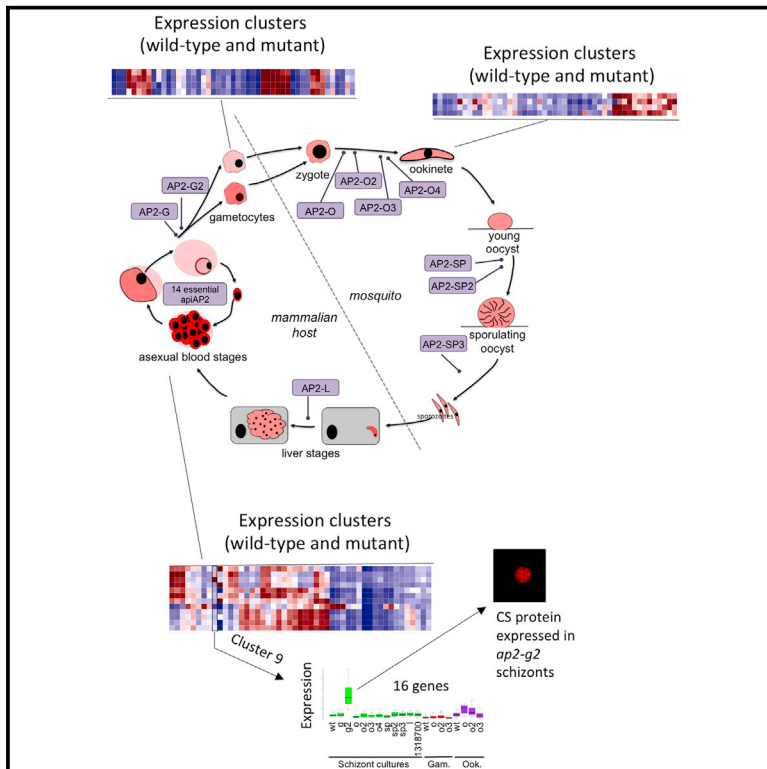# Cell Host & Microbe

A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the *Plasmodium* Life Cycle

Graphical Abstract

Authors

Katarzyna Modrzynska,
Claudia Pfander, Lia Chappell, ...,
Julian C. Rayner, Jyoti Choudhary,
Oliver Billker

Correspondence

km8@sanger.ac.uk (K.M.),
ob4@sanger.ac.uk (O.B.)

In Brief

Modrzynska et al. investigate 25 putative transcriptional regulators of the apiAP2 family in a malaria parasite using a systematic knockout screen. Cellular and molecular phenotyping of the 11 viable mutants obtained reveals complex interactions between positive and negative regulators controlling the *Plasmodium* life cycle at different stages.

Highlights

- Mutants in 11 of 26 apiAP2 genes reveal gene functions in mosquito transmission

- Co-expression clustering across mutants and stages reveals molecular phenotypes

- Multifunctional apiAP2 genes create complex regulatory networks in *Plasmodium*

- *Ap2-g2* is a transcriptional repressor in both asexual and sexual blood stages

Accession Numbers

GSE80634
PXD005180

CrossMark

CellPress

A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the *Plasmodium* Life Cycle

Katarzyna Modrzynska,[1,2,4,*] Claudia Pfander,[1] Lia Chappell,[1] Lu Yu,[1] Catherine Suarez,[1,3] Kirsten Dundas,[1] Ana Rita Gomes,[1] David Goulding,[1] Julian C. Rayner,[1] Jyoti Choudhary,[1] and Oliver Billker[1,*]
[1]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
[2]Present address: Wellcome Centre for Molecular Parasitology, Institute of Infection, Immunity, and Inflammation, University of Glasgow, 120 University Place, Glasgow G12 8TA, UK
[3]Present address: Université de Montpellier, CNRS, Dynamique des Interactions Membranaires Normales et Pathologiques, UMR 5235, F-34095 Montpellier, France
[4]Lead Contact
*Correspondence: km8@sanger.ac.uk (K.M.), ob4@sanger.ac.uk (O.B.)
http://dx.doi.org/10.1016/j.chom.2016.12.003

SUMMARY

A family of apicomplexa-specific proteins containing AP2 DNA-binding domains (ApiAP2s) was identified in malaria parasites. This family includes sequence-specific transcription factors that are key regulators of development. However, functions for the majority of ApiAP2 genes remain unknown. Here, a systematic knockout screen in *Plasmodium berghei* identified ten ApiAP2 genes that were essential for mosquito transmission: four were critical for the formation of infectious ookinetes, and three were required for sporogony. We describe non-essential functions for AP2-O and AP2-SP proteins in blood stages, and identify AP2-G2 as a repressor active in both asexual and sexual stages. Comparative transcriptomics across mutants and developmental stages revealed clusters of co-regulated genes with shared *cis* promoter elements, whose expression can be controlled positively or negatively by different ApiAP2 factors. We propose that stage-specific interactions between ApiAP2 proteins on partly overlapping sets of target genes generate the complex transcriptional network that controls the *Plasmodium* life cycle.

INTRODUCTION

Plasmodium parasites are the causative agents of malaria and have evolved complex life cycles to alternate between a vertebrate host and an anopheline mosquito vector. Asexual parasites replicating inside red blood cells (RBCs) are responsible for all disease symptoms, but successful transmission into a new host relies on a much smaller number of sexual blood stages, the gametocytes, which remain dormant until taken up by a female mosquito. In the blood meal male (micro-) and female (macro-) gametes emerge and fertilize. After ~20 hr, the

fertilized zygote transforms into a tetraploid, motile form, the ookinete, which penetrates the midgut wall and fixes itself to the basal lamina, where it turns into an oocyst. Within 2 weeks, the cyst undergoes multiple divisions generating sporozoites, small motile forms that invade the salivary glands, from which they can be transmitted to the next mammalian host. Following one round of replication in parenchymal cells of the liver, the parasites emerge into the bloodstream, thus completing the life cycle.

Stage transitions in the life cycle are accompanied by changes in gene expression, such that the majority of *Plasmodium* genes are only induced in a narrow window (e.g., López-Barragán et al., 2011; Otto et al., 2010; Zhou et al., 2008), reflecting the marked differences in cell morphology, function, and environment. The mechanisms linking gene expression and stage transitions remain poorly understood. In higher eukaryotes, the bulk of expression control happens at the level of transcription initiation and relies on sequence-specific transcription factors, which recognize their binding sequences in a promoter, thereby influencing the recruitment of the transcription preinitiation complex and RNA polymerase. The overall positive correlation between transcription and translation dynamics (Caro et al., 2014) and between mRNA and protein levels at various stages of the life cycle (Foth et al., 2011; Tarun et al., 2008) suggests that the same is likely true for *Plasmodium*. Additionally, some promoters can drive expression of a transgene with an intensity and pattern comparable to their endogenous gene, even when removed from the native genomic context (Janse et al., 2006a; Laurentino et al., 2011; Sebastian et al., 2012). However, a variety of alternative control mechanisms have been described in *Plasmodium*, which at different life cycle stages include changes in nuclear organization, chromatin modifications (Ay et al., 2015), dynamics of RNA degradation (Shock et al., 2007), and posttranscriptional repression (Mair et al., 2006). These are also likely to contribute to both the complexity and flexibility of the gene expression network.

Intriguingly, most sequence-specific transcription factor families found in other eukaryotes seem to be absent from *Plasmodium*. Instead, an expansion of a protein family containing one or more *apetala2* (AP2) DNA-binding domains was observed
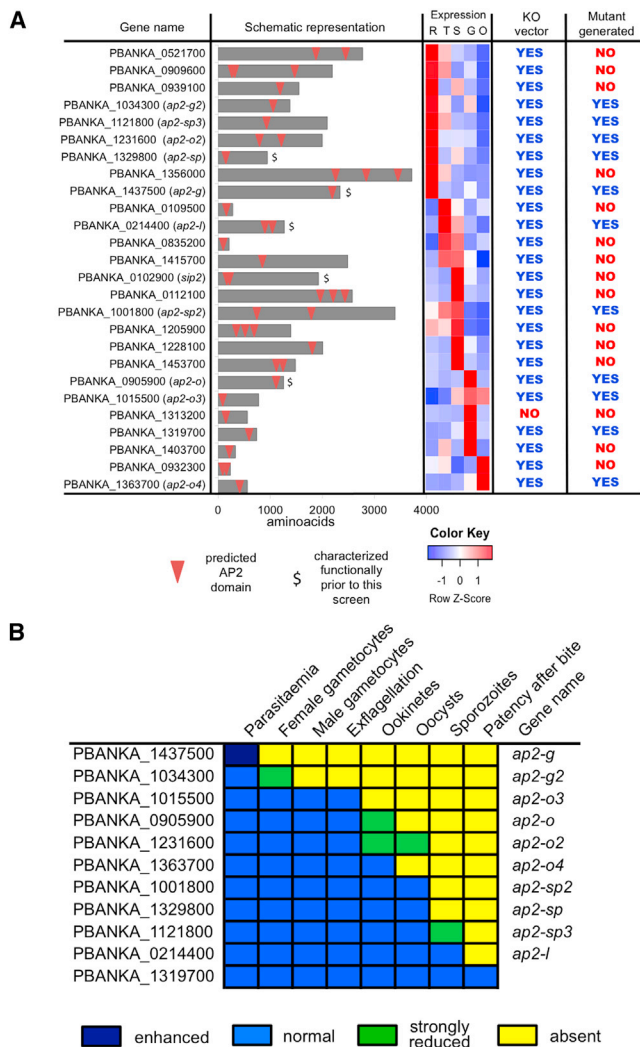
Figure 1. Systematic Phenotyping of ApiAP2 Gene KOs in *P. berghei*
(A) Schematic representation of *P. berghei* ApiAP2 genes and KO clones obtained in this study. Expression data for ring (R), trophozoite (T), schizont (S), gametocyte (G), and ookinete (O) stage are from Otto et al. (2014).
(B) Summary of a high-level phenotypic screen of eleven ApiAP2 KOs. Phenotypes were called if parasite numbers were reduced >80% or increased >20% compared to a wild-type control studied in parallel. Experiments were performed as triplicates on at least two separate occasions. Previously undescribed phenotypes were confirmed with independently created KO clones. For the complete data, see Data S1.

across the phylum apicomplexa (Balaji et al., 2005). The AP2 domain was originally identified in plants, where it is a defining feature of a major family of AP2/ERF proteins, which control many key aspects of plant biology by acting as repressors or activators of transcription (Licausi et al., 2013). AP2 domain-containing (ApiAP2) proteins are present in all apicomplexa studied so far (Oberstaller et al., 2014), and besides one or more 60 aa AP2 DNA-binding domains, they rarely contain other known functional features. In total, 27 members of this family have been found in the human malaria parasite *Plasmodium falciparum* (although a possible 28th member of the family may be present; Oberstaller et al., 2014). In total, 26 of these have syntenic

orthologs in rodent malaria species, each with its unique stage-specific expression profile (Figure S1, available online).

Recombinantly expressed *P. falciparum* AP2 domains bind specifically to a large variety of DNA sequences. Putative binding motifs were identified upstream of the majority of genes, making ApiAP2s the main candidates for generating stage-specific patterns of gene expression (Campbell et al., 2010). So far, six members of this family have been studied functionally in malaria parasites of either humans or rodents. Five were shown to play key roles in parasite progression through the life cycle and were crucial for gametocytogenesis (*ap2-g* and *ap2-g2*; Kafsack et al., 2014; Sinha et al., 2014; Yuda et al., 2015), ookinete development (*ap2-o*; Yuda et al., 2009), sporozoite formation (*ap2-sp*; Yuda et al., 2010), and liver stage maturation (*ap2-l*; Iwanaga et al., 2012). A sixth member of the family (*pfsip2*) was identified as a chromatin tethering protein, with a likely role in telomere biology in *P. falciparum*, and the overexpression of its N-terminal part containing the DNA-binding domains had no effect on gene expression (Flueck et al., 2010). However, the functions of most members of the family remain unknown.

Here we present a systematic knockout (KO) screen targeting the ApiAP2 family in the rodent malaria parasite *Plasmodium berghei*. Phenotyping of eleven viable ApiAP2 KO mutants reveals ten critical gene functions at different points in the life cycle. RNA sequencing (RNA-seq) analysis of the selected mutants at different life stages identifies changes in the transcriptomes associated with the early transmission phenotypes, and reveals clusters of functionally connected and co-regulated genes that likely share mechanisms of stage-specific control, and whose promoters we investigated for putative regulatory motifs. These data generate a deeper understanding of individual ApiAP2 proteins and how positive and negative regulators likely interact to generate complex patterns of gene expression. To test just one of the hypotheses arising from the screen, we present data supporting a general function *ap2-g2* as a repressor of transmission-specific genes, i.e., not only in gametocytes (Yuda et al., 2015) but also in asexual blood stages.

RESULTS

ApiAP2 Mutants Reveal Stage-Specific Gene Functions
All 26 ApiAP2 genes of *P. berghei* possess syntenic orthologs in *P. falciparum*, *P. vivax* (main malaria species infecting humans), and *P. yoelii* (another rodent malaria model). While across the family there is much variation in gene size, numbers of predicted AP2 domains, and expression patterns, these parameters are largely conserved within orthologous groups, suggesting a high degree of functional conservation (Figures 1A and S1).

To target *P. berghei* ApiAP2 genes systematically, we succeeded in producing deletion vectors for all but one member of the family (Data S1) and transfected each of them into a reporter line constitutively expressing GFP to facilitate phenotyping. Fourteen ApiAP2 genes resisted at least four disruption attempts with up to two different vector designs, providing tentative evidence that more than half of the genes in this family are potentially essential for asexual blood stage growth in vivo. For the remaining eleven, a KO line could be generated. These included six genes that had not previously been studied in *Plasmodium* (Figures 1A and S2; Data S1).

Ten of these lines showed complete developmental lethality at some stage during sexual development or mosquito transmission (Figure 1B; Data S1). Mutants in five of these genes have been described previously. For the remainder, we followed the convention of naming genes according to the life cycle stage at which the lethal phenotype was observed (ap2-o2, ap2-o3, and ap2-o4 for ookinete phenotypes and ap2-sp2 and ap2-sp3 for sporogony phenotypes). PBANKA_1319700 was the only gene whose deletion did not reveal a function, indicating the overall level of redundancy within the ApiAP2 family is low.

To measure effects of ApiAP2 gene deletions on asexual growth rate, we co-transfected multiple barcoded gene deletion vectors and determined the average relative growth rate of mutants during days 5–8 post-transfection by barcode counting (barseq) (Figure 2A). We used a competitive design with seven control vectors that cause known changes to parasite growth (Gomes et al., 2015). The ap2-g KO vector conveyed enhanced growth, as was expected, since wild-type parasites sacrifice some of their asexual reproductive potential to make gametocytes (Sinha et al., 2014). Four other mutants showed attenuated growth in asexual blood stages, including ap2-o and ap2-sp, both recognized so far only for their roles in mosquito stages.

