Published in final edited form as:
Nature. 2014 March 13; 507(7491): 248–252. doi:10.1038/nature12920.

A transcriptional switch underlies commitment to sexual development in human malaria parasites

Björn F.C. Kafsack[1,*], **Núria Rovira-Graells**[2,3], **Taane G. Clark**[4,5], **Cristina Bancells**[2], **Valerie M. Crowley**[1,3,#], **Susana G. Campino**[6], **April E. Williams**[7], **Laura G. Drought**[4], **Dominic P. Kwiatkowski**[6,8], **David A. Baker**[4], **Alfred Cortés**[2,3,9], and **Manuel Llinás**[1,7,#]

[1]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, USA

[2]Barcelona Centre for International Health Research (CRESIB, Hospital Clínic-Universitat de Barcelona), Barcelona, Spain

[3]Institute for Research in Biomedicine (IRB), Barcelona, Spain

[4]Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK

[5]Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

[6]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK

[7]Department of Molecular Biology, Princeton University, Princeton, USA

[8]Wellcome Trust Sanger Centre for Human Genetics, Oxford, UK

[9]Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

Abstract

The life cycles of many parasites involve transitions between disparate host species, requiring these parasites to go through multiple developmental stages adapted to each of these specialized niches. Transmission of malaria parasites (*Plasmodium* spp.) from humans to the mosquito vector requires differentiation from asexual stages replicating within red blood cells into non-dividing male and female gametocytes. Although gametocytes were first described in 1880, our understanding of the molecular mechanisms involved in commitment to gametocyte formation is extremely limited and disrupting this critical developmental transition remains a long-standing

[*]Currently: Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[#]Currently: Department of Molecular Biology and Center for Infectious Disease Dynamics, The Pennsylvania State University, State College PA USA

Author Contributions: ML managed the overall project with input from BFCK, DAB, and AC. BFCK generated the Δ*pfap2-g* knockout, PfAP2-G-ddFKBP, and luciferase lines and designed, performed and analyzed the microarray, gel shift, luciferase, and ligand-regulatable gametocytogenesis experiments. VMC performed qRT-PCR validation. AEW prepared Δ*pfap2-g* sequencing libraries and together with BFCK analyzed the sequencing data. DAB, TGC and SGC conceived the sequencing of gametocyte non-producer lines F12 and GNP-A4. TGC analyzed the gametocyte non-producer sequencing data. SGC and DPK carried out and supervised sequencing of gametocyte non-producer lines, respectively. AC and NRG generated E5 and other 3D7-B subclones and respectively supervised and performed the experiments presented in Fig. 1 and 2B, and provided the analysis presented in Supplementary Fig. 1. VMC and NRG performed and AC supervised ChIP experiments. CB and AC generated the PfAP2-G-HAx3 line and carried out IFAs and correlations with gametocyte formation. BFCK wrote the manuscript with major input from ML, DAB, and AC.

Europe PMC Funders Author Manuscripts

goal[1]. We show here that expression levels of the DNA-binding protein PfAP2-G correlate strongly with levels of gametocyte formation. Using independent forward and reverse genetics approaches, we demonstrate that PfAP2-G function is essential for parasite sexual differentiation. By combining genome-wide PfAP2-G cognate motif occurrence with global transcriptional changes resulting from PfAP2-G ablation, we identify early gametocyte genes as likely targets of PfAP2-G and show that their regulation by PfAP2-G is critical for their wild-type level expression. In the asexual blood-stage parasites *pfap2-g* appears to be among a set of epigenetically silenced loci[2,3] prone to spontaneous activation[4]. Stochastic activation presents a simple mechanism for a low baseline of gametocyte production. Overall, these findings identify PfAP2-G as a master regulator of sexual-stage development in malaria parasites and mark the first identification of a transcriptional switch controlling a differentiation decision in protozoan parasites.

From its uptake in a mosquito blood meal to initial infection of red blood cells in the subsequent host the malaria parasite *P. falciparum* goes through at least seven major developmental changes (asexual red cell stage -> gametocyte -> gamete -> ookinete -> oocyst -> sporozoite -> liver stage -> asexual red cell stage). In all but one case, as the parasite reaches its subsequent niche within the host, differentiation into the appropriate developmental stage is a necessity for continuation of the lifecycle. The lone exception occurs once the parasite has started replicating in red blood cells (RBCs). During the 48-hour lytic cycle following each new RBC invasion, a developmental decision is made that determines whether daughter parasites will continue replicating asexually and maintain the infection of the current host or differentiate into non-dividing male or female gametocytes. While the latter decision is a dead-end for replication within the current host it is essential for infection of mosquitoes and thus transmission to the next host[5,6].

A recent study on transcriptional variation identified differentially expressed genes linked to early gametocyte development in two stocks (3D7-A, 3D7-B) of the common 3D7 *P. falciparum* parasite line[7]. Within this expression cluster of early gametocyte markers, we noted the presence of a potential transcriptional regulator, PfAP2-G (PFL1085w/ PF3D7_1222600), which belongs to the apicomplexan AP2 (ApiAP2) family of DNA binding proteins (Supplementary Fig. 1) and that is conserved among most members of the phylum (Supplementary Figure 2). ApiAP2 proteins represent the major family of transcriptional regulators in malaria parasites[1,8] and have thus far been found to regulate several of the parasite's developmental transitions, including ookinete formation[2,3,9] and oocyst sporozoite maturation[4,10] within the mosquito, and development in the mammalian liver[11,12]. Follow-up quantitative RT-PCR analysis in blood-stage parasites confirmed higher PfAP2-G transcript abundance in 3D7-B compared to 3D7-A and also revealed significant variation in expression levels between individual 3D7-B subclones (Fig. 1A). Strikingly, when gametocyte formation was measured in these lines, PfAP2-G transcript levels were highly predictive ($R^2 > 0.99$) of relative gametocyte production (Fig. 1B).

