΔPHD::GFP lines during the IDC were compared with the WT 3D7 parasites as described [53]. To measure cell cycle progression, highly synchronous rings were obtained by incubation of schizonts with RBCs for 3 h. Progression of parasites through the IDC was monitored using Giemsa-stained smears every 2 h. Cycle time was determined as the duration between the peak ring parasitemias of two consecutive cycles. To measure parasite proliferation, synchronous cultures after two rounds of consecutive synchronization by sorbitol were initiated at 0.1% rings, and parasitemia was monitored daily for 7 days without replenishment of the RBCs. The number of merozoites produced per schizont was determined from mature segmenters. Three independent biological replications were done for each parasite line. Merozoite invasion assay was performed as described earlier [113]. The same numbers of purified merozoites from the WT, GCN5-ΔBrD::GFP and PHD1-ΔPHD::GFP lines were mixed with fresh RBCs, and the parasitemia of culture was determined 24 h later. The invasion rate was calculated as the percentage of merozoites invaded into RBCs. To measure the gametocyte development, gametocyte induction was conducted by using an established method [110, 114] and the gametocytemia was determined by counting gametocytes in Giemsa-stained thin blood smears at the middle developmental stage (stage III).

**Histone modifications**

To estimate histone modifications in domain deletion mutants, histones were purified from the WT, GCN5-ΔBrD::GFP and PHD1-ΔPHD::GFP lines (Miao *et al.*, 2006). Equal amounts of the histones at each developmental stage were separated by 15% SDS/PAGE and transferred to nitrocellulose membranes. Western blotting was performed using a standard procedure with anti-acetyl histone H3, H3K9Ac (Catalog no. 07-352, RRID:AB_310544, Millipore), anti-tri methyl histone H3, H3K4me3 (catalog no. 07-473, RRID:AB_1977252, Millipore) and anti-acetyl histone H4, H4Ac (catalog no.06-598, RRID:AB_2295074, Millipore) at 1:1000 dilution as the primary antibodies and horseradish peroxidase-conjugated goat anti-rabbit IgG (diluted at

1:2000) as the secondary antibodies. The detected proteins were visualized using an enhanced chemiluminescence (ECL) kit (Invitrogen).

**Immunofluorescence assay (IFA)**

IFA was performed as described [115, 116]. The parasitized RBCs were washed once with PBS and the cell pellet (~100 µl) was fixed with 1 ml of 4% (v/v) paraformaldehyde and 0.0075% (v/v) glutaraldehyde in PBS for 30 min followed by 10 min quenching with 50 mM glycine in PBS. Fixed cells were washed twice with PBS and treated with 0.5% (v/v) Triton X-100 in PBS for 10 min. Then, cells were washed twice with PBS and blocked in 3% (v/v) BSA for 1 h at room temperature. The anti-PfPHD1 (1 µg/ml), PfPHD2 antibodies (1 µg/ml), goat anti-GFP (1:2000; ab6673; Abcam, RRID:AB_305643, USA) and rabbit anti-H3K9ac (1:1000; 06-942, RRID:AB_310308, Millipore, USA) antibodies in PBS containing 1% BSA were added and incubated for another 1.5 h. After washing the cells three times with PBS, FITC-conjugated goat anti-rabbit IgG antibodies (Cat# F6005, RRID:AB_259682, Sigma, USA), Alexa fluor® 488-conjugated secondary donkey anti-goat IgG antibody or Alexa fluor® 594-conjugated secondary goat anti-rabbit antibody IgG antibody (A32814 RRID:AB_2762838 and R37117 RRID:AB_2556545, Thermo Fisher Scientific, USA) were added at 1:2000 dilution in 3% (v/v) BSA and incubated for 45 min. Nuclear staining was performed by incubating slides with 4',6-diamidino-2-phenylindole (DAPI, final 0.5 µg/mL; Invitrogen). Images were captured using an epifluorescence microscope (Nikon Eclipse Ni, USA; 100x/1.4 oil immersion lens) and were processed by Adobe Photoshop CS (Adobe Systems Inc. San José, CA). To quantitate co-localizations, images from at least 20 parasites were randomly selected, analyzed by ImageJ (1.52a; http://imagej.nih.gov/ij), and Pearson's coefficients were calculated.

**Transcriptome analysis**

To compare the transcriptomes during the IDC among the WT, GCN5-ΔBrD::GFP and PHD1-ΔPHD::GFP lines, RNA-seq was performed. Three replicates of total RNA from parasites at ring, early trophozoite, late trophozoite and schizont stages were harvested by using the ZYMO RNA purification kit, and used to generate the sequencing libraries using the KAPA Stranded mRNA Seq kit for the Illumina sequencing platform according to the manufacturer's protocol (KAPA biosystems). Libraries were sequenced on an Illumina HiSeq 2500 in the Rapid Run mode using 100 nt single read sequencing. Reads from Illumina sequencing were mapped to the

*P. falciparum* genome sequence (Genedb v3.1) using HISAT2 [117]. The coverage was analyzed by using the bedtools [118]. The expression levels and the differential expression were calculated by FeatureCounts and DESeq2 [119, 120] with the criteria of ≥ 2 fold of alteration and P-adjustment <0.01. The GO enrichment was performed on PlasmDB (https://plasmodb.org/plasmo/). RNA-Seq data were submitted to NCBI GEO repository (accession number GSE164070).

**Phaseogram of the transcriptomes of *P. falciparum* IDC**

The sine wave model was utilized here to model the gene expression timing [121]. The gene transcription level from RNA-seq was first normalized as TPM (transcripts per million). Only the differential expressed genes in PfGCN5-ΔBrd or PfPHD1-ΔPHD as compared with WT were considered for the analysis. The TPM of each gene *E(t)* was modeled as

$$E(t) = A \times \sin(\omega t - \alpha) + C \quad (1)$$

Where $E(t) = [TPM_{12h}, TPM_{24h}, TPM_{36h}, TPM_{48h}]$ is the TPM at the $t$ = [12, 24, 36, 48] hours of sample collection, $\omega$ is the angular frequency and given by $\omega = 2\pi/48$, $A$ is the amplitude of the expression profile, and $C$ is the vertical offset of the profile from zero. To identify the parameter $\alpha$ and $A$, $A \times \sin(\omega t - \alpha)$ are changed to

$$A \times \sin(\omega t - \alpha) = \sin(\omega t) \times A \times \cos(-\alpha) + \cos(\omega t) \times A \times \sin(-\alpha) \quad (2)$$

Then the R command *lm* was used to fit a linear regression model between $E(t)$ and $\sin(\omega t) + \cos(\omega t)$. The fitting coefficient from the *lm* result indicates $A \times \cos(-\alpha)$ and $A \times \sin(-\alpha)$. The $\alpha$ and $A$ were calculated as

$$\alpha = \text{arc-tangent}\left(\frac{A \times \sin(-\alpha)}{A \times \cos(-\alpha)}\right) \quad (3)$$

$$A = (A \times \cos(-\alpha))^2 + (A \times \sin(-\alpha))^2 \quad (4)$$

The $\alpha$ indicates the horizontal offset of the profile from zero, which is used in the phaseogram to order the gene in the heatmap.

**Association between chromatin structures and transcriptomic changes upon domain deletions**

To quantify the association between open chromatin accessibility and transcriptional changes upon BrD deletion in PfGCN5 or PHD domain deletion in PfPHD1, we retrieved the ATAC-seq profile showing ATAC-seq peaks upstream the TSSs [122]. Each TSS was assigned to the nearest ATAC-seq peak with a distance restriction lower than 1 kb. The values of chromatin accessibility (ATAC-seq RPM + 0.1)/(gDNA RPM + 0.1) were then compared to the altered expression from the up- or down-regulated genes after domain deletions, where RPM represents the scaled reads per million reads. To investigate the association between the PfHP1 occupancy and transcriptional changes in the domain deletion mutants, the PfHP1 values (ChIP/input ratio) along the coding sequence were downloaded [123] and compared with the altered expression from the up and down-regulated gene after domain deletions.

**ChIP quantitative PCR**

ChIP-qPCR was performed as described [52, 86, 124] with some modifications. Synchronized GCN5-ΔBrD::GFP and PHD1-ΔPHD::GFP parasite lines at the ring stage [10–16 h post-invasion (hpi), ~$5 \times 10^9$ infected RBCs (iRBCs)] and schizont stage (40–46 hpi, ~$1.5 \times 10^9$ iRBCs) were harvested and crosslinked with paraformaldehyde (1% final concentration; EMS, USA) at 37°C for 15 min with agitation and then immediately neutralized by adding glycine (0.125 M final concentration) on ice for 5 min with agitation. The fixed iRBCs were lysed with saponin (0.06% final concentration; Sigma, USA) on ice for 5–10 min. Parasites were treated with a lysis buffer (10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 10 mM Hepes pH 7.9, 1 x Protease inhibitor) and then gently homogenized using a douncer to free nuclei. Pelleted nuclei were suspended in a shearing buffer (0.1% SDS, 5 mM EDTA, 50 mM Tris-HCl, pH 8.1, 1X Protease inhibitor) [124]. Sonication was performed using a rod bioruptor (Microson ultrasonic cell disruptor, Misonix, Inc. USA) at high power for 20 cycles of 30 sec ON/30 sec OFF, resulting in sheared chromatin of approximately 100–1000 bps. 50µl of input samples was set aside before the remaining chromatin was diluted in incubation buffer (0.01% SDS, 1.5% Triton X-100, 0.5 mM EDTA, 200 mM NaCl, 5 mM Tris-HCl, pH 8.1). Chromatin (75 µl/400 ng) was incubated with GFP-Trap® (Cat# gta-20, RRID:AB_2631357, ChromoTek, Germany) overnight at 4°C while rotating. Beads were then washed for 5 min at 4°C while rotating with the following: buffer 1 (0.1% SDS, 1% Triton X-100, 150 mM NaCl, 2 mM EDTA, 20 mM Tris HCl, pH 8.1); buffer 2 (0.1% SDS, 1% Triton X-100, 500 mM NaCl, 2 mM EDTA, 20 mM Tris

HCl, pH 8.1), buffer 3 (250 mM LiCl, 1% NP-40, 1% Na-deoxycholate, 1 mM EDTA, 10 mM Tris HCl, pH 8.1) and finally twice with buffer 4 (10 mM EDTA, 10 mM Tris HCl, pH 8). The immunoprecipitated chromatin was eluted with the elution buffer (1% SDS, 0.1M $NaHCO_3$) at room temperature for 15 min with rotation. The eluted chromatin and input samples were reverse cross-linked in 10% SDS, 1 M $NaHCO_3$, 5 M NaCl, 10% Triton X-100 at 45°C overnight while shaking and purified by the phenol:chloroform method. For qPCR, the concentration of immunoprecipitated gDNA was determined by Qubit dsDNA Broad-Range Assay Kit (Invitrogen, USA), and 10 ng per well in triplicate were used for qPCR using the FastStart™ Universal SYBR® Green Master [Rox] (Sigma-Aldrich, USA) as described [52]. Primer pairs targeting 5′UTRs were designed to amplify fragments less than 200 bp (**Table S9B**). Fold enrichment relative to constitutively expressed reference gene *seryl-tRNA synthetase* (PF3D7_0717700) was calculated by using the $2^{-\Delta\Delta Ct}$ method [125].

**RNA fluorescent in-situ hybridization (FISH)**

RNA FISH was performed as described [31]. Briefly, purified ring-stage parasites were lysed with saponin and released parasites fixed in suspension with ice-cold 4% paraformaldehyde. Parasites were then deposited on Teflon coated microscope slides and hybridized with denatured *var* probes at 42℃ for at least 16 h. All the FISH probes were PCR amplified from genomic DNA using the primers listed in Table S9B. The slides were then washed three times in 2×SSC at 42℃. Finally, the slides were incubated with streptavidin-488 antibody at room temperature for 30 min. Images were taken using a Nikon ECLIPSE E600 epifluorescence microscope. NIS Elements 3.0 software was used for acquisition and ImageJ for composition.

**Statistical analysis**

For all experiments, three or more independent biological replicates were performed. The results are presented as mean ± SD. Results are regarded significant if $P < 0.05$ as established by ANOVA, Fisher's exact test, paired Mann Whitney *U* test or paired Wilcoxon test, and the respective analysis was shown in the figure legends. To analyze the schizont numbers containing different numbers of merozoites, a $\chi^2$ goodness of fit test was first used to evaluate if the number of schizonts that contain a certain number of merozoites was independent of the parasite lines. Then the proportions of schizonts with a certain number of merozoites were compared among these cell lines based on ANOVA for each merozoite number.

## Acknowledgments

We thank Dale Chaput for proteomic studies; Xiaolian Li for parasite culture and generation of transgenic lines. We are grateful to Jacobus Pharmaceuticals for providing the drug WR99210. This study was supported by grants U19AI 089672 and R01AI128940 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

## Author Contributions

JM and LC designed the project. JM, LC and KK supervised the project. JM executed most experiments and interpreted data. XL performed RNA-seq experiment. SA generated the RNA-seq data and CW performed analysis of the sequence data. RJ supervised bioinformatic analysis. AL conducted IFA of H3K9Ac and chromatin immunoprecipitation, and HM performed Western blot for GCN5 expression and FISH. JM and LC wrote the manuscript. All authors read and agreed to the content of this manuscript.