Gametocyte number is a difficult phenotype to measure precisely in *P. berghei*, since the rate at which sexual stages are produced differs substantially between infections by the same clone and between days of the same infection (see, e.g., Sinha et al., 2014). Unsurprisingly, both parameters differed somewhat between most mutants, although they remained within the normal range in all but two mutants, ap2-g and ap2-g2, whose complete loss or severe reduction in gametocytemia has been described previously (Sinha et al., 2014). We also found no evidence for an ApiAP2 gene specifically regulating sex ratio, which in haploid *Plasmodium* blood stages is variable, but generally female biased, and not the result of chromosomally encoded sex-determining genes (Paul et al., 2000).

ApiAP2 Genes Required for Infectious Ookinetes
Four mutants were characterized by completely or strongly reduced numbers of oocysts on the midgut epithelium of *Anopheles stephensi* mosquitoes (Figure 2B). All had gametocyte numbers and sex ratios that would not limit transmission (Figure 2C). Male gametocytes appeared mature since they readily differentiated into flagellated gametes when their development was triggered in vitro, a process termed exflagellation (Data S1), and light microscopy of activated gametocytes showed evidence of fertilization. However, zygotes of these mutants failed to varying degrees to differentiate into fully functional ookinetes in vitro (Figures 2D, 2E, S3A, and S3B). This group included ap2-o, in which zygotes were arrested at an intermediate stage of ookinete formation, as described previously (Yuda et al., 2009). Two additional mutants, ap2-o2 and ap2-o3, also gave rise to retort-shaped intermediate ookinete forms (Figures 2D and 2E). Mutant ap2-o3 showed the most severe defect in the ookinete group, with the majority of zygotes failing entirely to show apical complex protrusions. By comparison, ap2-o2 produced at least some morphologically normal ookinetes in vitro and was presumably responsible for the few oocysts seen in vivo. These cysts were, however, unusually small (Figure 2B) and unable to sporulate (data not shown).

In marked contrast, ap2-o4 mutant clones produced normal numbers of morphologically mature ookinetes in vitro (Figures 2D and 2E), but they were still entirely unable to infect mosquitoes and form oocysts in vivo (Figure 2B). Neither the gliding speed of ap2-o4 ookinetes nor the shape of their trajectories in matrigel motility assays (Moon et al., 2009) differed from wild-type (Figures S3A and S3B), raising the possibility that their developmental defect is due to either a block at the point of midgut traversal, as has been observed with mutants in secreted traversal proteins (e.g., Dessens et al., 2001), or to reduced survival of ap2-o4 ookinetes under in vivo conditions in the blood meal, where oxidative stress is high (Turturice et al., 2013) and availability of nutrients presumably reduced compared with optimized culture conditions (Sturm et al., 2015). None of the four ookinete stage mutants produced sporozoites in mosquito salivary glands; neither could they be transmitted back to mice by mosquito bite (Data S1).

ApiAP2 Genes Essential for Sporogony
Three mutants produced oocysts that were indistinguishable from wild-type in size and numbers, but failed to colonize the salivary glands with sporozoites (Figure 3A). The earliest phenotype occurred in ap2-sp oocysts, which were unable to enter sporogony, as described previously (Yuda et al., 2010; Figure S3C). In contrast, ap2-sp2 oocysts were blocked later, at the sporoblast stage (Figures 3B and 3C). These oocysts ceased to mature at a point when sporozoites were beginning to form around the edges of cytoplasmic islands (Figure 3D) and remained arrested at this point until at least day 22 post-infection, resulting in the complete absence of salivary gland colonization (Data S1).

Mutants in ap2-sp3 completed sporogony normally, as judged by cytosolic GFP fluorescence and transmission electron microscopy (TEM) (Figure S3D), yet sporozoites failed to appear in the salivary glands (Figure S3E). In contrast to a known egress mutant lacking the cysteine protease ECP1/SERA8 (Aly and Matuschewski, 2005), ap2-sp3 sporozoites remained non-motile within cysts, and when released mechanically and injected into mice, were non-infectious (Figure 3E), suggesting ap2-sp3 is required before egress to ensure full maturation of oocyst sporozoites. Interestingly, the small number of ap2-sp3 sporozoites that reached the salivary glands still remained non-infectious when mosquitoes fed on naive C57BL/6 mice (Figure 3E). The only mutant that produced normal salivary gland infections but subsequently failed to infect mice was ap2-l (Data S1), consistent with its known essential role at the liver stage (Iwanaga et al., 2012).

Gene Expression Changes in ApiAP2 Mutants
To reveal molecular functions of ApiAP2 genes, we studied the transcriptomes of all viable mutants by directional RNA-seq. *P. berghei* blood stage infections are asynchronous, and we therefore focused our analysis on schizonts synchronized in 22 hr in vitro cultures. It is important to note that schizont cultures also contained a small number of gametocytes (2%–10%). For three mutants blocked at the ookinete stage, we additionally investigated transcriptomes of gametocytes purified from peripheral blood, and of ookinete cultures harvested 20 hr after the initiation of fertilization in vitro, expecting these samples to
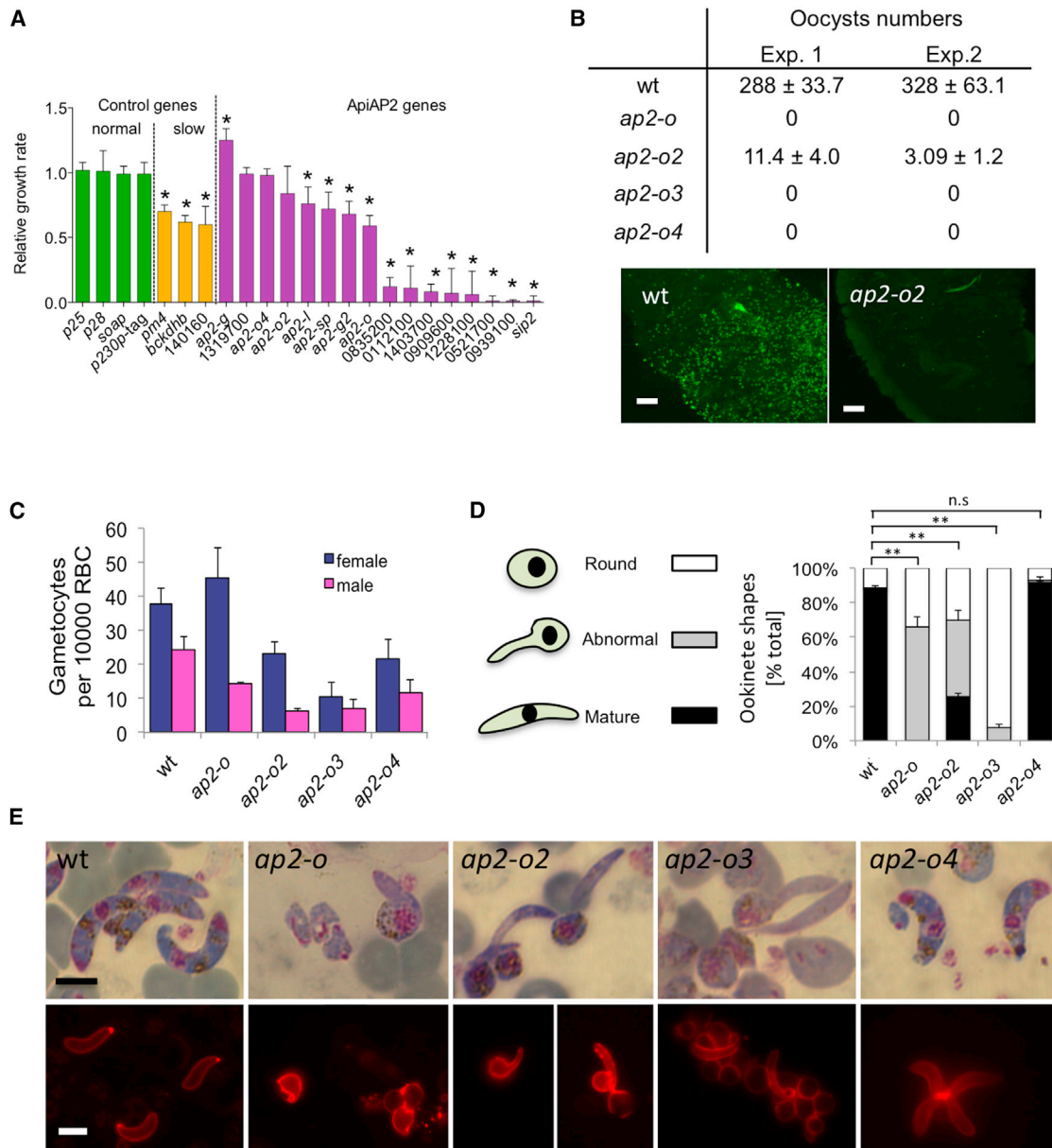
Figure 2. STM Analysis of ApiAP2 Family and Detailed Phenotyping of Four Mutants that Fail to Produce Infectious Ookinetes

(A) Competitive fitness of 16 ApiAP2 KOs (purple) as determined by barcode counting following co-transfection of barcoded vectors. Growth rates were determined from barcode counts and are expressed relative to four mutants with wild-type growth (green). Known attenuated control mutants are shown in orange. Relative growth rates are given as the arithmetic mean of measurements taken on days 5, 6, 7, and 8 post-transfection and error bars show SDs from three transfections. Asterisk (*) indicates significantly different from the pool of reference vectors with no growth effect ($p < 0.05$).

(B) Oocysts per midgut 7 days after mosquitoes had fed on infected mice. At least 15 mosquitoes were dissected per group and arithmetic means ± SE are shown.

(C) Male and female gametocytemia determined on Giemsa-stained blood films on day 3 post-infection. Error bars show SDs from three infected mice.

(D) Distribution in 20 hr ookinete cultures of different parasite forms expressing the macrogamete/ookinete marker P28. At least 100 P28-positive cells were counted in each sample. Error bars show SDs from three cultures.

(E) Sexual stage morphology 20 hr after gametocyte activation in culture as seen on Giemsa-stained blood films (top panels) and by live IFA using a Cy3-conjugated P28 antibody (bottom panels). Scale bar, 5 μm.

provide insights into the molecular basis of the developmental phenotypes of these mutants.

In a principal component (PC) analysis, transcriptomes clustered primarily by life cycle stage, and 76% of the variance was captured in the first two components (Figure 4A). Schizont samples grouped together with some spread along PC1, presumably reflecting natural variations in gametocyte numbers. Outliers were gametocyte non-producers (situated farthest
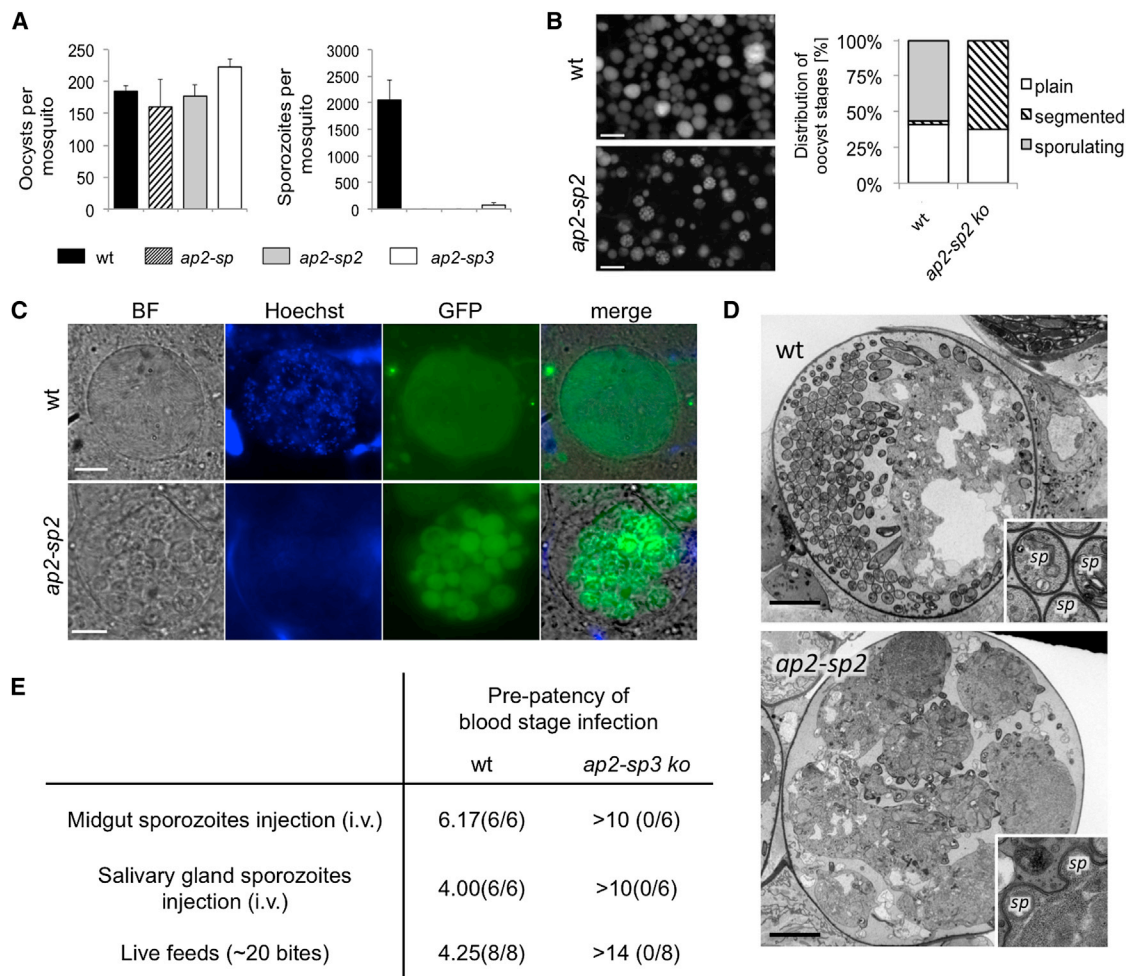
Figure 3. In-Depth Phenotypic Analysis of Three ApiAP2 Genes Required for Sporozoite Production

(A) Midgut oocysts and salivary gland sporozoites in mosquitoes fed on infected mice. Error bars show SDs from 15 midguts (day 7 post-feeding) or 3 batches of 10 pairs of glands (day 21 post-feeding), respectively.

(B) Representative fluorescence micrographs of wild-type and *ap2-sp2*-infected mosquito midguts showing different stages of sporogony as patterns of cytosolic GFP expression in oocysts on day 12 post-feeding (left; scale bar, 20 μm). The bar chart shows a quantitation of sporogonic stages averaged from ten midguts from two independent experiments.

(C) Light and fluorescence micrographs of typical oocysts showing incomplete sporogony in *ap2-sp2*. Hoechst 33342 dye shows DNA. Scale bar, 5 μm.