In a parallel line of inquiry, we screened the well-studied gametocyte non-producer line F12[5,6,13], as well as a second parasite line (GNP-A4) that had also spontaneously lost its ability to produce gametocytes, for mutations in protein coding regions. Whole genome

sequencing of these lines revealed that the only gene containing mutations in both F12 and GNP-A4 was *pfap2-g* (Supplementary Table 1), resulting in the introduction of stop codons upstream of or within the AP2 DNA-binding domain (Fig. 2A & Supplementary Fig. 3). Previous studies identified sub-telomeric deletions in the right arm of *P. falciparum* chromosome 9 that are associated with defective gametocyte production[7,14,15]. The F12 and GNP-A4 clones do not have coding sequence mutations or deletions within this region, nor within any of the 16 genes recently implicated in gametocyte development by random transposon mutatgenesis[16]. The presence of *pfap2-g* mutations in two independently derived gametocyte non-producer lines provides a second, independent connection between PfAP2-G and gametocyte formation, pointing to this locus as a key determinant of sexual differentiation. While spontaneous inactivation of *pfap2-g* has occurred repeatedly *in vitro*, no loss-of-function mutations could be found in the genomes of nearly 300 distict field isolates[17], further underlining its potential importance to transmission.

To directly test the contribution of PfAP2-G function to gametocyte formation, we generated a PfAP2-G null mutant (Δ*pfap2-g*) via double homologous recombination in the high gametocyte producing 3D7-B subclone E5 (Fig. 2A-B and Supplementary Fig. 4). As predicted based on our earlier sequencing results, the Δ*pfap2-g* mutant completely lost the ability to produce gametocytes (Fig. 2C). In order to identify any additional mutations that may have been acquired in the extended process of generating the PfAP2-G knockout, we sequenced the genomes of both Δ*pfap2-g* and its E5 parent. Apart from the targeted deletion, we found only a limited number of additional mutations within coding regions, none of which are shared with the other non-producer lines that we sequenced or found in genes previously linked to gametocyte development[14-16] (Supplementary Table 1). This combination of forward and reverse genetic evidence strongly implies the essentiality of PfAP2-G for the production of gametocytes in *P. falciparum*. In direct competition cultures Δ*pfap2-g* consistently outgrew its parent E5, consistent with the fact that PfAP2-G action occurs at or prior to the asexual/sexual decision but not thereafter as only a failure to initiate gametocytogenesis would provide an *in vitro* growth advantage (Supplementary Fig. 5).

Attempts at generating full-length complementation expression constructs were unsuccessful, likely due the considerable length (7.3kb) of the coding sequence and its very low complexity (21.8% GC and long repeat sequences). As an alternative confirmation for the role of PfAP2-G in gametocyte formation, we made PfAP2-G function ligand-regulatable by appending the ddFKBP destabilization domain to the 3′ end of the endogenous coding sequence, (*pfap2-g-ddfkbp*, Supplementary Fig. 6A-B). In the absence of its ligand Shield-1 (Shld1) the ddFKBP domain is unstable and targets fusion proteins for proteolytic degradation[18,19], thus making PfAP2-G protein levels regulatable by the addition of Shld1 (Supplementary Fig. 6C). Indeed, in the *pfap2-g-ddfkbp* line gametocyte formation was completely dependent on the addition of Shld1, while its presence had no effect on gametocyte production by the E5 parent (Fig. 2D-E), demonstrating that PfAP2-G function is essential for gametocyte formation.

Based on the localization of HA-tagged PfAP2-G to the parasite nucleus (Fig. 3A, Supplementary Fig. 7) and the fact that several ApiAP2 proteins act as transcriptional regulators, we aimed to identify possible regulatory targets of PfAP2-G. To do this, we

compared the global transcriptional pattern over the 48-hour intraerythrocytic cycle for the gametocyte producing parent E5 to those of the mutant non-producers Δ*pfap2-g* and F12. As expected, only a small number of transcripts changed by greater than 2-fold in both mutants; with four transcripts increasing and 24 transcripts decreasing in abundance (Fig. 3B and Supplementary Table 2). All four up-regulated genes are located in subtelomeric regions and have previously been shown to undergo spontaneous transcriptional variation and were therefore not considered further[7]. However, the cluster of down-regulated genes is highly enriched for genes expressed during the first stages of gametocyte formation ($p < 0.003$), including some of the earliest known markers of sexual commitment: Pfs16, Pfg27/25, and Pfg14.744[20,21]. Quantitative RT-PCR measurements of early gametocyte markers confirmed the lower relative abundance levels in Δ*pfap2-g* (Supplementary Fig. 8). Analysis of the upstream regions of most down-regulated genes showed that they were also enriched in the DNA motif recognized by PfAP2-G[22] ($p < 0.017$). These results implicate PfAP2-G as a transcriptional switch that controls sexual differentiation by activating the transcription of early gametocyte genes.

Using electrophoretic mobility shift assays we confirmed that recombinant PfAP2-G DNA-binding domain could interact with three gametocyte promoters in a motif-dependent manner *in vitro* (Fig. 3C). To test whether this interaction occurs within the parasite, we transfected E5 and Δ*pfap2-g* with luciferase reporter constructs under the control of these gametocyte promoters (Fig. 3D). There was a significant reduction in luciferase activity in the Δ*pfap2-g* background compared to its E5 parent for all three constructs. Additionally, luciferase levels were also significantly diminished in the parental E5 line when we altered the PfAP2-G recognition sequence in the two promoters tested, indicating that PfAP2-G likely acts as a direct transcriptional activator of the earliest gametocyte genes.

