The PfAP2-G2 transcription factor is a critical regulator of gametocyte maturation

Suprita Singh¹, Joana M. Santos^{1,*}, Lindsey M. Orchard¹, Naomi Yamada^{3\$}, Riëtte van Biljon¹, Heather J. Painter^{1,#}, Shaun Mahony², Manuel Llinás^{1,2,3,‡}

¹Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA 16802

Huck Center for Malaria Research, The Pennsylvania State University, University Park, PA, USA 16802

²Center for Eukaryotic Gene Regulation, Department of Biochemistry & Molecular Biology, The Pennsylvania State University, University Park, PA, USA 16802

³Department of Chemistry, The Pennsylvania State University, University Park, PA, USA 16802

*Present address: Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.

\$Present address: GSK R&D Functional Genomics, 1250 S. Collegeville Rd., Collegeville, PA, USA 19426

#Present address: Division of Bacterial, Parasitic, and Allergenic Products, Office of Vaccines Research and Review, Center for Biologics Evaluations and Research, Food and Drug Administration, Silver Spring, MD, USA

‡To whom correspondence should be addressed: manuel@psu.edu

Abstract

Differentiation from asexual blood stages to sexual gametocytes is required for transmission of malaria parasites from the human to the mosquito host. Preventing gametocyte commitment and development would block parasite transmission, but the underlying molecular mechanisms behind these processes remain poorly understood. Here, we report that the ApiAP2 transcription factor, PfAP2-G2 (PF3D7_1408200) plays a critical role in the maturation of *Plasmodium falciparum* gametocytes. PfAP2-G2 binds to the promoters of a wide array of genes that are expressed at many stages of the parasite life cycle. Interestingly, we also find binding of PfAP2-G2 within the gene body of almost 3000 genes, which strongly correlates with the location of H3K36me3 and several other histone modifications as well as Heterochromatin Protein 1 (HP1), suggesting that occupancy of PfAP2-G2 in gene bodies may serve as an alternative regulatory mechanism. Disruption of *pfap2-g2* does not impact asexual development, parasite multiplication rate, or commitment to sexual development but the majority of sexual parasites are unable to mature beyond stage III gametocytes. The absence of *pfap2-g2* leads to overexpression of 28% of the genes bound by PfAP2-G2 and none of the PfAP2-g2 bound are downregulated, suggesting that it is a repressor. We also find that PfAP2-G2 interacts with chromatin remodeling proteins, a microrchidia (MORC) protein, and another ApiAP2 protein (PF3D7_1139300). Overall our data demonstrate that PfAP2-G2 is an important transcription factor that establishes an essential gametocyte maturation program in association with other chromatin-related proteins.

Introduction

Malaria is a life-threatening disease that continues to impact the lives of millions of people worldwide. According to the latest WHO estimates, there were 228 million cases of malaria in 2018, resulting in 405,000 deaths (World Malaria Report, WHO, 2019). Malaria is caused by unicellular protozoan parasites belonging to the genus *Plasmodium*. Of the six species that infect humans, *Plasmodium falciparum* has the highest mortality rate (Weiss et al. 2019; WHO 2019). *P. falciparum* exhibits a complex life cycle with asexual and sexual erythrocytic phases

in the human host, followed by development in *Anopheles* mosquitoes which transmit the parasite back to a new host, initiating the pre-erythrocytic liver stage of infection. Although the cyclic 48-hour asexual blood-stage is responsible for symptomatic disease and severe malaria, this form of the parasite cannot be transmitted. Rather, in each round of replication, a fraction of the parasites <10% exit the asexual pathway and undergo sexual differentiation to form male and female gametocytes, which are transmission competent (Bruce et al. 1989). Inhibition of gametocyte development is of great interest, because it would prevent malaria parasite transmission, which is one of the major goals in the effort to achieve disease eradication (Rabinovich et al. 2017).

Gametocyte development in *P. falciparum* is a 9–12 day process, which is substantially longer than that of other human infecting *Plasmodium* species such as *P. vivax* and the rodent malaria parasites *P. berghei* (26–30 hours) and *P. yoelii* (36 hours) (Gautret and Motard, 1999; Liu, Miao, & Cui, 2011; Armistead et al., 2018). *P. falciparum* gametocyte development is divided into five (stage I to V) morphologically distinct stages (Sinden 1982). The earliest phase of gametocyte development, stage Ia, occurs around 24 to 30 hours post-invasion (hpi) and is morphologically indistinguishable from the young trophozoite stage. However, there are several well-defined markers of early stage I gametocytes such as Pfs16 and Pfg27 and PfGEXP-5 (Alano, Premawansa, Bruce, & Carter, 1991; Marian C. Bruce, Carter, Nakamura, Aikawa, & Carter, 1994; Silvestrini et al., 2010; Poran et al., 2017; Josling et al., 2020; Llorà-Batlle et al., 2020). In the human host, stage Ib to stage IV sexually-developing parasites are sequestered in deep tissues like the bone marrow and only stage V gametocytes freely circulate in the blood and can be picked up by mosquitoes (Aguilar et al. 2014; Joice et al. 2014; Venugopal et al. 2020)

Regulation of *Plasmodium* development is driven by stage-specific transcription factors (TFs), such as the well-studied Apicomplexan AP2 (ApiAP2) family of DNA binding proteins. ApiAP2 proteins are found among all members of the phylum, and each ApiAP2 protein contains between one to three Apetala2 (AP2) DNA binding domains (Balaji et al. 2005). In *P. falciparum* there are 27 members of the ApiAP2 protein family. ApiAP2 proteins have been shown to control all

developmental transitions in *Plasmodium* (Jeninga, Quinn, and Petter 2019; Modrzynska et al. 2017; Zhang et al. 2017)

A master regulator of sexual (gametocyte) commitment, AP2-G, has been identified in both *P. falciparum* and *Plasmodium berghei*, a rodent malaria parasite (Kafsack et al. 2014; Sinha et al. 2014). Expression levels of PfAP2-G (PF3D7_1222600) strongly correlate with the formation of gametocytes, and targeted disruption of the *pfap2-g* locus results in a complete block in gametocytogenesis and downregulation of many gametocyte-associated genes (Kafsack et al. 2014). Another regulator of *Plasmodium* gametocyte development is AP2-G3. Disruption of *pyap2-g3* (PY17X_1417400) in *P. yoelii* resulted in significant reduction in the numbers of male and female gametocytes, day 8 oocysts, and sporozoites, but it did not affect the asexual growth of the parasites (Zhang et al. 2017). The *P. berghei* orthologue, PBANKA_1415700, has also been knocked out and shown to be female-specific and essential for female gametocyte development, and was thus renamed PbAP2-FG (Yuda et al. 2019). The *P. falciparum* orthologue, PF3D7_1317200, was also found to be associated with gametocytogenesis (Ikadai et al. 2013), although it has not been extensively characterized.

AP2-G2 has been shown to play a role in gametocytogenesis in both *P. berghei* and *P. yoelii* (Sinha et al., 2014; Yuda, Iwanaga, Kaneko, & Kato, 2015; Modrzynska et al., 2017). Knockout of *pbap2-g2* does not inhibit sexual stage conversion but rather results in the nearly complete loss of gametocyte maturation and a block in transmission to mosquitoes (Yuda et al. 2015). Yuda *et al.* reported that the *P. berghei* transcription factor is bound to roughly 1,500 genes, or slightly over 1/3 of the genome during asexual development, and a number of these genes were up-regulated by more than two-fold in *pbap2-g2* knockout lines. In another study, disruption of *pbap2-g2* also caused premature expression of liver stage and sporozoite stage genes during asexual development in red blood cells (Modrzynska et al. 2017). Moreover, a *P. berghei* liver-stage (LS) transcriptome of *P. berghei* reported strong upregulation of *pbap2-g2* in early LS that is negatively correlated with the expression of liver stage-specific transcripts (Derbyshire 2019). Together, these findings suggest that AP2-G2 acts as a repressor in both the asexual and sexual stages and during early liver-stage

infection in *P. berghei*. Accordingly, *P. yoelii* parasites lacking the *pyap2-g2* gene had greatly reduced numbers of gametocytes and oocysts (Zhang et al. 2017). A recent genome-wide knockout screen suggests that AP2-G2 is also not essential for blood-stage development in *P. falciparum* (Zhang et al. 2018), and may have a role in a later stage of development.

In this study we explore the function of AP2-G2 in *P. falciparum* during parasite blood-stage development. The period of gametocyte maturation for rodent malaria parasites is shorter than in *P. falciparum*, making it difficult to discern the developmental phenotypes associated with *ap2-g2* knockouts. Therefore, disrupting this gene in *P. falciparum* is of interest due to the longer maturation period. Furthermore, two studies have reported differences in asexual blood-stage development resulting from *pbap2-g2* deletion. On one hand, Modrzynska *et al.* observed reduced growth and premature expression of liver- and sporozoite-stage genes (Modrzynska et al. 2017). Conversely, Yuda *et al.* reported that *pbap2-g2* KO parasites proliferate normally in blood (Yuda et al. 2015). Another discrepancy regards the impact of the KO on the sex ratio. Sinha *et al.* reported a difference in the ratio of male to female gametocytes was disrupted in *pbap2-g2* KO lines (Sinha et al. 2014) while Yuda et al. observed no such differences (Yuda et al. 2015).

We also find that the *P. falciparum* orthologue of AP2-G2 also plays a critical role in the maturation of gametocytes. Disruption of *pfap2-g2* does not impact asexual development, parasite multiplication rate, or commitment to sexual development, but parasites are unable to develop normally beyond stage III. Using transcriptomic analysis and ChIP-seq we have identified a number of candidate genes that are bound and regulated by PfAP2-G2. We also identify interacting partners using protein immunoprecipitation followed by mass spectrometry. Overall our work suggests that PfAP2-G2 is a transcriptional repressor that likely recruits additional transcription factors and chromatin remodeling machinery to the genome to control gene expression. PfAP2-G2 plays a critical role in the regulation of the development of malaria parasites as they transition from asexual to sexual parasites, allowing for proper gametocyte maturation.

Results

PfAP2-G2 is expressed during the trophozoite and schizont stages of asexual development

The *P. falciparum* orthologue of AP2-G2, PF3D7_1408200, encodes a 189kDa protein that contains a single AP2 DNA binding domain (**Figure 1A**). *pfap2-g2* is maximally transcribed at the ring and early trophozoite stages (Bozdech et al., 2003; Painter et al., 2018) and proteomics data indicates protein expression at the trophozoite and schizont stages (Oehring et al. 2012). To precisely determine the specific stage of expression of PfAP2-G2 and its subcellular localization, we tagged the gene at the C-terminus with GFP (**Supplementary Fig. 1**). Using live-cell fluorescence microscopy, we imaged a highly synchronized PfAP2-G2::GFP parasite population for one complete asexual replication cycle (48-hours), every 7 h beginning with the newly invaded ring stage (5–7 hpi) revealed that PfAP2-G2 was localized to the nucleus, as expected, and was expressed from the early trophozoite to the late schizont stages (**Figure 1B**). Further confirming the nuclear localization of PfAP2-G2, nuclear fractionation of trophozoite stage parasites (~30 hpi) showed that full length PfAP2-G2::GFP (~250 kDa) was only detected in the nuclear fraction of the parasite lysates (**Figure 1C**). We also detected other smaller sized peptides indicating that PfAP2-G2 may be proteolytically cleaved or is unstable in parasite lysates.

PfAP2-G2 is not required for proliferation during the asexual stages of the life cycle

To determine the role of *pfap2-g2*, we generated a genetic disruption (KO) line using selection linked integration (SLI) (Birnbaum et al. 2017) to truncate the *pfap2-g2* coding sequence (**Supplementary Fig. 2A**). As expected, *pfap2-g2* KO parasites were readily obtained. Live-cell fluorescence imaging of the transgenic parasites showed diffuse cytoplasmic staining suggesting that the remaining non AP2 domain-containing protein fragment no longer localized to the nucleus (**Supplementary Fig. 2B**). We next determined if truncation of PfAP2-G2 impacted asexual stage parasite growth or multiplication rate. Using a SYBR green growth assay (Vossen et al. 2010), we find that *pfap2-g2* KO parasites develop similarly to that of WT (**Figure 2A**), and

have similar multiplication rates (**Figure 2B**). Therefore, although PfAP2-G2 is expressed in the asexual blood stages, it is not required for normal asexual development.

PfAP2-G2 is expressed in the gametocyte stages

Recent transcriptomic data has shown that the mRNA abundance of *pfap2-g2* in gametocytes is low relative to the asexual stages, with a broad peak seen only in the early gametocyte stages followed by low expression during the later stages (**Supplementary Fig. 3**) (Biljon et al. 2019). To characterize protein expression of PfAP2-G2 during gametocyte development we imaged PfAP2-G2::GFP parasites (see methods). PfAP2-G2 was expressed in all stages of gametocyte development, but it was not confined to the nucleus in later stages (**Figure 3A**). Indeed nuclear fractionation of stage III gametocytes detected expression of full-length protein in both fractions, unlike in asexual parasites (**Figure 3B**). It also seems that the nuclear protein undergoes different proteolytical processing in gametocytes versus the asexual stages.

PfAP2-G2 is essential for gametocyte maturation

To determine if sexually committed PfAP2-G2 gametocytes mature fully to Stage V, we induced PfAP2-G2 KO and WT parasite lines to undergo gametocytogenesis and followed their development for 14 days, monitoring their maturation via morphological analysis of Giemsa-stained thin-blood smears. Whereas WT E5 parasites matured into Stage V gametocytes, most *pfap2-g2* KO parasites committed to sexual development could not progress beyond Stage III (**Figure 3C**). PfAP2-G2 is thus a critical regulator of gametocyte maturation and development.