(D) TEMs showing normal sporogony in wild-type oocysts (upper panel) and incomplete budding of *ap2-sp2* sporozoites from sporogonic islands (lower panel). Scale bar, 5 μm.

(E) Prepatency of blood stage infections in days following administration of sporozoites by intravenous injection, or by mosquito bite. Pooled data from independent experiments using two different *ap2-sp3* clones are shown. Prevalence is shown in brackets (infected mice/total).

away from purified gametocytes) and *ap2-sp* cultures, which were shifted toward the purified gametocytes, although gametocyte numbers in this mutant were not increased (Data S1). Gametocyte transcriptomes of the ookinete mutants grouped together, but the ookinete cultures were clearly set apart from wild-type and from each other, recapitulating when phenotypes first became apparent morphologically.

Pairwise comparisons confirmed the PC analysis, showing gametocyte genes were significantly downregulated in blood stage cultures of *ap2-g* and *ap2-g2*, but up in *ap2-sp* (Data S2). Deletion of *ap2-o* had a remarkable and unexpected effect on the transcriptome of schizont cultures, with more than 265 differentially expressed (DE) genes, potentially related to the

slow growth of this mutant in mice (Figure 2A). The deregulation signature did not come from gametocytes, which had normal transcriptomes (Figure 4B). It was not enriched in known *ap2-o* targets (Kaneko et al., 2015) but was, e.g., enriched in putative RNA-binding factors (p = 0.043; Data S2). In *ap2-o2* and *ap2-o3*, the numbers of DE genes increased with developmental stage. Interestingly, in these two mutants changes in gene expression become apparent already in gametocytes, suggesting that the developmental defects observed after fertilization may have their origins in transcriptome deregulation at the earlier stage. As expected, the transcriptomes of all three ookinete mutants were significantly deregulated in the ookinete cultures (5%–15% of DE genes; Data S2; Figure 4B), but despite their
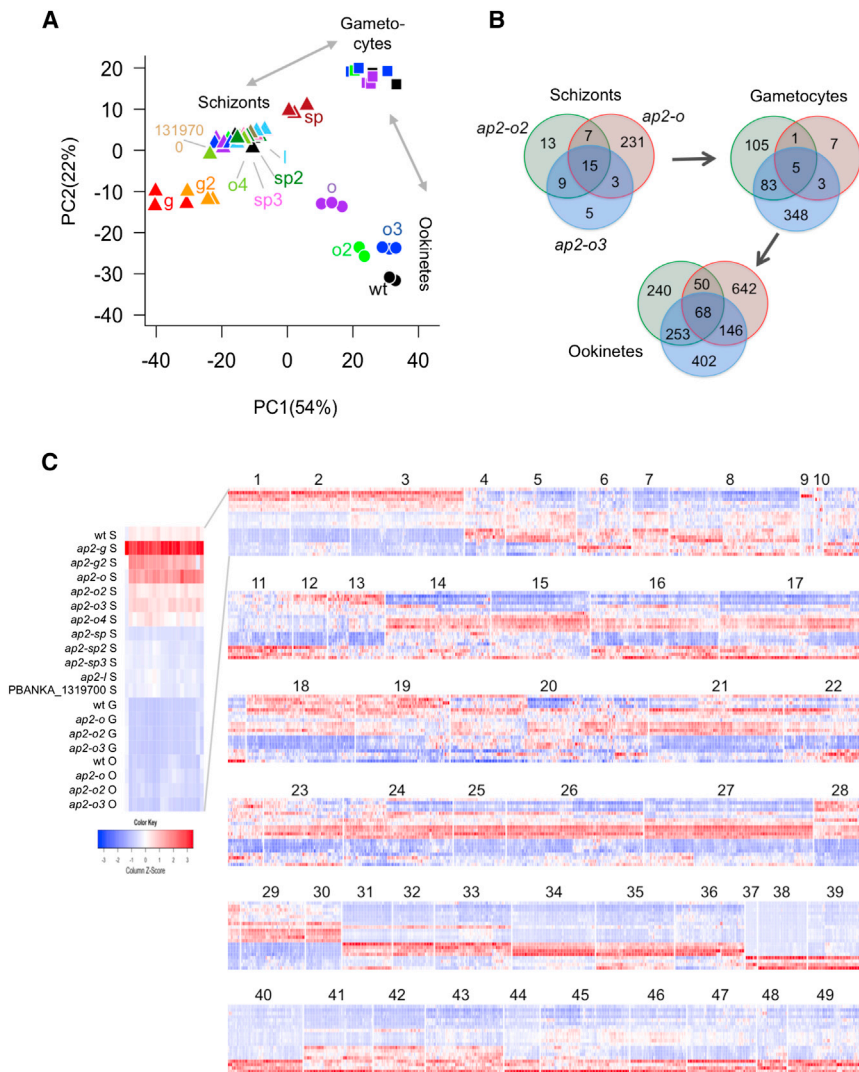
Figure 4. Comparative Transcriptomics of ApiAP2 KOs at Different Life Cycle Stages

(A) PC analysis of all transcriptome samples. Colors indicate strains, and shapes indicate life stages.

(B) Number of differentially expressed genes (absolute fold change > 2 and adjusted p value < 0.01) in three ookinete mutants at different life stages.

(C) Heatmap showing the expression of 4751 *P. berghei* genes through different strain/stage combinations. The gene transcripts are ordered according to their expression patterns using Ward hierarchical clustering. The borders of 49 gene expression clusters are marked.

similar cellular phenotypes, there was very little overlap in deregulated genes and molecular pathways (Figure 4B; Data S2).

Co-expression Clustering Identifies Coordinated Responses to Perturbations in the ApiAP2 Network

Co-expression clustering in *P. falciparum* has previously identified groups of functionally related genes that share conserved upstream sequences (e.g., Elemento et al., 2007; Young et al., 2008). Some of these motifs were later shown experimentally to serve as *cis*-regulatory elements recognized by specific AP2 domains (Campbell et al., 2010). Extending this concept, we wished to identify gene sets through their shared response to genetic perturbation of the ApiAP2 repertoire. Following Ward hierarchical clustering, 49 co-expression clusters were defined, each containing between 9 and 249 genes (Figure 4C; Data S3). Nearly all clusters contained functionally related genes, as evidenced by significant enrichment of gene ontology (GO) terms or metabolic pathways in 48 out of 49 clusters (Data S3 and S4). In addition, in 18 clusters we identified a total of 27 putative *cis*-regulatory elements of 8–12 bp that were significantly enriched

(E value < 0.05) within 2 kb upstream of the start codon (Figure 5; Data S4).

Differential gene expression in the ookinete mutants proved a major driver for clustering. A total of 227 genes with ookinete-specific expression were distributed among four clusters (37, 38, 39, and 47) according to their degree of dependence on *ap2-o*, *ap2-o2*, and *ap2-o3* (Figure 5; Data S4). Cluster 37, for instance, contained 14 genes encoding mostly well-known ookinete microneme proteins, whose expression was entirely dependent on *ap2-o*. All four clusters were significantly enriched for variants of the known AP2-O-binding motif (E = 1.9×10^{-4} to 2.4×10^{-12}) and for genes with published chromatin immunoprecipitation (ChIP) evidence (Kaneko et al., 2015) of AP2-O binding in ookinetes (p = 7.8×10^{-7} to 1.6×10^{-29}). Although not all genes upregulated in ookinetes were *ap2-o* dependent (see cluster 40 in Figure 5), these data generally confirmed the central role for *ap2-o* in ookinete gene expression (Yuda et al., 2009).

By comparison, neither *ap2-o2* nor *ap2-o3* was strictly essential for any cluster of genes. A moderate downregulation of most ookinete-specific genes in the *ap2-o2* KO may only reflect the reduced number of mature ookinetes in these cultures. In contrast, *ap2-o3* clearly differentiated between groups of genes, some with well-characterized functions in ookinete gliding and invasion (compare, for instance, clusters 37 and 38 in Figure 5). Unfortunately, the known binding motif of one of the two AP2-O2 DNA-binding domains (TGACATCA) was not associated with any of the affected gene expression clusters, and the binding specificity of AP2-O3 is not known, making it impossible to predict direct targets of AP2-O2 or AP2-O3 proteins.

One clue for a function for *ap2-o3* in ookinetes is provided by cluster 31 (Figure 5), which is highly enriched (p = 4.6×10^{-44}, 53 of 62 elements in the cluster) in genes known to be transcribed in macrogametocytes, but whose mRNAs are translationally repressed until gametocytes enter the mosquito (Mair et al., 2006). This cluster includes genes for the major ookinete surface
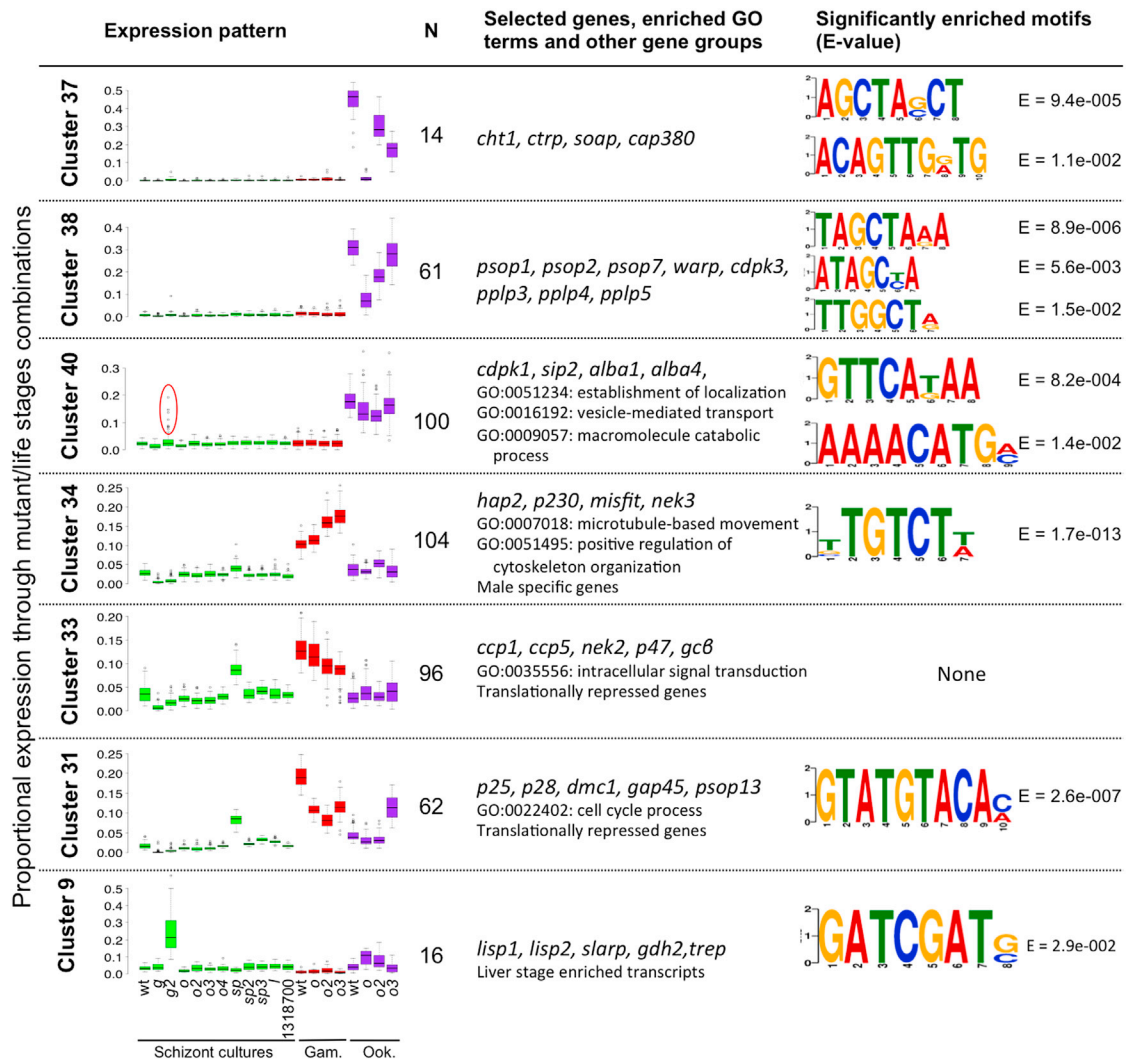
Figure 5. Characterization of Selected Co-expression Clusters
For each stage and mutant, the average relative expression across all genes in a cluster is given. N gives the number of genes in a cluster. GO terms and functional gene groups are shown if significantly enriched (false discovery rate [FDR]-adjusted p value < 0.05). Sequence motifs shown are significantly enriched within 2 kb upstream of the start codon, as compared with all genes. See Data S3 and S4 for a comprehensive analysis of all clusters.

antigens P25 and P28, but also the putative meiotic recombination protein Dmc1. For these genes, mRNA abundance usually peaks in gametocytes, but in ookinete cultures of *ap2-o3* they fail to get downregulated, which could be the result of reduced transcript turnover or a failure to repress transcription. Other genes failing to be repressed in *ap2-o3* ookinete cultures are involved in energy metabolism (cluster 45), which in the ookinete gets reorganized, and in nuclear division (cluster 31), suggesting *ap2-o3* may be required for zygotes to progress successfully beyond meiosis.

Importantly, the developmental phenotypes of *ap2-o2* or *ap2-o3* mutants at the point of ookinete formation were associated with perturbed transcriptomes already at the preceding gametocyte stage (Figure 4B). Genes upregulated in *ap2-o2* or *ap2-o3* gametocytes fell into three expression clusters (34, 35, and 36), all strongly enriched in genes encoding the male-specific gametocyte and gamete proteomes (Khan et al., 2005;

Tao et al., 2014). Genes in these clusters included the male gamete fertilization factors P48/45 and HAP2, as well as putative components of the flagellar axoneme, which were all independent of *ap2-o*. Sex ratios observed on stained blood films did not support a marked shift toward male gametocytes in *ap2-o2* and *ap2-o3* mutants (Figure 2C), and these data therefore raise the possibility that dysregulation of male-specific genes translates into differences in fertility or post-fertilization development. An AGACA motif enriched in two of the male gene clusters (Figure 5; Data S4) was found previously in male genes (Young et al., 2008), but does not resemble any of the motifs known to be recognized by recombinant AP2 domains.