The *pfap2-g* locus shares many features that have been associated with the epigenetic silencing of multi-gene families in *P. falciparum*[4,23] such as high levels of the H3K9me3 histone modification[3], associated binding of heterochromatin protein 1 (PfHP1)[2,24] and perinuclear localization[3]. Based on these data PfAP2-G expression is likely regulated epigenetically by reversible formation of repressive chromatin structures. Interestingly, we find that the pattern of histone modifications at this locus is typical of heterochromatin-silenced genes in both the high gametocyte producer E5 and its low producing A7 sibling clone (Supplementary Fig. 9A-C). This finding suggests that in predominantly asexual blood stage cultures, the *pfap2-g* locus is found in a heterochromatic (silenced) state in the majority of parasites and that the transcriptionally permissive state may only occur in a small number of sexually committed parasites. Indeed, the vast majority of asexually growing PfAP2-G-HAx3 parasites contained no detectable levels of PfAP2-G by immunofluorescence, while a small subpopulation exhibited clear nuclear PfAP2-G staining (Fig. 4A). As expected, every newly-formed merozoite within PfAP2-G expressing schizonts stained positive for PfAP2-G, lending further support to the previous findings that all daughter parasites from a given schizont are commited to the same developmental fate[25]. Furthermore, while the PfAP2-G-positive fraction varied between experiments, it was highly predictive of subsequent gametocyte formation in commitment assays (R^2=0.94, Fig. 4B).

Stochastic, low frequency activation would provide a simple mechanism for baseline gametocyte production, which may be modulated in response to environmental stimuli. Furthermore, the presence of insulator-like pairing element sequences, which have been suggested to play an important role in the silencing of *var* genes[26], flanking the *pfap2-g* locus (Supplementary Fig. 9D) raises the intriguing possibility that the expression of *pfap2-g* may be mutually exclusive with that of the *var* gene family[23]. In addition to chromatin-mediated control, PfAP2-G expression may be auto-regulated via binding to the eight instances of the PfAP2-G cognate motifs located 2.1kb to 3.6kb upstream of the PfAP2-G locus (Supplementary Fig. 10). We have integrated these various regulatory mechanisms into a model of how PfAP2-G expression controls the decision of individual cells to commit to gametocyte formation or to continue along the default pathway of asexual replication (Fig. 4C).

Together with the work of Sinha et al. (accompanying manuscript), our results demonstrate that AP2-G is an essential regulator of gametocyte formation in malaria parasites and acts as a developmental switch by activating the transcription of early gametocyte genes. This provides the first insight into the molecular mechanisms controlling the asexual/sexual developmental decision in malaria parasites and unveils new targets in the long-standing aim of interrupting malaria transmission by preventing the formation and/or maturation of the parasite's sexual stages.[1] Lastly, ligand-regulatable PfAP2-G is not only a powerful new tool for future studies of sexual stage development in malaria parasites but also holds great potential for inducible gene expression in general.

Full Methods

Parasites and Strains

Parasite lines 3D7-A[31], 3D7-B[31] and F12[13] have been previously described. Note that 3D7-A is not the same line as the competent gametocyte producer 3D7A[32], which was not used in this study. 3D7-B subclones E5, A7, and B11 were generated by limiting dilution. The gametocyte non-producer line GNP-A4 was generated during an attempt to knockout a phosphodiesterase gene (PfPDE∂, PF14_0672). Integration of the knockout construct by single crossover homologous recombination occurred at the targeted locus but this event was not responsible for the clone's inability to produce gametocytes (Catherine Taylor, PhD thesis 2007). A subsequent successful knockout of PfPDE∂ produced gametocytes at normal rates and the true phenotype was a significantly lower exflagellation rate than parental parasites due to a reduced ability of gametes to egress from red blood cells[33]. Δ*pfap2-g* knockout parasites were generated by transfection of 3D7 E5 with pHHT-FCU-PFL1085w (Supplementary Figure 4) followed by positive (*hdhfr*)/negative selection (*fcu*) using WR99210 and 5-fluoro-cytosine as previously described[34]. Resistant parasites were sub-cloned and verified by PCR & Southern blot. *pfap2-g-ddfkbp* parasites were generated by transfection of 3D7-B E5 with pJDD145-pfap2-g and selected on WR99210. After sub-cloning, integration was verified by PCR using a forward primer at position +4269 and a ddFKBP reverse primer. Displacement of the endogenous downstream sequence was verified using primers at +4269 and +7490 with respect to the translation initiation site (Supplementary Figure 6). Parasites expressing a HAx3-tagged version of PfAP2-G were

obtained by transfecting 3D7-B E5 with the plasmid pHH1inv_1085_HAx3 and cycling twice on/off WR99210 to select for parasites where the plasmid has integrated into the genome. After subcloning by limiting dilution and Southern blot analysis (Supplementary Fig. S7B), a subclones with a single copy of the plasmid integrated at the *pfap2-g* locus (E5-AP2-G-HAx3-9A) was selected for IFA analysis. All parasites were grown in media containing 0.25% Albumax II and synchronized by standard methods[35].

DNA Constructs

Knockout Construct—The region from −126bp to +366bp and +6945 to +7379 with respect to the *pfap2-g* initiation codon were cloned into the *Nco*I/*Eco*RI and *Spe*I/*Sac*II sites of pHHT-FCU[34] respectively to generate pHHT-FCU-PFL1085w.

ddFKBP C-terminal Tagging Construct—*pfap2-g* coding sequence positions +4740 to +7296 were cloned into with *Not*I/*Xho*I sites of pJDD145[37] (generous gift from M. Duraisingh).

Luciferase expression constructs—The *hdhfr* selectable marker of pVLbIDh[36] was replaced with blasticidin-S deaminase using the *Sac*I/*Not*I sites to generate pVL-BSD. The *var7b* promoter was excised with *Hpa*I/*Kpn*I, blunted and re-ligated, destroying these sites. The 1445bp, 1226bp, and 1159bp upstream of the pf11-1, pfg27/25 and pfpeg4 start codons were cloned into the *Aat*II/*Nco*I sites. The PfAP2-G cognate motifs in the upstream sequences of *pf11-1* (-328 to −323) and *pfpeg4* (-1138 to 1131) were converted to adenines using site directed mutagenesis.

