Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*

Cristina Bancells[1], Oriol Llorà-Batlle[1], Asaf Poran[2,3,4], Christopher Nötzel[5,6], Núria Rovira-Graells[1], Olivier Elemento[2,3], Björn F.C. Kafsack[5,6] and Alfred Cortés[1,7,*]

[1] ISGlobal, Hospital Clínic - Universitat de Barcelona, 08036 Barcelona, Catalonia, Spain

[2] Institute for Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

[3] Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA

[4] Physiology Biophysics and Systems Biology Graduate Program, Weill Cornell Medicine, New York, NY, USA

[5] Biochemistry, Cell & Molecular Biology Graduate Program, Weill Cornell Medicine, New York, NY, USA

[6] Department of Microbiology & Immunology, Weill Cornell Medicine, New York, NY, USA

[7] ICREA, 08010 Barcelona, Catalonia, Spain

[*]Corresponding author. E-mail: alfred.cortes@isglobal.org, Tel.: +34 93 2275400

ABSTRACT

Human to vector transmission of malaria requires that some blood stage parasites abandon asexual growth and convert into non-replicating sexual forms called gametocytes. The initial steps of gametocytogenesis remain largely uncharacterized. Here we studied this part of the malaria life cycle in *Plasmodium falciparum* using PfAP2-G, the master regulator of sexual conversion, as a marker of commitment. We demonstrate the existence of PfAP2-G-positive sexually-committed parasite stages preceding the previously known committed schizont stage. We also found that sexual conversion can occur by two different routes: the previously described route where PfAP2-G-expressing parasites complete a replicative cycle as committed forms before converting into gametocytes upon reinvasion, or a direct route with conversion within the same cycle as initial PfAP2-G expression. The latter route is linked to early PfAP2-G expression in ring stages. Re-analysis of published single-cell RNA-seq data confirmed the presence of both routes. Consistent with these results, using plaque assays we observed that, in contrast to the prevailing model, many schizonts produced mixed plaques containing both asexual parasites and gametocytes. Altogether, our results reveal unexpected features of the initial steps of sexual development and extend the current view of this part of the malaria life cycle.

Malaria symptoms result from repeated cycles of asexual replication of *Plasmodium* spp. parasites inside of erythrocytes. During each 48 h replicative cycle, a small fraction of the parasites abandon asexual growth and differentiate into non-replicating gametocytes that mediate human to vector transmission. In the case of *P. falciparum*, which is responsible for the deadliest form of human malaria, gametocyte maturation proceeds over ~10 days until they are infective to mosquito vectors[1-6].

The process by which parasites commit to sexual development and subsequently convert into gametocytes is not completely understood. Two models for gametocytogenesis were put forward almost forty years ago, proposing that gametocyte formation may take place either within the same cycle of commitment, or after one additional round of multiplication[7]. Plaque assays, in which the progeny of individual schizonts is visualized in immobilized erythrocytes, demonstrated a non-random distribution of sexual and asexual parasites, such that some schizonts predominantly produced sexual forms whereas others only produced asexual forms[8]. Later studies combining plaque assays with immunofluorescence assay (IFA) analysis of sexual and asexual markers concluded that the progeny of a single schizont produces only sexual or only asexual parasites, implying commitment at the cycle before differentiation[9]. Based on these results, the currently accepted model for gametocytogenesis holds that parasites that commit to sexual development go through an additional cycle of replication as committed forms before sexual conversion/differentiation[1,3-6]. Since there are discrepancies in the terminology used for the initial steps and stages of sexual development, we propose the nomenclature shown in Fig. 1a, in which sexual commitment and conversion are defined as distinct, sequential steps of the process.

The transcription factor AP2-G was recently identified as the master regulator of sexual conversion in malaria parasites[10,11]. This protein, which belongs to the ApiAP2 family of DNA binding proteins[12], is essential for gametocyte production in *P. falciparum*, as it regulates the transcriptional program that mediates sexual conversion[10,13]. The *pfap2-g* gene is the only member of the ApiAP2 family in *P. falciparum* that carries the heterochromatin marks of epigenetic silencing H3K9me3 and PfHP1[10,14-16], and it belongs to a subset of genes that show clonally variant expression[17,18]. Depletion of factors involved in heterochromatin formation results in activation of the gene and enhanced rates of sexual conversion[19,20]. Altogether, these observations indicate that sexual conversion is regulated at the epigenetic level: the *pfap2-g* locus is silenced by

H3K9me3-based heterochromatin in asexually-growing parasites, and transition to a permissive chromatin state, which requires the protein GDV1[21,22], leads to activation of the gene and subsequent sexual conversion[10]. Once PfAP2-G is expressed it locks in commitment by driving its own transcription[13].