PfAP2-G2 makes extensive genome-wide interactions

Because PfAP2-G2 is predicted to be a DNA-binding protein (Campbell, de Silva, Olszewski, Elemento, & Llinás, 2010; Yuda et al., 2015) we sought to determine the genome-wide localization of PfAP2-G2 in both asexual and gametocyte lifecycle

stages. To do this we first performed ChIP-seq at the trophozoite stage using the PfAP2-G2::GFP parasites. Peak calling identified ~5000 peaks in at least 2 of the 3 biological replicates (FDR 0.05) (**Supplementary Fig. 4A**). All replicates showed a high correlation with each other with respect to the overall peaks identified (**Figure 4A**). To our surprise only 401 of the predicted binding sites, corresponding to 120 genes, were found in the non-coding upstream regions of genes (**Supplementary Table 1**), while 4,600 peaks, corresponding to 2,932 genes, were located in exonic gene bodies (**Supplementary Table 2**). This differs from what has been previously observed for other characterized *P. falciparum* ApiAP2 proteins, such as SIP2 (Flueck et al. 2010), AP2-I (Santos et al. 2017) and AP2-G (Josling et al. 2020) which were found to predominantly bind to non-coding upstream regions. Although widescale binding to gene bodies was not previously reported for PbAP2-G2 (Yuda et al. 2015), a re-analysis of the ChIP-seq data from this study using the same pipeline that we used for our ChIP-seq data analysis, revealed that PbAP2-G2 is also broadly distributed across gene bodies in the *P. berghei* genome with 4,345 binding sites.

An analysis of all the regions (both upstream regions and ORFs) bound by PfAP2-G2, using DREME (Bailey 2011) identified two significantly enriched DNA sequence motifs. The first is an AGAA sequence motif which is related to a previously reported DNA motif found to associate with one (of the three) AP2 domains of the ApiAP2 protein PF3D7_1139300 (Campbell et al. 2010). The second is an ACCA core motif (**Figure 4B**) closely resembling the previously identified PfAP2-G2 binding PBM motif. (Campbell et al. 2010), demonstrating the success of ChIP experiments.

In order to characterize possible regulatory targets of PfAP2-G2, we first focused on the 120 genes that contain PfAP2-G2 peaks in their upstream regions. Nearly 87% of these 120 putative target genes are bound within 2.5kb upstream of the start codon (**Supplementary Fig. 4B**). 278 of the remaining intergenic peaks were localized to the subtelomeric ends of each chromosome (**Supplementary Fig. 4C**). For further analysis we only considered genes with upstream binding interactions less than 2.5 kb from the start codon, leaving 82 candidate genes (**Supplementary Table 1**). Interestingly, PfAP2-G2 at the asexual trophozoite stage binds to the promoters of genes known to be expressed later in the gametocyte and

mosquito lifecycle stages. These include 6-cysteine protein (*p36* PF3D7_0404400), cysteine repeat modular protein 2 (*crmp2* PF3D7_0718300), sporozoite protein essential for cell traversal (*spect* PF3D7_1342500), circumsporozoite protein (*csp* PF3D7_0304600), cell traversal protein for ookinetes and sporozoites (*celts* PF3D7_1216600), secreted protein altered thrombospondin repeat protein (*spatr* PF3D7_0212600), stearoyl-CoA desaturase (*scd* PF3D7_0511200), and *cap380* oocyst capsule protein (PF3D7_0320400) among others (**Supplementary Table 1**) (Chattopadhyay et al. 2003; van Dijk et al. 2010; Espinosa et al. 2017; Gratraud et al. 2009; Ishino et al. 2004; Itsara et al. 2018; Singh et al. 2007; Thompson et al. 2007; Zhao et al. 2016). Representative PfAP2-G2 peaks upstream of two target genes are shown in (**Figure 4C**). We also found PfAP2-G2 associated with the upstream regions of genes that are highly transcribed in the asexual ring stage (as analyzed in (López-Barragán et al. 2011)) including the knob-associated histidine rich protein (*kahrp* PF3D7_0202000), *Plasmodium* helical interspersed subtelomeric proteins (*phista* PF3D7_0115100), several variant family genes including *vars* and *rifins*, and genes encoding proteins involved in egress and invasion such as the Rh5 interacting protein (*ripr* PF3D7_0323400) and serine repeat antigen 7 (*sera7* PF3D7_0207400) (Miller et al. 2002; Pei et al. 2005; Sargeant et al. 2006; Volz et al. 2016). Therefore, PfAP2-G2 binds to the promoters and ORFs of genes that are expressed at a different developmental stage than that in which AP2-G2 is expressed, suggesting that, like in the rodent malaria species, this is a transcriptional repressor protein.

We next examined the genome-wide occupancy of PfAP2-G2 in stage III gametocytes by ChIP-seq. We identified 947 peaks in common across both replicates, corresponding to 674 genes at an FDR of 0.05 (**Figure 4D**) (**Supplementary Table 3, Supplementary Table 4**). PfAP2-G2 binding, once again, was found in both upstream regions and within gene bodies. Although PfAP2-G2 bound many fewer regions in gametocytes compared to asexual trophozoites, which may be associated with the difficulty in obtaining high amounts of chromatin in gametocytes, virtually all stage III PfAP2-G2 target genes were also targets during the trophozoite stages (**Figure 4E**). This suggests that there are a large number of genes that are always

bound by PfAP2-G2 during both asexual and sexual parasite development (**Figure 4F**). DNA motif analysis yet again identified the ACCA (**Figure 4G**).

Genetic disruption of PfAP2-G2 leads to aberrant gene expression in both asexual and sexual life cycle stages

Using synchronized parasites, we collected seven total RNA samples from WT or KO parasites throughout the 48-hour asexual life cycle for transcriptome analysis. In a separate experiment, total RNA samples were collected during sexual differentiation every 12 hours for 7 continuous days starting at 2 days post-gametocyte induction (**Supplementary Fig. 5**). Gene expression analysis of the asexual stage timecourse using Significance Analysis of Microarray (SAM) (Tusher, Tibshirani, and Chu 2001) identified 327 differentially expressed transcripts (>1.5 log₂FC, 0.30 FDR) in the PfAP2-G2 KO line. Out of these, 237 showed enhanced transcript abundance in the PfAP2G2 KO line and 90 showed decreased abundance (**Supplementary Fig. 6, Supplementary Table 5**). This result is quite surprising given the lack of any asexual growth phenotype between the PfAP2-G2 KO and the WT (**Figure 2**), implying that these changes in gene expression do not impact asexual parasite fitness or gametocyte commitment. A comparison of transcript abundance in the KO and WT sexual stages identified increased abundance for 274 transcripts and reduced abundance for 169 transcripts in the PfAP2-G2 KO line (>1.5 log₂FC, 0.30 FDR) (**Supplementary Fig. 6, Supplementary Table 6**). These differential transcription patterns presumably lead to the stall in development at Stage III gametocytes in parasites lacking PfAP2-G2.

To determine at which stage the significantly changing genes (by SAM) are maximally transcribed in wild-type parasites, we used a published RNA-seq dataset covering four different asexual stages (rings, early trophozoite, late trophozoite, and schizonts), two gametocyte stages (stage II and stage V), and the ookinete stage using the 3D7 *P. falciparum* strain (López-Barragán et al. 2011). Using this dataset, we identified the 3D7 parasite lifecycle stage at which each gene has the highest reads per kilobase of transcript, per million mapped sequencing reads (RPKM). For the asexual transcriptome, most of the differentially expressed genes we measured as

differentially regulated in the PfAP2-G2 KO are normally transcribed maximally in stage V gametocyte (and not in asexual stages), followed by the ookinete and ring stages (**Figure 5A**). This suggests that loss of PfAP2-G2 impacts the expression of genes from late-stage gametocytes, ookinetes, and asexual stages. However the number of genes showing enhanced transcript abundance is strikingly high for genes that are normally expressed in stage V gametocytes (**Figure 5A**). Therefore, PfAP2-G2 may prevent the premature expression of late-stage gametocyte genes during asexual development.

In the gametocyte transcriptome, we observed the opposite trend when comparing to the López-Barragán data, where 70% of the transcripts that showed increased abundance in the KO line were usually maximally abundant in the asexual stages, especially at the ring stage (**Figure 5A**). Examples of genes whose mRNA abundance was higher in gametocytes include knob-associated histidine-rich protein (*kahrp* PF3D7_0202000), 10 different *Plasmodium* helical interspersed subtelomeric proteins (*phista*, *phistb*, *phistc*), 7 different serine/threonine protein kinase (*FIKK* family) proteins, 12 genes encoding Plasmodium exported proteins (*hyp2*, *hyp8*, *hyp9*, *hyp10*, *hyp11*, *hyp12*, *hyp16*) merozoite surface proteins (*msp1*, *msp2*, *msp6*, *msp7*, *msp11*), the ApiAP2 transcription factor *ap2-l* (PF3D7_0730300), serine repeat antigen (*sera5*, PF3D7_0207600) among others (**Supplementary Table 6**) (Pei et al., 2005; Sargeant et al., 2006; Nunes, Goldring, Doerig, & Scherf, 2007; Beeson et al., 2016). These results indicate that PfAP2-G2 represses a subset of genes throughout both asexual and sexual development, and the absence of PfAP2-G2 leads to aberrant timing of gene expression.

To identify enriched DNA sequence motifs associated with gene expression changes, we analyzed the 5' upstream region (1000bp upstream of the start codon) of the genes showing significant change in transcript abundance in the asexual and sexual stages using the Finding Informative Regulatory Elements (FIRE) algorithm (Elemento, Slonim, and Tavazoie 2007). The most significant enriched sequence motif was the known DNA-interacting motif for PfAP2-G2, TGCAACCA (p-value 1.24 e-21) in genes with enhanced transcript abundance in the asexual stage (**Supplementary Fig. 6**). Interestingly, we also found an AGAACA (p-value 3.3574e-

15) DNA-binding motif that is recognized by the ApiAP2 protein, PF3D7_1139300 (**Supplementary Fig. 6**). Similar motifs were found for genes with increased transcript abundance in gametocytes (**Supplementary Fig. 6**). Intriguingly both *pfap2-g2* and *pf3d7_1139300* (*pf11_0404*) are expressed at similar times throughout asexual development based on transcriptomic data, suggesting that the two proteins may be acting together in some manner to regulate transcription (**Supplementary Fig. 7**). Based on earlier evidence, genetic disruption of the *pf3d7_1139300* orthologue in *P. berghei* (Gene ID) and *P. yoelii* (Gene ID) was refractory indicating its essentiality (Modrzynska et al., 2017; Zhang et al., 2017). Single-cell RNA-seq has shown a sharp upregulation of the gene encoding the ApiAP2 protein PF3D7_1139300 when *ap2-g* expression peaks just before egress in committed schizonts (Poran et al. 2017). Although we couldn't detect the ACCA PfAP2-G2 binding motif associated with genes showing decreased transcript abundance (**Supplementary Fig. 8**), we found enrichment of the AGACA motif, which has been associated with gametocyte commitment and development (**Supplementary Fig. 8**) (Young et al., 2005; Bischoff & Vaquero, 2010; Painter, Carrasquilla, & Llinás, 2017), in genes showing decreased transcript abundance in gametocytes.

Genes with decreased mRNA abundance in the gametocyte transcriptome from the PfAP2-G2 KO line (170 genes) overlap significantly with genes previously reported as late gametocyte markers (Van Biljon et al. 2019) (cluster 9, 208 genes) (**Figure 5B**). Some example genes are alpha tubulin 2 (PF3D7_0422300), TRAP-like protein (*tlp* PF3D7_0616500), secreted ookinete protein (*psop13* PF3D7_0518800, *psop20* PF3D7_0715400), dynein heavy chain (PF3D7_0729900), ookinete surface protein P28 (PF3D7_1031000), ookinete surface protein P25 (PF3D7_1031000) (Rawlings et al., 1992; Heiss et al., 2008); Ecker, Bushell, Tewari, & Sinden, 2008; Villard et al., 2007; Duffy & Kaslow, 1997). Therefore, one reason that PfAP2-G2 KO parasites may fail to progress beyond stage III may be due to the aberrant expression of mRNAs required during this and subsequent stages of gametocyte development. The inhibition of gametocyte maturation may also be due to the decreased expression of a few critical genes required for gametocyte development, such as male development gene 1 (*mdv1* PF3D7_1216500), gametocyte erythrocyte cytosolic

protein (*geco*, PF3D7_1253000), gametocyte exported protein (*gexp06* PF3D7_0114000), and gametocyte specific protein (*pf11-1* PF3D7_1038400) during the asexual stage. Interestingly, we also see overlap between the genes with decreased transcript abundance in both the asexual and gametocyte stages. These overlapping genes are enriched mostly for members of the *var*, *stevor*, and *rifin* variant gene families, *Pfmc-2tm* among others (**Figure 5C**). An early gametocyte proteome has revealed that exported and erythrocyte remodeling proteins are the most overrepresented proteins in gametocytes (Silvestrini et al. 2010). Although members of the PfEMP1, STEVOR and RIFIN protein repertoire are expressed in gametocyte stages, their role is not well established (Sharp et al., 2006; Petter, Bonow, & Klinkert, 2008; Tibúrcio et al., 2012, Mwakalinga et al., 2012; Tibúrcio et al., 2013; Neveu et al., 2018; Neveu & Lavazec, 2019;), although decreased abundance of these genes may contribute to the phenotypic effect.