HAx3-tagging construct—The plasmid pHH1inv_1085_HAx3 was derived from the plasmid E140-0[39]. A triple HA tag (HAx3) was cloned into *Kpn*I-*Xho*I sites of E140-0, replacing the *eba-140* ORF and introducing several new restriction sites (plasmid pHH1inv_HAx3). A fragment of the *pfap2-g* ORF from position +6685 to the stop codon was PCR-amplified and cloned in frame into *Kpn*I-*Pst*I sites of pHH1inv_HAx3, such that upon integration of the plasmid by single homologous recombination PfAP2-G is expressed as a fusion protein with the HAx3 tag, separated by the sequence YLQ.

Gametocytogenesis

Gametocyte conversion rates for the 3D7-A, 3D7-B and 3D7-B subclones (Fig. 1) were measured by treating synchronized ring-stage parasite cultures at 5% parasitemia with N-acetylglucosamine (NAG) and counting gametocytemia 3-4 days later. Gametocyte induction of E5, F12, GNP-A4 and Δ*pfap2-g* (Fig. 2C) was performed according to published methods[27] in media containing 5% AB+ heat-inactivated human serum and 0.25% Albumax II. For ligand-regulatable gametocytogenesis (Fig. 2D-E), synchronized parasites were set up at 0.5-1.0% late trophozoites in 3% hematocrit on day 0. Cultures were split in two and treated with 0.5μM Shld1 (generous gift from D. Goldberg) or an equal volume of ethanol solvent control for the remainder of the experiment. Parasitemia was determined on day 3 and 50mM N-acetyl-D-glucosamine was added to all cultures for the remainder of the experiment. Gametocytemia was determined on day 9 and converted into percent commitment by dividing by day 3 parasitemia. Statistical significance in gametocyte

production was determined using unpaired two-sided t-tests. Replicates were biological not technical.

Growth Competition

E5 and Δ*pfap2-g* parasites were mixed at 1:1 ratio and grown for 7+ weeks in triplicate. Parasites were diluted 1:20 with uninfected erythrocytes and gDNA was isolated whenever parasitemia exceeded 10%. For each time point, gDNA was isolated from the technical replicates, pooled at equal concentration, used for PCR amplification of a 414nt (1718-2131) region of PFF0275c covering the SNP described in Supplementary Table 1) and sequenced using the reverse primer. Difference in the growth rate was determined using the relative sequencing read peak height of A & C across at least 16 replication cycles (32+ days). Growth rate was fit to the data using:

$$\begin{aligned} WT_{t=x} &= WT_{\mathrm{t=0}} / \left(WT_{\mathrm{t=0}} + KO_{\mathrm{t=0}} {}^{*} (1+\Delta \mathrm{g})^{\mathrm{t}/2} \right) \\ KO_{t=x} &= 1 - WT_{t=x} \end{aligned}$$

where $WT_{\mathrm{t=x}}$ is the relative peak height of cytosine at day *x*, $KO_{t=\mathrm{x}}$ is the relative peak height of adenine at day *x*, and Δg is the percent difference in growth rate between Δ*pfap2-g* and E5. 95% confidence intervals were determined using 1.96*Standard error of the mean for the difference in growth rate. Replicates were biological not technical.

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed as previously described[37]. Briefly, Cultures were synchronized to late trophozoites/ schizont stage, saponin-lysed and cross-linked using formaldehyde. Nuclei were released using a Dounce homogenizer (Kimble Chase) and DNA was subsequently fragmented using a Bioruptor (Diagenode). Immunoprecipitations were carried out using commercial antibodies against H3K9ac (Millipore 07-352) and H3K9me3 (Millipore 07-442) and analyzed by qPCR using the relative standard curve method. The primers used for ChIP analysis of the *pfap2-g* locus amplify positions (relative to the start codon) −4954 to −4875 (5′-1), −1412 to −1302 (5′-2), −449 to −351 (5′-3), +3874 to +3979 (ORF-1), +5318 to +5433 (ORF-2) and +8492 to +8632 (3′-1). Primers for the control genes *clag3.1* (primer pair 5, beginning of the ORF), *clag3.2* (primer pair 5, beginning of the ORF), *ama-1* (primer pair 2, beginning of the ORF) and the *var* gene PFL1950w (upstream region, presumably 5′UTR) have been described before[37,38]. Replicates were biological not technical.

Western Blots

At the late trophozoite stage synchronized parasites were treated with 0.5μM Shld1 or ethanol solvent control for 36h. Proteins were sequentially extracted as previously described[2], separated on 4-12% polyacrylamide gels (Life Technologies), and assayed using anti-HA tag (Roche Diagnostics 11 867 423 001), anti-Histone 3 (Abcam ab1791), and anti-PfPP2c (generous gift from C. Ben Mamoun). Replicates were biological not technical.

Europe PMC Funders Author Manuscripts

Quantitative RT-PCR

PfAP2-G transcript abundance measurements were carried out as previously described[7] using a primers set designed to amplify positions 3874-3979 and normalized to seryl-tRNA synthetase abundance. Statistical significance was determined using a two-sided t-test. Replicates were biological not technical.

Immunofluorescence Assays (IFA)

Immunofluorescence assays (IFA) were performed on smears of E5-AP2-G-HAx3-9A cultures synchronized to different stages. Air-dried smears were fixed for 10 min with 1% formaldehyde and permeabilized for 10 min in 0.1% Triton X-100 in PBS. Experiments performed on smears fixed with 90% acetone / 10% methanol yielded identical results (not shown). Smears were incubated with rabbit anti-HA (1:100; Life technologies 71-5500) or rabbit anti-H3K4me3 (1:10,000; Millipore 05-745) antibodies. Secondary anti-rabbit antibodies were conjugated with Alexa Fluor 488 (Life technologies A-11034). Nuclei were stained with DAPI. Importantly, no wild-type E5 parasites were positive for staining with anti-HA antibody, and secondary antibody controls also yielded no signal. Preparations were observed under a confocal Leica TCS-SP5 microscope with LAS-AF image acquisition software and were processed using ImageJ software. The proportion of HA-positive schizonts was determined by counting >3000 schizonts (identified by DAPI staining) for each experiment. The gametocyte conversion rate was measured for each of the same parasite cultures used to quantify the proportion of schizonts positive for anti-HA by IFA. Replicates were biological not technical.