## Declaration of Interests

The authors declare no competing interests.

## References

1. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21(3):381-95. Epub 2011/02/16. doi: 10.1038/cr.2011.22. PubMed PMID: 21321607; PubMed Central PMCID: PMCPMC3193420.
2. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000;403(6765):41-5. Epub 2000/01/19. doi: 10.1038/47412. PubMed PMID: 10638745.
3. Jenuwein T, Allis CD. Translating the histone code. Science. 2001;293(5532):1074-80. Epub 2001/08/11. doi: 10.1126/science.1063127. PubMed PMID: 11498575.
4. Winter S, Fischle W. Epigenetic markers and their cross-talk. Essays Biochem. 2010;48(1):45-61. Epub 2010/09/09. doi: 10.1042/bse0480045. PubMed PMID: 20822485.
5. Lee JS, Smith E, Shilatifard A. The language of histone crosstalk. Cell. 2010;142(5):682-5. Epub 2010/09/04. doi: 10.1016/j.cell.2010.08.011. PubMed PMID: 20813257; PubMed Central PMCID: PMCPMC3711869.
6. Murr R. Interplay between different epigenetic modifications and mechanisms. Adv Genet. 2010;70:101-41. Epub 2010/10/06. doi: 10.1016/B978-0-12-380866-0.60005-8. PubMed PMID: 20920747.
7. Marmorstein R, Zhou MM. Writers and readers of histone acetylation: structure, mechanism, and inhibition. Cold Spring Harbor perspectives in biology. 2014;6(7):a018762. Epub 2014/07/06. doi: 10.1101/cshperspect.a018762. PubMed PMID: 24984779; PubMed Central PMCID: PMCPMC4067988.

8. Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, et al. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes & development. 1997;11(13):1640-50. Epub 1997/07/01. doi: 10.1101/gad.11.13.1640. PubMed PMID: 9224714.
9. Helmlinger D, Tora L. Sharing the SAGA. Trends Biochem Sci. 2017;42(11):850-61. Epub 2017/10/02. doi: 10.1016/j.tibs.2017.09.001. PubMed PMID: 28964624; PubMed Central PMCID: PMCPMC5660625.
10. Spedale G, Timmers HT, Pijnappel WW. ATAC-king the complexity of SAGA during evolution. Genes Dev. 2012;26(6):527-41. Epub 2012/03/20. doi: 10.1101/gad.184705.111. PubMed PMID: 22426530; PubMed Central PMCID: PMCPMC3315114.
11. Huisinga KL, Pugh BF. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Mol Cell. 2004;13(4):573-85. Epub 2004/03/03. PubMed PMID: 14992726.
12. Vlachonasios KE, Thomashow MF, Triezenberg SJ. Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. Plant Cell. 2003;15(3):626-38. Epub 2003/03/05. doi: 10.1105/tpc.007922. PubMed PMID: 12615937; PubMed Central PMCID: PMCPMC150018.
13. Bonnet J, Wang CY, Baptista T, Vincent SD, Hsiao WC, Stierle M, et al. The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. Genes Dev. 2014;28(18):1999-2012. Epub 2014/09/18. doi: 10.1101/gad.250225.114. PubMed PMID: 25228644; PubMed Central PMCID: PMCPMC4173158.
14. Baptista T, Grunberg S, Minoungou N, Koster MJE, Timmers HTM, Hahn S, et al. SAGA Is a General Cofactor for RNA Polymerase II Transcription. Mol Cell. 2017;68(1):130-43 e5. Epub 2017/09/19. doi: 10.1016/j.molcel.2017.08.016. PubMed PMID: 28918903; PubMed Central PMCID: PMCPMC5632562.
15. Koutelou E, Hirsch CL, Dent SY. Multiple faces of the SAGA complex. Curr Opin Cell Biol. 2010;22(3):374-82. Epub 2010/04/07. doi: 10.1016/j.ceb.2010.03.005. PubMed PMID: 20363118; PubMed Central PMCID: PMCPMC2900470.
16. WHO. Would Malaria Report 2020. Geneva: WHO. 2020.
17. Hall N, Karras M, Raine JD, Carlton JM, Kooij TW, Berriman M, et al. A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. Science. 2005;307(5706):82-6. Epub 2005/01/08. doi: 10.1126/science.1103717. PubMed PMID: 15637271.
18. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol. 2003;1(1):E5. Epub 2003/08/21. doi: 10.1371/journal.pbio.0000005. PubMed PMID: 12929205; PubMed Central PMCID: PMCPMC176545.
19. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. Science. 2003;301(5639):1503-8. Epub 2003/08/02. doi: 10.1126/science.1087025. PubMed PMID: 12893887.
20. Cortes A, Deitsch KW. Malaria Epigenetics. Cold Spring Harb Perspect Med. 2017;7(7). Epub 2017/03/23. doi: 10.1101/cshperspect.a025528. PubMed PMID: 28320828; PubMed Central PMCID: PMCPMC5495052.

21. Gupta AP, Bozdech Z. Epigenetic landscapes underlining global patterns of gene expression in the human malaria parasite, Plasmodium falciparum. Int J Parasitol. 2017;47(7):399-407. Epub 2017/04/18. doi: 10.1016/j.ijpara.2016.10.008. PubMed PMID: 28414071.
22. Batugedara G, Lu XM, Bunnik EM, Le Roch KG. The Role of Chromatin Structure in Gene Regulation of the Human Malaria Parasite. Trends Parasitol. 2017;33(5):364-77. Epub 2017/01/10. doi: 10.1016/j.pt.2016.12.004. PubMed PMID: 28065669; PubMed Central PMCID: PMCPMC5410391.
23. Voss TS, Bozdech Z, Bartfai R. Epigenetic memory takes center stage in the survival strategy of malaria parasites. Curr Opin Microbiol. 2014;20:88-95. Epub 2014/06/20. doi: 10.1016/j.mib.2014.05.007. PubMed PMID: 24945736.
24. Vembar SS, Scherf A, Siegel TN. Noncoding RNAs as emerging regulators of Plasmodium falciparum virulence gene expression. Curr Opin Microbiol. 2014;20:153-61. Epub 2014/07/16. doi: 10.1016/j.mib.2014.06.013. PubMed PMID: 25022240; PubMed Central PMCID: PMCPMC4157322.
25. Callebaut I, Prat K, Meurice E, Mornon JP, Tomavo S. Prediction of the general transcription factors associated with RNA polymerase II in Plasmodium falciparum: conserved features and differences relative to other eukaryotes. BMC Genomics. 2005;6:100. Epub 2005/07/27. doi: 10.1186/1471-2164-6-100. PubMed PMID: 16042788; PubMed Central PMCID: PMCPMC1199594.
26. Coulson RM, Hall N, Ouzounis CA. Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. Genome Res. 2004;14(8):1548-54. Epub 2004/07/17. doi: 10.1101/gr.2218604. PubMed PMID: 15256513; PubMed Central PMCID: PMCPMC509263.
27. de Boer CG, Hughes TR. YeTFaSCo: a database of evaluated yeast transcription factor sequence specificities. Nucleic acids research. 2012;40(Database issue):D169-79. Epub 2011/11/22. doi: 10.1093/nar/gkr993. PubMed PMID: 22102575; PubMed Central PMCID: PMCPMC3245003.
28. Balaji S, Babu MM, Iyer LM, Aravind L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. Nucleic Acids Res. 2005;33(13):3994-4006. Epub 2005/07/26. doi: 10.1093/nar/gki709. PubMed PMID: 16040597; PubMed Central PMCID: PMCPMC1178005.
29. Campbell TL, De Silva EK, Olszewski KL, Elemento O, Llinas M. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS Pathog. 2010;6(10):e1001165. Epub 2010/11/10. doi: 10.1371/journal.ppat.1001165. PubMed PMID: 21060817; PubMed Central PMCID: PMCPMC2965767.
30. Cui L, Miao J. Chromatin-mediated epigenetic regulation in the malaria parasite Plasmodium falciparum. Eukaryot Cell. 2010;9(8):1138-49. Epub 2010/05/11. doi: 10.1128/EC.00036-10. PubMed PMID: 20453074; PubMed Central PMCID: PMCPMC2918932.
31. Lopez-Rubio JJ, Mancio-Silva L, Scherf A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. Cell Host Microbe. 2009;5(2):179-90. Epub 2009/02/17. doi: 10.1016/j.chom.2008.12.012. PubMed PMID: 19218088.

32. Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AM, Cohen AM, et al. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proc Natl Acad Sci U S A. 2009;106(24):9655-60. Epub 2009/06/06. doi: 10.1073/pnas.0902515106. PubMed PMID: 19497874; PubMed Central PMCID: PMCPMC2701018.
33. Ay F, Bunnik EM, Varoquaux N, Bol SM, Prudhomme J, Vert JP, et al. Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. Genome Res. 2014;24(6):974-88. Epub 2014/03/29. doi: 10.1101/gr.169417.113. PubMed PMID: 24671853; PubMed Central PMCID: PMCPMC4032861.
34. Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, Guinet F, et al. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature. 2000;407(6807):1018-22. Epub 2000/11/09. doi: 10.1038/35039531. PubMed PMID: 11069183.
35. Bunnik EM, Venkat A, Shao J, McGovern KE, Batugedara G, Worth D, et al. Comparative 3D genome organization in apicomplexan parasites. Proc Natl Acad Sci U S A. 2019;116(8):3183-92. Epub 2019/02/07. doi: 10.1073/pnas.1810815116. PubMed PMID: 30723152; PubMed Central PMCID: PMCPMC6386730.
36. Bunnik EM, Cook KB, Varoquaux N, Batugedara G, Prudhomme J, Cort A, et al. Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. Nat Commun. 2018;9(1):1910. Epub 2018/05/17. doi: 10.1038/s41467-018-04295-5. PubMed PMID: 29765020; PubMed Central PMCID: PMCPMC5954139.
37. Crowley VM, Rovira-Graells N, Ribas de Pouplana L, Cortes A. Heterochromatin formation in bistable chromatin domains controls the epigenetic repression of clonally variant Plasmodium falciparum genes linked to erythrocyte invasion. Mol Microbiol. 2011;80(2):391-406. Epub 2011/02/11. doi: 10.1111/j.1365-2958.2011.07574.x. PubMed PMID: 21306446.
38. Freitas-Junior LH, Hernandez-Rivas R, Ralph SA, Montiel-Condado D, Ruvalcaba-Salazar OK, Rojas-Meza AP, et al. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. Cell. 2005;121(1):25-36. Epub 2005/04/12. doi: 10.1016/j.cell.2005.01.037. PubMed PMID: 15820676.
39. Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, Hartl DL, et al. Epigenetic memory at malaria virulence genes. Proc Natl Acad Sci U S A. 2007;104(3):899-902. Epub 2007/01/09. doi: 10.1073/pnas.0609084103. PubMed PMID: 17209011; PubMed Central PMCID: PMCPMC1764221.
40. Perez-Toledo K, Rojas-Meza AP, Mancio-Silva L, Hernandez-Cuevas NA, Delgadillo DM, Vargas M, et al. Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. Nucleic Acids Res. 2009;37(8):2596-606. Epub 2009/03/10. doi: 10.1093/nar/gkp115. PubMed PMID: 19270070; PubMed Central PMCID: PMCPMC2677873.
41. Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, Alako BT, et al. Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. PLoS Pathog. 2009;5(9):e1000569. Epub 2009/09/05. doi: 10.1371/journal.ppat.1000569. PubMed PMID: 19730695; PubMed Central PMCID: PMCPMC2731224.