PfAP2-G2 binding in the promoter and in gene bodies is associated with differential regulation of transcription

Out of 82 genes with upstream PfAP2-G2 peaks, only 23 displayed significantly elevated mRNA abundance (as analyzed by SAM, >1.5 log₂FC, 0.30 FDR) (**Figure 5D**). This suggests that the transcription of other genes may be impacted indirectly in the PfAP2-G2 knockout. On the other hand, none of the genes with significantly reduced mRNA abundance were bound by PfAP2-G2 in the upstream regions. However, considering the stage at which ChIP was performed we see that the mRNA abundance of many of the PfAP2-G2 ChIP-seq targets do not change in the absence of PfAP2-G2 indicating that binding of PfAP2-G2 alone might not be sufficient to affect transcription (**Figure 5E**). PfAP2-G2 binding to gene bodies results in both repression and activation in the *pfap2-g2* KO. Open reading frames bound by PfAP2-G2 in the gene-body in gametocytes showed similar effects (**Supplementary Fig. 9**). Overall these results indicate that PfAP2-G2 binding to the upstream promoter region is largely associated with repression, but PfAP2-G2 alone is not sufficient to repress transcription.

PfAP2-G2 shares occupancy with exonic epigenetic marks during asexual development

To understand the role of the pervasive genome-wide PfAP2-G2 binding in exonic gene bodies detected by ChIP-seq (**Supplementary Table 2**), we investigated whether PfAP2-G2 is associated with any known epigenetic marks. We calculated a pairwise Pearson correlation coefficient by comparing the read counts per 1000bp bins of the PfAP2-G2 ChIP-seq results against an array of histone marks for which ChIP data is available (H3K9me3, H4ac, H4K20me3, H3K26ac, H3K9ac, H3K4me3 (Karmodiya et al., 2015), H3K36me2 and H3K36me3 (Jiang et al., 2013), as well as heterochromatin protein 1 (PfHP1) (Brancucci et al., 2014)). Interestingly, we found that PfAP2-G2 occupancy has the strongest correlation with H3K36me3 ($R=0.933$), followed by H4K20me3 ($R=0.911$), H3K27ac ($R=0.826$), and H4ac ($R=0.765$) (**Figure 6A**), which are all histone marks that show moderate positive correlation with *P. falciparum* transcription and are associated with transcriptionally poised genes (Karmodiya et al. 2015). In total, 3,633 (80%) PfAP2-G2 binding sites co-occur with H3K36me3 marks (**Figure 6B, 6C**). The functional role of the association between PfAP2-G2 and H3K36me3 is, and what the potential role of PfAP2-G2 in H3K36me3 recruitment remains to be determined. PfAP2-G2 also co-occurs with PfHP1 towards the end of chromosomes although this correlation is moderate ($R=0.6$) (**Figure 6D, 6E**).

PfAP2-G2 interacts with the chromatin remodeling machinery

To identify proteins interacting with PfAP2-G2, we performed immunoprecipitations (IP) from PfAP2-G2::GFP parasites at the trophozoite stage using an anti-GFP antibody. Western blot analysis confirmed that the full length PfAP2-G2 protein was purified and eluted (**Figure 7A**), although smaller GFP-positive band(s) were also seen. The resulting IP was separated by SDS PAGE and two regions of the gel (upper and lower band, **Figure 7A**) were subjected to protein analysis by mass spectrometry. The purpose of analyzing these bands separately was to ensure that we could detect all of the proteins including the low abundance protein which otherwise would get masked when analyzing all of the bands together (**Figure 7B**).

The proteomic IP data was analyzed using SAINT (Choi et al. 2011) with a score of 0.9% and 1% FDR across all three replicates (**Supplementary Table 7**). Interestingly, we identified HP1 as an associated protein, which is consistent with the observed co-occurrence of PfAP2-G2 peaks with HP1 (**Figure 6D**) (**Figure 7C**). The upper band contained larger chromatin remodeling proteins such as ISWI, SNF2L, GCN5, FACT complex proteins (FACT-S and FACT-L), a MORC family protein, and several ApiAP2 family proteins, including PF3D7_1139300 (PF11_0404) whose motif was found to be enriched in both the transcription and ChIP-seq data (**Figure 7D**). This suggests that direct interaction of PF3D7_1139300 and PfAP2-G2 may be required for the regulation of a subset of target genes.

Discussion

Gametocytogenesis in *Plasmodium falciparum* parasites is a long 10-12 day developmental progression resulting in the formation of fertile male and female gametes that will form an oocyte only upon transmission to the mosquito host. Recent studies have shed light on the commitment and differentiation of asexual parasites into sexual parasites, which is driven by the AP2-G master regulator (Josling et al. 2020), (Kafsack et al. 2014). However, how asexual cells are reprogrammed for gametocyte commitment, and the downstream post-commitment regulation of sexual development is not well characterized. There are known examples of gametocyte stage genes such as *pfsegxp*, *mdv1*, and *pfs16* that are expressed very early on in the asexual stages suggesting that programming for gametocytogenesis may begin early in asexual development, and disruption of this program could result in profound effects post-commitment during sexual development (Furuya et al., 2005; Joice et al., 2014; Painter et al., 2017; Nixon et al., 2018).

In this study we show that the ApiAP2 protein PfAP2-G2 is essential for the development and maturation of *P. falciparum* gametocytes. PfAP2-G is expressed from the very early trophozoite stage through to stage V gametocytes. This is in contrast to *P. berghei*, where the protein is expressed from 16hpi, during the schizont stage of development (Yuda et al. 2015). Live fluorescence microscopy and nuclear fractionation show that PfAP2-G2 is localized to the nucleus of the trophozoite,

schizont, and early gametocyte stages (**Figure 1B**). In later stage gametocytes PfAP2-G2 protein levels are reduced and expressed throughout the cytoplasm (**Figure 3A**). This is consistent with the mRNA abundance profile, demonstrating peak *pfap2-g2* expression in the early asexual stages, whereas in gametocytes its expression is broad and weak (Biljon et al. 2019). The nuclear localization and strength of protein expression indicate that the major role for PfAP2-G2 is during the asexual stages and in early gametocytes.

By performing genome-wide ChIP-seq analysis at the asexual trophozoite stage and stage III gametocytes, we show that PfAP2-G2 extensively binds to both intergenic and genic regions. Genes bound by PfAP2-G2 in the upstream regions are largely expressed later in the gametocyte and mosquito stages. However, PfAP2-G2 also binds the promoters of genes expressed in the asexual ring stage and genes that are transcribed in schizont stages like *msps* encoding the merozoite surface proteins, which are regulated by PfAP2-I (Santos et al. 2017). Of course MSPs are unlikely to be required in committed gametocytes, as they no longer egress from the red blood cell for subsequent re-invasion. Interestingly, genes bound by PfAP2-G2 in stage III gametocytes are essentially the same as those bound in trophozoites. An analysis of the DNA motifs enriched in the regions bound by PfAP2-G2 identified two motifs: an AGAA motif and an ACCA motif (**Figure 4B** and **5B**). The latter resembles the PBM-based predicted motif for PfAP2-G2, but it is different from that identified in *P. berghei* (GTTGT), even though the AP2 DNA-binding domain is highly conserved (97% identical). We also found enrichment for a second motif in the PfAP2-G2-bound regions, which resembles the *in vitro* DNA sequence motif bound by the ApiAP2 protein PF3D7_1139300 (PF11_0404) (Campbell et al. 2010). Supporting the hypothesis that these ApiAP2 proteins may function together in a complex, our immunoprecipitation experiments against PfAP2-G2 GFP detected AP2 PF3D7_1139300 (PF11_0404) as an interaction partner (**Figure 7, Supplementary Table 7**).

Parasites expressing a truncated PfAP2-G2 version proliferate normally in the asexual stages. These results are surprising given the many global transcriptional (327 transcripts, >1.5 log₂FC, 0.30 FDR) changes occurring in the asexual stages, and

suggests that the parasite is tolerant to large fluctuations in gene expression. Compared with wild-type parasites, the expression levels of 90 genes (>1.5 log₂FC, 0.30 FDR) were significantly decreased during asexual development in PfAP2-G2 KO parasites, including genes that have been reported to have functional roles in gametocytes, perhaps leading to the developmental stall observed in Stage III gametocytes (Painter, Carrasquilla, & Llinás, 2017; Van Biljon et al., 2019). This suggests that the program for sexual progression is established very early, in the asexual stage. KO of PbAP2-G2 also resulted in the downregulation of gametocyte genes (Yuda et al. 2015). Genes upregulated in the PfAP2-G2 KO were mostly those normally expressed at other later developmental stages including the gametocyte and ookinete stages (**Figure 5A**). In the upstream promoter regions of genes showing increased transcript abundance in the PfAP2-G2 KO, we found an enrichment of the PfAP2-G2 ACCA motif. PfAP2-G2 therefore is involved in the repression of a number of genes that are not required for asexual development.

A previous study that examined the genome-wide binding of AP2-G2 from *P. berghei* gametocytes using ChIP-seq reported PbAP2-G2 binding at over 1,500 genes that were mostly associated with roles in asexual-stage proliferation (Yuda et al. 2015). Transcriptome analysis in gametocytes revealed the upregulation of 927 genes by more than twofold in *pbap2-g2* KO lines, suggesting that PbAP2-G2 acts as a repressor for asexual-stage genes during sexual development. However, only 397 (26.5%) of the PbAP2-G2 bound genes were upregulated in the *pbap2-g2* KO, leading Yuda *et al.* to hypothesize that relieving repression was not sufficient for the upregulation of the rest of the genes. Also, in our data, not all genes that are bound by PfAP2-G2 show enhanced transcript abundance (only 23/82 genes with upstream PfAP2-G2 peaks (28%)) displayed significantly elevated mRNA abundance. In fact, the PfAP2-G2 motif is only present in the promoter of genes upregulated in both developmental stages indicating that, as in *P. berghei*, PfAP2-G2 occupancy is not always predictive of changes in gene expression, perhaps because it requires interaction with other proteins. Once more reflecting the repressive nature of PfAP2-G2, none of the genes bound by PfAP2-G2 in the upstream region are downregulated in the KO.

The effect of PfAP2-G2 truncation in transcript abundance is more pronounced in gametocytes. Unfortunately, due to the massive dysregulation of mRNA transcripts, it is difficult to ascertain which genes cause the stall in development observed at Stage III gametocytes. The significantly upregulated genes in gametocytes are mostly thought to be important for asexual stage development, similar to what was shown in *P. berghei* (Yuda et al. 2015). Interestingly, the ApiAP2 protein AP2-L, which plays a critical role in liver-stage development, is upregulated in the gametocyte stage in both the *P. falciparum* and *P. berghei* *ap2-g2* knockouts, and parasites lacking PbAP2-G2 are unable to cause liver infection.

We observed a moderate to strong correlation between PfAP2-G2 occupancy and that of H3K9me3 ($R=0.933$), followed by H4K20me3 ($R=0.911$), H3K27ac ($R=0.826$), and H4ac ($R=0.765$) as well as PfHP1 ($R=0.6$) (**Figure 6A**). H3K36 methylation is a well-studied mark in model organisms, and carries out several roles. Studies in metazoans have shown that H3K36me3 is enriched at gene exons and plays a role in the regulation of alternative splicing (Kolasinska-Zwierz et al. 2009). Although H3K36me3 correlates with active transcription, another study showed the association of this modification with facultative heterochromatin (Chantalat et al. 2011). Thus, H3K36me3 is involved with both actively transcribed as well as silenced regions and may contribute to the formation of heterochromatin in combination with other histone modifications. At a subset of heterochromatic genes, binding to HP1a is required for enzyme-mediated H3K36me3 demethylation suggesting that there are different ways in which H3K36me3 regulates chromatin (Lin et al. 2012). In *P. falciparum*, H3K36me3 is present along the entire gene body of silent *var* genes including the transcription start site (TSS), and it is deposited by the *P. falciparum* variant-silencing methyltransferase containing a SET domain (PfSETvs) (Jiang et al. 2013). However, in that study enrichment of H3K36me3 was found at the 3' end of other ring-stage active genes besides the variable antigen *var*, *rifin* and *stevor* genes in PfSET KO parasites. This result indicates there is an alternative methyltransferase resulting in H3K36 methylation. We speculate that PfAP2-G2 plays a role in regulating regions of the genome which need to be silenced to prevent spurious transcription.

We also found that PfHP1 localized with PfAP2-G2 towards the end of the chromosomes and is associated with both sub-telomeric and chromosome-internal *var* genes and Plasmodium-exported proteins. A direct interaction between PfHP1 and PfAP2-G2 was supported by IP-MS (**Figure 7C**). Another association of PfAP2-G2 that was significant is with the MORC protein (**Figure 7D**) which has been identified to play important roles in chromatin compaction in plants and animals and was recently established as the upstream transcriptional repressor of sexual commitment in *Toxoplasma gondii* (Farhat et al. 2020). This study found that MORC forms a complex with *T. gondii* ApiAP2 transcription factors at sexual stage genes. Two of the *T. gondii* ApiAP2 proteins found to associate with MORC by IP-MS are TgAP2IV-2 and TgAP2IX-9, whose AP2 DNA-binding domains are most identical to the PfAP2-G2 AP2 domain (Farhat et al. 2020). We also found the *P. falciparum* ISWI chromatin-remodeling protein associated with PfAP2-G2, which was recently identified in a complex with the MORC protein at *var* gene promoters (Bryant et al. 2020)(**Figure 7D**).