Gel Shifts

Electrophoretic mobility shift assays were performed using Light Shift EMSA kits (Thermo Scientific) as previously described using 2μg of protein and 20fmol of probe[22]. Biotinylated double-stranded probes were designed using the 24nt flanking the PfAP2-G motif of the indicated upstream sequence. Probe sequences were as follows with capital letters indicating the AP2-G motifs: *pfg27/25*: 5′-ttattagtatctGTACACattggtatttgt-3′, *pf11-1* 5′-tatatatatattGTACACatacatgtagtt-3′, *pfpeg4*: 5′-gacaataaagaaGTGTACACatatatcaataa-3′. The motifs were replaced by an equal number of adenines for “no motif” probes. Replicates were conducted using the same materials on separate days.

Microarrays

Starting at 3 hours post-invasion, tightly synchronized parasites were collected at eight time-points with 6-hour intervals. RNA isolation, cDNA generation/labeling, array hybridization, and feature extraction was performed as previously described[28]. Cy5-labeled cDNA was hybridized with a common Cy3-labeled reference pool on the *P. falciparum* 8×15K Agilent (Santa Clara, USA) nuclear expression array (GEO platform ID GPL17880). Relative transcript abundance was determined using a shared Cy3-labeled reference pool. All microarray data was submitted to the NCBI Gene Expression Ontology (GEO) repository (Series accession number: GSE52030). Genes were ordered by their average relative transcript abundance differences across the eight time-points between the WT (E5) and mutant (F12/Δ*ap2-g*). Occurrences of the trimmed (6nt) PfAP2-G motif were mapped using

ScanACE to intergenic regions up to 2000bp upstream of the start codon as previously described[22] (see Supplementary Figure 10 for motif). Significant enrichments of proteomic evidence and PfAP2-G motif occurrence was calculated using a unpaired two-sided t-test comparing the occurrences within the cluster of down-regulated genes and their frequency genome-wide. Results were validated by qRT-PCR for a subset of down-regulated genes using the primers in Supplementary Table 3 and methods described above. Statistically significant differences in relative expression levels were determined by two-sided t-test.

Luciferase assays

Equal numbers of synchronized, stably-transfected parasites were isolated and saponin-lysed (0.05% in PBS) at ~18-30 hours post-invasion and assayed using Bright-Glo Luciferase Assay System (Promega) on a Synergy H1 (Bio-Tek) plate reader as previously described[22]. Statistical significance was determined using unpaired two-sided t-tests. Replicates were biological not technical.

Next-generation sequencing and analysis

Genomic DNA was extracted (10 μg each) from E5, Δ*pfap2-g*, GNP-A4, and F12 parasite lines. This genomic DNA was used to generate barcoded sequencing libraries for an Illumina TruSeq single-end sequencing run, analyzed, and visualized as previously described[29]. Genomic DNA for 3D7A, F12 and GNP-A4 was also used for whole genome sequencing at the Sanger Institute using Illumina GA II technology with 76-base paired end reads. The raw sequence data were processed as described previously[30]. In brief, the raw data for each isolate was mapped onto the 3D7 reference genome (version 3) using the SMALT short read alignment algorithm[40]. High quality SNPs and insertions and deletions (supported by bi-directional reads, and error rates less than one per thousand base pairs) in unique genomic regions were called using *samtools* [samtools.sourceforge.net]. Regions of interest were inspected using the Artemis alignment viewer [www.sanger.ac.uk/resources/software/artemis], and polymorphisms compared to publically available sequence data[17,30,39] processed as described above. Experimental confirmation of informative genomic variants was performed using capillary sequencing methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Christian Klein, Tracey Campbell, and Ariel Schieler for technical assistance and are grateful to Oliver Billker, Christian Flueck, John Kelly, Colin Sutherland Akhil Vaidya, and Andy Waters, for helpful discussion and critical reading of the manuscript. We would also like to thank Pietro Alano for providing *P. falciparum* clone F12, Catherine Taylor for providing the *P. falciparum* GNP-A4 clone, Eloise Thompson for isolating *P. falciparum* DNA for whole genome analysis, Zaria Gorvett for assistance with confirming SNPs in gametocyte non-producing clones and generation of the complementation construct, Manoj Duraisingh for the ddFKBP tagging construct pJDD145, Choukri Ben Mamoun for the anti-PP2c antibody, and Dan Goldberg for Shld1.

ML is funded by NIH R01 AI076276 with support from the Centre for Quantitative Biology (P50GM071508). BFCK was supported by a HHMI fellowship of the Damon Runyon Cancer Research Foundation. DAB is funded by Wellcome Trust grant Ref: 094752 and European Commission FP7 ‘MALSIG’ (Ref 223044). LGD is supported

by a BBSRC Case PhD studentship with Pfizer as the Industrial partner. AC is funded by the Spanish Ministry of Science and Innovation grant SAF2010-20111. VMC was supported by a fellowship from IRB Barcelona. DPK and SGC are supported through the Wellcome Trust (098051; 090532/Z/09/Z) and the Medical Research Council UK (G0600230). CB is supported by the Catalan Government fellowship 20011-BP-B00060 (AGAUR, Spain)