42. Brancucci NMB, Bertschi NL, Zhu L, Niederwieser I, Chin WH, Wampfler R, et al. Heterochromatin protein 1 secures survival and transmission of malaria parasites. Cell Host Microbe. 2014;16(2):165-76. Epub 2014/08/15. doi: 10.1016/j.chom.2014.07.004. PubMed PMID: 25121746.
43. Ponts N, Harris EY, Lonardi S, Le Roch KG. Nucleosome occupancy at transcription start sites in the human malaria parasite: a hard-wired evolution of virulence? Infect Genet Evol. 2011;11(4):716-24. Epub 2010/08/17. doi: 10.1016/j.meegid.2010.08.002. PubMed PMID: 20708104; PubMed Central PMCID: PMCPMC3466077.
44. Bunnik EM, Polishko A, Prudhomme J, Ponts N, Gill SS, Lonardi S, et al. DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite Plasmodium falciparum. BMC Genomics. 2014;15:347. Epub 2014/06/03. doi: 10.1186/1471-2164-15-347. PubMed PMID: 24885191; PubMed Central PMCID: PMCPMC4035074.
45. Westenberger SJ, Cui L, Dharia N, Winzeler E, Cui L. Genome-wide nucleosome mapping of Plasmodium falciparum reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes. BMC Genomics. 2009;10:610. Epub 2009/12/18. doi: 10.1186/1471-2164-10-610. PubMed PMID: 20015349; PubMed Central PMCID: PMCPMC2801526.
46. Ponts N, Harris EY, Prudhomme J, Wick I, Eckhardt-Ludka C, Hicks GR, et al. Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res. 2010;20(2):228-38. Epub 2010/01/08. doi: 10.1101/gr.101063.109. PubMed PMID: 20054063; PubMed Central PMCID: PMCPMC2813478.
47. Kensche PR, Hoeijmakers WA, Toenhake CG, Bras M, Chappell L, Berriman M, et al. The nucleosome landscape of Plasmodium falciparum reveals chromatin architecture and dynamics of regulatory sequences. Nucleic Acids Res. 2016;44(5):2110-24. Epub 2015/11/19. doi: 10.1093/nar/gkv1214. PubMed PMID: 26578577; PubMed Central PMCID: PMCPMC4797266.
48. Toenhake CG, Fraschka SA, Vijayabaskar MS, Westhead DR, van Heeringen SJ, Bartfai R. Chromatin Accessibility-Based Characterization of the Gene Regulatory Network Underlying Plasmodium falciparum Blood-Stage Development. Cell Host Microbe. 2018;23(4):557-69 e9. Epub 2018/04/13. doi: 10.1016/j.chom.2018.03.007. PubMed PMID: 29649445; PubMed Central PMCID: PMCPMC5899830.
49. Gupta AP, Chin WH, Zhu L, Mok S, Luah YH, Lim EH, et al. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite Plasmodium falciparum. PLoS Pathog. 2013;9(2):e1003170. Epub 2013/03/08. doi: 10.1371/journal.ppat.1003170. PubMed PMID: 23468622; PubMed Central PMCID: PMCPMC3585154.
50. Bartfai R, Hoeijmakers WA, Salcedo-Amaya AM, Smits AH, Janssen-Megens E, Kaan A, et al. H2A.Z demarcates intergenic regions of the plasmodium falciparum epigenome that are dynamically marked by H3K9ac and H3K4me3. PLoS Pathog. 2010;6(12):e1001223. Epub 2010/12/29. doi: 10.1371/journal.ppat.1001223. PubMed PMID: 21187892; PubMed Central PMCID: PMCPMC3002978.
51. Fan Q, An L, Cui L. Plasmodium falciparum histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. Eukaryot Cell. 2004;3(2):264-76. Epub 2004/04/13. doi: 10.1128/ec.3.2.264-276.2004. PubMed PMID: 15075257; PubMed Central PMCID: PMCPMC387650.
52. Cui L, Miao J, Furuya T, Li X, Su XZ, Cui L. PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in Plasmodium falciparum. Eukaryot Cell. 2007;6(7):1219-

27. Epub 2007/04/24. doi: 10.1128/EC.00062-07. PubMed PMID: 17449656; PubMed Central PMCID: PMCPMC1951105.
53. Miao J, Fan Q, Cui L, Li X, Wang H, Ning G, et al. The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite Plasmodium falciparum. Mol Microbiol. 2010;78(4):883-902. Epub 2010/09/03. doi: 10.1111/j.1365-2958.2010.07371.x. PubMed PMID: 20807207; PubMed Central PMCID: PMCPMC2978264.
54. Cui L, Fan Q, Miao J. Histone lysine methyltransferases and demethylases in Plasmodium falciparum. International journal for parasitology. 2008;38(10):1083-97. Epub 2008/02/27. doi: 10.1016/j.ijpara.2008.01.002. PubMed PMID: 18299133; PubMed Central PMCID: PMC4566933.
55. Xiao B, Yin S, Hu Y, Sun M, Wei J, Huang Z, et al. Epigenetic editing by CRISPR/dCas9 in Plasmodium falciparum. Proc Natl Acad Sci U S A. 2019;116(1):255-60. Epub 2018/12/26. doi: 10.1073/pnas.1813542116. PubMed PMID: 30584102; PubMed Central PMCID: PMCPMC6320497.
56. Bhowmick K, Tehlan A, Sunita, Sudhakar R, Kaur I, Sijwali PS, et al. Plasmodium falciparum GCN5 acetyltransferase follows a novel proteolytic processing pathway that is essential for its function. J Cell Sci. 2020;133(1). Epub 2019/12/22. doi: 10.1242/jcs.236489. PubMed PMID: 31862795.
57. Cui L, Miao J, Furuya T, Fan Q, Li X, Rathod PK, et al. Histone acetyltransferase inhibitor anacardic acid causes changes in global gene expression during in vitro Plasmodium falciparum development. Eukaryot Cell. 2008;7(7):1200-10. Epub 2008/05/20. doi: 10.1128/EC.00063-08. PubMed PMID: 18487348; PubMed Central PMCID: PMCPMC2446667.
58. Hoeijmakers WAM, Miao J, Schmidt S, Toenhake CG, Shrestha S, Venhuizen J, et al. Epigenetic reader complexes of the human malaria parasite, Plasmodium falciparum. Nucleic Acids Res. 2019;47(22):11574-88. Epub 2019/11/16. doi: 10.1093/nar/gkz1044. PubMed PMID: 31728527.
59. Fan Q, An L, Cui L. PfADA2, a Plasmodium falciparum homologue of the transcriptional coactivator ADA2 and its in vivo association with the histone acetyltransferase PfGCN5. Gene. 2004;336(2):251-61. Epub 2004/07/13. doi: 10.1016/j.gene.2004.04.005. PubMed PMID: 15246536.
60. Bischoff E, Vaquero C. In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of Plasmodium falciparum. BMC genomics. 2010;11:34. Epub 2010/01/19. doi: 10.1186/1471-2164-11-34. PubMed PMID: 20078850; PubMed Central PMCID: PMC2821373.
61. Schimanski B, Nguyen TN, Gunzl A. Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. Eukaryot Cell. 2005;4(11):1942-50. Epub 2005/11/10. doi: 10.1128/EC.4.11.1942-1950.2005. PubMed PMID: 16278461; PubMed Central PMCID: PMCPMC1287860.
62. Takebe S, Witola WH, Schimanski B, Gunzl A, Ben Mamoun C. Purification of components of the translation elongation factor complex of Plasmodium falciparum by tandem affinity purification. Eukaryotic cell. 2007;6(4):584-91. Epub 2007/02/20. doi: 10.1128/EC.00376-06. PubMed PMID: 17307963; PubMed Central PMCID: PMC1865644.
63. Miao J, Li X, Cui L. Cloning of Plasmodium falciparum by single-cell sorting. Experimental parasitology. 2010;126(2):198-202. Epub 2010/05/04. doi:

10.1016/j.exppara.2010.04.022. PubMed PMID: 20435038; PubMed Central PMCID: PMC2914851.
64. Teo G, Liu G, Zhang J, Nesvizhskii AI, Gingras AC, Choi H. SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. J Proteomics. 2014;100:37-43. Epub 2014/02/12. doi: 10.1016/j.jprot.2013.10.023. PubMed PMID: 24513533; PubMed Central PMCID: PMCPMC4102138.
65. Singh M, D'Silva L, Holak TA. DNA-binding properties of the recombinant high-mobility-group-like AT-hook-containing region from human BRG1 protein. Biol Chem. 2006;387(10-11):1469-78. Epub 2006/11/04. doi: 10.1515/BC.2006.184. PubMed PMID: 17081121.
66. Darvekar S, Johnsen SS, Eriksen AB, Johansen T, Sjottem E. Identification of two independent nucleosome-binding domains in the transcriptional co-activator SPBP. Biochem J. 2012;442(1):65-75. Epub 2011/11/16. doi: 10.1042/BJ20111230. PubMed PMID: 22081970.
67. Sjottem E, Rekdal C, Svineng G, Johnsen SS, Klenow H, Uglehus RD, et al. The ePHD protein SPBP interacts with TopBP1 and together they co-operate to stimulate Ets1-mediated transcription. Nucleic Acids Res. 2007;35(19):6648-62. Epub 2007/10/05. doi: 10.1093/nar/gkm739. PubMed PMID: 17913746; PubMed Central PMCID: PMCPMC2095823.
68. Avgousti DC, Cecere G, Grishok A. The conserved PHD1-PHD2 domain of ZFP-1/AF10 is a discrete functional module essential for viability in Caenorhabditis elegans. Mol Cell Biol. 2013;33(5):999-1015. Epub 2012/12/25. doi: 10.1128/MCB.01462-12. PubMed PMID: 23263989; PubMed Central PMCID: PMCPMC3623080.
69. Liu Z, Li F, Ruan K, Zhang J, Mei Y, Wu J, et al. Structural and functional insights into the human Borjeson-Forssman-Lehmann syndrome-associated protein PHF6. J Biol Chem. 2014;289(14):10069-83. Epub 2014/02/21. doi: 10.1074/jbc.M113.535351. PubMed PMID: 24554700; PubMed Central PMCID: PMCPMC3974978.
70. Sanchez R, Zhou MM. The PHD finger: a versatile epigenome reader. Trends Biochem Sci. 2011;36(7):364-72. Epub 2011/04/26. doi: 10.1016/j.tibs.2011.03.005. PubMed PMID: 21514168; PubMed Central PMCID: PMCPMC3130114.
71. Miao J, Lawrence M, Jeffers V, Zhao F, Parker D, Ge Y, et al. Extensive lysine acetylation occurs in evolutionarily conserved metabolic pathways and parasite-specific functions during Plasmodium falciparum intraerythrocytic development. Mol Microbiol. 2013;89(4):660-75. Epub 2013/06/26. doi: 10.1111/mmi.12303. PubMed PMID: 23796209; PubMed Central PMCID: PMCPMC3757501.
72. LaCount DJ, Vignali M, Chettier R, Phansalkar A, Bell R, Hesselberth JR, et al. A protein interaction network of the malaria parasite Plasmodium falciparum. Nature. 2005;438(7064):103-7. Epub 2005/11/04. doi: 10.1038/nature04104. PubMed PMID: 16267556.
73. Zhang M, Wang C, Otto TD, Oberstaller J, Liao X, Adapa SR, et al. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. Science. 2018;360(6388). Epub 2018/05/05. doi: 10.1126/science.aap7847. PubMed PMID: 29724925; PubMed Central PMCID: PMCPMC6360947.
74. Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, et al. Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. EMBO J. 2011;30(14):2829-42. Epub 2011/06/21. doi: 10.1038/emboj.2011.193. PubMed PMID: 21685874; PubMed Central PMCID: PMCPMC3160252.
75. Volz J, Carvalho TG, Ralph SA, Gilson P, Thompson J, Tonkin CJ, et al. Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in Plasmodium

falciparum. International journal for parasitology. 2010;40(1):109-21. Epub 2009/09/22. doi: 10.1016/j.ijpara.2009.09.002. PubMed PMID: 19765590.
76. Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, Carrozza MJ, et al. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. Cell. 2002;111(3):369-79. Epub 2002/11/07. doi: 10.1016/s0092-8674(02)01005-x. PubMed PMID: 12419247.
77. Vosnakis N, Koch M, Scheer E, Kessler P, Mely Y, Didier P, et al. Coactivators and general transcription factors have two distinct dynamic populations dependent on transcription. EMBO J. 2017;36(18):2710-25. Epub 2017/07/21. doi: 10.15252/embj.201696035. PubMed PMID: 28724529; PubMed Central PMCID: PMCPMC5599802.
78. Ralph SA, Scheidig-Benatar C, Scherf A. Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(15):5414-9. Epub 2005/03/31. doi: 10.1073/pnas.0408883102. PubMed PMID: 15797990; PubMed Central PMCID: PMC556247.
79. Dzikowski R, Li F, Amulic B, Eisberg A, Frank M, Patel S, et al. Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. EMBO reports. 2007;8(10):959-65. Epub 2007/09/01. doi: 10.1038/sj.embor.7401063. PubMed PMID: 17762879; PubMed Central PMCID: PMC2002552.
80. Cowman AF, Berry D, Baum J. The cellular and molecular basis for malaria parasite invasion of the human red blood cell. J Cell Biol. 2012;198(6):961-71. Epub 2012/09/19. doi: 10.1083/jcb.201206112. PubMed PMID: 22986493; PubMed Central PMCID: PMCPMC3444787.
81. Painter HJ, Campbell TL, Llinas M. The Apicomplexan AP2 family: integral factors regulating Plasmodium development. Molecular and biochemical parasitology. 2011;176(1):1-7. Epub 2010/12/04. doi: 10.1016/j.molbiopara.2010.11.014. PubMed PMID: 21126543; PubMed Central PMCID: PMC3026892.
82. Jeninga MD, Quinn JE, Petter M. ApiAP2 Transcription Factors in Apicomplexan Parasites. Pathogens. 2019;8(2). Epub 2019/04/10. doi: 10.3390/pathogens8020047. PubMed PMID: 30959972.
83. Flueck C, Bartfai R, Niederwieser I, Witmer K, Alako BT, Moes S, et al. A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in chromosome end biology. PLoS pathogens. 2010;6(2):e1000784. Epub 2010/03/03. doi: 10.1371/journal.ppat.1000784. PubMed PMID: 20195509; PubMed Central PMCID: PMC2829057.
84. Josling GA, Russell TJ, Venezia J, Orchard L, van Biljon R, Painter HJ, et al. Dissecting the role of PfAP2-G in malaria gametocytogenesis. Nat Commun. 2020;11(1):1503. Epub 2020/03/22. doi: 10.1038/s41467-020-15026-0. PubMed PMID: 32198457; PubMed Central PMCID: PMCPMC7083873.
85. Kafsack BF, Rovira-Graells N, Clark TG, Bancells C, Crowley VM, Campino SG, et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. Nature. 2014;507(7491):248-52. Epub 2014/02/28. doi: 10.1038/nature12920. PubMed PMID: 24572369; PubMed Central PMCID: PMC4040541.
86. Santos JM, Josling G, Ross P, Joshi P, Orchard L, Campbell T, et al. Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor. Cell Host Microbe. 2017;21(6):731-41 e10. Epub 2017/06/16. doi: 10.1016/j.chom.2017.05.006. PubMed PMID: 28618269; PubMed Central PMCID: PMCPMC5855115.