Previous studies have shown expression of the master gametocyte regulator *pfap2-g* peaks at two times of asexual development (Poran et al. 2017). The initial expression peak in trophozoites is associated with the expression of genes encoding ISWI, SNF2L expression and the ApiAP2 encoding gene *pf3d7_1222400*. The second increase in abundance occurs just before egress in committed schizonts and leads to the expression of genes encoding LSD2, HDA1 and the AP2-G2 associated ApiAP2 protein PF3D7_1139300 (Poran et al. 2017). We propose that PF3D7_1139300 may act in a complex with PfAP2-G2 to compress chromatin structure. In summary, we propose that PfAP2-G2 plays an important role in gene repression during the maturation of the *P. falciparum* parasite sexual stages by associating with other chromatin-related proteins.

Materials & Methods:

Construction of plasmid for tagging 3' end of *pfap2-g2* and *pfap2-g2* gene disruption

The pDC2-cam-CRT-GFP (Fidock 2000) plasmid was used to tag the 3' end of the gene *pfap2-g2* with *gfp*. A 948 bp homologous fragment from the 3' end of the gene was amplified from 3D7 genomic DNA using the oligonucleotide pairs (Forward Primer P7 and Reverse Primer P8, see primer list), which also encoded restriction sites BglII and XhoI for cloning. This fragment was ligated into the BglII-and XhoI-digested pDC2-cam-CRT-GFP resulting in the final vector pDC2-cam-AP2-G2-GFP.

To create the *pfap2-g2* disrupted line, the *ap2-g2* gene locus was targeted using the selection linked integration (SLI) method (Birnbaum et al. 2017). To do this a 500 bp fragment of the 5' end of *pfap2-g2* was amplified using a forward primer with a NotI restriction site as well as a single point mutation in the homologous sequence to introduce a stop codon (Forward Primer P9) and a reverse primer with an MluI site (Reverse Primer P10) to be ligated into the final vector. This fragment was cloned (in frame) into pSLI-TGD using NotI and MluI restriction sites resulting in the final vector, pSLI-TGD-AP2-G2.

Culturing, synchronization, transfection, and cloning of *Plasmodium falciparum* parasites

Plasmodium falciparum parasites were cultured as previously described (Trager & Jensen, 1976) and maintained in 6% O₂ and 5% CO₂ and grown in sterile filtered RPMI 1640 media (Gibco) supplemented with 2 M HEPES, NaHCO₃ (2g/L), 0.1 M hypoxanthine, 0.25%(w/v) Albumax II (Gibco), and gentamycin (50 ug/ml). The parasites were maintained at a hematocrit of 3% and were kept between 2-5% parasitemia or greater based on the needs of the individual experiments. Parasite synchronization was performed using 5% sorbitol (Lambros and Vandenberg 1979).

To generate transgenic AP2-G2::GFP and AP2-G2 KO lines, transfections were carried out according to a standard protocol (Deitsch, Driskill, and Wellem's 2001). The plasmid pDC2-cam-AP2-G2-GFP was transfected into the 3D7 strain of *P. falciparum* whereas pSLI-TGD-AP2-G2 was transfected into the 3D7 clone, E5

(Rovira-Graells et al. 2012). In brief, 100 µg of plasmid DNA (maxi prepped using Qiagen kit) was preloaded into O+ RBCs at 50% hematocrit in cytomix prior to culturing with parasite-infected RBCs. After reinvasion, media containing 2.5 nM WR99210 was added to select for transgenic parasites. Stable integrants were maintained under constant drug pressure and were PCR-verified using specific primers (P1, P2, P4, P5, P6) outside the homology region after the isolation of gDNA (DNeasy Blood and Tissue kit, Qiagen). The resulting PfAP2G2::GFP and *pfap2-g2* ko parasites were cloned by limiting dilution (Rosario 2008) and verified by diagnostic PCR (Supplementary Fig. 1 & 2A) and whole genome sequencing. Sequencing analysis was done by creating a reference genome with GFP and rest of the plasmid at the expected location and reads were aligned to it.

***Plasmodium falciparum* gametocyte production**

Gametocytes were generated using the nutrient deprivation induction method as described (Miao et al. 2013). Briefly, synchronized trophozoites were cultured at ~2% parasitemia using 50% fresh RBC at 4% hematocrit. When the parasitemia reached 7–10%, gametocyte production was induced by nutrient deprivation. The following day, stressed schizonts were split 50% into two flasks by adding fresh blood and fresh media. The following day, ring-stage committed gametocytes were counted as Day 1 of gametocytogenesis. To prevent the reinvasion and propagation of asexual parasites, parasites were treated with media with heparin (20 U/ml) from Day 1 post induction through Day 4. The media was changed daily to ensure the proper development and growth of the gametocytes.

Nuclear and cytoplasmic protein extraction from parasites and western blotting

Nuclear and cytoplasmic extracts from parasites were prepared as previously described (Voss et al. 2002) using 5–7% of the synchronized parasites. The proteins from the nuclear and cytoplasmic fractions were run on a gel (mini PROTEAN precast TGX gels) and then transferred to a nitrocellulose membrane. The membrane was then blocked using 5% non-fat milk powder in 1X PBS-0.05% Tween for 45 minutes

at room temperature. The membrane was then incubated overnight at 4°C with an anti-GFP (ROCHE Anti-GFP, from mouse IgG1K, Millipore Sigma) (1:1000), anti-histone (Anti-Histone H3 antibody, AbCam) (1:3000), anti-aldolase (Anti-Plasmodium aldolase antibody (HRP), AbCam) (1:1000) primary antibody in the blocking buffer. After incubation, the membrane was washed three times with 1X PBS-0.05% Tween and was incubated with the appropriate secondary antibodies, then washed. The secondary antibodies used were as follows: peroxidase-conjugated goat anti-rat HRP conjugate (Millipore) / goat anti-rabbit HRP conjugate (Millipore) / goat anti-mouse HRP conjugate (Pierce) was used (1:3000). Bound antibodies were then visualized on a film after enhancing the signal with ECL (Pierce) as the substrate.

Parasite growth assay

To compare the growth rates of *pfap2-g2* KO and WT parasites we used a SYBR green growth assay (Vossen et al., 2010). Parasites were sampled every 24 hours for 10 days, starting at 0.1% parasitemia at the trophozoite stage. Both the WT and PfAP2-G2 knockout strains were synchronized using 5% sorbitol (described above) 3x before the start of the timecourse. Parasites were seeded in a 25 cm² culture flask at 0.1% parasitemia and 3% hematocrit. To evaluate growth differences every 24 hours for 10 days, 100 µl of parasite culture was collected in triplicate, transferred to a 96-well plate, and stored at -80°C. Following the completion of the timecourse, the parasites were thawed at room temperature, 100 µl of SYBR green I (Molecular Probes, Eugene Oregon) was diluted into parasite lysis buffer (0.2 µl of SYBR Green I/ml of lysis buffer) and added to each well and mixed. The plates were incubated at 37°C for 2 hours. Cell growth was measured as quantification of DNA by SYBR green I using a Tecan GENios microplate detection device at the excitation wavelength and emission wavelength of 485 nm and 535 nm, respectively. All values were plotted as an average of three technical replicates (±S.D.) using GraphPad Prism (version 7).

RNA purification and cDNA synthesis for DNA microarrays

Total RNA was extracted from tightly synchronized PfAP2-G2::KO and WT parasites every 6 hours, starting at 3 hpi, throughout the 48 hour asexual life cycle. A

separate timecourse was designed to collect RNA from gametocytes starting at Day 1 and every 12 hours for the following 7 days. RNA was extracted from parasite-infected RBCs collected at each timepoint using TRIzol (ThermoFisher Scientific), following the manufacturer's protocol. cDNA synthesis and DNA microarray analysis was carried out as previously described (Painter et al. 2013). Two-channel Agilent DNA microarrays (AMADID 037237) were hybridized, washed, and scanned on an Axon 4200A scanner. Signal intensities for each gene were extracted from the scanned image using Agilent Feature Extraction Software version 9.5. Detailed microarray protocols can be found at <http://linaslab.psu.edu/protocols/>. The data were analyzed by Significance Analysis of Microarrays (SAM) (Tusher et al. 2001) for the significantly ($>1.5 \log_2FC$, 0.30 FDR) changing genes between the WT and AP2-G2 KO lines using two-class paired t-test throughout the time course. All heat maps were clustered using Cluster 4.0 and was visualized using Java Tree View.

Chromatin immunoprecipitation and Library preparation for sequencing (ChIP-Seq)

ChIP was performed as described (Josling et al. 2020) using synchronized trophozoite and gametocyte stage PfA2-G2-GFP parasites (around 7% parasitemia). Extracted chromatin was sheared in SDS lysis buffer (200 μ l 1×10^9 trophozoites and 5×10^8 gametocytes) to obtain a fragment size of 100-150bp using an M220 focused-ultrasonicator (Covaris Inc.) using the following settings: peak power 75W, 2% duty factor, 200 cycles per burst and total treatment time of 600 s. Sonication time was optimized for PfAP2-G2 using the crosslinked DNA and shearing it for 0, 4, 8, 10, 12 and 15 minutes and then running it on agarose gel for the size estimate. For immunoprecipitations, sheared chromatin was pre-cleared using 20 μ l/ml of magnetic beads A/G (Millipore 16-663) for 2 hours at 4°C with gentle agitation. The supernatant was collected and 50 μ l of the aliquot was removed as input. The rest of the chromatin was incubated overnight at 4°C with 1 μ g of polyclonal anti-GFP antibody (Abcam ChIP grade, Abcam 290) or, as control, the same amount of IgG. The immunoprecipitated antibody/chromatin complex was collected using Protein A/G magnetic beads (Millipore 16-663). The input and ChIP samples were reverse cross-

linked overnight at 45°C using 0.2M final concentration of NaCl. The samples were then treated with RNAase (30min at 37°C) and proteinase K (3 µL of 20 mg/mL with 2 hours incubation at 45°C) and purified by using the QIAGEN MinElute PCR purification kit. Library was prepared as described earlier (Santos et al. 2017). The final library was quantified using a Qubit fluorometer HD DNA kit and analyzed using an Agilent DNA 1000 Bioanalyzer to assess the quality, size distribution, and detection of any artifacts. Sequencing was performed using an Illumina Hiseq 2500 to obtain 150 bp single-end reads.

ChIP-seq data analysis

ChIP-seq data analysis was done using tools in Galaxy (usegalaxy.org). Before starting the analysis, the quality of the sequencing reads were assessed using FastQC and were trimmed to remove adapter and low quality bases using Trimmomatic (Bolger, Lohse, and Usadel 2014). The resulting reads were mapped to the *P. falciparum* genome (Pf 3D7 v28, obtained from PlasmoDB) using BWA-MEM (Li and Durbin 2009) and duplicates reads were removed using SAMtools (Li et al. 2009). Mapped sequences were converted into bigwig files using bamCoverage (Ramírez et al. 2016) and were viewed in using the Integrative Genomics Viewer (IGV) (Thorvaldsdóttir, Robinson, and Mesirov 2013) for each input and treatment. MACS2 (Feng et al. 2012) was used to call peaks with q-value cutoff of 0.05. The common overlapping intervals between replicates were established using Multiple Intersect function of BEDtools, and the overlapping intervals were combined into a single file using MergeBED (Quinlan and Hall 2010). The genes closest to the peak summits were identified using closestBED. Peaks that were farther than 2.5Kb were removed from the analysis. When the peak was present between two genes then only the gene closest to the peak was considered for the analysis.

Peak sequences were extracted using extract genomic DNA tool using genomic coordinates, and motifs associated with called peaks were identified using DREME (Bailey 2011) and compared to identified ApiAP2 motifs using TOMTOM (Gupta et al. 2007). The correlation heatmap of the replicates performed on GFP tagged PfAP2-G2

ChIP-seq was made using multiBigwigSummary and plotCorrelation in the deepTools suite (Ramírez et al. 2016).

Correlation analysis of PfAP2-G2 with other histone marks

Correlations were calculated between the occupancy of PfAP2-G2 and that of various histone marks previously reported in the literature, including H3K9me3, H4ac, H4K20me3, H3K26ac, H3K9ac, H3K4me3 (Karmodiya et al. 2015), H3K36me2 and H3K36me3 (Jiang et al. 2013) and heterochromatin protein 1 (Nicolas M B Brancucci et al. 2014). First, read counts were obtained from PfAP2-G2 and nine histone marks in the gene body of all *P. falciparum* genes (n=5735; PlasmoDB annotations, release 28.0). NCIS (Liang and Keleş 2012) was used to scale a control input experiment to each analyzed signal ChIP-seq experiment. Pairwise Pearson correlation coefficient of normalized read counts across 5735 regions for PfAP2-G2 and nine histone marks. Then, proteins were ordered using complete-linkage hierarchical clustering with Euclidean distance.

PfAP2-G2-GFP Immunoprecipitation and Mass Spectrometry

Nuclear extraction was performed from the PfAP2-G2::GFP parasites as previously described (Voss et al. 2002). A day before the immunoprecipitation ~1mg of M-270 Epoxy Dynabeads (Invitrogen) per 100 ml of culture were conjugated with polyclonal anti-GFP antibodies at concentration of 5ug/mg (Abcam ab290) or with IgG (Abcam ab46540) overnight at 30°C as previously described (Joshi et al. 2013). Immunoprecipitation was carried out after washing the conjugated beads with dilution buffer and incubating with diluted nuclear extracts for 90 minutes at 4°C. The beads were washed extensively with dilution buffer and 1X PBS and protein was eluted using 20 µl loading buffer by shaking for 10 minutes at 70°C at 1000 rpm and protein was stored at -20°C. Samples were quality controlled by western blotting and subjected to in-gel tryptic digest.