References

1. Wells TNC, Alonso PL, Gutteridge WE. New medicines to improve control and contribute to the eradication of malaria. Nature reviews Drug discovery. 2009; 8:879–891.
2. Flueck C, et al. Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. PLoS Pathog. 2009; 5:e1000569. [PubMed: 19730695]
3. Lopez-Rubio JJ, Mancio-Silva L, Scherf A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. Cell Host Microbe. 2009; 5:179–190. [PubMed: 19218088]
4. Cortés A, Crowley VM, Vaquero A, Voss TS. A view on the role of epigenetics in the biology of malaria parasites. PLoS Pathog. 2012; 8:e1002943. [PubMed: 23271963]
5. Alano P. Plasmodium falciparum gametocytes: still many secrets of a hidden life. Mol Microbiol. 2007; 66:291–302. [PubMed: 17784927]
6. Dixon MWA, Thompson J, Gardiner DL, Trenholme KR. Sex in Plasmodium: a sign of commitment. Trends in Parasitology. 2008; 24:168–175. [PubMed: 18342574]
7. Rovira-Graells N, et al. Transcriptional variation in the malaria parasite Plasmodium falciparum. Genome Res. 2012; 22:925–938. [PubMed: 22415456]
8. Painter HJ, Campbell TL, Llinás M. The Apicomplexan AP2 family: integral factors regulating Plasmodium development. Mol Biochem Parasitol. 2011; 176:1–7. [PubMed: 21126543]
9. Yuda M, et al. Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. Mol Microbiol. 2009; 71:1402–1414. [PubMed: 19220746]
10. Yuda M, Iwanaga S, Shigenobu S, Kato T, Kaneko I. Transcription factor AP2-Sp and its target genes in malarial sporozoites. Mol Microbiol. 2010; 75:854–863. [PubMed: 20025671]
11. Sherman, IW. Malaria. Amer Society for Microbiology; 1998.
12. Iwanaga S, Kaneko I, Kato T, Yuda M. Identification of an AP2-family protein that is critical for malaria liver stage development. PLoS ONE. 2012; 7:e47557. [PubMed: 23144823]
13. Alano P, et al. Plasmodium falciparum: parasites defective in early stages of gametocytogenesis. Experimental Parasitology. 1995; 81:227–235. [PubMed: 7556565]
14. Day KP, et al. Genes necessary for expression of a virulence determinant and for transmission of Plasmodium falciparum are located on a 0.3-megabase region of chromosome 9. Proc Natl Acad Sci USA. 1993; 90:8292–8296. [PubMed: 8367496]
15. Eksi S, et al. Plasmodium falciparum Gametocyte Development 1 (Pfgdv1) and Gametocytogenesis Early Gene Identification and Commitment to Sexual Development. PLoS Pathog. 2012; 8:e1002964. [PubMed: 23093935]
16. Ikadai H, et al. Transposon mutagenesis identifies genes essential for Plasmodium falciparum gametocytogenesis. Proceedings of the National Academy of Sciences of the United States of America. 2013 doi:10.1073/pnas.1217712110.
17. Manske M, et al. Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. Nature. 2012; 487:375–379. [PubMed: 22722859]
18. Armstrong CM, Goldberg DE. An FKBP destabilization domain modulates protein levels in Plasmodium falciparum. Nat Methods. 2007; 4:1007–1009. [PubMed: 17994030]
19. Banaszynski LA, Chen L-C, Maynard-Smith LA, Ooi AGL, Wandless TJ. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell. 2006; 126:995–1004. [PubMed: 16959577]
20. Pradel G. Proteins of the malaria parasite sexual stages: expression, function and potential for transmission blocking strategies. Parasitology. 2007; 134:1911–1929. [PubMed: 17714601]

21. Silvestrini F, et al. Protein export marks the early phase of gametocytogenesis of the human malaria parasite Plasmodium falciparum. Molecular & Cellular Proteomics. 2010; 9:1437–1448. [PubMed: 20332084]
22. Campbell TL, De Silva EK, Olszewski KL, Elemento O, Llinás M. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS Pathog. 2010; 6:e1001165. [PubMed: 21060817]
23. Guizetti J, Scherf A. Silence, activate, poise, and switch! Mechanisms of antigenic variation in Plasmodium falciparum. Cell Microbiol. 2013 doi:10.1111/cmi.12115.
24. Pérez-Toledo K, et al. Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. Nucleic Acids Res. 2009; 37:2596–2606. [PubMed: 19270070]
25. Bruce MC, Alano P, Duthie S, Carter R. Commitment of the malaria parasite Plasmodium falciparum to sexual and asexual development. Parasitology. 1990; 100(Pt 2):191–200. [PubMed: 2189114]
26. Avraham I, Schreier J, Dzikowski R. Insulator-like pairing elements regulate silencing and mutually exclusive expression in the malaria parasite Plasmodium falciparum. Proc Natl Acad Sci USA. 2012 doi:10.1073/pnas.1214572109.
27. Fivelman QL, et al. Improved synchronous production of Plasmodium falciparum gametocytes in vitro. Mol Biochem Parasitol. 2007; 154:119–123. [PubMed: 17521751]
28. Kafsack BFC, Painter HJ, Llinás M. New Agilent platform DNA microarrays for transcriptome analysis of Plasmodium falciparum and Plasmodium berghei for the malaria research community. Malar J. 2012; 11:187. [PubMed: 22681930]
29. Straimer J, et al. Site-specific genome editing in Plasmodium falciparum using engineered zinc-finger nucleases. Nat Methods. 2012; 9:993–998. [PubMed: 22922501]
30. Robinson T, et al. Drug-resistant genotypes and multi-clonality in Plasmodium falciparum analysed by direct genome sequencing from peripheral blood of malaria patients. PLoS ONE. 2011; 6:e23204. [PubMed: 21853089]
31. Cortés A, Benet A, Cooke BM, Barnwell JW, Reeder JC. Ability of Plasmodium falciparum to invade Southeast Asian ovalocytes varies between parasite lines. Blood. 2004; 104:2961–2966. [PubMed: 15265796]
32. Walliker D, et al. Genetic analysis of the human malaria parasite Plasmodium falciparum. Science. 1987; 236:1661–1666. [PubMed: 3299700]
33. Taylor CJ, McRobert L, Baker DA. Disruption of a Plasmodium falciparumcyclic nucleotide phosphodiesterase gene causes aberrant gametogenesis. Mol Microbiol. 2008; 69:110–118. [PubMed: 18452584]
34. Maier AG, Braks JAM, Waters AP, Cowman AF. Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in Plasmodium falciparum for targeted gene deletion by double crossover recombination. Mol Biochem Parasitol. 2006; 150:118–121. [PubMed: 16901558]
35. Malaria. Vol. 923. Humana Press; 2013.
36. Calderwood MS, Gannoun-Zaki L, Wellems TE, Deitsch KW. Plasmodium falciparum var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. J Biol Chem. 2003; 278:34125–34132. [PubMed: 12832422]
37. Crowley VM, Rovira-Graells N, Ribas de Pouplana L, Cortés A. Heterochromatin formation in bistable chromatin domains controls the epigenetic repression of clonally variant Plasmodium falciparum genes linked to erythrocyte invasion. Mol Microbiol. 2011; 80:391–406. [PubMed: 21306446]
38. Jiang L, et al. Epigenetic control of the variable expression of a Plasmodium falciparum receptor protein for erythrocyte invasion. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:2224–2229. [PubMed: 20080673]
39. The Wellcome Trust Sanger Institute SRA Study ERP000190. http://www.ebi.ac.uk/ena/data/view/ERP000190
40. SMALT. Wellcome Trust Sanger Institute; http://www.sanger.ac.uk/resources/software/smalt/