PfAP2-G provides a much needed specific marker for the highly unexplored sexually-committed parasite stages[1]. Here we used PfAP2-G to characterize the initial steps of sexual differentiation in *P. falciparum* and re-examined the idea that an additional round of replication after commitment is an obligate step.

RESULTS

When PfAP2-G is stabilized at the early ring stage, sexual conversion can occur within the same cycle of stabilization.

In the E5 PfAP2-G-DD parasite line[10] (E5-HA-DD hereafter), in which endogenous PfAP2-G is fused to the FKBP destabilization domain, PfAP2-G is stable only in the presence of the Shield-1 (Shld) ligand, whereas in its absence PfAP2-G is degraded and gametocytes never form[10]. When adding Shld to synchronized E5-HA-DD cultures at the ring stage, we observed single-nucleated Pfg27-positive parasites within as little as ~30 h after Shld addition, when the majority of parasites were still at the schizont stage (Fig. 1b). As Pfg27 is a well-established marker of early gametocytes from stage I onwards[1,2,4,23], this result suggests that asexual parasites can convert directly into gametocytes without going through an additional round of multiplication after PfAP2-G stabilization. To test this idea, we treated E5-HA-DD cultures at the ring stage simultaneously with Shld and N-acetyl-D-glucosamine (GlcNAc), which at the concentration used (50 mM) inhibits schizont maturation[24,25]. Indeed, mature gametocytes emerged in these cultures (Fig. 1c), demonstrating that gametocytes can form without going through the PfAP2-G-positive committed schizont stage. This suggests that sexual conversion can occur through a same cycle conversion (SCC) route, in addition to the canonical next cycle conversion (NCC) route involving an additional round of replication after commitment. We define the new SCC route as conversion to sexual gametocyte in the progeny of PfAP2-G-negative schizonts, without additional replication. This implies that commitment, as marked by PfAP2-G expression, and conversion to gametocyte can occur within the same cycle.

Next, we investigated whether the timing of PfAP2-G stabilization during the intraerythrocytic cycle affects the frequency of sexual conversion via the NCC or SCC routes. E5-HA-DD cultures were tightly synchronized to a 5 h window and Shld was added at 0-5, 10-15, 20-25 or 30-35 h post-invasion (hpi). When we added GlcNAc at the ring stage (~5-10 hpi) of the next cycle to measure sexual conversion by either route, we observed high conversion rates (≥25%) regardless of the timing of PfAP2-G stabilization (Fig. 1d). In sharp contrast, when we measured conversion within the same cycle of PfAP2-G stabilization by adding GlcNAc simultaneously with Shld, conversion was maximal when Shld was added shortly after invasion (0-5 hpi) and nearly negligible when PfAP2-G was stabilized at 20-25 hpi or later (Fig. 1e). These results indicate that when PfAP2-G is stabilized at the early ring stage, sexual conversion frequently occurs within the same growth cycle through the SCC route, whereas PfAP2-G stabilization at later stages results in conversion predominantly at the next cycle via the NCC route. Gametocyte activation and egress assays didn't reveal differences between gametocytes generated by the SCC or the NCC routes (Supplementary Fig. 1).

PfAP2-G-negative schizonts produce gametocytes within the first cycle after reinvasion.