For tryptic digest the samples in polyacrylamide gel slices were destained with 50mM AmBic/50% ACN (Acetonitrile) and dehydrated using 100% ACN. Disulfide bonds in protein were reduced with 10mM DTT and reduced cysteine residues were

then alkylated with 50mM Iodoacetamide. Finally, the processed samples were tryptic digested (Trypsin gold, 6ng/μl) overnight at 37°C on a thermomixer, and peptides were extracted from the gels. The dried peptide pellets were analyzed by LC-MS/MS using Q Exactive mass spectrometer (Indiana University Core Proteomics).

The data was processed using Trans-Proteomic Pipeline (TPP) as previously described with a few modifications (Deutsch et al. 2010). Spectra were searched against the *P. falciparum* proteome and contaminant repository database (CRAPome) using tandem and comet searches (Mellacheruvu et al. 2013). This search was combined in InterProphet and the proteins were searched with ProteinProphet. Only proteins with error rate less than 0.01 were reported. The immunopurification assay was performed 3 times and the data was combined to identify the significantly enriched proteins between the IP and control using SAINT (Choi et al. 2011). Proteins with probability score of 0.99 were considered for analysis.

Additional supplementary materials

Supplementary Table 1: PfAP2-G2::GFP ChIP-seq data in trophozoite stage First tab of the table contains all the upstream peaks common in two out of three biological replicates. Second tab has only 86 genes that were considered for the analysis. Indicated are coordinates of each peak, the orientation of the gene, genes downstream to the peaks and distance of peak and ATG of the gene.

Supplementary Table 2: PfAP2-G2::GFP ChIP-seq data in trophozoite stage Table contains peaks on the gene-body common in two out of three biological replicates. Also, indicated are coordinates of each peak and genes containing the peaks. .

Supplementary Table 3: PfAP2-G2::GFP ChIP-seq data in gametocyte stage III Table contains all the upstream peaks common in two biological replicates. Indicated are coordinates of each peak, the orientation of the gene, genes downstream to the peaks and distance of peak and ATG of the gene.

Supplementary Table 4: PfAP2-G2::GFP ChIP-seq data in gametocyte stage III

Table contains peaks on the gene-body common in two biological replicates. Also, indicated are coordinates of each peak and genes containing the peaks.

Supplementary Table 5: List of the differentially expressed genes in asexual blood stage

Significance analysis of microarray (SAM) of the asexual stage timecourse data showed that 327 transcripts were differentially expressed (>1.5 log₂FC, 0.30 FDR) in the PfAP2-G2 KO line compared to WT. Out of these, 237 showed enhanced transcript abundance in the PfAP2G2 KO line and 90 showed decreased abundance. The first tab of the sheet has upregulated genes and the second tab has downregulated genes along with fold change, q-value and local false discover rate.

Supplementary Table 6: List of the differentially expressed genes in gametocytes

Significance analysis of microarray (SAM) of the gametocyte timecourse data showed that 444 transcripts were differentially expressed (>1.5 log₂FC, 0.30 FDR) in the PfAP2-G2 KO line compared to WT. Out of these, 274 showed enhanced transcript abundance in the PfAP2G2 KO line and 170 showed decreased abundance. The first tab of the sheet has upregulated genes and the second tab has downregulated genes along with fold change, q value and local false discover rate.

Supplementary Table 7: List of PfAP2-G2 interacting partners

First tab contains the list of peptides obtained in the in the upper band of PfAP2- G2 and the second tab has the list of peptides obtained in the in the lower band of PfAP2- G2 along with the spectral counts, average and maximum probability and FDR. Only the genes with high probability of interaction measured by SAINT (pSAINT 0.9 and 1%FDR) was considered for the analysis.

Acknowledgements

This project was largely funded by NIH/NIAID 1R01 AI125565 (ML). NY and SM were supported by R01 GM121613 (SM). We thank Kelly Rios for assistance with proteomic data analysis.

References

- Aguilar, Ruth, Ariel Magallon-Tejada, Ariel H. Achtman, Cinta Moraleda, Regina Joice, Pau Cisteró, Connie S. N. Li Wai Suen, Augusto Nhabomba, Eusebio Macete, Ivo Mueller, Matthias Marti, Pedro L. Alonso, Clara Menéndez, Louis Schofield, and Alfredo Mayor. 2014. "Molecular Evidence for the Localization of *Plasmodium falciparum* Immature Gametocytes in Bone Marrow." *Blood* 123(7):959–66.
- Alano, Pietro, Sunil Premawansa, Marian C. Bruce, and Richard Carter. 1991. "A Stage Specific Gene Expressed at the Onset of Gametocytogenesis in *Plasmodium falciparum*." *Molecular and Biochemical Parasitology* 46(1):81–88.
- Armistead, Jennifer S., Roberto R. Moraes Barros, Tyler J. Gibson, Whitney A. Kite, J. Patrick Mershon, Lynn E. Lambert, Sachy E. Orr-Gonzalez, Juliana M. Sá, John H. Adams, and Thomas E. Wellems. 2018. "Infection of Mosquitoes from in Vitro Cultivated *Plasmodium knowlesi* H Strain." *International Journal for Parasitology* 48(8):601–10.
- Bailey, Timothy L. 2011. "DREME: Motif Discovery in Transcription Factor ChIP-Seq Data." *Bioinformatics* 27(12):1653–59.
- Balaji, S., M. Madan Babu, Lakshminarayan M. Iyer, and L. Aravind. 2005. "Discovery of the Principal Specific Transcription Factors of Apicomplexa and Their Implication for the Evolution of the AP2-Integrase DNA Binding Domains." *Nucleic Acids Research* 33(13):3994–4006.
- Beeson, James G., Damien R. Drew, Michelle J. Boyle, Gaoqian Feng, Freya J. I. Fowkes, and Jack S. Richards. 2016. "Merozoite Surface Proteins in Red Blood Cell Invasion, Immunity and Vaccines against Malaria." *FEMS Microbiology Reviews* 40(3):343–72.
- Van Biljon, Riëtte, Roelof Van Wyk, Heather J. Painter, Lindsey Orchard, Janette Reader, Jandeli Niemand, Manuel Llinás, and Lyn Marie Birkholtz. 2019. "Hierarchical Transcriptional Control Regulates *Plasmodium falciparum* Sexual Differentiation." *BMC Genomics* 20(1):1–16.
- Biljon, Riette van, Roelof van Wyk, Heather Painter, Lindsey Orchard, Janette Reader, Jandeli Niemand, Manuel Llinas, and Lyn-Marie Birkholtz. 2019. "Hierarchical Transcriptional Control Regulates *Plasmodium falciparum* Sexual

Differentiation.” *BioRxiv* 633222.

- Birnbaum, Jakob, Sven Flemming, Nick Reichard, Alexandra Blancke Soares, Paolo Mesén-Ramírez, Ernst Jonscher, Bärbel Bergmann, and Tobias Spielmann. 2017. “A Genetic System to Study *Plasmodium falciparum* Protein Function.” *Nature Methods* 14(4):450–56.
- Bischoff, Emmanuel and Catherine Vaquero. 2010. “In Silico and Biological Survey of Transcription-Associated Proteins Implicated in the Transcriptional Machinery during the Erythrocytic Development of *Plasmodium falciparum*.” *BMC Genomics* 11(1).
- Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30(15):2114–20.
- Bozdech, Zbynek, Manuel Llinás, Brian Lee Pulliam, Edith D. Wong, Jingchun Zhu, and Joseph L. DeRisi. 2003. “The Transcriptome of the Intraerythrocytic Developmental Cycle of *Plasmodium falciparum*.” *PLoS Biology* 1(1):85–100.
- Brancucci, Nicolas M.B., Nicole L. Bertschi, Lei Zhu, Igor Niederwieser, Wai Hoe Chin, Rahel Wampfler, Céline Freymond, Matthias Rottmann, Ingrid Felger, Zbynek Bozdech, and Till S. Voss. 2014. “Heterochromatin Protein 1 Secures Survival and Transmission of Malaria Parasites.” *Cell Host and Microbe* 16(2):165–76.
- Brancucci, Nicolas M B, Nicole L. Bertschi, Lei Zhu, Igor Niederwieser, Wai Hoe Chin, Rahel Wampfler, Matthias Rottmann, Ingrid Felger, Zbynek Bozdech, and Till S. Voss. 2014. “Article Heterochromatin Protein 1 Secures Survival and Transmission of Malaria Parasites.” 165–76.
- Bruce, M. C., P. Alano, S. Duthie, and R. Carter. 1989. “Commitment of the Malaria Parasite *Plasmodium falciparum* to Sexual and Asexual Development.” *Parasitology* 100(2):191–200.
- Bruce, Marian C., Roderick N. Carter, Kei ichiro Nakamura, Masamichi Aikawa, and Richard Carter. 1994. “Cellular Location and Temporal Expression of the *Plasmodium falciparum* Sexual Stage Antigen Pfs16.” *Molecular and Biochemical Parasitology* 65(1):11–22.
- Bryant, Jessica M., Sebastian Baumgarten, Florent Dingli, Damaris Loew, Ameya Sinha, Aurélie Claës, Peter R. Preiser, Peter C. Dedon, and Artur Scherf. 2020. “

- Exploring the Virulence Gene Interactome with CRISPR / DC As9 in the Human Malaria Parasite ." *Molecular Systems Biology* 16(8):1–22.
- Campbell, Tracey L., Erandi K. de Silva, Kellen L. Olszewski, Olivier Elemento, and Manuel Llinás. 2010. "Identification and Genome-Wide Prediction of DNA Binding Specificities for the ApiAP2 Family of Regulators from the Malaria Parasite." *PLoS Pathogens* 6(10).
- Chantalat, Sophie, Arnaud Depaux, Patrick Héry, Sophie Barral, Jean Yves Thuret, Stefan Dimitrov, and Matthieu Gérard. 2011. "Histone H3 Trimethylation at Lysine 36 Is Associated with Constitutive and Facultative Heterochromatin." *Genome Research* 21(9):1426–37.
- Chattopadhyay, Rana, Dharmendar Rathore, Hishasi Fujioka, Sanjai Kumar, Patricia De la Vega, David Haynes, Kathleen Moch, David Fryauff, Ruobing Wang, Daniel J. Carucci, and Stephen L. Hoffman. 2003. "PfSPATR, a *Plasmodium falciparum* Protein Containing an Altered Thrombospondin Type I Repeat Domain Is Expressed at Several Stages of the Parasite Life Cycle and Is the Target of Inhibitory Antibodies." *Journal of Biological Chemistry* 278(28):25977–81.
- Choi, Hyungwon, Brett Larsen, Zhen Yuan Lin, Ashton Breitzkreutz, Dattatreya Mellacheruvu, Damian Fermin, Zhaohui S. Qin, Mike Tyers, Anne Claude Gingras, and Alexey I. Nesvizhskii. 2011. "SAINT: Probabilistic Scoring of Affinity Purification-Mass Spectrometry Data." *Nature Methods* 8(1):70–73.
- Deitsch, Kirk W., Casey L. Driskill, and Thomas E. Wellems. 2001. "Transformation of Malaria Parasites by the Spontaneous Uptake and Expression of DNA from Human Erythrocytes." *Nucleic Acids Research* 29(3):850–53.
- Derbyshire, Emily. 2019. "RNA-Seq Analysis Illuminates the Early Stages of Plasmodium Liver Infection."
- Deutsch, Eric W., Luis Mendoza, David Shteynberg, Terry Farrah, Henry Lam, Natalie Tasman, Zhi Sun, Erik Nilsson, Brian Pratt, Bryan Prazen, Jimmy K. Eng, Daniel B. Martin, Alexey I. Nesvizhskii, and Ruedi Aebersold. 2010. "A Guided Tour of the Trans-Proteomic Pipeline." *Proteomics* 10(6):1150–59.
- van Dijk, Melissa R., Ben C. L. van Schaijk, Shahid M. Khan, Maaike W. van Dooren, Jai Ramesar, Szymon Kaczanowski, Geert Jan van van Gemert, Hans Kroeze,

- Hendrik G. Stunnenberg, Wijnand M. Eling, Robert W. Sauerwein, Andrew P. Waters, and Chris J. Janse. 2010. "Three Members of the 6-Cys Protein Family of Plasmodium Play a Role in Gamete Fertility." *PLoS Pathogens* 6(4):1–13.
- Duffy, P. E. and D. C. Kaslow. 1997. "A Novel Malaria Protein, Pfs28, and Pfs25 Are Genetically Linked and Synergistic as Falciparum Malaria Transmission-Blocking Vaccines." *Infection and Immunity* 65(3):1109–13.
- Ecker, Andrea, Ellen S. C. Bushell, Rita Tewari, and Robert E. Sinden. 2008. "Reverse Genetics Screen Identifies Six Proteins Important for Malaria Development in the Mosquito." *Molecular Microbiology* 70(1):209–20.
- Elemento, Olivier, Noam Slonim, and Saeed Tavazoie. 2007. "A Universal Framework for Regulatory Element Discovery across All Genomes and Data Types." *Molecular Cell* 28(2):337–50.
- Espinosa, Diego A., Joel Vega-Rodriguez, Yewel Flores-Garcia, Amy R. Noe, Christian Muñoz, Russell Coleman, Torben Bruck, Keith Haney, Alex Stevens, Diane Retallack, Jeff Allen, Thomas S. Vedvick, Christopher B. Fox, Steven G. Reed, Randall F. Howard, Ahmed M. Salman, Chris J. Janse, Shahid M. Khan, Fidel Zavala, and Gabriel M. Gutierrez. 2017. "The *Plasmodium falciparum* Cell-Traversal Protein for Ookinetes and Sporozoites as a Candidate for Preerythrocytic and Transmission-Blocking Vaccines." *Infection and Immunity* 85(2):1–10.
- Farhat, Dayana C., Christopher Swale, Céline Dard, Dominique Cannella, Philippe Ortet, Mohamed Barakat, Fabien Sindikubwabo, Lucid Belmudes, Pieter Jan De Bock, Yohann Couté, Alexandre Bougdour, and Mohamed Ali Hakimi. 2020. "A MORC-Driven Transcriptional Switch Controls Toxoplasma Developmental Trajectories and Sexual Commitment." *Nature Microbiology* 5(4):570–83.
- Feng, Jianxing, Tao Liu, Bo Qin, Yong Zhang, and Xiaole Shirley Liu. 2012. "Identifying ChIP-Seq Enrichment Using MACS." *Nature Protocols* 7(9):1728–40.
- Fidock, D. 2000. "Mutations in the *P. falciparum* Digestive Vacuole Transmembrane Protein PfCRT and Evidence for Their Role in Chloroquine Resistance." *Molecular Cell* 6(4):861–71.
- Flueck, Christian, Richard Bartfai, Igor Niederwieser, Kathrin Witmer, Blaise T. F.