Europe PMC Funders Author Manuscripts

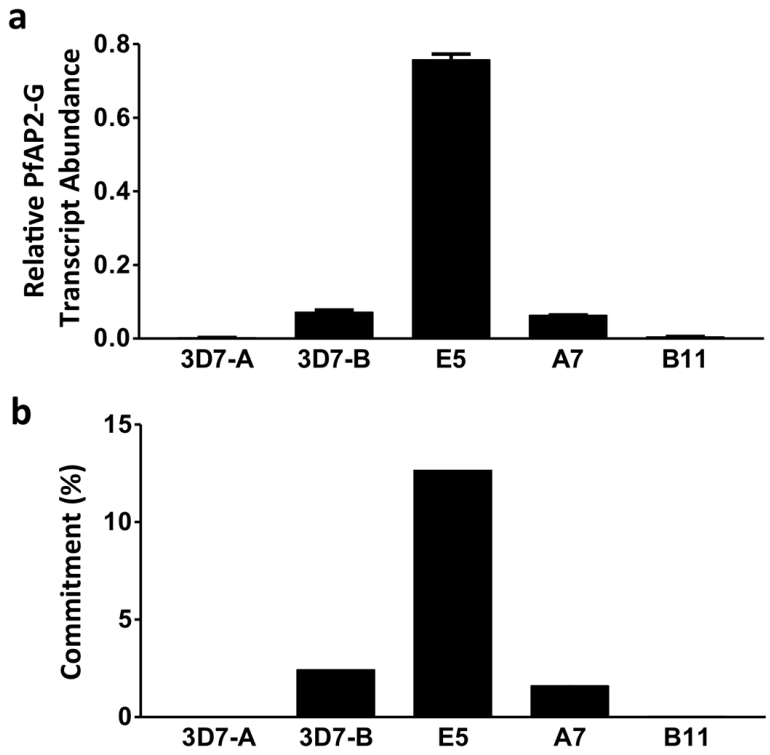

Figure 1. PfAP2-G transcript levels mirror gametocyte production

A) PfAP2-G relative transcript abundance in synchronized (early schizont stage) cultures as measured by qPCR varies significantly between 3D7-A and 3D7-B populations as well as the 3D7-B subclones E5, A7, and B11. Values are normalized against seryl tRNA synthetase (PF07_0073) (n=3, standard deviation shown). **B)** Percent commitment to gametocyte differentiation in these lines mirrors relative PfAP2-G transcript levels (mean of n=2).

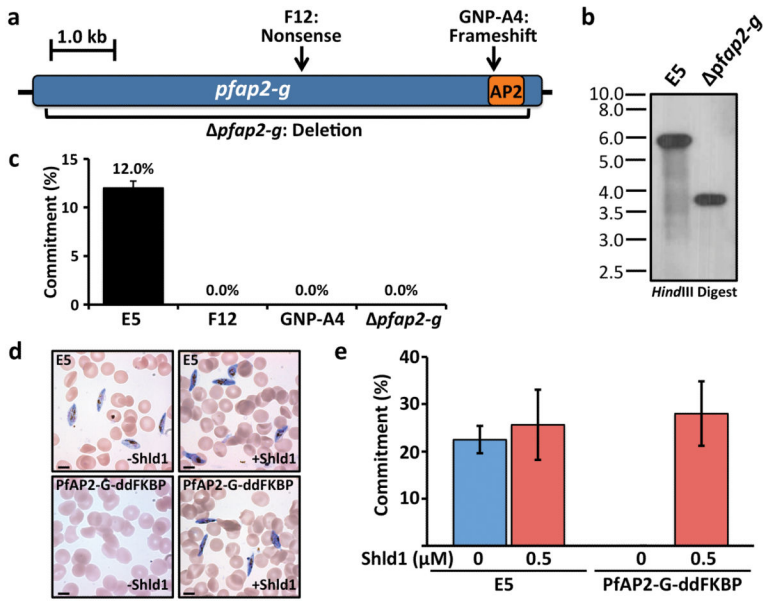

Figure 2. Disrupting PfAP2-G function results in loss of gametocyte production

A) Positions of *pfap2-g* mutations in the gametocyte non-producer lines F12 and GNP-A4 and the targeted deletion of Δpf*ap2-g*. **B)** Southern blot showing successful disruption of the *pfap2-g* locus by homologous recombination (also see Supplementary Fig. 4). Single replicate. **C)** *pfap2-g* mutants fail to produce gametocytes (n=3, standard error shown, scalebar = 5 μm). **D)** Ligand-regulatable gametocyte formation in PfAP2-G-ddFKBP (bottom row images) but not in the E5 parent (top row images) Representive of n=4. **E)** Quantification of ligand-regulatable gametocyte formation (n=4, standard error shown).