87. Wang J, Dixon SE, Ting LM, Liu TK, Jeffers V, Croken MM, et al. Lysine acetyltransferase GCN5b interacts with AP2 factors and is required for Toxoplasma gondii proliferation. PLoS Pathog. 2014;10(1):e1003830. Epub 2014/01/07. doi: 10.1371/journal.ppat.1003830. PubMed PMID: 24391497; PubMed Central PMCID: PMCPMC3879359.
88. Jiang L, Mu J, Zhang Q, Ni T, Srinivasan P, Rayavara K, et al. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. Nature. 2013;499(7457):223-7. Epub 2013/07/05. doi: 10.1038/nature12361. PubMed PMID: 23823717; PubMed Central PMCID: PMCPMC3770130.
89. Harris MT, Jeffers V, Martynowicz J, True JD, Mosley AL, Sullivan WJ, Jr. A novel GCN5b lysine acetyltransferase complex associates with distinct transcription factors in the protozoan parasite Toxoplasma gondii. Mol Biochem Parasitol. 2019;232:111203. Epub 2019/08/06. doi: 10.1016/j.molbiopara.2019.111203. PubMed PMID: 31381949; PubMed Central PMCID: PMCPMC6717552.
90. Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, et al. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell. 2010;142(6):967-80. Epub 2010/09/21. doi: 10.1016/j.cell.2010.08.020. PubMed PMID: 20850016.
91. Li S, Shogren-Knaak MA. The Gcn5 bromodomain of the SAGA complex facilitates cooperative and cross-tail acetylation of nucleosomes. J Biol Chem. 2009;284(14):9411-7. Epub 2009/02/17. doi: 10.1074/jbc.M809617200. PubMed PMID: 19218239; PubMed Central PMCID: PMCPMC2666593.
92. Govind CK, Zhang F, Qiu H, Hofmeyer K, Hinnebusch AG. Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. Mol Cell. 2007;25(1):31-42. Epub 2007/01/16. doi: 10.1016/j.molcel.2006.11.020. PubMed PMID: 17218269.
93. Nightingale KP, Gendreizig S, White DA, Bradbury C, Hollfelder F, Turner BM. Cross-talk between histone modifications in response to histone deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation. J Biol Chem. 2007;282(7):4408-16. Epub 2006/12/15. doi: 10.1074/jbc.M606773200. PubMed PMID: 17166833.
94. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, et al. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell. 2002;10(5):1107-17. Epub 2002/11/28. PubMed PMID: 12453418.
95. Chaal BK, Gupta AP, Wastuwidyaningtyas BD, Luah YH, Bozdech Z. Histone deacetylases play a major role in the transcriptional regulation of the Plasmodium falciparum life cycle. PLoS Pathog. 2010;6(1):e1000737. Epub 2010/01/29. doi: 10.1371/journal.ppat.1000737. PubMed PMID: 20107518; PubMed Central PMCID: PMCPMC2809759.
96. Deitsch KW, Dzikowski R. Variant gene expression and antigenic variation by malaria parasites. Annual review of microbiology. 2017;71:625-41. Epub 2017/07/13. doi: 10.1146/annurev-micro-090816-093841. PubMed PMID: 28697665.
97. Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R, Scherf A. 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol Microbiol. 2007;66(6):1296-305. Epub 2007/11/22. doi: 10.1111/j.1365-2958.2007.06009.x. PubMed PMID: 18028313; PubMed Central PMCID: PMCPMC2228885.

98. Brancucci NM, Bertschi NL, Zhu L, Niederwieser I, Chin WH, Wampfler R, et al. Heterochromatin protein 1 secures survival and transmission of malaria parasites. Cell host & microbe. 2014;16(2):165-76. Epub 2014/08/15. doi: 10.1016/j.chom.2014.07.004. PubMed PMID: 25121746.
99. Coleman BI, Skillman KM, Jiang RH, Childs LM, Altenhofen LM, Ganter M, et al. A Plasmodium falciparum histone deacetylase regulates antigenic variation and gametocyte conversion. Cell host & microbe. 2014;16(2):177-86. Epub 2014/08/15. doi: 10.1016/j.chom.2014.06.014. PubMed PMID: 25121747; PubMed Central PMCID: PMC4188636.
100. Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. Cell. 2005;121(1):13-24. Epub 2005/04/12. doi: 10.1016/j.cell.2005.01.036. PubMed PMID: 15820675.
101. Tonkin CJ, Carret CK, Duraisingh MT, Voss TS, Ralph SA, Hommel M, et al. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. PLoS Biol. 2009;7(4):e84. Epub 2009/05/01. doi: 10.1371/journal.pbio.1000084. PubMed PMID: 19402747; PubMed Central PMCID: PMCPMC2672602.
102. Dastidar EG, Dzeyk K, Krijgsveld J, Malmquist NA, Doerig C, Scherf A, et al. Comprehensive histone phosphorylation analysis and identification of Pf14-3-3 protein as a histone H3 phosphorylation reader in malaria parasites. PloS one. 2013;8(1):e53179. Epub 2013/01/12. doi: 10.1371/journal.pone.0053179. PubMed PMID: 23308157; PubMed Central PMCID: PMC3538786.
103. Lo WS, Duggan L, Emre NC, Belotserkovskya R, Lane WS, Shiekhattar R, et al. Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. Science. 2001;293(5532):1142-6. Epub 2001/08/11. doi: 10.1126/science.1062322. PubMed PMID: 11498592.
104. Kumar A, Bhowmick K, Vikramdeo KS, Mondal N, Subbarao N, Dhar SK. Designing novel inhibitors against histone acetyltransferase (HAT: GCN5) of Plasmodium falciparum. Eur J Med Chem. 2017;138:26-37. Epub 2017/06/24. doi: 10.1016/j.ejmech.2017.06.009. PubMed PMID: 28644986.
105. Moustakim M, Clark PG, Trulli L, Fuentes de Arriba AL, Ehebauer MT, Chaikuad A, et al. Discovery of a PCAF Bromodomain Chemical Probe. Angew Chem Int Ed Engl. 2017;56(3):827-31. Epub 2016/12/15. doi: 10.1002/anie.201610816. PubMed PMID: 27966810; PubMed Central PMCID: PMC5412877.
106. Trager W, Jensen JB. Human malaria parasites in continuous culture. Science. 1976;193(4254):673-5. Epub 1976/08/20. PubMed PMID: 781840.
107. Miao J, Cui L. Rapid isolation of single malaria parasite-infected red blood cells by cell sorting. Nat Protoc. 2011;6(2):140-6. Epub 2011/02/05. doi: 10.1038/nprot.2010.185. PubMed PMID: 21293455; PubMed Central PMCID: PMCPMC3150729.
108. Fan Q, Miao J, Cui L. Characterization of PRMT1 from Plasmodium falciparum. The Biochemical journal. 2009;421(1):107-18. Epub 2009/04/07. doi: 10.1042/BJ20090185. PubMed PMID: 19344311.
109. Deitsch K, Driskill C, Wellems T. Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. Nucleic acids research. 2001;29(3):850-3. Epub 2001/02/13. PubMed PMID: 11160909; PubMed Central PMCID: PMC30384.

110. Miao J, Li J, Fan Q, Li X, Li X, Cui L. The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite Plasmodium falciparum. J Cell Sci. 2010;123(Pt 7):1039-49. Epub 2010/03/04. doi: 10.1242/jcs.059824. PubMed PMID: 20197405; PubMed Central PMCID: PMCPMC2844316.
111. Miao J, Chen Z, Wang Z, Shrestha S, Li X, Li R, et al. Sex-Specific Biology of the Human Malaria Parasite Revealed from the Proteomes of Mature Male and Female Gametocytes. Mol Cell Proteomics. 2017;16(4):537-51. Epub 2017/01/28. doi: 10.1074/mcp.M116.061804. PubMed PMID: 28126901; PubMed Central PMCID: PMCPMC5383777.
112. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019;47(D1):D442-D50. Epub 2018/11/06. doi: 10.1093/nar/gky1106. PubMed PMID: 30395289; PubMed Central PMCID: PMCPMC6323896.
113. Boyle MJ, Wilson DW, Richards JS, Riglar DT, Tetteh KK, Conway DJ, et al. Isolation of viable Plasmodium falciparum merozoites to define erythrocyte invasion events and advance vaccine and drug development. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(32):14378-83. Epub 2010/07/28. doi: 10.1073/pnas.1009198107. PubMed PMID: 20660744; PubMed Central PMCID: PMC2922570.
114. Shrestha S, Li X, Ning G, Miao J, Cui L. The RNA-binding protein Puf1 functions in the maintenance of gametocytes in Plasmodium falciparum. J Cell Sci. 2016;129(16):3144-52. Epub 2016/07/08. doi: 10.1242/jcs.186908. PubMed PMID: 27383769; PubMed Central PMCID: PMCPMC5004898.
115. Tonkin CJ, van Dooren GG, Spurck TP, Struck NS, Good RT, Handman E, et al. Localization of organellar proteins in Plasmodium falciparum using a novel set of transfection vectors and a new immunofluorescence fixation method. Molecular and biochemical parasitology. 2004;137(1):13-21. Epub 2004/07/29. doi: 10.1016/j.molbiopara.2004.05.009. PubMed PMID: 15279947.
116. Lucky AB, Sakaguchi M, Katakai Y, Kawai S, Yahata K, Templeton TJ, et al. Plasmodium knowlesi Skeleton-Binding Protein 1 Localizes to the 'Sinton and Mulligan' Stipplings in the Cytoplasm of Monkey and Human Erythrocytes. PLoS One. 2016;11(10):e0164272. Epub 2016/10/13. doi: 10.1371/journal.pone.0164272. PubMed PMID: 27732628; PubMed Central PMCID: PMCPMC5061513.
117. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357-60. Epub 2015/03/10. doi: nmeth.3317 [pii]

10.1038/nmeth.3317. PubMed PMID: 25751142.
118. Quinlan AR. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinformatics. 2014;47:11 2 1-34. Epub 2014/09/10. doi: 10.1002/0471250953.bi1112s47. PubMed PMID: 25199790.
119. Anders S, Huber W. Differential expression analysis for sequence count data. Genome biology. 2010;11(10):R106. Epub 2010/10/29. doi: 10.1186/gb-2010-11-10-r106. PubMed PMID: 20979621; PubMed Central PMCID: PMC3218662.
120. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30. Epub 2013/11/15. doi: btt656 [pii]

10.1093/bioinformatics/btt656. PubMed PMID: 24227677.

121. Zhu L, Mok S, Imwong M, Jaidee A, Russell B, Nosten F, et al. New insights into the Plasmodium vivax transcriptome using RNA-Seq. Sci Rep. 2016;6:20498-. doi: 10.1038/srep20498. PubMed PMID: 26858037.
122. Toenhake CG, Fraschka SA-K, Vijayabaskar MS, Westhead DR, van Heeringen SJ, Bártfai R. Chromatin accessibility-based characterization of the gene regulatory network underlying Plasmodium falciparum blood-stage development. Cell Host Microbe. 2018;23(4):557-69.e9. doi: 10.1016/j.chom.2018.03.007. PubMed PMID: 29649445.
123. Fraschka SA, Filarsky M, Hoo R, Niederwieser I, Yam XY, Brancucci NMB, et al. Comparative Heterochromatin Profiling Reveals Conserved and Unique Epigenome Signatures Linked to Adaptation and Development of Malaria Parasites. Cell Host Microbe. 2018;23(3):407-20.e8. Epub 2018/03/01. doi: 10.1016/j.chom.2018.01.008. PubMed PMID: 29503181.
124. Fraschka SA, Filarsky M, Hoo R, Niederwieser I, Yam XY, Brancucci NMB, et al. Comparative Heterochromatin Profiling Reveals Conserved and Unique Epigenome Signatures Linked to Adaptation and Development of Malaria Parasites. Cell Host Microbe. 2018;23(3):407-20 e8. Epub 2018/03/06. doi: 10.1016/j.chom.2018.01.008. PubMed PMID: 29503181; PubMed Central PMCID: PMCPMC5853956.
125. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8. Epub 2002/02/16. doi: 10.1006/meth.2001.1262. PubMed PMID: 11846609.

## Figure legends

**Figure 1. Identification of the PfGCN5 Core Complex in *P. falciparum*.**

(A) Proteins identified from parasite nuclear extracts by IP and by LC-MS/MS. TAP procedure was performed using the PfGCN5::PTP line (three replicates R1 – R3), while single-step IP with the anti-Myc beads was done using the PfPHD1::Myc parasite line (two replicates, R1, and R2). The proteomic data were analyzed by SAINT using a threshold of probability >94% and 1% FDR. Nine proteins consistently identified are marked as the PfGCN5 complex core subunits. Gene ID and annotation are shown on the right.

(B) Schematic diagrams showing the features (putative domains and protein size) of the core subunits.