A Putative Role for ap2-sp in Macrogametocyte Gene Repression

A notable feature of macrogametocyte genes in clusters 33 and 42 is their marked upregulation in schizont cultures of *ap2-sp*,
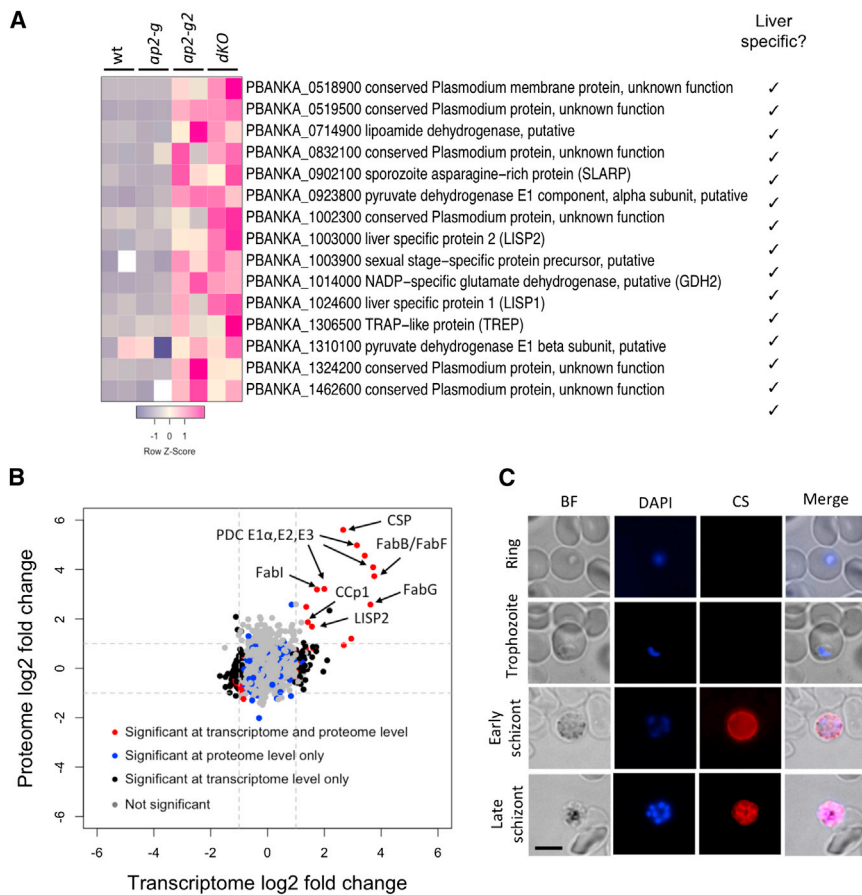
A

	wt	ap2-g	ap2-g2	dKO		Liver specific?
					PBANKA_0518900 conserved Plasmodium membrane protein, unknown function	✓
					PBANKA_0519500 conserved Plasmodium protein, unknown function	✓
					PBANKA_0714900 lipoamide dehydrogenase, putative	✓
					PBANKA_0832100 conserved Plasmodium protein, unknown function	✓
					PBANKA_0902100 sporozoite asparagine–rich protein (SLARP)	✓
					PBANKA_0923800 pyruvate dehydrogenase E1 component, alpha subunit, putative	✓
					PBANKA_1002300 conserved Plasmodium protein, unknown function	✓
					PBANKA_1003000 liver specific protein 2 (LISP2)	✓
					PBANKA_1003900 sexual stage–specific protein precursor, putative	✓
					PBANKA_1014000 NADP–specific glutamate dehydrogenase, putative (GDH2)	✓
					PBANKA_1024600 liver specific protein 1 (LISP1)	✓
					PBANKA_1306500 TRAP–like protein (TREP)	✓
					PBANKA_1310100 pyruvate dehydrogenase E1 beta subunit, putative	✓
					PBANKA_1324200 conserved Plasmodium protein, unknown function	✓
					PBANKA_1462600 conserved Plasmodium protein, unknown function	✓

-1 0 1
Row Z-Score

B

C

Figure 6. A Double Mutant with *ap2-g* Identifies *ap2-g2* as Repressor in Asexual Blood Stage

(A) Heatmap showing expression of cluster 9 genes in blood stage cultures. Row-scaled reads per kilobase of transcript per million mapped reads (RPKM) counts for two replicates are shown for each parasite line: wild-type, both single mutants, and the *ap2-g/ap2-g2* double mutant (dKO). Genes previously identified as strongly upregulated at the liver stage (Tarun et al., 2008) are marked.

(B) Assessing the impact on schizont proteome and transcriptome of deleting *ap2-g2* in a gametocyte-deficient *ap2-g* mutant. Normalized transcript and protein ratios for 2,169 genes detected as both transcript and protein are plotted against each other. See Data S5 for the underlying datasets.

(C) Fluorescence micrographs showing aberrant expression and localization of the sporozoite surface protein CSP in schizonts of the *ap2-g/ap2-g2* double mutant. Fixed and permeabilized infected RBCs were stained with Cy3-conjugated CSP monoclonal antibody. Schizonts from an ap2-g2 single mutant also expressed CSP, while no fluorescence was observed in *ap2-g* schizonts (Figure S6A). Scale bar, 5 μm.

which accounts for the more gametocyte-like transcriptomes of these samples (Figure 4A). In addition to its established role as direct activator of sporozoite genes in oocysts (Yuda et al., 2010), *ap2-sp* may function as a negative regulator of female-specific genes in gametocytes, or possibly even in asexual blood stages, which could explain the mild growth defect associated with the *ap2-sp* KO vector in the barseq experiment (Figure 2A). It remains to be established how *ap2-sp* could perform such additional roles, since the well-characterized binding motif targeted by its only AP2 domain is not enriched in gametocyte gene clusters (De Silva et al., 2008). Regulation of macrogametocyte genes by *ap2-sp* may be indirect, possibly involving another ApiAP2 gene. A strong candidate is PBANKA_1313200, whose expression peaks in gametocytes, where its mRNA associates with the translation repression complex until activation (Mair et al., 2006), and which is 5-fold overexpressed in *ap2-sp* blood stage cultures (Figure S4). PBANKA_1313200 has a single AP2 domain of unknown specificity, as the previous work failed to identify a putative binding motif for its *P. falciparum* ortholog, PF3D7_1449500 (Campbell et al., 2010). This is the only ApiAP2 member for which we failed to generate a targeting vector.

AP2-G2 Is a General Repressor of Transmission Genes in Blood Stages

Disruption of *ap2-g2* results in a loss of mature gametocytes (Figure 1B) and in reduced competitive growth of asexual blood stages (Figure 2A). It was also accompanied by a marked upregu-

lation of sporozoite and liver stage genes in blood stage cultures in two previously published datasets (Sinha et al., 2014; Yuda et al., 2015). Transcriptional phenotyping, chromatin-binding studies, and reporter assays all converged to show that in *P. berghei* gametocytes, AP2-G2 represses genes from other life cycle stages, including *lisp2* and *msp1*, by binding directly to the promoter region of its target genes (Yuda et al., 2015). The inappropriate derepression in the KO gametocytes may explain its developmental phenotype, if it causes gametocytes to degenerate before they become functionally and morphologically mature.

Consistent with this hypothesis, our current RNA-seq analysis shows sporozoite and liver stage genes repressed by *ap2-g2* strongly enriched in cluster 9, with additional deregulated genes appearing among the ookinete genes in cluster 40 (Figure 5, circle). One of the upregulated genes, however (encoding the circumsporozoite protein [CS], the major surface protein of sporozoites), was so abundant in *ap2-g2* schizont cultures that we suspected transcriptional dysregulation was not limited to the small number of "contaminating" gametocytes but more likely extended to the schizonts that accounted for the majority of parasites in the cultures. Importantly, this could help explain the reduced asexual growth rate of the *ap2-g2* mutant in vivo.

To address this possibility, we generated a double mutant with *ap2-g* (Figure S5), reasoning that this would block sexual development already at the point of commitment, such that all remaining deregulation would have to originate from asexual blood stages. The double mutant lacked gametocytes, as expected, but dysregulation of cluster 9 sporozoite and liver stage genes was reproduced (Figure 6A). Furthermore, despite the absence of gametocytes in the *ap2-g* genetic background, we additionally

observed upregulation of gametocyte-specific and ookinete-specific genes upon deleting *ap2-g2* (Data S5). Importantly, transcriptional derepression of these genes led to expression of the corresponding proteins in asexual cultures, as shown by a proteometric analysis (Figure 6B; Data S5). Ectopically expressed proteins in double KO schizonts included CS; the gametocyte surface protein CCP1; the liver stage exported protein, LISP2 (Orito et al., 2013); nearly the complete type II fatty acid biosynthesis pathway of the apicoplast, including three components of the pyruvate dehydrogenase complex (PDC); as well as FabI, FabG, and FabB/FabF proteins, which in wild-type parasites are essential only in liver stages (Yu et al., 2008). Double KO schizonts produced substantial amounts of CS protein, which ectopically localized to the merozoite surface (Figure 6C). CS was not expressed in the parental *ap2-g* line (Figure S6A). CS was also absent from ring and trophozoite stages of the double mutant (Figure 6C), showing that some stage-specific control was maintained, probably involving additional transcriptional regulators other than AP2-G2. Taken together, these data showed that AP2-G2, in addition to its role in the gametocyte development, is important for transcriptional repression in schizonts, but that its targets differ significantly at that stage.

DISCUSSION

In a systematic KO screen, we have here revealed new cellular functions for five ApiAP2 genes, doubling the number of family members known to be required for parasite transmission by mosquitoes. Additionally, 14 genes that have resisted disruption are likely to function in asexual replication and can now be investigated further using conditional knockdown approaches that have recently become available in *P. berghei* (Philip and Waters, 2015; Pino et al., 2012). Additional work will be required to define more precisely the primary targets of *ap2-o2*, *ap2-o3*, *ap2-o4*, *ap2-sp2*, and *ap2-sp3*; their DNA-binding specificities; and how these relate to the cellular phenotypes of the mutants. However, the comparative analysis of how stage-specific transcriptomes are perturbed by the disruption of different ApiAP2 genes already reveals an unexpected complexity in gene regulatory networks in *Plasmodium*.

AP2 Genes Act Together to Create Complex Patterns of Gene Expression

In contrast to what the single-gene KO studies seemed to indicate, drastic changes in gene expression in *Plasmodium* emerge not merely from a small number of stage-specifically expressed transcriptional activators, but rather from the net effect of different DNA-binding proteins acting in concert. This is perhaps best illustrated by the role of AP2-G2 as a versatile repressor in different parasite stages. In gametocytes, for instance, repression of one set of genes by AP2-G2 presumably needs to be overcome by AP2-G during sexual differentiation, but other activating ApiAP2 proteins need to overcome repression of other genes in asexual blood stages. Other ApiAP2 proteins with pleiotropic functions include *ap2-o* and *ap2-sp*, which we here show to contribute to normal gene expression and growth in asexual blood stages, in addition to their published functions as direct transcriptional activators in ookinetes and sporozoites, respectively (Yuda et al., 2009, 2010). This is reminiscent of model eu-

karyotes, where complex gene expression patterns are the result of hierarchical networks, within which transcription factors function in combination and sometimes synergistically (Gerstein et al., 2012; Levo and Segal, 2014). In such networks, both recognition of the binding site and the direction and magnitude of the resulting regulation depend strongly on the context provided by other binding sites (Stampfel et al., 2015), adding to the plasticity of the whole system. It appears that the *Plasmodium* transcriptome is regulated in a similar way and that ApiAP2 proteins are key factors involved in this process.

A particularly striking example of a set of transcripts that change in a large number of ApiAP2 gene KOs is provided by cluster 31 (Figure 5). Importantly, this cluster is made up mostly of mRNAs, whose increased abundance in macrogametocytes is known to rely critically on posttranscriptional stabilization (Mair et al., 2006). The relatively high abundance of these transcripts in ookinete cultures of the *ap2-o3* mutant may therefore not reflect increased expression but reduced turnover, possibly due to developmental arrest. First, this serves as a reminder that transcript abundance is an imperfect proxy for promoter activity. Second, it illustrates that RNA-seq cannot distinguish deregulation events that are due to the direct action of an ApiAP2 protein on a target gene from indirect effects. That such indirect effects must strongly impact the transcriptomes of ApiAP2 mutants is illustrated by a network representation of how disrupting individual ApiAP2 genes affects many other members of this gene family, thereby probably affecting their target genes indirectly (Figure S6B). Hypotheses generated by the current analysis will therefore need to be tested in vitro and in vivo using binding studies and reporter assays. These limitations notwithstanding, comparative phenotyping by RNA-seq has revealed an unexpected complexity of gene expression control in *Plasmodium*, which can now be investigated further to gain a deeper understanding of how the *Plasmodium* life cycle is regulated and to assess how parasites can respond to environmental conditions, such as drug treatment.

EXPERIMENTAL PROCEDURES

Animals and Parasite Strains
Parasites were maintained and phenotyped in female Theiler's Original (TO) outbred mice of 6–12 weeks of age. All animal procedures were conducted under a project license issued by the UK Home Office and with local ethical approval. All mutants were generated on a *P.berghei* ANKA background stably expressing a GFP/luciferase fusion protein throughout the life cycle (RMgm-29) (Franke-Fayard et al., 2008).

KO Vector Generation and Transfections
KO vectors were designed and produced as described previously (Pfander et al., 2011) and are available as part of the *Plasmo*GEM resource (http://plasmogem.sanger.ac.uk). For transfections, 1 µg vector was released from the plasmid backbone by restriction with an *Not*I and electroporated into purified schizont-stage parasites using a standard protocol (Janse et al., 2006b). Following selection with pyrimethamine, resistant parasites were initially genotyped for the presence of the selection cassette by PCR, cloned by limiting dilution, and then genotyped in-depth by PCR and southern hybridization of separated chromosomes, as appropriate.

Phenotyping of the Mutant Clones
For blood stage and early sexual stage phenotyping, three mice were injected intraperitoneally with 150 µLof 6 mg/mL phenylhydrazine to induce reticulocyte formation, followed by an intraperitoneal injection of 5 × 10e7 parasitized

RBCs 2 days later. Asexual parasitemia and male and female gametocytemias were assessed using Giemsa-stained thin blood smears on day 3 post-infection. Exflagellation efficiency was determined in the same animals by mixing 5 μL blood from the tail vein into 500 μL exflagellation medium (RPMI 1640 containing 25 mM HEPES, 4 mM sodium bicarbonate, 5% FCS, and 100 μM XA [pH 7.4]) and counting exflagellation centers per 1,000 RBCs 12–15 min afterward in a Neubauer heamocytometer. Ookinete conversion assays were performed by adding 500 μL blood to 10 mL full ookinete medium (RPMI1640 containing 25 mM HEPES, 10% FCS, and 100 μM XA [pH 7.5]) and incubating at 19°C for 20 hr. The efficiency of conversion and morphology of the cells were assessed on Giemsa-stained thin films and by live microscopy using Cy3-conjugated 13.1 monoclonal anti-P28 antibody staining.

To investigate mosquito stages, ~100 female *Anopheles stephensi* mosquitoes, 6–9 days old, were allowed to feed for 20 min on 2 mice that 3 days earlier had been injected with 5 × 10e7 parasitized RBCs. Seven days later, midguts of 20 mosquitoes were dissected and photographed under fluorescence illumination. The pictures were then processed using Fiji software with particle counter plug-in to quantify the number and size of oocysts. On day 14, oocysts were additionally imaged using bright field illumination to monitor sporulation within the oocysts. On day 21, salivary glands of two batches of 10 mosquitoes were dissected and homogenized, and the number of sporozoites was quantified in a hemocytometer. The capacity of sporozoites to infect a new host was assessed using the same batch of mosquitoes on day 24 post-infection by allowing all remaining mosquitoes to feed for 10 min on three naive mice that were then monitored daily for the appearance of blood stage parasites.