- Alako, Suzette Moes, Zbynek Bozdech, Paul Jenoe, Hendrik G. Stunnenberg, and Till S. Voss. 2010. "A Major Role for the *Plasmodium falciparum* ApiAP2 Protein PfSIP2 in Chromosome End Biology." *PLoS Pathogens* 6(2).
- Furuya, Tetsuya, Jianbing Mu, Karen Hayton, Anna Liu, Junhui Duan, Louis Nkrumah, Deirdre A. Joy, David A. Fidock, Hisashi Fujioka, Akhil B. Vaidya, Thomas E. Wellems, and Xin Zhuan Su. 2005. "Disruption of a *Plasmodium falciparum* Gene Linked to Male Sexual Development Causes Early Arrest in Gametocytogenesis." *Proceedings of the National Academy of Sciences of the United States of America* 102(46):16813–18.
- Gratraud, Paul, Enlli Huws, Brie Falkard, Sophie Adjalley, David A. Fidock, Laurence Berry, William R. Jacobs, Mark S. Baird, Henri Vial, and Laurent Kremer. 2009. "Oleic Acid Biosynthesis in *Plasmodium falciparum*: Characterization of the Stearoyl-CoA Desaturase and Investigation as a Potential Therapeutic Target." *PLoS ONE* 4(9).
- Gupta, Shobhit, John A. Stamatoyannopoulos, Timothy L. Bailey, and William Stafford Noble. 2007. "Quantifying Similarity between Motifs." *Genome Biology* 8(2).
- Heiss, Kirsten, Hui Nie, Sumit Kumar, Thomas M. Daly, Lawrence W. Bergman, and Kai Matuschewski. 2008. "Functional Characterization of a Redundant Plasmodium TRAP Family Invasin, TRAP-like Protein, by Aldolase Binding and a Genetic Complementation Test." *Eukaryotic Cell* 7(6):1062–70.
- Ikadai, Hiromi, Kathryn Shaw Saliba, Stefan M. Kanzok, Kyle J. McLean, Takeshi Q. Tanaka, Jun Cao, Kim C. Williamson, and Marcelo Jacobs-Lorena. 2013. "Transposon Mutagenesis Identifies Genes Essential for *Plasmodium falciparum* Gametocytogenesis." *Proceedings of the National Academy of Sciences of the United States of America* 110(18).
- Ishino, Tomoko, Kazuhiko Yano, Yasuo Chinzei, and Masao Yuda. 2004. "Cell-Passage Activity Is Required for the Malarial Parasite to Cross the Liver Sinusoidal Cell Layer." *PLoS Biology* 2(1):77–84.
- Itsara, Leslie S., Yaxian Zhou, Julie Do, Samrita Dungel, Matthew E. Fishbaugher, Will W. Betz, Thao Nguyen, Mary Jane Navarro, Erika L. Flannery, Ashley M.

- Vaughan, Stefan H. I. Kappe, and Anil K. Ghosh. 2018. "PfCap380 as a Marker for *Plasmodium falciparum* Oocyst Development in Vivo and in Vitro." *Malaria Journal* 17(1):1–13.
- Jeninga, Myriam, Jennifer Quinn, and Michaela Petter. 2019. "ApiAP2 Transcription Factors in Apicomplexan Parasites." *Pathogens* 8(2):47.
- Jiang, Lubin, Jianbing Mu, Qingfeng Zhang, Ting Ni, Prakash Srinivasan, Kempaiah Rayavara, Wenjing Yang, Louise Turner, Thomas Lavstsen, Thor G. Theander, Weiqun Peng, Guiying Wei, Qingqing Jing, Yoshiyuki Wakabayashi, Abhisheka Bansal, Yan Luo, José M. C. Ribeiro, Artur Scherf, L. Aravind, Jun Zhu, Keji Zhao, and Louis H. Miller. 2013. "PfSETvs Methylation of Histone H3K36 Represses Virulence Genes in *Plasmodium falciparum*." *Nature* 499(7457):223–27.
- Joice, Regina, Sandra K. Nilsson, Jacqui Montgomery, Selasi Dankwa, Elizabeth Egan, Belinda Morahan, Karl B. Seydel, Lucia Bertuccini, Pietro Alano, Kim C. Williamson, Manoj T. Duraisingh, Terrie E. Taylor, Danny A. Milner, and Matthias Marti. 2014. "*Plasmodium falciparum* Transmission Stages Accumulate in the Human Bone Marrow." *Science Translational Medicine* 6(244):1–9.
- Joshi, Preeti, Todd M. Greco, Amanda J. Guise, Yang Luo, Fang Yu, Alexey I. Nesvizhskii, and Ileana M. Cristea. 2013. "The Functional Interactome Landscape of the Human Histone Deacetylase Family." *Molecular Systems Biology* 9(1):1–21.
- Josling, Gabrielle A., Timothy J. Russell, Jarrett Venezia, Lindsey Orchard, Riëtte van Biljon, Heather J. Painter, and Manuel Llinás. 2020. "Dissecting the Role of PfAP2-G in Malaria Gametocytogenesis." *Nature Communications* 11(1):1503.
- Kafsack, Björn F. C., Núria Rovira-Graells, Taane G. Clark, Cristina Bancells, Valerie M. Crowley, Susana G. Campino, April E. Williams, Laura G. Drought, Dominic P. Kwiatkowski, David A. Baker, Alfred Cortés, and Manuel Llinás. 2014. "A Transcriptional Switch Underlies Commitment to Sexual Development in Malaria Parasites." *Nature* 507(7491):248–52.
- Karmodiya, Krishanpal, Saurabh J. Pradhan, Bhagyashree Joshi, Rahul Jangid, Puli Chandramouli Reddy, and Sanjeev Galande. 2015. "A Comprehensive Epigenome Map of *Plasmodium falciparum* Reveals Unique Mechanisms of

- Transcriptional Regulation and Identifies H3K36me2 as a Global Mark of Gene Suppression." *Epigenetics and Chromatin* 8(1):1–18.
- Kolasinska-Zwierz, Paulina, Thomas Down, Isabel Latorre, Tao Liu, X. Shirley Liu, and Julie Ahringer. 2009. "Differential Chromatin Marking of Introns and Expressed Exons by H3K36me3." *Nature Genetics* 41(3):376–81.
- Lambros, Chris and Jerome P. Vandenberg. 1979. "Synchronization of *Plasmodium falciparum* Erythrocytic Stages in Culture Author (s): Chris Lambros and Jerome P. Vanderberg Published by : Allen Press on Behalf of The American Society of Parasitologists Stable URL : <https://www.jstor.org/stable/3280287>." *Journal Parasitology* 65(3):418–20.
- Li, Heng and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics* 25(14):1754–60.
- Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25(16):2078–79.
- Liang, Kun and Sündüz Keleş. 2012. "Normalization of ChIP-Seq Data with Control." *BMC Bioinformatics* 13(1).
- Lin, Chia Hui, Ariel Paulson, Susan M. Abmayr, and Jerry L. Workman. 2012. "HP1a Targets the Drosophila KDM4A Demethylase to a Subset of Heterochromatic Genes to Regulate H3K36me3 Levels." *PLoS ONE* 7(6).
- Liu, Zhenyu, Jun Miao, and Liwang Cui. 2011. "Gametocytogenesis in Malaria Parasite: Commitment, Development and Regulation." *Future Microbiology* 6(11):1351–69.
- Llorà-Batlle, Oriol, Lucas Michel-Todó, Kathrin Witmer, Haruka Toda, Carmen Fernández-Becerra, Jake Baum, and Alfred Cortés. 2020. "Conditional Expression of PfAP2-G for Controlled Massive Sexual Conversion in *Plasmodium falciparum*." *Science Advances* 6(24).
- López-Barragán, María J., Jacob Lemieux, Mariam Quiñones, Kim C. Williamson, Alvaro Molina-Cruz, Kairong Cui, Carolina Barillas-Mury, Keji Zhao, and Xinzhuan Su. 2011. "Directional Gene Expression and Antisense Transcripts in Sexual and Asexual Stages of *Plasmodium falciparum*." *BMC Genomics* 12.

- Mellacheruvu, Dattatreya, Zachary Wright, Amber L. Couzens, Jean Philippe Lambert, Nicole A. St-Denis, Tuo Li, Yana V. Miteva, Simon Hauri, Mihaela E. Sardu, Teck Yew Low, Vincentius A. Halim, Richard D. Bagshaw, Nina C. Hubner, Abdallah Al-Hakim, Annie Bouchard, Denis Faubert, Damian Fermin, Wade H. Dunham, Marilyn Goudreault, Zhen Yuan Lin, Beatriz Gonzalez Badillo, Tony Pawson, Daniel Durocher, Benoit Coulombe, Ruedi Aebersold, Giulio Superti-Furga, Jacques Colinge, Albert J. R. Heck, Hyungwon Choi, Matthias Gstaiger, Shabaz Mohammed, Ileana M. Cristea, Keiryn L. Bennett, Mike P. Washburn, Brian Raught, Rob M. Ewing, Anne Claude Gingras, and Alexey I. Nesvizhskii. 2013. "The CRAPome: A Contaminant Repository for Affinity Purification-Mass Spectrometry Data." *Nature Methods* 10(8):730–36.
- Miao, Jun, Zenglei Wang, Min Liu, Daniel Parker, Xiaolian Li, Xiaoguang Chen, and Liwang Cui. 2013. "*Plasmodium falciparum*: Generation of Pure Gametocyte Culture by Heparin Treatment." *Experimental Parasitology* 135(3):541–45.
- Miller, Susanne K., Robert T. Good, Damien R. Drew, Mauro Delorenzi, Paul R. Sanders, Anthony N. Hodder, Terence P. Speed, Alan F. Cowman, Tania F. De Koning-Ward, and Brendan S. Crabb. 2002. "A Subset of *Plasmodium falciparum* SERA Genes Are Expressed and Appear to Play an Important Role in the Erythrocytic Cycle." *Journal of Biological Chemistry* 277(49):47524–32.
- Modrzynska, Katarzyna, Claudia Pfander, Lia Chappell, Lu Yu, Catherine Suarez, Kirsten Dundas, Ana Rita Gomes, David Goulding, Julian C. Rayner, Jyoti Choudhary, and Oliver Billker. 2017. "A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle." *Cell Host and Microbe* 21(1):11–22.
- Neveu, Gaëlle and Catherine Lavazec. 2019. "Erythrocyte Membrane Makeover by *Plasmodium falciparum* Gametocytes." *Frontiers in Microbiology* 10(November):1–8.
- Nixon, Christian P., Christina E. Nixon, Ian C. Michelow, Rayna A. Silva-Viera, Bonnie Colantuono, Aisha S. Obeidallah, Ambrish Jha, Dominique Dockery, Dipak Raj, Sangshin Park, Patrick E. Duffy, and Jonathan D. Kurtis. 2018. "Antibodies to PfsEGXP, an Early Gametocyte-Enriched Phosphoprotein, Predict Decreased

- Plasmodium falciparum* Gametocyte Density in Humans.” *Journal of Infectious Diseases* 218(11):1792–1801.
- Nunes, Marta C., J. P. Dea. Goldring, Christian Doerig, and Artur Scherf. 2007. “A Novel Protein Kinase Family in *Plasmodium falciparum* Is Differentially Transcribed and Secreted to Various Cellular Compartments of the Host Cell.” *Molecular Microbiology* 63(2):391–403.
- Oehring, Sophie C., Ben J. Woodcroft, Suzette Moes, Johanna Wetzel, Olivier Dietz, Andreas Pulfer, Chaitali Dekiwadia, Pascal Maeser, Christian Flueck, Kathrin Witmer, Nicolas M. B. Brancucci, Igor Niederwieser, Paul Jenoe, Stuart A. Ralph, and Till S. Voss. 2012. “Organellar Proteomics Reveals Hundreds of Novel Nuclear Proteins in the Malaria Parasite *Plasmodium falciparum*.” *Genome Biology* 13(11).
- Painter, Heather J., Lindsey M. Altenhofen, Björn F. C. Kafsack, and Manuel Llinás. 2013. Whole-Genome Analysis of *Plasmodium spp.* Utilizing a New Agilent Technologies DNA Microarray Platform. Vol. 923.
- Painter, Heather J., Manuela Carrasquilla, and Manuel Llinás. 2017. “Capturing in Vivo RNA Transcriptional Dynamics from the Malaria Parasite *Plasmodium falciparum*.” *Genome Research* 27(6):1074–86.
- Painter, Heather J., Neo Christopher Chung, Aswathy Sebastian, Istvan Albert, John D. Storey, and Manuel Llinás. 2018. “Genome-Wide Real-Time in Vivo Transcriptional Dynamics during *Plasmodium falciparum* Blood-Stage Development.” *Nature Communications* 9(1):1–12.
- Pei, Xinhong, Xiuli An, Xinhua Guo, Michal Tarnawski, Ross Coppel, and Narla Mohandas. 2005. “Structural and Functional Studies of Interaction between *Plasmodium falciparum* Knob-Associated Histidine-Rich Protein (KAHRP) and Erythrocyte Spectrin.” *Journal of Biological Chemistry* 280(35):31166–71.
- Petter, Michaela, Insa Bonow, and Mo Quen Klinkert. 2008. “Diverse Expression Patterns of Subgroups of the Rif Multigene Family during *Plasmodium falciparum* Gametocytogenesis.” *PLoS ONE* 3(11).
- Poran, Asaf, Christopher Nötzel, Omar Aly, Nuria Mencia-Trinchant, Chantal T. Harris, Monica L. Guzman, Duane C. Hassane, Olivier Elemento, and Björn F. C.