(C) Gel filtration analysis of the PfGCN5 complex. Aliquots of different fractions were used for Western blots with anti-PTP, PfPHD1 and PfPHD2 antibodies, and for the HAT assay using recombinant histone H3.

**Figure 2. Growth phenotypes in PfGCN5 and PfPHD1 domain deletion mutants.**

(A) Asexual growth rates of WT 3D7, PfGCN5-ΔBrD::GFP, and PfPHD1-ΔPHD::GFP. ** indicate $P < 0.01$ (ANOVA) at days 5 and 7.

(B) The distribution of the number of mature schizonts with a variable number of merozoite. No differences were identified among the three parasite lines ($P > 0.05$, ANOVA).

(C) Merozoite invasion rates showing significantly reductions in GCN5-ΔBrD::GFP and PHD1-ΔPHD::GFP parasite lines ($P < 0.05$, paired Wilcoxon test).

(D) The duration of the IDC showing significantly increased lengths in GCN5-ΔBrD::GFP and PHD1-ΔPHD::GFP parasite lines ($P < 0.05$, paired Wilcoxon test).

(E, F) Detailed analysis of the IDC showing extended ring stage in the GCN5-ΔBrD::GFP (E) and PHD1-ΔPHD::GFP (F).

(G) Gametocytemias at day 6 after induction of gametocytogenesis showing significantly increased gametocytemia in the two domain deletion mutants ($P < 0.05$, paired Wilcoxon test).

**Figure 3. Domain deletions affect the abundance and localization of active histone marks and the integrity of the PfGCn5 complex.**

(A) The levels of active histone marks in 3D7, PfGCN5-ΔBrD::GFP (ΔBrD), and PfPHD1-ΔPHD::GFP (ΔPHD) parasite lines. Histones were purified from the ring, trophozoite, and

schizont stages, and detected by Western blots with specific antibodies against the modified histones H3K9ac, H3K14ac, H3K4me3 and H4Acs. Anti-H3 antibodies were used for loading control.

(B, C) Co-localization of full-length PfGCN5 (GCN5::GFP) (B) or truncated PfGCN5 (GCN5-ΔBrd::GFP) (C) with H3K9ac and DAPI by IFA with anti-GFP and H3K9Ac antibodies. Note the expansion of the truncated PfGCN5-ΔBrD::GFP and H3K9ac beyond the periphery of the euchromatin areas demarcated by DAPI staining.

(D) Effects of domain deletions in PfGCN5 and PfPHD1 on complex integrity. Proteins were pulled down from the trophozoite nuclear extracts of the PfGCN5::GFP, GCN5-ΔBrD::GFP, and PHD1-ΔPHD::GFP parasite lines and identified by LC-MS/MS. R1, R2, R3 and R4 indicate individual repeats of the experiment. Shown here are proteins passing the threshold of SAINT (probability >94% and FDR <1%). The nine PfGCN5 complex core subunits were all detected in the IPs of PfGCN5::GFP and GCN5-ΔBrD::GFP, whereas only four of the core subunits were identified in the IPs of PHD1-ΔPHD::GFP.

**Figure 4. Global transcriptomic changes up domain deletions in PfGCN5 and PfPHD1.**

(A) The phaseograms of transcriptome from the WT 3D7, PfGCN5-ΔBrD::GFP (ΔBrd), PfPHD1-ΔPHD::GFP (ΔPHD) showing the disturbance of the cascade-like gene expression pattern in the deletion mutants at different developmental stages. R, ring; ET, early trophozoite; LT, late trophozoite; S, schizont.

(B, C) Volcano plots showing altered gene expression at the ring (B) and schizont (C) stages in PfGCN5-ΔBrD::GFP compared to the WT 3D7. The x axis indicates $\log_2$ (Fold change) of the transcript level in PfGCN5-ΔBrD::GFP compared to WT 3D7, while the y axis indicates $-\log_{10}$ of the *P* values.

(D) Number of genes with altered expression at different developmental stages of the IDC in the two domain deletion mutants. The up- and down- regulated genes are labeled in red and blue, respectively.

(E, F) Volcano plots showing altered gene expression at the ring (E) and schizont (F) stages in PfPHD1-ΔPHD::GFP compared to the WT 3D7. The x axis indicates $\log_2$ (Fold change) of the transcript level in PfPHD1-ΔPHD::GFP compared to WT 3D7, while the y axis indicates $\log_{10}$ of the *P* values.

(G, H) Overlaps of down-regulated (G) and up-regulated (H) genes between the PfGCN5-ΔBrD::GFP and PfPHD1-ΔPHD::GFP parasite lines at different stages.

**Figure 5. Biological processes and virulence gene expression altered upon domain deletions.**
(A, B) Gene ontology enrichment analysis of up-regulated (A) and down-regulated (B) genes in PfGCN5-ΔBrD::GFP parasites compared to WT 3D7.
(C) Pie graphs showing the overall levels of the *var* gene transcripts in the WT 3D7, PfGCN5-ΔBrD::GFP, and PfPHD1-ΔPHD::GFP parasite lines at the ring stage. The numbers in parentheses are the total numbers of reads of all *var* genes identified by RNA-seq analysis.
(D) Western blot showing PfEMP1 protein levels in the iRBC membranes of the WT 3D7, PfGCN5-ΔBrD::GFP, and PfPHD1-ΔPHD::GFP parasite lines with the anti-ATS antibodies.
(E) Representative images of RNA FISH analysis showing single locus of the B-type *var* gene expression in the WT 3D7 and more than one B-type *var* gene locus in the two deletion mutants.

**Figure 6. Down-regulation of invasion-related pathway and alteration of AP2 genes in domain deletion mutants.**
(A) Heatmaps displaying down-regulation of genes involved in the invasion of the RBC in the two deletion mutants. * and # indicate the AP2-I and putative AP2-LT target genes, respectively.
(B) Heatmaps showing altered expression of the AP2 genes in the deletion mutants.
(C) Putative AP2-LT target genes (with AP2-LT binding motifs) are significantly enriched in the down-regulated genes in PfGCN5-ΔBrD::GFP at the late stages. R, ring; ET, early trophozoite; LT, late trophozoite; S, schizont. ***, $P < 0.001$ (Fisher's exact test).

**Figure 7. Correlation of genes showing altered expression in domain deletion mutants with promoter accessibility and chromatin states.**
(A, B) Changed levels of gene expression in PfGCN5-ΔBrd and PfPHD1-ΔPHD are negatively correlated with the accessibility of the promoters (from the ATAC-seq analysis) (A), but are positively correlated with the heterochromatin state (represented by the HP1 occupancy) (B). U, upregulation; D, downregulation. *, $P < 0.05$; ****, $P < 0.0001$ (Wilcoxon rank-sum test).
(C) Heatmaps displaying the transcriptional activation of gametocyte and ookinete genes in the domain deletion mutants. R, ring; ET, early trophozoite; LT, late trophozoite; S, schizont.

(D) Overlaps of activated gametocyte- or ookinete-specific genes between PfGCN5-ΔBrd and PfPHD1-ΔPHD.

## Supplementary Figures

**Figure S1. PfGCN5:PTP tagging and TAP purification. A.** Schematic diagram of PTP tagging at C-terminal of PfGCN5. P1 and P2 are primers used for verification of integration by PCR. **B.** Cartoon shows the TAP procedure for purification of PfGCN5 complex. A, B, and C are subunits of the GCN5 complex. TEV: tobacco etch virus protease. **C**. PCR verification of positive clones from two clones (C1 and C3) from transfected parasite. **D**. Western blot detecting PfGCN5:PTP in the recombinant parasite clone C3 at different developmental stages (R: ring; ET: early trophozoite; LT: late trophozoite; S: schizont). The blot was probed with antibodies against protein C. Molecular markers in kDa are shown on the left. The expression of aldolase was used for equal loading control. The protein bands are indicated by asterisks. **E.** Western blot detecting PfGCN5::PTP at late trophozoite stage with or without 10 μM E64 treatment for 12 h. E64 blocks processing of PfGCN5.

**Figure S2. PfGCN5 expression and localization during the IDC.** Live cell imaging shows the localization of PfGCN5 in GCN5::GFP parasite under a fluorescence microscope. DAPI was used to stain nucleus. BF, bright field.

**Figure S3. PHD domains in PfPHD1 and PfPHD2 proteins. A**. Sequence of four PHD domains in PfPHD1, C and H amino acid residues in the PHD domain are highlighted underneath the sequence. PHD-SF: PHD superfamily; ePHD: elongated PHD domain, PHD_TAF3: TAF3 type PHD domain. **B**. Sequences of four PHD domains in PfPHD2. **C**. Alignment of PfPHD1 PHD_TAF3 domain with other known authentic PHD domains which bind H3K4me3/2. Alignment shows the conserved Zinc-binding residues in light gray for Zinc 1 and dark gray for Zinc 2, and the two core β-strands in green. The residues involved in H3K4me3 recognition are labeled I through V (forming the aromatic cages) and the aromatic residues in the recognition cage are shadowed in green. MLL1: mixed-lineage leukemia-1; JARID1A: jumonji, AT-rich interactive domain 1A; PYGO: pygopus homolog 1; BPTF:

bromodomain PHD finger transcription factor; PHF2: PHD finger protein 2; Yng1: yeast homolog of mammalian ING1; ING4: inhibitor of growth protein 4; TAF3: transcription initiation factor TFIID subunit 3.

**Figure S4. Association of PfPHD1 or PfPHD2 with PfGCN5. A**. Nuclear localizations of PfPHD1 and PfPHD2 were interrogated by IFA using anti-PfPHD1 and PfPHD2 antibodies. Nuclei were counter-stained by DAPI. **B**. IPs of proteins from lysates of synchronized trophozoites of the PfGCN5::PTP parasite line using agarose conjugated with either anti-PfPHD1 or PfPHD2 antibodies. Immunoprecipitated proteins were separated by SDS-PAGE and probed with anti-Protein C antibodies recognizing the PTP tagged GCN5. Pre-immune sera were used as controls.

**Figure S5. Tagging of PfPHD1 (PF3D7_1008100) with c-Myc. A.** Schematic diagram of Myc tagging at C-terminal of PfPHD1. A, AvrII; S, StuI. **B.** Southern blot of 3D7 and three transgenic clones (1-3). Genomic DNA was digested with AvrII and StuI and hybridized with labeled DNA shown as "Probe" in **A**. **C.** Western blot analysis of nuclear (Nu) and cytoplasmic (Cyto) protein extracts with antibodies against the Myc tag, aldolase (for cytoplasmic compartment) and histone H3 (for nuclear compartment).

**Figure S6. Tagging of PF3D7_1019700) with GFP. A.** Diagram shows GFP tagging of the PF3D7_1019700 at its C-terminus by single-crossover homologous recombination. Purple blocks show the fragment used for homologous recombination. **B.** Integration-specific PCR using primers P1 and P2. WT, Wildtype 3D7; C1 and C2 are two transgenic clones**. C.** Live cell imaging shows the localization of PF3D7_1019700::GFP in the nuclei by fluorescence microcopy. DAPI was used to stain the nuclei.

**Figure S7. Deletion of PfGCN5 Bromodomain (BrD) and the PfPHD1 PHD-TAF3 domain. A.** Schematic showing BrD deletion by single crossover homologous recombination. **B.** Southern blot analysis of three positive clones from transfected parasites. Genomic DNA was digested with AvrII and SpeI, and hybridized to the probe marked in **A**. **C.** Schematic showing the deletion of PHD-TAF3 domain. **D.** Southern blot of two positive clones from transfected

parasites. Genomic DNA was digested with SpeI and XhoI, and hybridized to the probe marked in C. In both cases, GFP was tagged at the ends of truncated genes. Probes for Southern blots are marked. **E**. Images of Giemsa-stained films of parasite cultures synchronized at the ring stage to show the extended IDC of the two domain deletion parasite lines.

**Figure S8. Expression of truncated PfGCN5 and PfPHD1 after domain deletions**. **A**. Western blot shows the size changes of the truncated GCN5 protein bands and the reduced expression levels after BrD deletion. **B.** Live cell imaging shows GFP signals in parasite with truncated GCN5 in GCN5-ΔBrd::GFP parasite line. Compared to Figure S2, the GCN5-ΔBrd-GFP protein shows weaker fluorescence and a more diffused nuclear localization pattern. **C.** Localization of truncated PfPHD1 in PHD1-ΔPHD::GFP parasite line. Compared to Figure S4A, the PfPHD1-ΔPHD::GFP protein also shows a more diffused nuclear localization pattern.

**Figure S9. The effect of domain deletions in PfGCN5 and PfPHD1 on transcription. A–D.** Volcano plots show the genes with altered transcription at the early trophozoite **(A)** and late trophozoite **(B)** stages in PfGCN5-ΔBrD, and at the early trophozoite **(C)** and late trophozoite **(D)** stages in PfPHD1-ΔPHD. **E.** Pearson correlation in fold change between PfGCN5-ΔBrD and PfPHD1-ΔPHD in different developmental stages.

**Figure S10. Transcriptional alteration upon domain deletions.** Gene ontology enrichment analysis of up- (**A**) and down-regulated (**B**) genes in PfPHD1-ΔPHD parasites. **C.** Heatmaps display the alterations of gene transcription associated with protein translation and gene transcription upon domain deletions. **D.** Putative target genes of AP2-LT were significantly enriched in those that down-regulated in late stages of PHD1-ΔPHD::GFP. R, ring; ET, early trophozoite; LT, late trophozoite; S, schizont. ***, $P$ <0.001 (Fisher's exact test).