In all phenotyping experiments, a wild-type parasite was included alongside the mutant to control for biological and technical variability between the experiments. At least two blood stage and two mosquito stage phenotyping experiments were performed for each of the mutants. For the parasites not described previously, this was repeated with two independent KOs. See Supplemental Experimental Procedures for additional phenotyping assays.

Transcriptome Sequencing and Analysis

Schizonts, gametocytes, and ookinetes were purified and used for RNA-seq library production as described in the Supplemental Experimental Procedures. Pools of indexed libraries were sequenced on an Illumina HiSeq2500 using 100 bp paired-end reads, and the read counts for all predicted *P. berghei* transcripts were generated. R 3.0 software with the DEseq2 package was used for PC and differential expression analyses. To cluster expression profiles, reads were first normalized for library size and genes with <50 reads across all samples were removed. For each remaining gene and experimental condition, the mean percentage of the normalized reads for all replicates was then expressed as a proportion of all reads for that gene. Clusters of co-regulated genes were identified by Ward's hierarchical clustering with a cut-off set to 49 clusters.

The enrichment for GO terms, metabolic pathways, and stage-specific datasets was performed using Fisher's exact test. For motif analysis, a command line version of DREME (Discriminative Regular Expression Motif Elicitation) (Bailey, 2011) was used with the predicted motif size of 4–10 bp and cut-off E value of 0.05. The 2 kb sequence upstream of all genes from the cluster was used as a test sequence and the remaining promoters in the genome as a reference dataset. Further details of data analysis are available in the Supplemental Experimental Procedures section, and the full dataset was submitted to the Gene Expression Omnibus (GEO: GSE80634).

Proteome Analysis

Purified schizont samples were prepared as for the transcriptome analysis, total protein was extracted from parasite pallet of three replicates per clone and digested with trypsin, and samples were labeled using TMT10plex technology (Thermo Fisher). The *ap2-g* replicates served as reference samples (see Supplemental Experimental Procedures for further details).

ACCESSION NUMBERS

The Gene Expression Omnibus accession number for the RNA-seq data from the apiAP2 knockouts reported in this paper is GEO: GSE80634. The ProteomeXchange Consortium accession number for the proteome sequences reported in this paper is ProteomeXchange: PXD005180.

AUTHOR CONTRIBUTIONS

K.M. and O.B. conceived of and designed the study, with initial input from C.P. K.M. and C.P. generated the KO vectors and parasites lines. K.M., A.R.G., K.D., and C.S. conducted the phenotyping assays. D.G. performed TEM. K.M. and L.C. generated and analyzed RNA-seq data. L.Y. and J.C. generated and analyzed mass spectroscopy data. O.B. and J.C.R. supervised the work. K.M. and O.B. wrote the manuscript with contributions from all authors.

REFERENCES

Aly, A.S.I., and Matuschewski, K. (2005). A malarial cysteine protease is necessary for Plasmodium sporozoite egress from oocysts. J. Exp. Med. *202*, 225–230.

Ay, F., Bunnik, E.M., Varoquaux, N., Vert, J.-P., Noble, W.S., and Le Roch, K.G. (2015). Multiple dimensions of epigenetic gene regulation in the malaria parasite Plasmodium falciparum: gene regulation via histone modifications, nucleosome positioning and nuclear architecture in P. falciparum. BioEssays *37*, 182–194.

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics *27*, 1653–1659.

Balaji, S., Babu, M.M., Iyer, L.M., and Aravind, L. (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 Res. *33*, 3994–4006.

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

Caro, F., Ahyong, V., Betegon, M., and DeRisi, J.L. (2014). Genome-wide regulatory dynamics of translation in the Plasmodium falciparum asexual blood stages. eLife *3*, e04106.

De Silva, E.K., Gehrke, A.R., Olszewski, K., León, I., Chahal, J.S., Bulyk, M.L., and Llinás, M. (2008). Specific DNA-binding by apicomplexan AP2 transcription factors. Proc. Natl. Acad. Sci. USA *105*, 8393–8398.

Dessens, J.T., Mendoza, J., Claudianos, C., Vinetz, J.M., Khater, E., Hassard, S., Ranawaka, G.R., and Sinden, R.E. (2001). Knockout of the rodent malaria parasite chitinase pbCHT1 reduces infectivity to mosquitoes. Infect. Immun. *69*, 4041–4047.

Elemento, O., Slonim, N., and Tavazoie, S. (2007). A universal framework for regulatory element discovery across all genomes and data types. Mol. Cell *28*, 337–350.

Flueck, C., Bartfai, R., Niederwieser, I., Witmer, K., Alako, B.T.F., Moes, S., Bozdech, Z., Jenoe, P., Stunnenberg, H.G., and Voss, T.S. (2010). A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in chromosome end biology. PLoS Pathog. *6*, e1000784.

Foth, B.J., Zhang, N., Chaal, B.K., Sze, S.K., Preiser, P.R., and Bozdech, Z. (2011). Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite Plasmodium falciparum. Mol. Cell. Proteomics *10*, 006411.

Franke-Fayard, B., Djokovic, D., Dooren, M.W., Ramesar, J., Waters, A.P., Falade, M.O., Kranendonk, M., Martinelli, A., Cravo, P., and Janse, C.J. (2008). Simple and sensitive antimalarial drug screening in vitro and in vivo using transgenic luciferase expressing Plasmodium berghei parasites. Int. J. Parasitol. *38*, 1651–1662.

Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. Nature *489*, 91–100.

Gomes, A.R., Bushell, E., Schwach, F., Girling, G., Anar, B., Quail, M.A., Herd, C., Pfander, C., Modrzynska, K., Rayner, J.C., and Billker, O. (2015). A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. Cell Host Microbe *17*, 404–413.

Iwanaga, S., Kaneko, I., Kato, T., and Yuda, M. (2012). Identification of an AP2-family protein that is critical for malaria liver stage development. PLoS ONE *7*, e47557.

Janse, C.J., Franke-Fayard, B., and Waters, A.P. (2006a). Selection by flow-sorting of genetically transformed, GFP-expressing blood stages of the rodent malaria parasite, Plasmodium berghei. Nat. Protoc. *1*, 614–623.

Janse, C.J., Ramesar, J., and Waters, A.P. (2006b). High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei. Nat. Protoc. *1*, 346–356.

Kafsack, B.F.C., Rovira-Graells, N., Clark, T.G., Bancells, C., Crowley, V.M., Campino, S.G., Williams, A.E., Drought, L.G., Kwiatkowski, D.P., Baker, D.A., et al. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. Nature *507*, 248–252.

Kaneko, I., Iwanaga, S., Kato, T., Kobayashi, I., and Yuda, M. (2015). Genome-Wide Identification of the Target Genes of AP2-O, a Plasmodium AP2-Family Transcription Factor. PLoS Pathog. *11*, e1004905.

Khan, S.M., Franke-Fayard, B., Mair, G.R., Lasonder, E., Janse, C.J., Mann, M., and Waters, A.P. (2005). Proteome analysis of separated male and female gametocytes reveals novel sex-specific Plasmodium biology. Cell *121*, 675–687.

Laurentino, E.C., Taylor, S., Mair, G.R., Lasonder, E., Bartfai, R., Stunnenberg, H.G., Kroeze, H., Ramesar, J., Franke-Fayard, B., Khan, S.M., et al. (2011). Experimentally controlled downregulation of the histone chaperone FACT in Plasmodium berghei reveals that it is critical to male gamete fertility. Cell. Microbiol. *13*, 1956–1974.

Levo, M., and Segal, E. (2014). In pursuit of design principles of regulatory sequences. Nat. Rev. Genet. *15*, 453–468.

Licausi, F., Ohme-Takagi, M., and Perata, P. (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. New Phytol. *199*, 639–649.

López-Barragán, M.J., Lemieux, J., Quiñones, M., Williamson, K.C., Molina-Cruz, A., Cui, K., Barillas-Mury, C., Zhao, K., and Su, X.Z. (2011). Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. BMC Genomics *12*, 587.

Mair, G.R., Braks, J.A.M., Garver, L.S., Wiegant, J.C.A.G., Hall, N., Dirks, R.W., Khan, S.M., Dimopoulos, G., Janse, C.J., and Waters, A.P. (2006). Regulation of sexual development of Plasmodium by translational repression. Science *313*, 667–669.

Moon, R.W., Taylor, C.J., Bex, C., Schepers, R., Goulding, D., Janse, C.J., Waters, A.P., Baker, D.A., and Billker, O. (2009). A cyclic GMP signalling mod-

ule that regulates gliding motility in a malaria parasite. PLoS Pathog. *5*, e1000599.

Oberstaller, J., Pumpalova, Y., Schieler, A., Llinás, M., and Kissinger, J.C. (2014). The Cryptosporidium parvum ApiAP2 gene family: insights into the evolution of apicomplexan AP2 regulatory systems. Nucleic Acids Res. *42*, 8271–8284.

Orito, Y., Ishino, T., Iwanaga, S., Kaneko, I., Kato, T., Menard, R., Chinzei, Y., and Yuda, M. (2013). Liver-specific protein 2: a Plasmodium protein exported to the hepatocyte cytoplasm and required for merozoite formation. Mol. Microbiol. *87*, 66–79.

Otto, T.D., Wilinski, D., Assefa, S., Keane, T.M., Sarry, L.R., Böhme, U., Lemieux, J., Barrell, B., Pain, A., Berriman, M., et al. (2010). New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq. Mol. Microbiol. *76*, 12–24.

Otto, T.D., Böhme, U., Jackson, A.P., Hunt, M., Franke-Fayard, B., Hoeijmakers, W.A., Religa, A.A., Robertson, L., Sanders, M., Ogun, S.A., et al. (2014). A comprehensive evaluation of rodent malaria parasite genomes and gene expression. BMC Biol. *12*, 86.

Paul, R.E., Coulson, T.N., Raibaud, A., and Brey, P.T. (2000). Sex determination in malaria parasites. Science *287*, 128–131.

Pfander, C., Anar, B., Schwach, F., Otto, T.D., Brochet, M., Volkmann, K., Quail, M.A., Pain, A., Rosen, B., Skarnes, W., et al. (2011). A scalable pipeline for highly effective genetic modification of a malaria parasite. Nat. Methods *8*, 1078–1082.

Philip, N., and Waters, A.P. (2015). Conditional Degradation of Plasmodium Calcineurin Reveals Functions in Parasite Colonization of both Host and Vector. Cell Host Microbe *18*, 122–131.

Pino, P., Sebastian, S., Kim, E.A., Bush, E., Brochet, M., Volkmann, K., Kozlowski, E., Llinás, M., Billker, O., and Soldati-Favre, D. (2012). A tetracycline-repressible transactivator system to study essential genes in malaria parasites. Cell Host Microbe *12*, 824–834.

Sebastian, S., Brochet, M., Collins, M.O., Schwach, F., Jones, M.L., Goulding, D., Rayner, J.C., Choudhary, J.S., and Billker, O. (2012). A Plasmodium calcium-dependent protein kinase controls zygote development and transmission by translationally activating repressed mRNAs. Cell Host Microbe *12*, 9–19.

Shock, J.L., Fischer, K.F., and DeRisi, J.L. (2007). Whole-genome analysis of mRNA decay in Plasmodium falciparum reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle. Genome Biol. *8*, R134.

Sinha, A., Hughes, K.R., Modrzynska, K.K., Otto, T.D., Pfander, C., Dickens, N.J., Religa, A.A., Bushell, E., Graham, A.L., Cameron, R., et al. (2014). A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. Nature *507*, 253–257.

Stampfel, G., Kazmar, T., Frank, O., Wienerroither, S., Reiter, F., and Stark, A. (2015). Transcriptional regulators form diverse groups with context-dependent regulatory functions. Nature *528*, 147–151.

Sturm, A., Mollard, V., Cozijnsen, A., Goodman, C.D., and McFadden, G.I. (2015). Mitochondrial ATP synthase is dispensable in blood-stage Plasmodium berghei rodent malaria but essential in the mosquito phase. Proc. Natl. Acad. Sci. USA *112*, 10216–10223.

Tao, D., Ubaida-Mohien, C., Mathias, D.K., King, J.G., Pastrana-Mena, R., Tripathi, A., Goldowitz, I., Graham, D.R., Moss, E., Marti, M., and Dinglasan, R.R. (2014). Sex-partitioning of the Plasmodium falciparum stage V gametocyte proteome provides insight into falciparum-specific cell biology. Mol. Cell. Proteomics *13*, 2705–2724.

Tarun, A.S., Peng, X., Dumpit, R.F., Ogata, Y., Silva-Rivera, H., Camargo, N., Daly, T.M., Bergman, L.W., and Kappe, S.H.I. (2008). A combined transcriptome and proteome survey of malaria parasite liver stages. Proc. Natl. Acad. Sci. USA *105*, 305–310.

Turturice, B.A., Lamm, M.A., Tasch, J.J., Zalewski, A., Kooistra, R., Schroeter, E.H., Sharma, S., Kawazu, S., and Kanzok, S.M. (2013). Expression of cytosolic peroxiredoxins in Plasmodium berghei ookinetes is regulated by environmental factors in the mosquito bloodmeal. PLoS Pathog. *9*, e1003136.

Young, J.A., Johnson, J.R., Benner, C., Yan, S.F., Chen, K., Le Roch, K.G., Zhou, Y., and Winzeler, E.A. (2008). In silico discovery of transcription regulatory elements in Plasmodium falciparum. BMC Genomics *9*, 70.

Yu, M., Kumar, T.R.S., Nkrumah, L.J., Coppi, A., Retzlaff, S., Li, C.D., Kelly, B.J., Moura, P.A., Lakshmanan, V., Freundlich, J.S., et al. (2008). The fatty acid biosynthesis enzyme FabI plays a key role in the development of liver-stage malarial parasites. Cell Host Microbe *4*, 567–578.

Yuda, M., Iwanaga, S., Shigenobu, S., Mair, G.R., Janse, C.J., Waters, A.P., Kato, T., and Kaneko, I. (2009). Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. Mol. Microbiol. *71*, 1402–1414.

Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T., and Kaneko, I. (2010). Transcription factor AP2-Sp and its target genes in malarial sporozoites. Mol. Microbiol. *75*, 854–863.

Yuda, M., Iwanaga, S., Kaneko, I., and Kato, T. (2015). Global transcriptional repression: An initial and essential step for Plasmodium sexual development. Proc. Natl. Acad. Sci. USA *112*, 12824–12829.

Zhou, Y., Ramachandran, V., Kumar, K.A., Westenberger, S., Refour, P., Zhou, B., Li, F., Young, J.A., Chen, K., Plouffe, D., et al. (2008). Evidence-based annotation of the malaria parasite's genome using comparative expression profiling. PLoS ONE *3*, e1570.