The E5-HA-DD line shows an unusually high sexual conversion rate when Shld is added (Fig. 1d-e). In other parasite lines, high rates of *pfap2-g* activation would result in high sexual conversion and lower multiplication rates, posing a fitness cost, but this doesn't occur in the E5-HA-DD line maintained in the absence of Shld. This raises the possibility that the regulation of *pfap2-g* expression may be altered in this line. To determine whether parasites in which *pfap2-g* is regulated normally can convert via the SCC route, we tagged endogenous PfAP2-G with the fluorescent marker eYFP using the CRISPR-Cas9 system (E5-eYFP line, Supplementary Fig. 2) and used tightly synchronized E5-eYFP late schizont cultures to FACS-sort PfAP2-G-eYFP-negative parasites (Supplementary Fig. 3). The almost complete absence of PfAP2-G-eYFP-positive schizonts after sorting was validated by IFA (Fig. 1f). IFA analysis with antibodies against the very early gametocyte marker Pfs16[1-4,26] conducted 48 h after sorting, when most parasites that continued asexual growth were at the schizont stage, revealed the presence of new Pfs16-positive gametocytes (Fig. 1f). For these experiments, we used Pfs16 instead of Pfg27 as a gametocyte marker because we found that it is an absolutely specific marker that is never expressed in PfAP2-G-deficient parasites and its expression

starts earlier during gametocyte development (Supplementary Fig. 4), as previously reported[26]. These results show that gametocytes can arise from PfAP2-G-negative asexual schizonts in the first cycle after reinvasion, confirming conversion via the SCC route.

Schizonts can produce pure sexual or pure asexual plaques as well as mixed plaques.

Previous reports described that in immobilized cultures individual schizonts produce plaques comprised of either only gametocytes or only asexual forms[9], which led to the generally accepted view that sexual commitment always occurs in the cycle before conversion[1,3,4,6]. In the light of our identification of the SCC route we wanted to re-examine this observation: if commitment and conversion can occur within the same cycle, parasites arising from an asexual schizont could become gametocytes independently of the choice made by their sibling progeny, resulting in mixed plaques (Fig. 2a). One-cycle plaque assays (Fig. 2) performed with tightly synchronized cultures of the wild type 3D7 subclone E5 indeed revealed that almost 40% of gametocyte-containing plaques consisted of a mixture of Pfs16-positive gametocytes and multinucleated asexual parasites (Fig. 2d). This proportion of mixed plaques cannot be explained by multiply-infected erythrocytes in the overlaid schizonts preparation (Supplementary Table 1 and Supplementary Fig. 5). To confirm that the levels of mixed plaques reflect the use of the SCC route, in these experiments we included E5-HA-DD cultures treated with Shld at times that result in conversion predominantly via the SCC or the NCC routes, which revealed a much higher proportion of mixed plaques in SCC cultures (Fig. 2c-d). Additional characterization of the plaques is provided in Supplementary Fig. 5b-c. Of note, IFA analysis of PfAP2-G in >200 PfAP2-G-positive mature schizonts of the E5-HA-DD and E5-HA lines[10] (the latter expressing HA-tagged endogenous PfAP2-G) did not identify any schizont containing a mixture of HA-positive and HA-negative merozoites, supporting the view that mixed plaques arise from direct conversion via the SCC route and not from single schizonts containing sexual and asexual merozoites (Fig. 2a). Abundant mixed plaques were also observed in assays performed with a different 3D7 stock (Supplementary Table 2 and Supplementary Fig. 5d). Altogether, these results confirm that sexual conversion can occur at the same cycle of commitment in wild type parasites.

Single-cell transcriptomics confirms same cycle sexual conversion.

In a previous study, E5-HA-DD cultures with Shld added at the ring stage (~4-16 hpi) were used for single-cell RNA-seq analysis either during the commitment cycle (~30, ~36 and ~42 hpi of the same cycle of Shld treatment, cycle 1) or at the next cycle (stage I gametocytes, cycle 2)[13]. Cluster analysis of transcriptional patterns of individual parasites revealed that the majority of cycle 2 gametocytes fall within clusters 13 and 14. Notably, some parasites from the commitment cycle (cycle 1) also fall within these two clusters (Fig. 3a-b). Comparing the relative abundance of Shld-treated versus untreated cells from cycle 1 shows an especially strong enrichment of treated cells for cluster 13 (Fig. 3c and Supplementary Fig. 6). These results suggest that cycle 1 parasites that fall within cluster 13 (4.8% of the total cycle 1 cells, Fig. 3b) converted into stage I gametocytes at the same cycle of Shld treatment using the SCC route.