**Figure S11. Enrichment of PfGCN5 or PfPHD1 at the promoters of genes was correlated with the activation status of the genes.** The enrichment of PfGCN5 or PfPHD1 with (+) or without (-) domain deletion at the promoters of genes at the ring (R) and schizont (S) stages. **A.** *MSP1*; **B.** *VAR2CSA*; **C.** *Pfg27/25*. Enrichment was determined by chromatin-immunoprecipitation (ChIP) followed by qPCR using primer pairs marked as 1, 2, and 3 located

in the promoters of the respective genes. Upon domain deletion, PfGCN5-ΔBrD and PfPHD1-ΔPHD were depleted in the promoters of *MSP1* at schizont but highly enriched at ring stage as comparing to control parasites (**A**). They were enriched at the promoters of *var2csa* (**B**) and the sexual stage gene Pfg27/25 (**C**) at the ring stage. (*, **, and *** donate P<0.05, 0.01 and 0.001, paired Mann Whitney u test)

## Supplementary Tables

**Table S1. Proteomic analyses of GCN5 associated complex**. **A**. Proteomic data from GCN5-PTP TAPs. **B**. SAINT analysis of GCN5-PTP TAPs. **C**. Proteomic data from PHD1-Myc IPs. **D**. SAINT analysis of PHD1-Myc IPs. **E**. Proteomic data from PF3D7_1019700-GFP IPs. **F**. SAINT analysis of PF3D7_1019700-GFP IPs.

**Table S2. Proteomic analyses of GCN5 associated complex after domain deletions. A**. Proteomic data from GCN5-GFP IPs. **B**. SAINT analysis of GCN5-GFP IPs. **C**. Proteomic data from GCN5-ΔBrD-GFP IPs. **D**. SAINT analysis of GCN5-ΔBrD-GFP IPs IPs. **E**. Proteomic data from PHD1-ΔPHD-GFP IPs. **F**. SAINT analysis of PHD1-ΔPHD-GFP IPs.

**Table S3. Transcriptome data of GCN5-ΔBrD::GFP line as compared to 3D7 wildtype.** Deseq2 analysis of three replicates of RNAseq data at ring (**A**), early trophozoite (**B**), late trophozoite (**C**) and schizont (**D**) stages.

**Table S4. Transcriptome data of PHD1-ΔPHD::GFP line as compared to 3D7 wildtype.** Deseq2 analysis of three replicates of RNAseq data at ring (**A**), early trophozoite (**B**), late trophozoite (**C**) and schizont (**D**) stages.

**Table S5. GO enrichment analyses of altered genes upon domain deletions. A**. GO enrichment analyses of altered genes upon BrD domain deletion in GCN5. **B**. GO enrichment analyses of altered genes upon PHD domain deletion in PHD1.

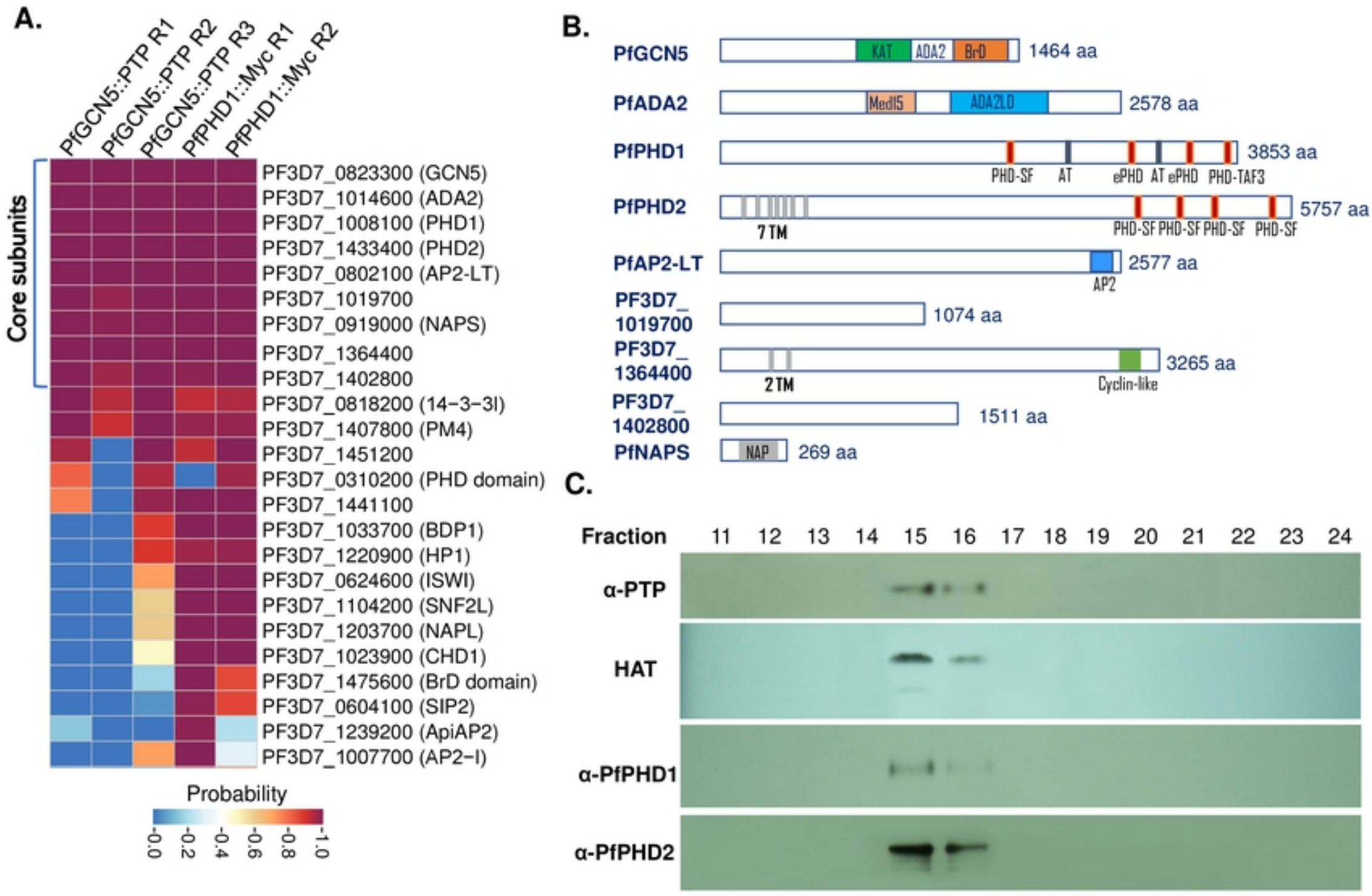
**Table S6. Transcriptional alteration of different biological pathways upon domain deletions. A**. Up-regulation of *var* gene expression at early asexual stage upon domain deletions. **B**. Down-regulation of invasion related genes upon domain deletions. **C**. Down-regulation of translation related genes upon domain deletions. **D**. Downregulation of transcription related genes upon domain deletions. **E**. Alteration of AP2 gene expression upon domain deletions.

**Table S7. Transcriptional escalation of HP1 controlled genes upon domain deletions.**

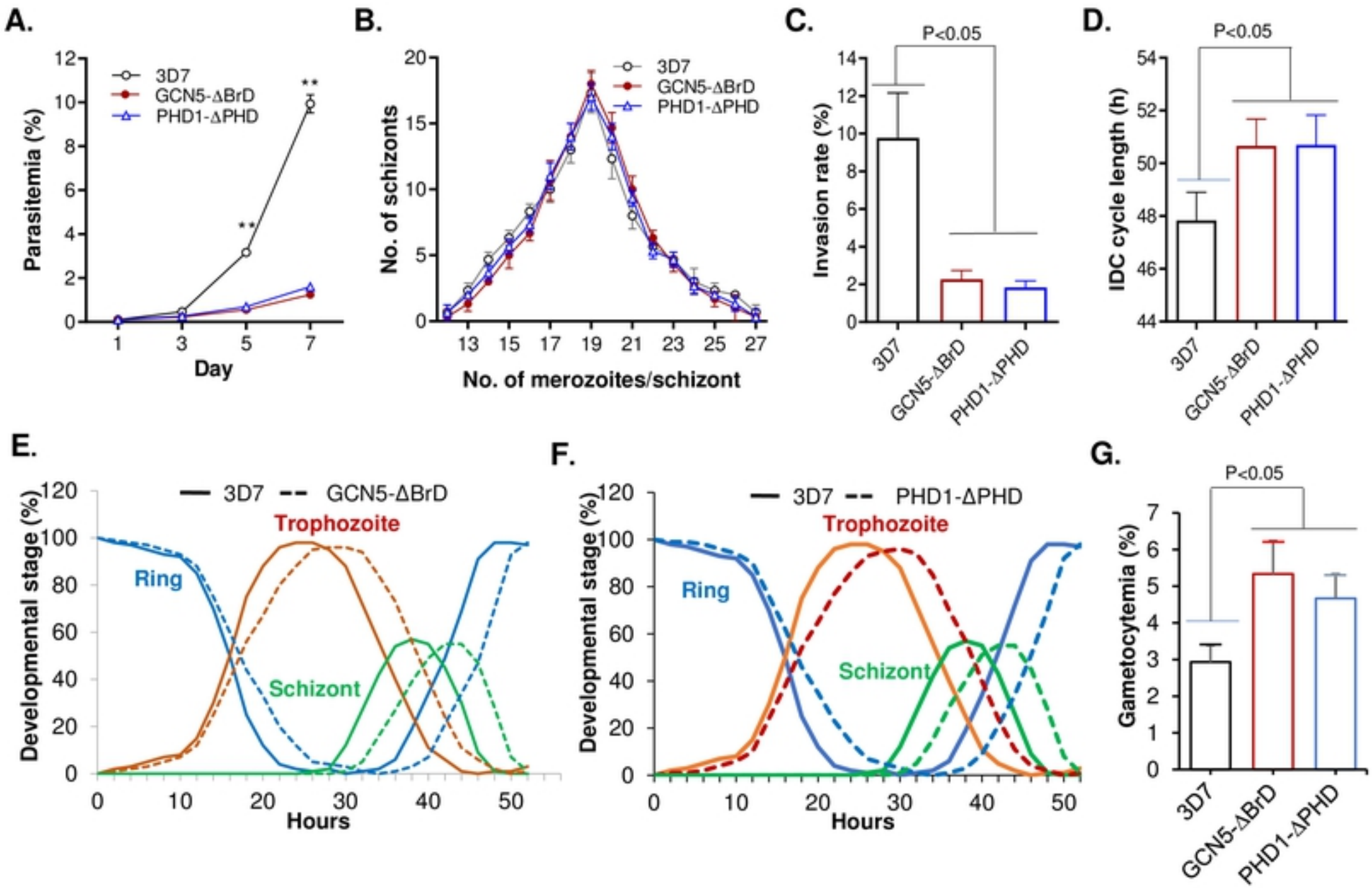
**Table S8. Transcriptional escalation of gametocyte and ookinete specific genes upon domain deletions.**

**Table S9. Primers list. A**. for tagging, domain deletion, integration checking and probe. **B**. for ChIP-qPCR and FISH