Supplemental Information

A Knockout Screen of ApiAP2 Genes Reveals

Networks of Interacting Transcriptional

Regulators Controlling the *Plasmodium* Life Cycle

Katarzyna Modrzynska, Claudia Pfander, Lia Chappell, Lu Yu, Catherine Suarez, Kirsten Dundas, Ana Rita Gomes, David Goulding, Julian C. Rayner, Jyoti Choudhary, and Oliver Billker
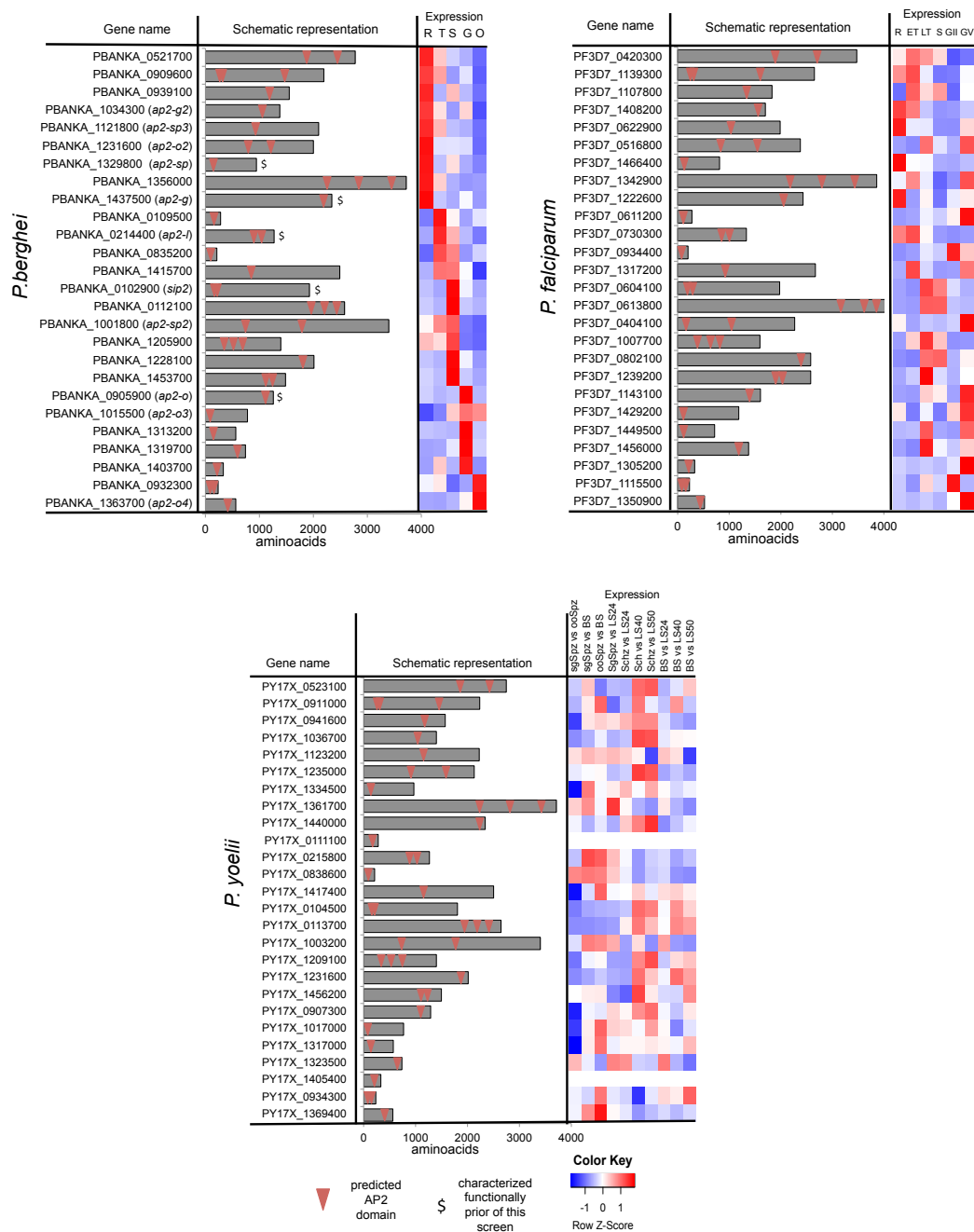
Supplemental figures:

Figure S1, related to Figure 1. Comparison of gene structure and expression profiles of the members of ApiAP2 family across three different *Plasmodium* species.

Gene names and structures presented according for the latest annotation for each of the species (http://plasmodb.org). One *P.falciparum* gene with no *P.berghei/P.yoelii* ortholog has been omitted. RNA-seq (*Pb* and *Pf*) and microarray (*Py*) expression data extracted from Otto et al., 2014 (*Pb*),López-Barragán et al., 2011(*Pf*) and Tarun et al., 2008(*Py*). R=rings, T= trophozoites, S/Schz = schizonts, G = gametocytes, O=ookinetes, ET= early schizonts, LT= late schizonts, GII= gametocytes stage II, GV = gametocytes stage V, BS= mixed blood stages, ooSpz = oocyst sporozoites, sgSpz= salivary gland sporozoites, LS24/LS40/LS50 = liver stages 24, 40 and 50h after invasion.
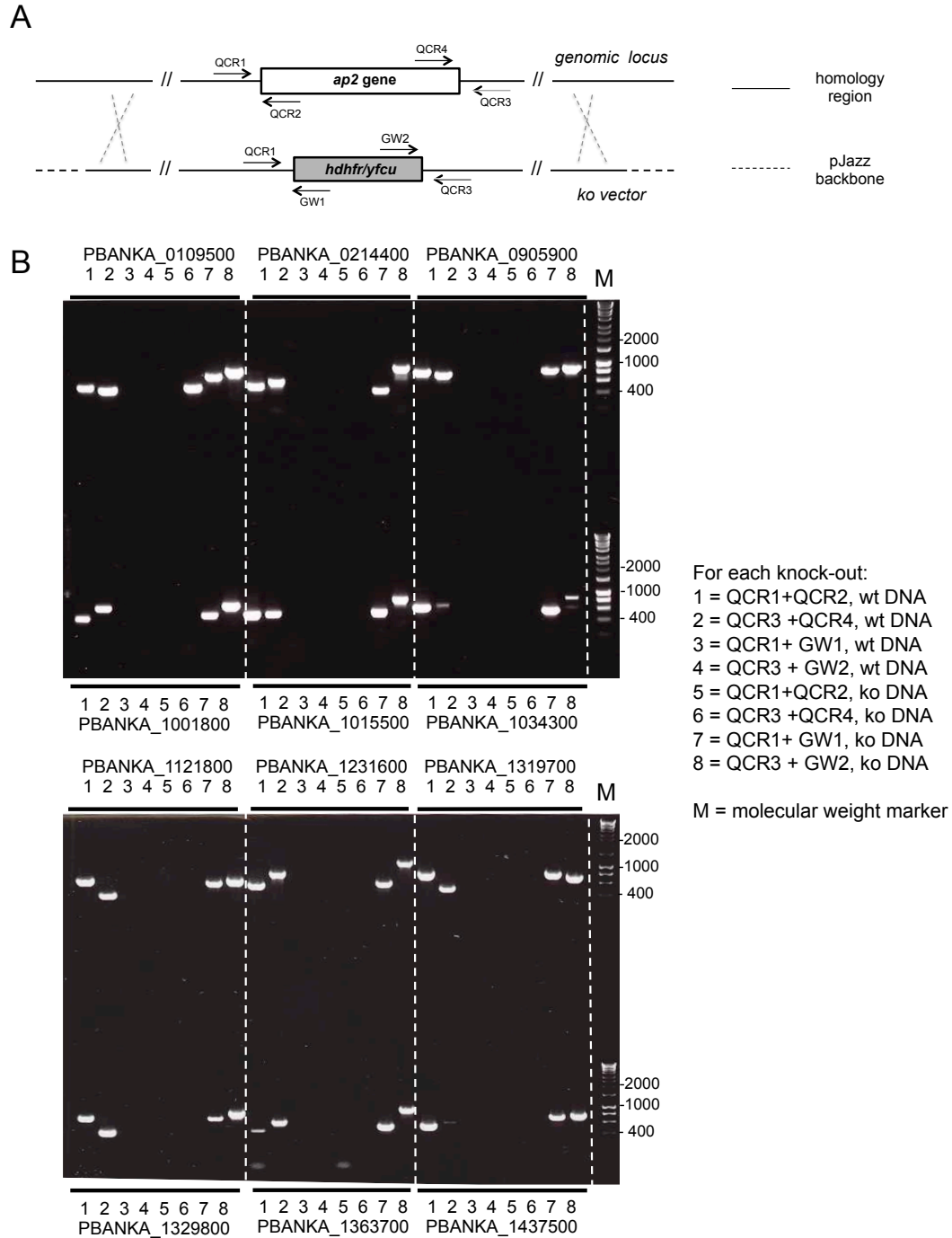
Figure S2, related to Figure 1. Generation and genotyping of ApiAP2 knockout mutants.

A) Schematic representation of a generic ApiAP2 locus, the corresponding transfection vector and annealing sites for genotyping primers. *hdhfr/yfcu* = positive/negative selection cassette.

B) Agarose gel images showing PCR products from genotyping a representative knockout clone for each gene disrupted in this study. The genotyping panel for each gene consists of four PCR reactions performed on wt and mutant gDNA as described in the side panel. Transfections with a knockout vector for PBANKA_0109500 also repeatedly generated PCR-positive parasites, but cloning was unsuccessful and a PCR product indicative of an undisrupted gene locus was retained (lane 6).
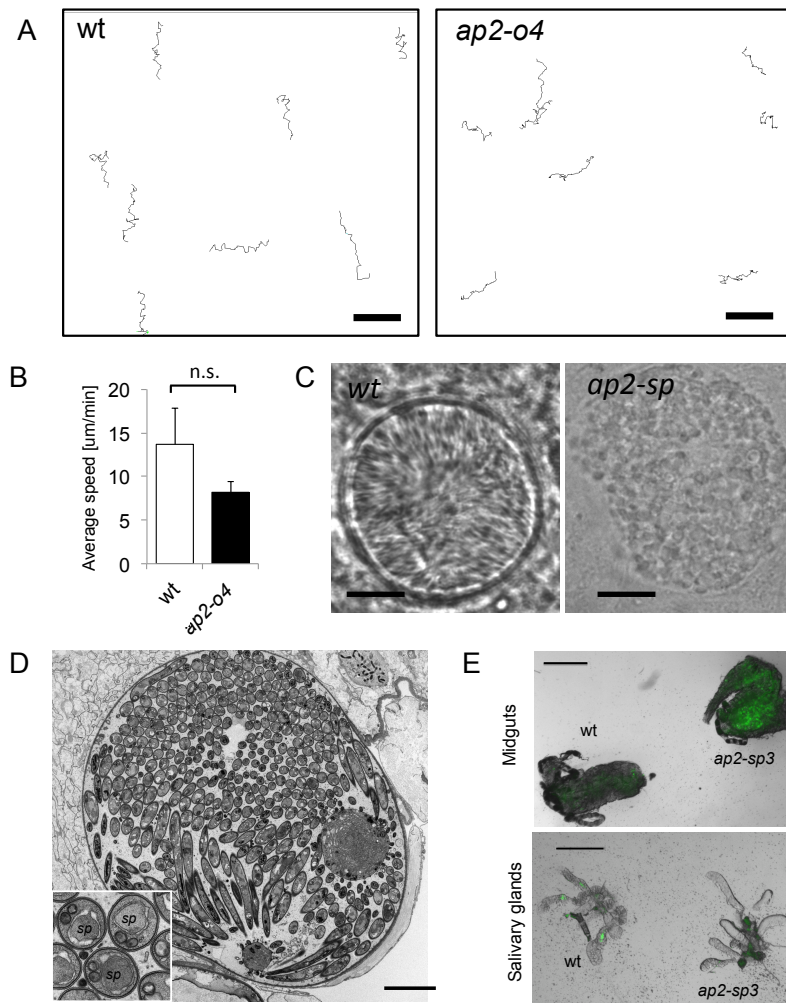
Figure S3, related to Figures 2 and 3. Additional phenotyping of *ap2-o4*, *ap2-sp* and *ap2-sp3* parasites.

A) Motility traces generated by manual tracking of ookinetes from time lapse movies acquired for 10 min at 1 frame every 10s. Typical helical gliding patterns of ookinetes in matrigel are shared by wild type and *ap2-o4* ookinetes. Scale bar = 75 μm.

B) Gliding speed of wt and *ap2-o4* ookinetes in Matrigel determined by time-lapse video microscopy. Error bars show standard deviations for 25 wild type and 19 mutant ookinetes.

C) Comparison of the development of the *ap2-sp* and wt oocysts on day 14 post infections. While the developing sporozoites can be observed in the wt oocysts, the *ap2-sp* ones are undergoing the abnormal granulation as described previously (Yuda et al., 2010). Scale bar = 10 μm.

D) Transmission electron micrograph of a representative *ap2-sp3* oocyst on a mosquito midgut on day 12 post feeding, showing normal sporozoite budding (compare to wild type oocysts in Fig. 3D). Scale bar = 5 μm.