- Kafsack. 2017. "Single-Cell RNA Sequencing Reveals a Signature of Sexual Commitment in Malaria Parasites." *Nature* 551(7678):95–99.
- Quinlan, Aaron R. and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26(6):841–42.
- Rabinovich, Regina N., Chris Drakeley, Abdoulaye A. Djimde, B. Fenton Hall, Simon I. Hay, Janet Hemingway, David C. Kaslow, Abdisalan Noor, Fredros Okumu, Richard Steketee, Marcel Tanner, Timothy N. C. Wells, Maxine A. Whittaker, Elizabeth A. Winzeler, Dyann F. Wirth, Kate Whitfield, and Pedro L. Alonso. 2017. "MalERA: An Updated Research Agenda for Malaria Elimination and Eradication." *PLoS Medicine* 14(11):1–17.
- Ramírez, Fidel, Devon P. Ryan, Björn Grüning, Vivek Bhardwaj, Fabian Kilpert, Andreas S. Richter, Steffen Heyne, Friederike Dünder, and Thomas Manke. 2016. "DeepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis." *Nucleic Acids Research* 44(W1):W160–65.
- Rawlings, David J., Hisashi Fujioka, Michal Fried, David B. Keister, Masamichi Aikawa, and David C. Kaslow. 1992. "α-Tubulin II Is a Male-Specific Protein in *Plasmodium falciparum*." *Molecular and Biochemical Parasitology* 56(2):239–50.
- Rosario, Virgilio. 2008. "Cloning of Naturally Occurring Mixed Infections of Malaria Parasites Author (s): Virgilio Rosario Published by : American Association for the Advancement of Science Stable URL : <http://www.jstor.org/stable/1685487> Cloning of Naturally Occurring Mixed In." *Advancement Of Science* 212(4498):1037–38.
- Rovira-Graells, Núria, Archana P. Gupta, Evarist Planet, Valerie M. Crowley, Sachel Mok, Lluís Ribas De Pouplana, Peter R. Preiser, Zbynek Bozdech, and Alfred Cortés. 2012. "Transcriptional Variation in the Malaria Parasite *Plasmodium falciparum*." *Genome Research* 22(5):925–38.
- Santos, Joana Mendonca, Gabrielle Josling, Philipp Ross, Preeti Joshi, Lindsey Orchard, Tracey Campbell, Ariel Schieler, Ileana M. Cristea, and Manuel Llinás. 2017. "Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor." *Cell Host and Microbe* 21(6):731-741.e10.
- Sargeant, Tobias J., Matthias Marti, Elisabet Caler, Jane M. Carlton, Ken Simpson,

- Terence P. Speed, and Alan F. Cowman. 2006. "Lineage-Specific Expansion of Proteins Exported to Erythrocytes in Malaria Parasites." *Genome Biology* 7(2).
- Silvestrini, Francesco, Edwin Lasonder, Anna Olivieri, Grazia Camarda, Ben Van Schaijk, Massimo Sanchez, Sumera Younis Younis, Robert Sauerwein, and Pietro Alano. 2010. "Protein Export Marks the Early Phase of Gametocytogenesis of the Human Malaria Parasite *Plasmodium falciparum*." *Molecular and Cellular Proteomics* 9(7):1437–48.
- Sinden, R. E. 1982. "Gametocytogenesis of *Plasmodium falciparum* in Vitro: An Electron Microscopic Study." *Parasitology* 84(1):1–11.
- Singh, Agam Prasad, Carlos A. Buscaglia, Qian Wang, Agata Levay, Daniel R. Nussenzweig, John R. Walker, Elizabeth A. Winzeler, Hodaka Fujii, Beatriz M. A. Fontoura, and Victor Nussenzweig. 2007. "Plasmodium Circumsporozoite Protein Promotes the Development of the Liver Stages of the Parasite." *Cell* 131(3):492–504.
- Sinha, Abhinav, Katie R. Hughes, Katarzyna K. Modrzynska, Thomas D. Otto, Claudia Pfander, Nicholas J. Dickens, Agnieszka a Religa, Ellen Bushell, Anne L. Graham, Rachael Cameron, Bjorn F. C. Kafsack, April E. Williams, Manuel Llinás, Matthew Berriman, Oliver Billker, and Andrew P. Waters. 2014. "A Cascade of DNA-Binding Proteins for Sexual Commitment and Development in Plasmodium." *Nature* 507(7491):253–57.
- Thompson, Joanne, Delmiro Fernandez-Reyes, Lisa Sharling, Sally G. Moore, Wijnand M. Eling, Sue A. Kyes, Christopher I. Newbold, Fotis C. Kafatos, Chris J. Janse, and Andrew P. Waters. 2007. "Plasmodium Cysteine Repeat Modular Proteins 1-4: Complex Proteins with Roles throughout the Malaria Parasite Life Cycle." *Cellular Microbiology* 9(6):1466–80.
- Thorvaldsdóttir, Helga, James T. Robinson, and Jill P. Mesirov. 2013. "Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration." *Briefings in Bioinformatics* 14(2):178–92.
- Tibúrcio, Marta, Makhtar Niang, Guillaume Deplaine, Sylvie Perrot, Emmanuel Bischoff, Papa Alioune Ndour, Francesco Silvestrini, Ayman Khattab, Geneviève Milon, Peter H. David, Max Hardeman, Kenneth D. Vernick, Robert W.

- Sauerwein, Peter R. Preiser, Odile Mercereau-Puijalon, Pierre Buffet, Pietro Alano, and Catherine Lavazec. 2012. "A Switch in Infected Erythrocyte Deformability at the Maturation and Blood Circulation of *Plasmodium falciparum* Transmission Stages." *Blood* 119(24):172–80.
- Tibúrcio, Marta, Francesco Silvestrini, Lucia Bertuccini, Adam Frederik Sander, Louise Turner, Thomas Lavstsen, and Pietro Alano. 2013. "Early Gametocytes of the Malaria Parasite *Plasmodium falciparum* Specifically Remodel the Adhesive Properties of Infected Erythrocyte Surface." *Cellular Microbiology* 15(4):647–59.
- Trager, William and James B. Jensen. 2005. "Human Malaria Parasites in Continuous Culture." *Journal of Parasitology* 91(3):484–86.
- Tusher, Virginia Goss, Robert Tibshirani, and Gilbert Chu. 2001. "Significance Analysis of Microarrays Applied to the Ionizing Radiation Response." *Proceedings of the National Academy of Sciences of the United States of America* 98(9):5116–21.
- Venugopal, Kannan, Franziska Hentzschel, Gediminas Valkiūnas, and Matthias Marti. 2020. "Plasmodium Asexual Growth and Sexual Development in the Haematopoietic Niche of the Host." *Nature Reviews Microbiology* 18(3):177–89.
- Villard, Viviane, George W. Agak, Géraldine Frank, Ali Jafarshad, Catherine Servis, Issa Nébié, Sodlomon B. Sirima, Ingrid Felger, Myriam Arevalo-Herrera, Socrates Herrera, Frederic Heitz, Volker Bäcker, Pierre Druilhe, Andrey V. Kajava, and Giampietro Corradin. 2007. "Rapid Identification of Malaria Vaccine Candidates Based on α -Helical Coiled Coil Protein Motif." *PLoS ONE* 2(7).
- Volz, Jennifer C., Alan Yap, Xavier Sisquella, Jenn K. Thompson, Nicholas T. Y. Lim, Lachlan W. Whitehead, Lin Chen, Marko Lampe, Wai Hong Tham, Danny Wilson, Thomas Nebl, Danushka Marapana, Tony Triglia, Wilson Wong, Kelly L. Rogers, and Alan F. Cowman. 2016. "Essential Role of the PfRh5/PfRipr/CyRPA Complex during *Plasmodium falciparum* Invasion of Erythrocytes." *Cell Host and Microbe* 20(1):60–71.
- Voss, Till S., Thierry Mini, Paul Jenoe, and Hans Peter Beck. 2002. "*Plasmodium falciparum* Possesses a Cell Cycle-Regulated Short Type Replication Protein A

- Large Subunit Encoded by an Unusual Transcript.” *Journal of Biological Chemistry* 277(20):17493–501.
- Vossen, Matthias G., Sandra Pferschy, Peter Chiba, and Harald Noedl. 2010. “The SYBR Green I Malaria Drug Sensitivity Assay: Performance in Low Parasitemia Samples.” *American Journal of Tropical Medicine and Hygiene* 82(3):398–401.
- Weiss, Daniel J., Tim C. D. Lucas, Michele Nguyen, Anita K. Nandi, D. Bisanzio, Katherine E. Battle, Ewan Cameron, Katherine A. Twohig, Daniel A. Pfeffer, Jennifer A. Rozier, Harry S. Gibson, Puja C. Rao, Daniel Casey, Amelia Bertozzi-Villa, Emma L. Collins, Ursula Dalrymple, N. Gray, Joseph R. Harris, Rosalind E. Howes, Sun Yun Kang, Suzanne H. Keddie, Daniel May, S. Rumisha, Michael P. Thorn, Ryan Barber, N. Fullman, Chantal K. Huynh, Xie Kulikoff, Michael J. Kutz, Alan D. Lopez, A. H. Mokdad, Mohsen Naghavi, G. Nguyen, Katya Anne Shackelford, Theo Vos, Haidong Wang, David L. Smith, Stephen S. Lim, Christopher J. L. Murray, S. Bhatt, Simon I. Hay, and Peter W. Gething. 2019. “Mapping the Global Prevalence, Incidence, and Mortality of *Plasmodium falciparum*, 2000–17: A Spatial and Temporal Modelling Study.” *The Lancet* 394(10195):322–31.
- WHO. 2019. *World Malaria Report*.
- Young, Jason A., Quinton L. Fivelman, Peter L. Blair, Patricia De La Vega, Karine G. Le Roch, Yingyao Zhou, Daniel J. Carucci, David A. Baker, and Elizabeth A. Winzeler. 2005. “The *Plasmodium falciparum* Sexual Development Transcriptome: A Microarray Analysis Using Ontology-Based Pattern Identification.” *Molecular and Biochemical Parasitology* 143(1):67–79.
- Yuda, Masao, Shiroh Iwanaga, Izumi Kaneko, and Tomomi Kato. 2015. “Global Transcriptional Repression: An Initial and Essential Step for *Plasmodium* Sexual Development.” *Proceedings of the National Academy of Sciences of the United States of America* 112(41):12824–29.
- Yuda, Masao, Izumi Kaneko, Shiroh Iwanaga, Yuho Murata, and Tomomi Kato. 2019. “Female-specific Gene Regulation in Malaria Parasites by an AP2-family Transcription Factor.” *Molecular Microbiology* 0–2.
- Zhang, Cui, Zhenkui Li, Huiting Cui, Yuanyuan Jiang, Zhenke Yang, and Xu Wang.

2017. “Systematic CRISPR-Cas9-Mediated ApiAP2 Genes Reveal Functional Insights into Parasite Development.” *MBio* 8(6):1–17.

Zhang, Min, Chengqi Wang, Thomas D. Otto, Jenna Oberstaller, Xiangyun Liao, Swamy R. Adapa, Kenneth Udenze, Iraad F. Bronner, Deborah Casandra, Matthew Mayho, Jacqueline Brown, Suzanne Li, Justin Swanson, Julian C. Rayner, Rays H. Y. Jiang, and John H. Adams. 2018. “Uncovering the Essential Genes of the Human Malaria Parasite *Plasmodium falciparum* by Saturation Mutagenesis.” *Science* 360(6388).