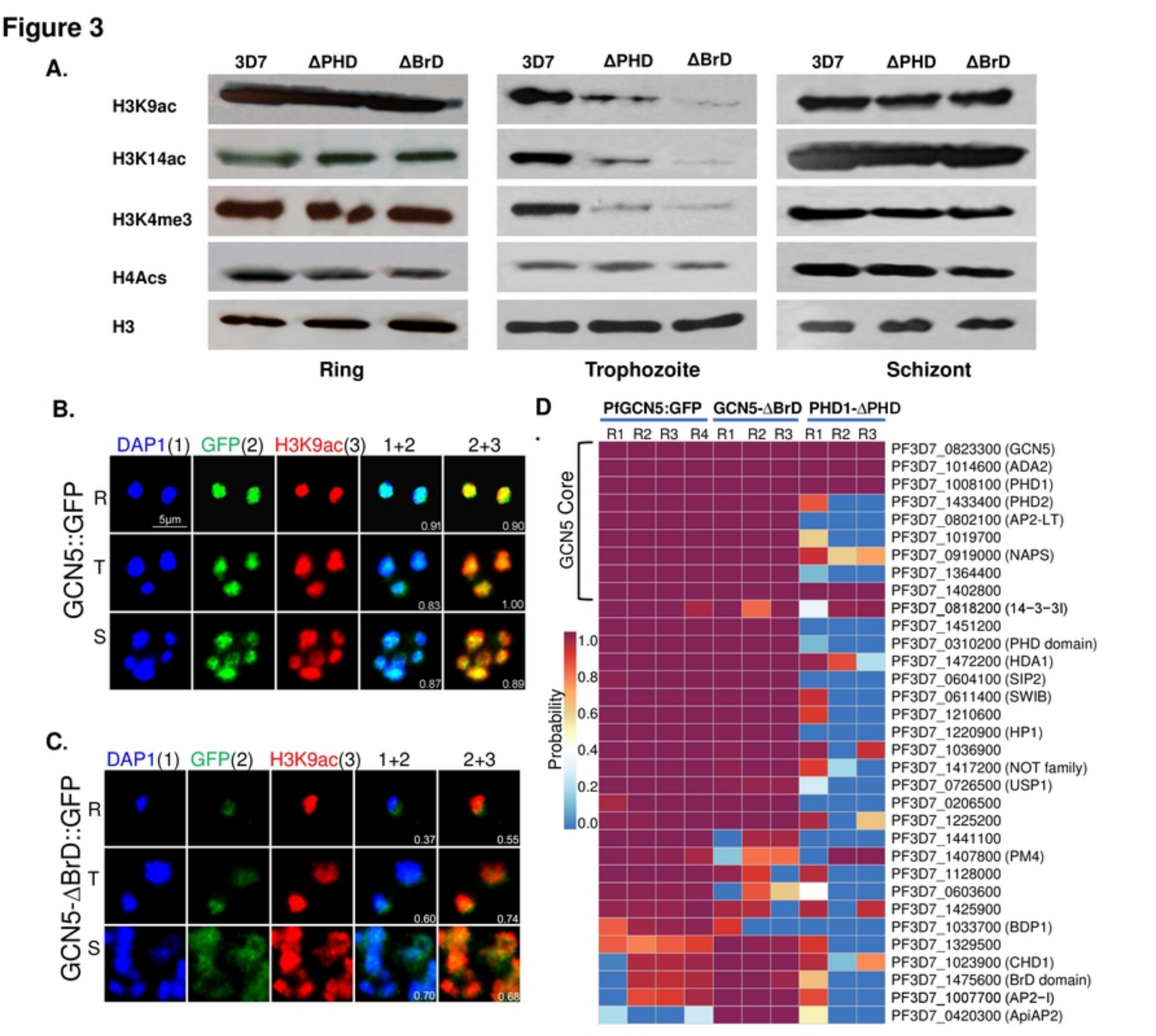
# Figure 1


A.
PfGCN5::PTP R1
PfGCN5::PTP R2
PfGCN5::PTP R3
PfPHD1::Myc R1
PfPHD1::Myc R2
Core subunits
PF3D7_0823300 (GCN5)
PF3D7_1014600 (ADA2)
PF3D7_1008100 (PHD1)
PF3D7_1433400 (PHD2)
PF3D7_0802100 (AP2-LT)
PF3D7_1019700
PF3D7_0919000 (NAPS)
PF3D7_1364400
PF3D7_1402800
PF3D7_0818200 (14-3-3I)
PF3D7_1407800 (PM4)
PF3D7_1451200
PF3D7_0310200 (PHD domain)
PF3D7_1441100
PF3D7_1033700 (BDP1)
PF3D7_1220900 (HP1)
PF3D7_0624600 (ISWI)
PF3D7_1104200 (SNF2L)
PF3D7_1203700 (NAPL)
PF3D7_1023900 (CHD1)
PF3D7_1475600 (BrD domain)
PF3D7_0604100 (SIP2)
PF3D7_1239200 (ApiAP2)
PF3D7_1007700 (AP2-I)
Probability
-0.0
-0.2
-0.4
-0.6
-0.8
-1.0
B.
PfGCN5
KAT
ADA2
BrD
1464 aa
PfADA2
Med15
ADA2LD
2578 aa
PfPHD1
PHD-SF
AT
ePHD
AT
ePHD
PHD-TAF3
3853 aa
PfPHD2
7 TM
PHD-SF
PHD-SF
PHD-SF
PHD-SF
5757 aa
PfAP2-LT
AP2
2577 aa
PF3D7_1019700
1074 aa
PF3D7_1364400
2 TM
Cyclin-like
3265 aa
PF3D7_1402800
1511 aa
PfNAPS
NAP
269 aa
C.
Fraction
11
12
13
14
15
16
17
18
19
20
21
22
23
24
α-PTP
HAT
α-PfPHD1
α-PfPHD2

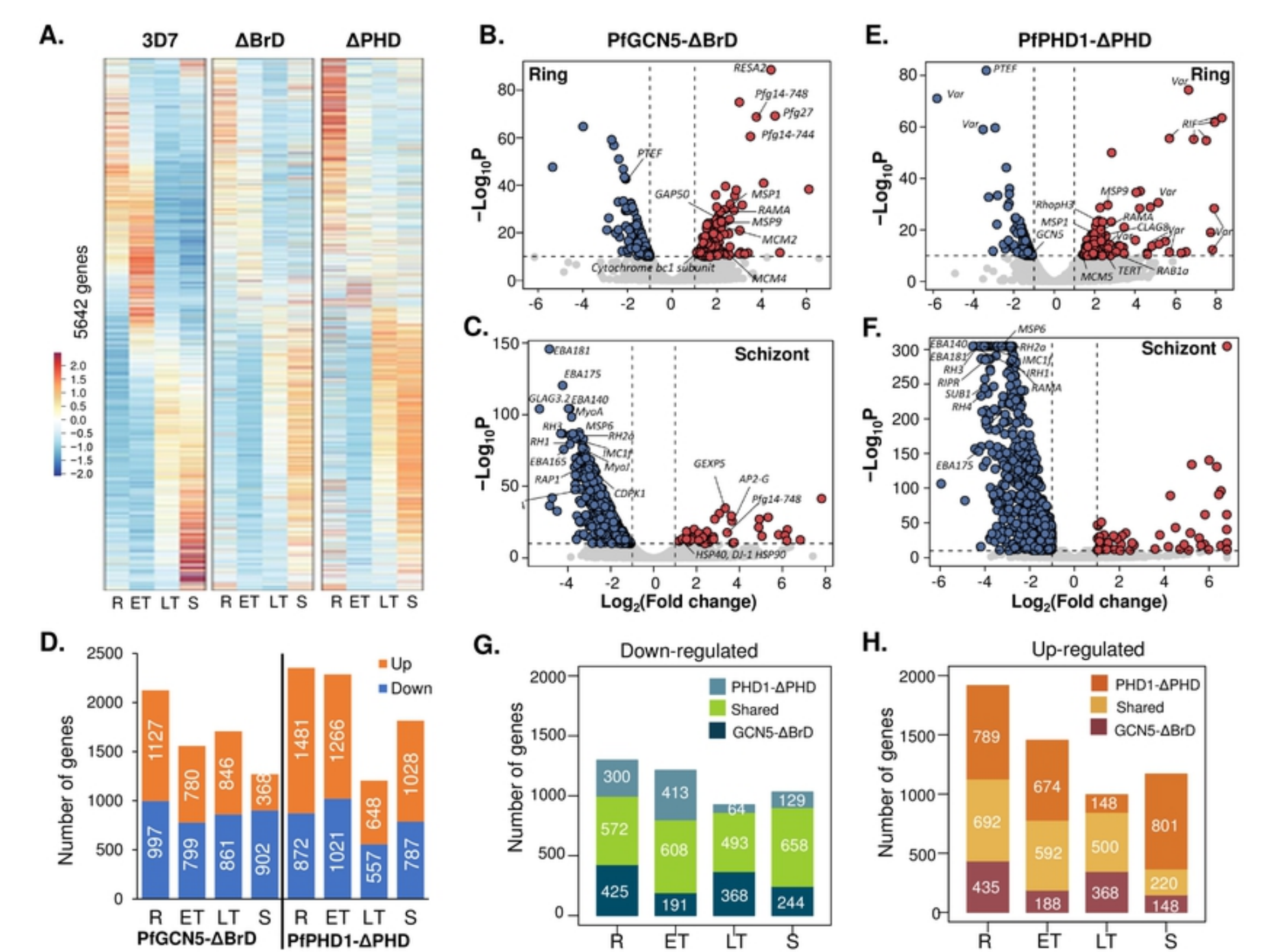
# Figure 2


A.
Parasitemia (%)
Day
3D7
GCN5-ΔBrD
PHD1-ΔPHD
**
**
B.
No. of schizonts
No. of merozoites/schizont
3D7
GCN5-ΔBrD
PHD1-ΔPHD
C.
P<0.05
Invasion rate (%)
3D7
GCN5-ΔBrD
PHD1-ΔPHD
D.
P<0.05
IDC cycle length (h)
3D7
GCN5-ΔBrD
PHD1-ΔPHD
E.
Developmental stage (%)
Hours
3D7
GCN5-ΔBrD
Trophozoite
Ring
Schizont
F.
Developmental stage (%)
Hours
3D7
PHD1-ΔPHD
Trophozoite
Ring
Schizont
G.
P<0.05
Gametocytemia (%)
3D7
GCN5-ΔBrD
PHD1-ΔPHD

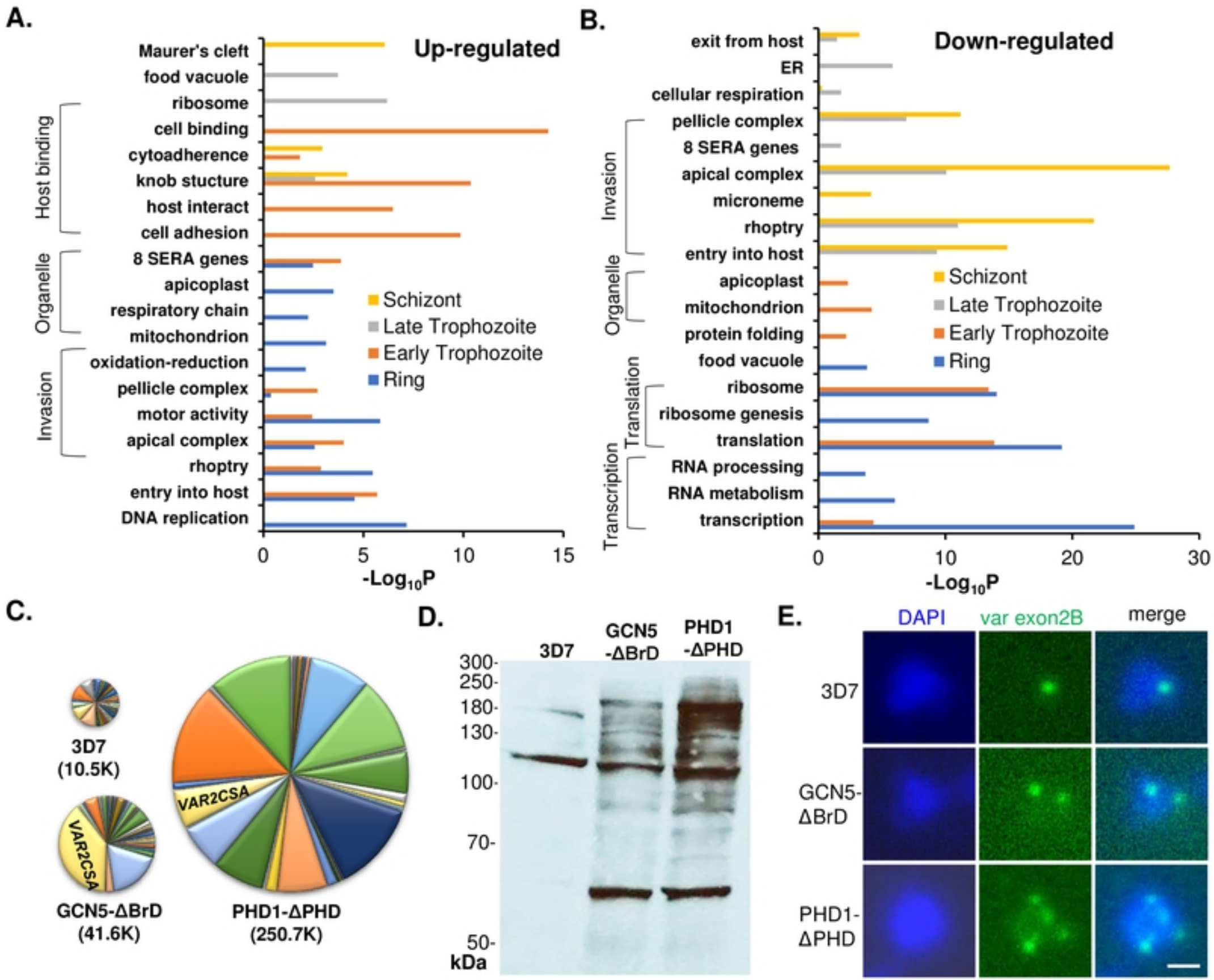
# Figure 3


A.
3D7
ΔPHD
ΔBrD
H3K9ac
H3K14ac
H3K4me3
H4Acs
H3
Ring
Trophozoite
Schizont
B.
DAP1(1)
GFP(2)
H3K9ac(3)
1+2
2+3
GCN5::GFP
R
T
S
5μm
0.91
0.90
0.83
1.00
0.87
0.89
C.
GCN5-ΔBrD::GFP
0.37
0.55
0.60
0.74
0.70
0.68
D
PfGCN5:GFP
GCN5-ΔBrD
PHD1-ΔPHD
R1 R2 R3 R4 R1 R2 R3 R1 R2 R3
GCN5 Core
Probability
1.0
0.8
0.6
0.4
0.2
0.0
PF3D7_0823300 (GCN5)
PF3D7_1014600 (ADA2)
PF3D7_1008100 (PHD1)
PF3D7_1433400 (PHD2)
PF3D7_0802100 (AP2-LT)
PF3D7_1019700
PF3D7_0919000 (NAPS)
PF3D7_1364400
PF3D7_1402800
PF3D7_0818200 (14-3-3I)
PF3D7_1451200
PF3D7_0310200 (PHD domain)
PF3D7_1472200 (HDA1)
PF3D7_0604100 (SIP2)
PF3D7_0611400 (SWIB)
PF3D7_1210600
PF3D7_1220900 (HP1)
PF3D7_1036900
PF3D7_1417200 (NOT family)
PF3D7_0726500 (USP1)
PF3D7_0206500
PF3D7_1225200
PF3D7_1441100
PF3D7_1407800 (PM4)
PF3D7_1128000
PF3D7_0603600
PF3D7_1425900
PF3D7_1033700 (BDP1)
PF3D7_1329500
PF3D7_1023900 (CHD1)
PF3D7_1475600 (BrD domain)
PF3D7_1007700 (AP2-I)
PF3D7_0420300 (ApiAP2)