E) Representative light microscopic images of midguts and salivary glands overlaid with fluorescence images to reveal parasites, which express GFP. Organs were dissected from the same two mosquitoes 21 days after infection with wild type or *ap2-sp3* parasites. Wild type parasites have progressed from the midgut to the salivary glands, while mutant parasites appear to be retained within intact oocysts on the midgut and fail to migrate to the salivary glands. Scale bar = 75 μm.
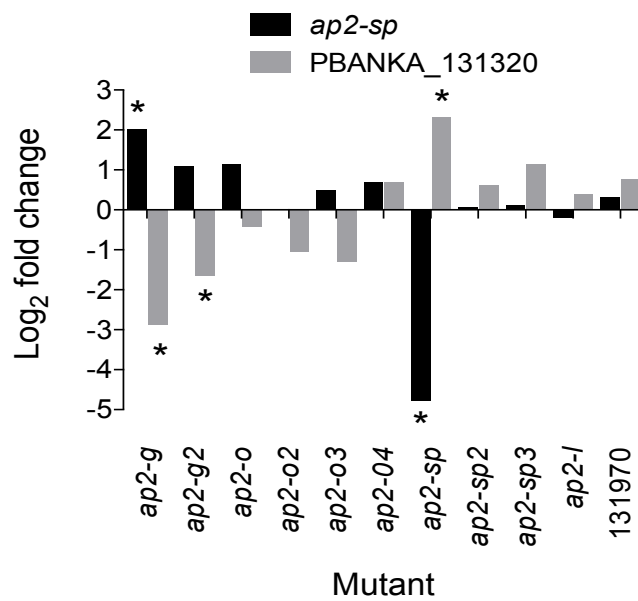
Figure S4 related to Figure 4. PBANKA_1313200 as potential secondary regulator connected with *ap2-sp* expression

Differential expression of *sp2-sp* and PBANKA_1313200 in synchronous blood stage cultures of mutant clones expressed relative to wild type. Asterisks indicate all instances of statistical significance (adjusted p-values < 0.05, n = 3 biological replicates).
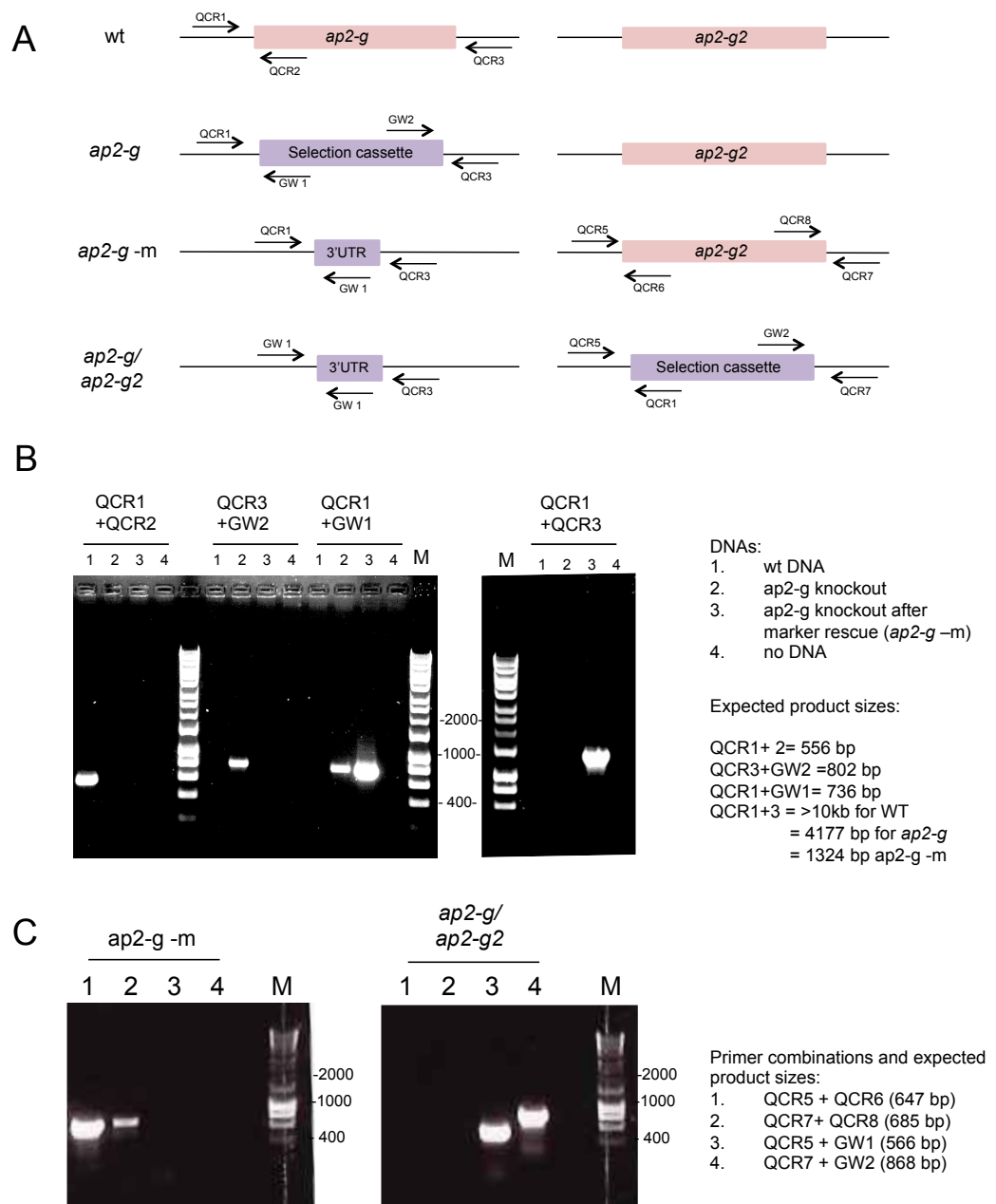
Figure S5, related to Figure 6. Generation and genotyping of a *ap2-g/ap2-g2* double mutant.

A)   Schematic representation of *ap2-g* and *ap2-g2* loci in wild type and mutant parasite including marker free intermediate (*ap2-g* -m), showing annealing sites of genotyping primers. QCR1-4 sequences are the same as for genotyping the *ap2-g* single knockout. QCR5-8 sequences correspond to QCR1-4 of *ap2-g2* genotyping set.

B)   Agarose gel images showing PCR products that confirm loss of the selection marker from an *ap2-g* –m clone following negative selection.

C)   PCR products confirming modification of the *ap2-g2* locus in an *ap2-g/ap2-g2* double knockout.
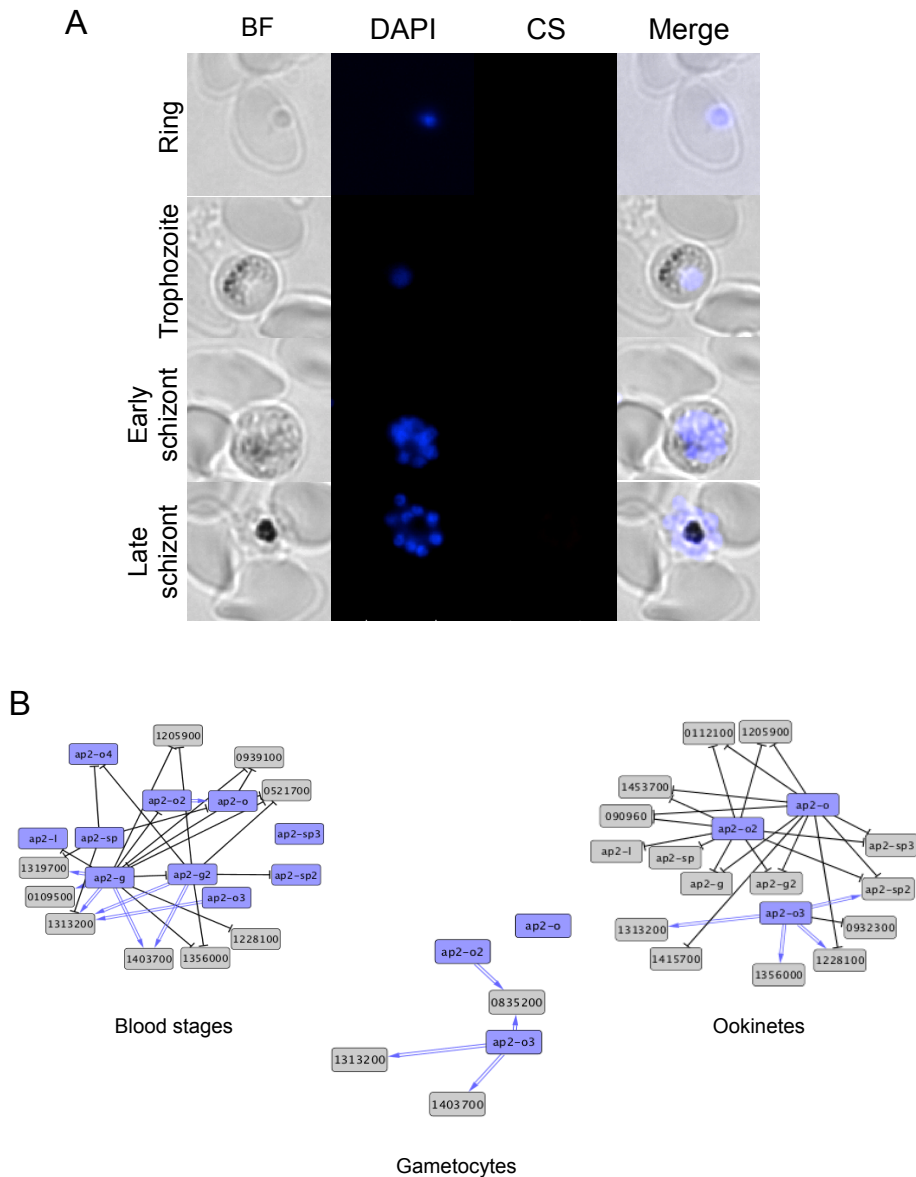
Figure S6, related to Figures 6 and 4. Lack of CSP expression in the *ap2-g* parasite line before the introduction of *ap2-g2 ko* mutation and network representation of ApiAP2 gene expression.

A) The *ap2-g* parasites processed and imagined in parallel with *the ap2-g/ap2-g2* parasites presented in Fig 6 C.

B) Changes in the expression of ApiAP2 genes in the generated mutants mutants. The Blue nodes show disrupted genes, edges show the resulting expression changes with p<0.05. Blue double arrows indicate positive interactions (i.e. the target gene is less expressed in the mutant) and black lines show negative interactions. The diagrams illustrate a strong interdependence of ApiAP2 genes, but interactions shown are not necessarily direct.

Supplemental data items:

Data S1, related to Figure 1. Summary of the apiAP2 knockout screen and phenotyping of the generated mutants.

Data S2, related to Figure 4. Transcriptome analysis of the apiAP2 mutants.

Data S3, related to Figure 4 and 5. Groups of genes co-expressed across different apiAP2 mutants.

Data S4, related to Figure 4 and 5. Characterization of the 49 co-expression clusters generated from ApiAP2 expression data.

Data S5, related to Figure 6. Transcriptome and proteome analysis of the double *ap2-g/ap2-g2* knockout.

Supplemental experimental protocols:

Growth rate quantification using barseq approach. The relative fitness associated with *Plasmo*GEM knockout vectors was quantified as described previously (Gomes et al., 2015). Briefly, a mixed pool containing 14 barcoded ApiAP2 ko vectors and 7 control vectors was prepared and used for three independent transfections of a schizont culture prepared from a Wistar rat (~1% parasitaemia). Electroporated parasites were injected intravenously into BALB/c mice, and resistant parasites were selected with 70 mg/ml of pyrimethamine in the drinking water. Infections were monitored daily using Giemsa stained thin blood smears. Blood samples were taken on days 4-8 and used for genomic DNA extractions with phenol/chloroform. Vector specific barcodes of 10-11 bp were PCR-amplified using a generic primer pair and the amplicons were converted into sequencing libraries by PCR-mediated Illumina adapter addition as described (see Gomes et al., 2015). Samples were pooled and sequenced on an Illumina MiSeq sequencer (150 bp paired-end reads). A perl script was used to extract and count the barcodes from the sequencing output. The barcodes corresponding to vectors not included in the input or accounting for less than 0.1% of barcode reads obtained from a given sample were excluded from further analysis. The number of reads corresponding to each barcode was corrected for library size and the relative fitness of each vector on each day was calculated by dividing the fold change of the relative abundance of the vector tag by the averaged results from four reference genes with normal growth rates.

PCR genotyping of the generated mutants. Each mutant clone was genotyped using four combinations of primers, specific for either wt or ko locus on both sides of the targetted gene (schematic design shown in Fig. S2). Wt DNA control was included in each genotyping panel. The details of DNA preparation and PCR program can be found on http://plasmogem.sanger.ac.uk/info/protocols.

Following primers were used:

QCR1	
PBANKA_0109500	ACTTAATATTTATACACTGCGTTTTCTC
PBANKA_0214400	TATAGTGCGGATGTATAGCACACG
PBANKA_0905900	TTACATATATGGAAAATAATGAGCACC
PBANKA_1001800	TATCATGTCAACATATCATCATATTGC
PBANKA_1015500	CACCAAATTGTGCATCTTTCG
PBANKA_1034300	ATTAATAAACTAGAGGTTCTTTACCAGG
PBANKA_1121800	TGTGGATGAAAATGTGGTTGC
PBANKA_1231600	ATTATCATTGTGGATTAACGGAAG
PBANKA_1319700	ATAGACTGTGTCAATAGATGATTTCAGC
PBANKA_1329800	TTATGCCTGTATTTGAAACGAGG
PBANKA_1363700	CGAACAAATATATAAATGCGAAGAGG
PBANKA_1437500	AGGGGCATTGAACAATAGCACCT

QCR2	
PBANKA_0109500	TAAGAAAAAGTTAGACCAAATAATGAGG
PBANKA_0214400	AAATAATATACACAATATGAGAATCCAAG
PBANKA_0905900	AGAGGCATCTATATTCAAATTGGC
PBANKA_1001800	TTTTCAAATAATGGGAATACAGTGG
PBANKA_1015500	AAAGATTCTATACATTCAGACACTTTACAC
PBANKA_1034300	TAAATCAGTGAAATCAACTTATGGATC
PBANKA_1121800	CAAAGCAATAATAGTTATATAACGGAAGG
PBANKA_1231600	CGAGACAATTGTTTATCAAATTTAGC
PBANKA_1319700	AGGATATTATGATCTTGAGATTGATGG
PBANKA_1329800	GGAAGTAGAACAGGAATAGTAACTGACC
PBANKA_1363700	TATTTCGTTCAAGGTTTTCCCTC
PBANKA_1437500	TGGCGTTGAAACTAGTCCCGA
QCR3	
PBANKA_0109500	ATAATCGTAAAACTCGTTCAAAGG
PBANKA_0214400	AGAAGCATAGTTGTAAATTCATCAAATACTG
PBANKA_0905900	AACAAATTAAATGCATTGATTAGACAG
PBANKA_1001800	CTGTACATATTTTGAAGCGATGATGTAC
PBANKA_1015500	TCAACGACAGACATTGAAAAGC
PBANKA_1034300	GTTATACTATTTCGATGTAA
PBANKA_1121800	CATTTGTTTTTATGTACATGAGTATAAGTGC
PBANKA_1231600	CCATGTGTATGTTATTATATAAGAAGAATG
PBANKA_1319700	TGATAATTATTATACAGTTGTGCCGC
PBANKA_1329800	CACATATAGTCATATATATACCGATAAGAAGG
PBANKA_1363700	TTGATTAAAGAAATCTCACAACATAAACAG
PBANKA_1437500	CATTTCAATTGCAACGGG
QCR4	
PBANKA_0109500	AATTTGTGTGCAAATAGCTATACATG
PBANKA_0214400	TCACATCGGTTATCTATTTCATTTC
PBANKA_0905900	TTAGAATTACCAGATAAAGATGGACC
PBANKA_1001800	TTTACAACGTACCACTCCTCCTG
PBANKA_1015500	TCGGACCACCTTGCTTCC
PBANKA_1034300	CTTTCAGAATGAAGGAAG
PBANKA_1121800	CCAAATACTTTTCCTTTTATCTTGG
PBANKA_1231600	GCTATATCATATGAATATGTATTAATAGGTGC
PBANKA_1319700	TTACATCTTATATCATCCTTATTCTGGAG
PBANKA_1329800	CTGTTAATAAGTTCGGCATTTATCG
PBANKA_1363700	TCTATCAAGGATCTTATGTAGCTAACTGG
PBANKA_1437500	AGAATATTTCATACCATATG
Cassette specific primers	
GW1	ACTTAATATTTATACACTGCGTTTTCTC
GW2	TATAGTGCGGATGTATAGCACACG

Ookinete motility assays. For the motility assay 30 μl of 22 h ookinete culture were mixed with 30 μl Matrigel (BD Biosciences) on ice. A drop of the mixture was spotted on a microscope slide, covered with a coverslip, sealed with nail polish and allowed to settle for ~30 min at 19°C. A Leica M205A microscope was used to acquire time-lapse movies (1 frame every 10 s for 10 min) of representative fields containing at least five ookinetes. The results were exported as *.tiff image series and analysed using Fiji software with the Manual Tracking plugin (http://pacific.mpi-cbg.de/wiki/index.php/Manual_Tracking).