Zhao, Jinghua, Purnima Bhanot, Junjie Hu, and Qian Wang. 2016. “A Comprehensive Analysis of Plasmodium Circumsporozoite Protein Binding to Hepatocytes.” *PLoS ONE* 11(8):1–13.

A

B

C

Figure 1 PfAP2-G2 is expressed from around 16-19 hpi in asexual stage parasite (A) Structure of PfAP2-G2 protein which contains single AP2 DNA-binding domain. (B) Live fluorescence microscopy performed on a highly synchronized population of AP2-G2::GFP parasites every 7 h of the asexual life cycle showed that PfAP2-G2::GFP colocalizes with the nucleus in both the trophozoite and schizont stages. DRAQ-5 was used as the nuclear stain. (C) Nuclear and cytoplasmic fractions from the GFP tagged PfAP2-G2 synchronized trophozoite stage parasites were subjected to western blot analysis using anti-GFP antibodies (1:1000). Full length (FL) GFP-tagged PfAP2-G2 (~250 kDa) was detected only in the nuclear fraction of the parasite lysate whereas cleavage product (CP, ~50kDa) was in both fractions. WT parasites were used as a negative control. Anti-histone H3 (1:3000) and anti-aldolase (1:1000) were used as loading controls for the nuclear and cytoplasmic fractions, respectively.

Figure 2 Absence of PfAP2-G2 has no effect on the growth of asexual stage parasite (A) Growth profiles of the WT and PfAP2-G2 KO parasites measured using an SYBR green assay every 24 hours over a period of 10 days starting with highly synchronous ring-stage parasites. The graph represents the plotted average $\pm$ S.D. of biological triplicates. Fluorescence units are plotted along the Y-axis and the number of days was plotted along the X-axis for the experiment done in triplicate. (B) Bar graph representing the measurement of the multiplication rate of the WT and PfAP2-G2 KO parasites. The values are the average of three replicates, and the error bars represent the standard deviation.

A

B

C

Figure 3 PfAP2-G2 is expressed throughout the gametocyte stages and is essential for maturation of gametocytes

(A) Images obtained by live fluorescence microscopy performed on Stage I through Stage V gametocytes. The DIC image was merged with the DRAQ5 nuclear stain imaging to confirm the nuclear localization of the signal. The results indicate that PfAP2-G2 was expressed in all stages of the gametocytes and was localized to the nucleus. (B) Nuclear (N) and cytoplasmic (C) fractions from the stage III gametocytes were subjected to western blot analysis using anti-GFP antibodies (1:1000). The GFP-tagged AP2-G2 of expected size (~250 kDa) was detected only in the nuclear fraction of the parasite lysate. The WT parasite was used as a negative control. Anti-histone H3 (1:3000) and anti-aldolase (1:1000) were used as the loading controls for the nuclear and cytosolic fractions, respectively. (C) Giemsa-stained smears of gametocytes from the WT and PfAP2-G2 KO lines showed that the morphology of the KO gametocytes appeared normal until Stage III. The KO gametocytes were unable to develop beyond Stage III and showed severe morphological defects.

A

B

Motif enriched in PfAP2-G2 binding regions

PBM motif

PF3D7_1139300
1790/5008
P-value 5.2e-129

PfAP2-G2 motif
1770/5008
P-value 8.7e-088

C

PF3D7_0304600 (CSP)
Motif AACCA

Input 1
AP2-G2 GFP 1
Input 2
AP2-G2 GFP 2
Input 3
AP2-G2 GFP 3

Gene

Called Peaks
Motif

PF3D7_1216600 (CelTOS)
Motif AACCA

D

Input 1
AP2-G2 GFP 1
Input 2
AP2-G2 GFP 2
Gene
(PF3D7_0501300)

Called Peaks
Motif AACCA

E

Genes with peaks in asexuals

Genes with peaks in gametocytes

F

Input Asexual
AP2-G2 GFP Asexual
Input Gametocytes
AP2-G2 GFP Gametocytes
Gene (msp2)
Called peak

G

Motif enriched in PfAP2-G2 binding regions

PF11_1139300
PBM motif

252/945
P-value 4.8e-034

Figure 4 Genome wide predictions of PfAP2-G2 targets and binding regions by ChIP seq (A) Pearson correlation analysis on three biological replicates of ChIP-seq including the Input (In) and AP2-G2 GFP (GFP) (B) DREME logo for the motif enriched in the genomic locations bound by PfAP2-G2 in the in the gene body and non-coding regions in asexual stage. Shown is the enriched PF11_1139300 motif AGAA and PfAP2-G2 motif ACCA obtained by DREME along with the PBM motif. The P-values are reported by DREME. (C) Example PfAP2-G2 peaks in asexual stages for all three replicates. Peaks are represented as red tracks below with the motif present is highlighted in blue. (D) Example PfAP2-G2 peaks in gametocytes for both replicates. (E) Venn diagram showing common genes with PfAP2-G2 peaks in asexuals and gametocytes (F) Representative peaks for the gene *msp2* enriched with PfAP2-G2 in both asexual and gametocyte stages. (G) DREME logo for the motifs enriched in the genomic locations bound by PfAP2-G2 in the in the gene body and non-coding regions in gametocytes. The P-values are reported by DREME.

D

Genes with enhanced transcript abundance in asexual stages (236)

ChIP targets in asexual stages (82)

List of the target PfAP2-G2 genes in asexual stages with enhanced transcript abundance

PF3D7_0212600	secreted protein with altered thrombospondin repeat domain
PF3D7_0207400	serine repeat antigen 7
PF3D7_0216800	TMEM121 domain-containing protein, putative
PF3D7_0202000	knob-associated histidine-rich protein
PF3D7_0200100	erythrocyte membrane protein 1, PfEMP1
PF3D7_0323400	Rh5 interacting protein
PF3D7_0320400	oocyst capsule protein Cap380
PF3D7_0304600	circumsporozoite (CS) protein
PF3D7_0301800	Plasmodium exported protein, unknown function
PF3D7_0511200	stearoyl-CoA desaturase
PF3D7_0632800	erythrocyte membrane protein 1, PfEMP1
PF3D7_0702100	Plasmodium exported protein (PHISTb), unknown function, pseudogene
PF3D7_0800100	erythrocyte membrane protein 1, PfEMP1
PF3D7_0931300	conserved Plasmodium protein, unknown function
PF3D7_0936100	early transcribed membrane protein
PF3D7_1016500	Plasmodium exported protein (PHISTc), unknown function
PF3D7_1102000	DnaJ protein, putative
PF3D7_1100100	erythrocyte membrane protein 1, PfEMP1
PF3D7_1252900	Plasmodium exported protein, unknown function
PF3D7_1327100	conserved protein, unknown function
PF3D7_1447500	inner membrane complex protein, putative
PF3D7_1455600	ferlin, putative
PF3D7_1477800	acyl-CoA binding protein

E

Figure 5 Transcriptional changes in absence of PfAP2-G2 (A) Classification of significantly changing genes in asexual and gametocyte stages based on its maximum transcription in either asexual, gametocytes or mosquito stages, using RNA-seq data (López-Barragán et al., 2011). (B) Overlap between the genes with decreased transcript abundance (>1.5 log₂FC, 0.30 FDR) in gametocytes and earlier established late gametocyte markers (van Biljon et al., 2019). (C) Venn diagram showing overlap between genes with decreased transcript abundance in gametocytes and asexual stages. (D) Venn diagram showing number of overlapping genes with PfAP2-G2 enrichment in the upstream regions and also increased transcript abundance (analyzed by SAM, >1.5 log₂FC, 0.30 FDR) in absence of PfAP2-G2. (E) Expression fold change (log₂ KO versus WT) for the genes that are bound by PfAP2-G2 in the upstream regions in asexual stages. The genes are clustered based on Pearson's correlation. Boxed is the timepoint at which ChIP-seq was performed.

A

B

C

D

E

Figure 6 PfAP2-G2 shows high correlation with H3K36me3 repressive histone modifications (A) Heatmap showing correlation between occupancy of PfAP2-G2 and nine histone modifications in the gene body of all *P. falciparum* genes (n=5735). PfAP2-G2 strongly correlates with H3K36me3 and H4K20me3 (B) Venn diagram comparing the peaks obtained by ChIP-seq on H3K36me3 marks (orange) (Jiang *et al.*, 2013) and PfAP2-G2-GFP (green) shows that 80% of times PfAP2-G2 shows co-occupancy with H3K36me3 marks. (C) IGV screenshot showing co-occupancy of PfAP2-G2 with H3K36me3 and H4K20me3. First gray track represents Input for PfAP2-G2 and second green track shows the regions bound by PfAP2-G2. Third gray track represents input for H3K36me3 and H4K20me3 and the last two orange tracks represent the regions bound by H3K36me3 and H4K20me3 as indicated in the figure. Blue track shows the genes. (D) IGV screenshot showing co-occupancy of PfAP2-G2 with PfHP1. First gray track represents Input for PfAP2-G2 and second green track shows the regions bound by PfAP2-G2. Third gray track represents input for PfHP1 and the last two pink tracks represent the regions bound by PfHP1 as indicated in the figure. Blue track shows the genes. (E) Venn diagram comparing the peaks obtained by ChIP-seq on PfHP1 marks (pink) (Brancucci *et al.*, 2014) and PfAP2-G2-GFP (green) shows that 80% of times PfAP2-G2 shows co-occupancy with H3K36me3 marks.

A

B

Upper band ~1940 aa

Number of PfAP2-G2 peptides detected in IP vs control
9|5|16 0|0|0

Lower band ~400 aa

Number of PfAP2-G2 peptides detected in IP vs control
9|5|16 0|0|0

400 aa

C

PfAP2-G2 interacting partners in the lower band

Gene ID	Gene Description	Peptides in replicate IPs	Peptides in control	Average Probability
PF3D7_1105100	histone H2B (H2B)	24 11 14	0 2 1	1
PF3D7_1441400	FACT complex subunit SSRP1, putative (FACT S)	17 11 11	0 2 2	1
PF3D7_0617900	histone H3 variant, putative (H3.3)	16 8 9	0 0 0	1
PF3D7_0917900	heat shock protein 70 (HSP70)	13 8 12	0 0 0	1
PF3D7_0517400	FACT complex subunit SPT16, putative (FACT L)	19 8 12	0 0 0	1
PF3D7_1358800	40S ribosomal protein S15 (RPS15)	5 5 5	0 0 0	0.9997
PF3D7_1105000	histone H4 (H4)	37 14 22	0 3 6	0.9993
PF3D7_1220900	heterochromatin protein 1 (HP1)	13 5 6	0 1 0	0.999

D

PfAP2-G2 interacting partners in the upper band

Gene ID	Gene Description	Peptides in replicate IPs	Peptides in control	Average Probability
PF3D7_0624600	SNF2 helicase, putative (ISWI)	28 14 27	0 1 5	1
PF3D7_1212700	eukaryotic translation initiation factor 3 subunit A, putative (EIF3A)	13 17 22	0 0 1	1
PF3D7_0607000	translation initiation factor IF-2, putative	11 5 14	0 0 0	1
PF3D7_0711000	AAA family ATPase, CDC48 subfamily (Cdc48)	7 10 10	0 2 0	0.9997
PF3D7_0710200	conserved Plasmodium protein, unknown function	6 6 23	0 0 0	0.9997
PF3D7_1104200	chromatin remodeling protein (SNF2L)	15 7 12	0 3 0	0.9993
PF3D7_1007700	transcription factor with AP2 domain(s) (ApiAP2)	13 6 6	0 0 1	0.9993
PF3D7_0517400	FACT complex subunit SPT16, putative (FACT-L)	7 4 8	0 0 0	0.999
PF3D7_1468100	MORC family protein, putative	14 8 17	0 1 3	0.9983
PF3D7_1433500	DNA topoisomerase 2 (TOP2)	9 8 16	0 0 4	0.9983
PF3D7_1008100	conserved Plasmodium protein, unknown function	4 5 14	0 0 0	0.998
PF3D7_0823300	histone acetyltransferase GCN5 (GCN5)	5 5 8	0 1 0	0.998
PF3D7_0500800	mature parasite-infected erythrocyte surface antigen (MESA)	36 28 42	0 8 7	0.9977
PF3D7_0811200	ER membrane protein complex subunit 1, putative (EMC1)	5 3 7	0 0 0	0.9977
PF3D7_1462300	conserved Plasmodium protein, unknown function	4 7 12	0 0 0	0.9973
PF3D7_0610900	transcription elongation factor SPT5, putative (SPT5)	5 7 8	0 1 0	0.9957
PF3D7_1023900	chromodomain-helicase-DNA-binding protein 1 homolog, putative (CHD1)	30 12 33	0 3 4	0.9933
PF3D7_1139300	transcription factor with AP2 domain(s) (ApiAP2)	2 2 3	0 0 0	0.967

Figure 7 Identification of potential interacting partners of PfAP2-G2 (A) SDS PAGE showing immunoprecipitation (without crosslinking) of the PfAP2-G2 GFP parasite line using antibodies against GFP. Full length PfAP2-G2 and the cleaved product was analyzed separately by mass spectrometry to identify the interacting partners of PfAP2-G2. (B) Schematic showing the peptides recovered from upper and lower band. Also shown are the number of PfAP2-G2 peptides in both the bands and in the controls. (C) List of PfAP2-G2 interacting partners in the lower band with high probability of interaction measured by SAINT (pSAINT 0.9 and 1%FDR). (D) List of PfAP2-G2 interacting partners in the upper band with high probability of interaction measured by SAINT (pSAINT 0.9 and 1%FDR) in the upper band.