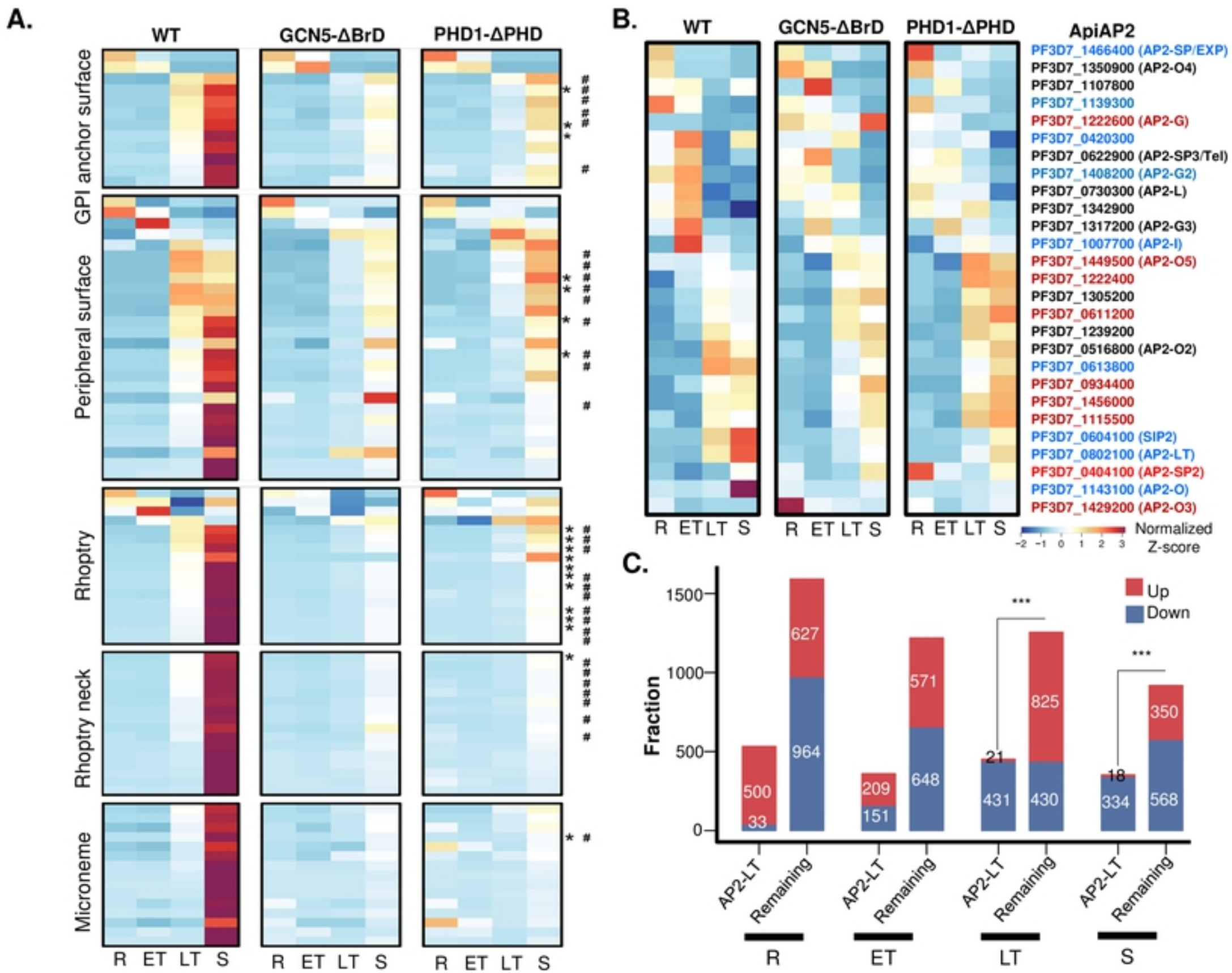
# Figure 4


A.
3D7
ΔBrD
ΔPHD
5642 genes
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
R ET LT S
B.
PfGCN5-ΔBrD
Ring
RESA2
Pfg14-748
Pfg27
Pfg14-744
PTEF
GAP50
MSP1
RAMA
MSP9
MCM2
Cytochrome bc1 subunit
MCM4
−Log10P
C.
Schizont
EBA181
EBA175
GLAG3.2
EBA140
MyoA
RH3
MSP6
RH2a
RH1
IMC1f
EBA165
MyoJ
RAP1
CDPK1
GEXP5
AP2-G
Pfg14-748
HSP40, DJ-1 HSP90
Log2(Fold change)
E.
PfPHD1-ΔPHD
Ring
PTEF
Var
RIF
MSP9
RhopH3
MSP1
GCN5
RAMA
CLAG8
MCM5
TERT
RAB1a
F.
Schizont
MSP6
EBA140
EBA181
RH2a
IMC1f
RH1
RH3
RIPR
RAMA
SUB1
RH4
EBA175
Log2(Fold change)
D.
Up
Down
Number of genes
1127
997
780
799
846
861
368
902
1481
872
1266
1021
648
557
1028
787
R ET LT S
PfGCN5-ΔBrD
PfPHD1-ΔPHD
G.
Down-regulated
PHD1-ΔPHD
Shared
GCN5-ΔBrD
Number of genes
300
572
425
413
608
191
64
493
368
129
658
244
R ET LT S
H.
Up-regulated
PHD1-ΔPHD
Shared
GCN5-ΔBrD
Number of genes
789
692
435
674
592
188
148
500
368
801
220
148
R ET LT S

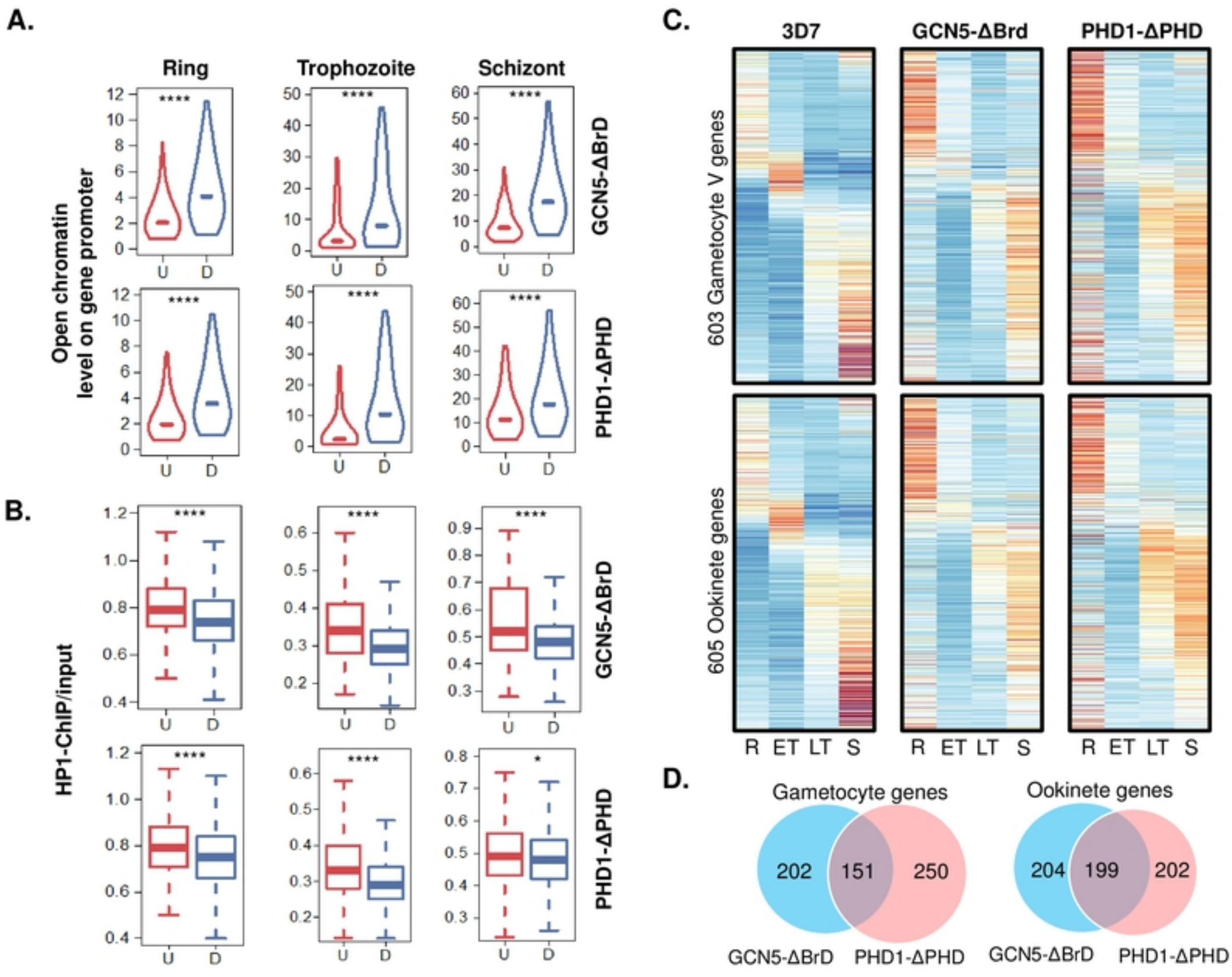
# Figure 5


A.
Up-regulated
Host binding
Organelle
Invasion
Maurer's cleft
food vacuole
ribosome
cell binding
cytoadherence
knob stucture
host interact
cell adhesion
8 SERA genes
apicoplast
respiratory chain
mitochondrion
oxidation-reduction
pellicle complex
motor activity
apical complex
rhoptry
entry into host
DNA replication
Schizont
Late Trophozoite
Early Trophozoite
Ring
0
5
10
15
-Log10P
B.
Down-regulated
Invasion
Organelle
Translation
Transcription
exit from host
ER
cellular respiration
pellicle complex
8 SERA genes
apical complex
microneme
rhoptry
entry into host
apicoplast
mitochondrion
protein folding
food vacuole
ribosome
ribosome genesis
translation
RNA processing
RNA metabolism
transcription
Schizont
Late Trophozoite
Early Trophozoite
Ring
0
10
20
30
-Log10P
C.
3D7
(10.5K)
VAR2CSA
GCN5-ΔBrD
(41.6K)
VAR2CSA
PHD1-ΔPHD
(250.7K)
D.
3D7
GCN5
-ΔBrD
PHD1
-ΔPHD
300-
250-
180-
130-
100-
70-
50-
kDa
E.
DAPI
var exon2B
merge
3D7
GCN5-
ΔBrD
PHD1-
ΔPHD

# Figure 6


A.
WT
GCN5-ΔBrD
PHD1-ΔPHD
GPI anchor surface
Peripheral surface
Rhoptry
Rhoptry neck
Microneme
R ET LT S
B.
ApiAP2
PF3D7_1466400 (AP2-SP/EXP)
PF3D7_1350900 (AP2-O4)
PF3D7_1107800
PF3D7_1139300
PF3D7_1222600 (AP2-G)
PF3D7_0420300
PF3D7_0622900 (AP2-SP3/Tel)
PF3D7_1408200 (AP2-G2)
PF3D7_0730300 (AP2-L)
PF3D7_1342900
PF3D7_1317200 (AP2-G3)
PF3D7_1007700 (AP2-I)
PF3D7_1449500 (AP2-O5)
PF3D7_1222400
PF3D7_1305200
PF3D7_0611200
PF3D7_1239200
PF3D7_0516800 (AP2-O2)
PF3D7_0613800
PF3D7_0934400
PF3D7_1456000
PF3D7_1115500
PF3D7_0604100 (SIP2)
PF3D7_0802100 (AP2-LT)
PF3D7_0404100 (AP2-SP2)
PF3D7_1143100 (AP2-O)
PF3D7_1429200 (AP2-O3)
Normalized Z-score
C.
Fraction
Up
Down
AP2-LT
Remaining
R
ET
LT
S

# Figure 7


A.
Ring
Trophozoite
Schizont
Open chromatin level on gene promoter
GCN5-ΔBrD
PHD1-ΔPHD
U
D
****
B.
HP1-ChIP/input
*
C.
3D7
GCN5-ΔBrd
PHD1-ΔPHD
603 Gametocyte V genes
605 Ookinete genes
R ET LT S
D.
Gametocyte genes
202
151
250
GCN5-ΔBrD
PHD1-ΔPHD
Ookinete genes
204
199
202
GCN5-ΔBrD
PHD1-ΔPHD