Immunofluorescence imaging. Infected red blood cells were washed with phosphate buffered saline (PBS), fixed with 4% paraformaldehyde in PBS (20 min), permeabilized with 0.01% Triton X-100/PBS solution (10 min) and blocked with 3% bovine serum albumin in PBS for 1 h. Cy3 conjugated anti-PbCSP D11 monoclonal antibody (Yoshida et al., 1980) diluted with blocking solution to 1 μg/ml was incubated with the sample for 1 h in the dark. The cells were washed 3 times with 1xPBS and mounted on a microscope slide using ProLong Gold Antifade Mountant with DAPI stain (Life Technologies). Samples were imaged using a fluorescence microscope (Leica DMi8) with a digital camera (Leica DFC265Fx).

Transmission electron microscopy. For each strain the midguts from 10 mosquitoes were dissected and the presence of oocysts verified using fluorescence microscopy. Five highly infected guts were fixed at 20 °C for 2 h in 2% PFA with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer at pH 7.42 and then in 1% buffered osmium tetroxide for another 2 h, mordanted with 1% tannic acid for 1 h and dehydrated through an ethanol series staining en bloc with 2% uranyl acetate at the 30% stage. The midguts were oriented and embedded in Epon resin for ultrathin sectioning on a Leica UC6 microtome, stained with uranyl acetate and lead citrate and analysed on a 120 kV Spirit Biotwin transmission electron microscope. Images were taken on a Tietz F4.15 CCD camera.

Parasite purification for transcriptome sequencing. To synchronise parasites at the schizont stage, phenylhydrazine-treated TO mice were injected intraperitoneally with ~1×10^7 infected red blood cells. On day 3 post infection the presence of parasitaemia (5-15%) was confirmed using Giemsa stained thin blood smears and blood was harvested by cardiac puncture. White blood cells were removed using Plasmodipur filters (Europroxima). The parasites were cultured for 22 h and red blood cells harbouring schizonts or gametocytes were purified on a cushion of 55% Nycodenz in culture medium, as described previously (Janse et al., 2006). The resulting pellet was washed in ice cold erythrocyte lysis buffer (0.15 M NH_4Cl, 0.01 M $KHCO_3$, 0.001 M EDTA) until the supernatant was clear to remove host cell RNA, and parasites were resuspended and lysed in TRIzol reagent (Invitrogen).

To obtain highly enriched gametocytes, phenylhydrazine-treated mice were infected with ~1×10^7 parasites, and treated with sulfadiazine by intraperitoneal injection (100 μl of 40 μg/ml solution in PBS) from day 2 post infection as well as by addition of drug to the drinking water (20 mg/l). On day 4, after verifying the absence of asexual stages, the blood was harvested by cardiac puncture and

transferred into suspended animation (SA) medium (RPMI1640 with L-glutamine, 25 mM HEPES, 4 mM NaHCO$_3$, 5% FCS, pH = 7.2), which prevents gametocyte activation. White blood cells were removed using Plasmodipur filters and the blood was layered on a cushion of 48% Nycodenz/SA buffer and spun for 15 min at 500 g (no breaks) to separate the gametocytes from the other blood cells. The gametocyte layer was harvested, pelleted and resuspended in TRIzol reagent.

For ookinete samples the animals were infected and drug treated as for gametocyte preparations. To initiated gamete formation and fertilization, gametocyte infected blood from 1 mouse was added to 30 ml of the ookinete medium (RPMI1640 containing 25 mM HEPES, 10% FCS, 100 µM xanthurenic acid, pH 7.5) and cultured at 19 °C. After 22 h parasites were purified using MACS CS magnetic separation columns (Miltenyi Biotec) according to manufactures instructions, with ookinete medium used for column conditioning, washes and elution. The ookinetes were spun down at 500 g and resuspended in TRIzol reagent. TRIzol samples from all parasite stages were left at room temperature for 10 min and then either frozen at -80 °C or used directly for RNA extraction.

RNA extraction and RNA-seq library preparation. Total RNA was extracted using standard chloroform extraction and isopropanol precipitation (Chomczynski, 1993). RNA concentration and integrity was measured on an Agilent Bioanalyzer using the RNA 6000 Nano kit and on a NanoDrop 1000 spectrophotometer. 1-2 µg of total RNA from each sample was used for mRNA isolation (Magnetic mRNA Isolation Kit, NEB). First strand cDNA synthesis was performed using the SuperScript III First-Strand Synthesis System and a 1:1 mix of Oligo(dT) and random primers (Invitrogen). The DNA/RNA hybrids were purified using Agencourt RNAClean XP beads (Beckman Coulter) and the second cDNA strand was synthetized using a 10 mM dUTP nucleotide mix, DNA Polymerase I (Invitrogen) and RNAseH (NEB) for 2.5 h at 16 °C. The long cDNA fragments were purified and fragmented using a Covaris S220 system (duty cycles = 20, intensity = 5, cycles/burst = 200, time = 30 s). The ~200 bp long fragments were end-repaired, dA-tailed and ligated to "PCR-free" adapters (Kozarewa et al., 2009) with index tags using NEBNext kits according to the manufacturer's instructions. Excess adapters were removed by two rounds of clean-up with 1x volume of Agencourt AMPure XP beads. Final libraries were eluted in 30 µl water, quality-controlled using Agilent High Sensitivity DNA chip, digested with USER enzyme (NEB) and quantified by qPCR. For some libraries additional 5 cycles of PCR amplification were performed, using KAPA HiFi HotStart PCR mix and Illumina tag-specific primers to obtain enough material for sequencing. Pools of ~12 indexed libraries were sequenced using an Illumina HiSeq2500 system (100 bp paired-end reads).

RNAseq data processing. Sequence reads underwent basic quality control (quality score, GC content, adapter contamination), adapter clipping and library segregation according to the multiplexed tags, using standard data processing pipelines at the Wellcome Trust Sanger Institute. The filtered reads were downloaded as *.cram files aligned to the *P. berghei* ANKA reference genome and transformed using Scramble (Bonfield, 2014) and Picard tools (http://broadinstitute.github.io/picard/). The resulting unaligned *.fastq files were mapped against the *P. berghei* transcriptome (PlasmoDB version 9.3)

using Tophat2 aligner (Kim et al., 2013) set to "fr-firststrand" for library type and a maximum intron size of 15 kb. The re-aligned *.bam files were sorted and indexed using SAMtools suite (Li et al., 2009) and used both for visual data inspection in the Artemis genome browser (Rutherford et al., 2000), and to generate read count tables for all annotated *P. berghei* transcripts with HTSeq Python package (Anders et al., 2014).

R 3.0 software with the DEseq2 package (Love et al., 2014) was used for principal component and differential expression analyses. To cluster analysis, reads were first normalized for library size and genes with <50 reads across all samples were removed. For each remaining gene the average number of reads at each strain/stage combination was divided the sum of averaged reads from all conditions, generating the proportional expression profile of the gene. The expression profiles were clustered using Ward's hierarchical clustering with a cutoff set to 49 clusters.

The Gene Ontology (GO) terms enrichment was performed using PlasmoDB (http://plasmodb.org). For metabolic pathways the data from the manually curated Malaria Parasite Metabolic Pathways database (mpmp.huji.ac.il); (Ginsburg, 2006) was translated to *P. berghei* orthologs, and Fisher's exact test with multiple testing correction was used to test for the enrichment. Additionally custom datasets were downloaded from the literature and tested in the same way. These included: genes encoding male- and female- specific proteins (Tao et al., 2014), transcripts overrepresented at more than one time points during liver development compared with asexual blood stages (translated from *P. yoelii* orthologs) (Tarun et al., 2008), targets for AP2-O binding in ookinetes (Kaneko et al., 2015) and mRNAs bound by DOZI/CITH translational repression complex in female gametocytes (Guerreiro et al., 2014).

LC-MS/MS analysis. Three independent schizonts cultures from *ap2-g/ap2-g2* and *ap2-g* parasites were prepared and parasites were isolated as described in the transcriptome analysis section. Each parasite pellet was lysed in 300 µL RIPA buffer with 1x phosphatase and proteinase inhibitor (Thermo Fisher) on ice, then TCEP was added to final ~40 mM and heated at 56 °C for 15 min. The suspension was processed in ultrasonic bath then centrifuged at 14,000 rpm to collect the supernatant. To the pellet formed, 200 µL of 8 M urea/40 mM TCEP/50 mM TEAB was added sonicated and centrifuged as above. Both supernatants were pooled, and the protein concentrations were measured by 660 nm Protein Assay (Thermo Fisher). Iodoacetamide was then added to the sample for alkylation. 90 µg of protein was taken from each sample and made up to the same volume and concentration with 100 mM TEAB (Fluka). Chloroform/methanol precipitation was performed to remove detergent. The protein pellet was resuspended in 90 µL of 100 mM TEAB and digested with 3 µg trypsin (Gold, Promega) at 37°C for 12 h, then labelled with TMT10plex (Thermo Fisher) as described in the product instruction. Ten labelled samples were then pooled to one and dried in a SpeedVac.

The peptides were fractionated on an XBridge BEH C18 column (4.6 mm i.d. x 250 mm, 130 Å, 3.5 µm, Waters) at a high pH reverse phase (pH ~ 10.5 by NH_4OH solution) at a 30 s intervals. Fractions were then pooled into 16 pools and dried in a SpeedVac. Peptides were resuspended in 0.5% formic

acid (FA)/100% H$_2$O, and ¼ was injected for the LC-MS/MS analysis on an Orbitrap Fusion coupled with an Ultimate 3000 RSLCnano instrument. The peptides were first loaded to a peptide trap (100 µm i. d. x 20 mm, 100 Å, 5 µm) then separated on a nano-analytical column (75 µm i. d. x 500 mm) at a linear gradient of 4-32% ACN/0.1% FA within 120 min. The Orbitrap Fusion was operated at a TopSpeed method with a 3 s cycle time, while the full MS scans were acquired at 120,000 resolution at m/z 200; the MS2 spectra were acquired in ion trap mode with CID fragmentation with the isolation window set to m/z 0.7; and finally the MS3 spectra were acquired in Orbitrap at a resolution of 60,000 (m/z 200) using a multinotch MS3 strategy and HCD fragmentation with the isolation window at m/z 0.5. The data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005180.

All LC-MS data were processed in Proteome Discoverer 1.4 using the SequestHT to search a combined protein database of *P. berghei* (http://plasmodb.org/) and mouse (UniprotKB). Peptide spectral matches (PSM) were filtered to a false discovery rate (FDR) of 1%, and isobaric mass tags were used to calculate the sample/reference ratios for each sample. Host and *Plasmodium* proteins with PSM ratio counts < 3 were removed from further analysis. For the remaining proteins, sample/reference ratios from all unique peptides were averaged and normalised within each sample. A two-tailed T-test on log2 transformed ratios was used to identify the significantly up- and down regulated proteins.

Supplemental references:

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq - A Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Bonfield, J.K. (2014). The Scramble conversion tool. Bioinformatics *30*, 2818–2819.

Chomczynski, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques *15*, 532–534, 536–537.

Ginsburg, H. (2006). Progress in in silico functional genomics: the malaria Metabolic Pathways database. Trends Parasitol. *22*, 238–240.

Gomes, A.R., Bushell, E., Schwach, F., Girling, G., Anar, B., Quail, M.A., Herd, C., Pfander, C., Modrzynska, K., Rayner, J.C., et al. (2015). A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. Cell Host Microbe *17*, 404–413.

Guerreiro, A., Deligianni, E., Santos, J.M., Silva, P.A., Louis, C., Pain, A., Janse, C.J., Franke-Fayard, B., Carret, C.K., Siden-Kiamos, I., et al. (2014). Genome-wide RIP-Chip analysis of translational repressor-bound mRNAs in the Plasmodium gametocyte. Genome Biol. *15*, 493.

Janse, C.J., Ramesar, J., and Waters, A.P. (2006). High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei. Nat. Protoc. *1*, 346–356.

Kaneko, I., Iwanaga, S., Kato, T., Kobayashi, I., and Yuda, M. (2015). Genome-Wide Identification of the Target Genes of AP2-O, a Plasmodium AP2-Family Transcription Factor. PLoS Pathog. *11*, e1004905.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kozarewa, I., Ning, Z., Quail, M.A., Sanders, M.J., Berriman, M., and Turner, D.J. (2009). Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. Nat. Methods *6*, 291–295.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

López-Barragán, M.J., Lemieux, J., Quiñones, M., Williamson, K.C., Molina-Cruz, A., Cui, K., Barillas-Mury, C., Zhao, K., and Su, X. (2011). Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. BMC Genomics *12*, 587.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Otto, T.D., Böhme, U., Jackson, A.P., Hunt, M., Franke-Fayard, B., Hoeijmakers, W.A.M., Religa, A.A., Robertson, L., Sanders, M., Ogun, S.A., et al. (2014). A comprehensive evaluation of rodent malaria parasite genomes and gene expression. BMC Biol. *12*, 86.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., and Barrell, B. (2000). Artemis: sequence visualization and annotation. Bioinformatics *16*, 944–945.

Tao, D., Ubaida-Mohien, C., Mathias, D.K., King, J.G., Pastrana-Mena, R., Tripathi, A., Goldowitz, I., Graham, D.R., Moss, E., Marti, M., et al. (2014). Sex-partitioning of the Plasmodium falciparum stage V gametocyte proteome provides insight into falciparum-specific cell biology. Mol. Cell. Proteomics *13*, 2705–2724.

Tarun, A.S., Peng, X., Dumpit, R.F., Ogata, Y., Silva-Rivera, H., Camargo, N., Daly, T.M., Bergman, L.W., and Kappe, S.H.I. (2008). A combined transcriptome and proteome survey of malaria parasite liver stages. Proc. Natl. Acad. Sci. U. S. A. *105*, 305–310.

Yoshida, N., Nussenzweig, R.S., Potocnjak, P., Nussenzweig, V., and Aikawa, M. (1980). Hybridoma produces protective antibodies directed against the sporozoite stage of malaria parasite. Science *207*, 71–73.

Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T., and Kaneko, I. (2010). Transcription factor AP2-Sp and its target genes in malarial sporozoites. Mol. Microbiol. *75*, 854–863.