Europe PMC Funders Author Manuscripts

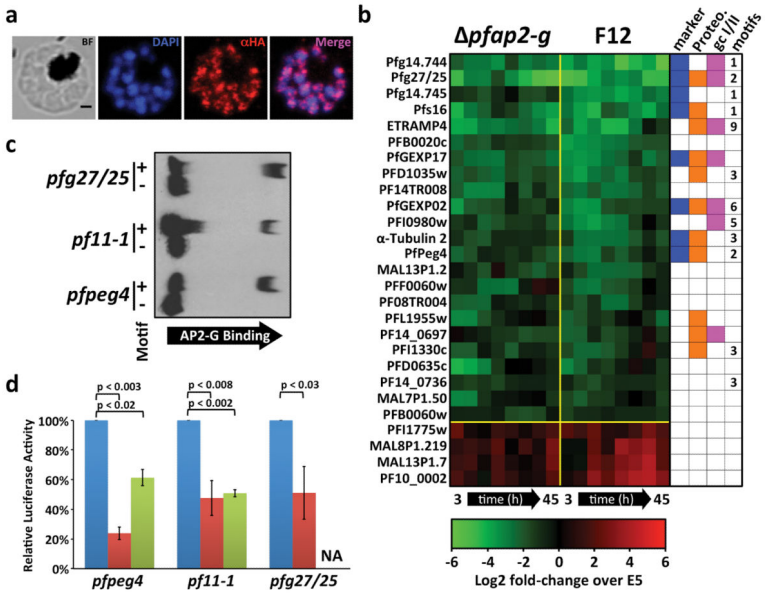

Figure 3. Identification of PfAP2-G targets

A) PfAP2-G-HAx3 localizes to the nuclei of schizonts in asexually growing parasites (see Fig S7 for additional stages, scale bar = 1μm). Representative of n=4. **B)** Relative abundance of transcripts with greater than 2-fold average difference in both Δ*pfap2-g* and F12 with respect to 3D7 clone E5 across the intra-erythrocytic developmental cycle at 6h intervals. Columns on the right indicate whether genes are known gametocyte markers (blue), detected in 2+ gametocyte proteomes (orange), enriched in early gametocyte proteome (purple), and the number of PfAP2-G cognate motifs within 2kB upstream of the start codon. **C)** Binding of recombinant PfAP2-G AP2 domain to three gametocyte promoters occurs only in the presence of the wild-type cognate motif (+). Representative of n=3. **D)** Relative luciferase activity under the control of wild-type gametocyte promoters in 3D7 E5 (blue) and Δ*pfap2-g* (red), or in 3D7 E5 under control of promoters lacking the PfAP2-G motif (green). (18-30h post-invasion, n=3, standard error is shown, two-sided t-test used, NA: Not tested.)

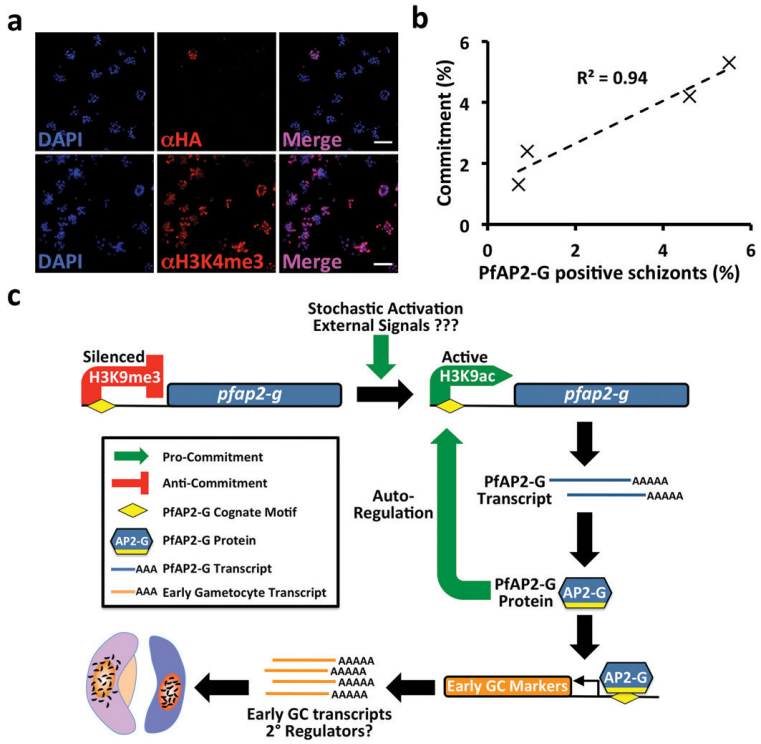

Figure 4. Activation of PfAP2-G

A) Only a small fraction (1-6%) of asexually growing subclone 9A schizonts (see Fig. S7 for details) express detectable levels of PfAP2-G-HAx3 (first row). H3K4me3 staining was performed in parallel to confirm full permeabilization (second row), scale bar = 10μm. Representative of n=3. **B)** The percentage of PfAP2-G-HAx3 positive cells is highly predictive (R^2 = 0.94) of subsequent gametocyte formation levels. **C)** Model of PfAP2-G activation and function.

Europe PMC Funders Author Manuscripts