To confirm that parasites from the commitment cycle (cycle 1) that fall in cluster 13 are actually stage I gametocytes, we identified the most up-regulated genes expressed by cycle 2 stage I gametocytes in cluster 13 as compared to expression in cells from other clusters. For all of these genes, which include known early gametocyte markers such as *pfg27* or *pfgexp02*[23,27], and also for the well-established early gametocyte markers *pfs16* and *pfg14.748*, transcript levels in cycle 1 cluster 13 cells were higher than in cycle 1 cells from other clusters (Fig. 3d-e, Supplementary Table 3 and Supplementary Fig. 6-7). However, average *pfap2-g* expression was similar between cluster 13 and other clusters. This is because in contrast to gametocyte-specific markers, *pfap2-g* is abundantly expressed in sexually-committed cells[13], which account for *pfap2-g* expression within other clusters (Fig. 3e and Supplementary Fig. 6-7).

In spite of high overall transcriptional similarity between cycle 1 and cycle 2 cluster 13 cells (Fig. 3a,d), we identified a small number of transcriptional differences that are likely associated with the route of conversion (Supplementary Fig. 8). The early gametocyte marker *pfg14.748*[28] and two other genes with maximal expression in gametocytes[29] showed >3-fold higher transcript levels in parasites that converted via the NCC route (cycle 2).

Expression of *pfap2-g* is maximal at the ring and mature schizont stages.

The discovery of the SCC route predicts that initial *pfap2-g* expression can be activated in rings. To test this idea, we employed the gametocyte non-producer F12 line, which contains a non-sense mutation that results in a non-functional PfAP2-G protein[10,30]. This makes it ideally suited to determine the temporal dynamics of *pfap2-g* transcription in the absence of gametocytes or the PfAP2-G transcriptional feedback-loop[10,13]. Time-course analysis of *pfap2-g* relative transcript levels in F12 revealed a clear peak of expression in 10-15 hpi rings. In contrast, in E5 cultures transcript levels were already high at 0-5 hpi and decreased until 30-35 hpi before increasing again at 40-45 hpi (Fig. 4a).

In addition to late schizonts, 40-45 hpi samples already contain early rings of the new growth cycle. To determine whether mature schizonts actually express *pfap2-g*, we collected RNA at 45-50 hpi from cultures treated with the cyclic GMP-dependent protein kinase (PfPKG) inhibitor ML10, which blocks merozoite egress and reinvasion[31]. These experiments revealed expression of *pfap2-g* in mature schizonts (Fig. 4b), which was confirmed by analysis of magnet-purified schizonts (Supplementary Fig. 9a). As a cautionary note, we observed that measuring *pfap2-g* transcripts from RNA extracted with Trizol directly from Percoll-purified schizonts yielded artifactual results (Supplementary Fig. 9a).

To directly assess the effect of PfAP2-G protein on *pfap2-g* expression, we used cultures of the E5-HA-DD line with or without Shld. As expected, the *pfap2-g* temporal expression profile in E5-HA-DD without Shld was similar to F12, with peak expression at 10-15 hpi, whereas in Shld-treated cultures expression was high at all the time points analyzed except at 30-35 hpi (Fig. 4c), similar to E5. The higher expression in Shld-treated versus untreated cultures is consistent with PfAP2-G auto-regulating the expression of its own gene[10,13]. Characterization of the expression of other early gametocyte markers[4] in the same time-course experiments showed that *pfg14.744* transcripts[28] are absolutely specific for gametocytes, whereas *pfgexp5* transcripts[32] and low levels of *pfs16* and *pfg27* transcripts[23,26] can be detected in the absence of functional PfAP2-G and gametocytes (Supplementary Fig. 10), consistent with previous reports[17,32,33].

Overall, these results show that basal *pfap2-g* transcript levels peak at the ring stage, suggesting that activation is possible at this stage and likely involves a ring-specific ApiAP2 transcription factor[12]. In contrast, in the presence of functional PfAP2-G,

transcript levels are high at all stages except for late trophozoites/ early schizonts, consistent with previous reports[13,17,19,34]. Our conclusions are largely independent of the gene used for normalization (Supplementary Fig. 9b). Analysis of *pfap2-g* transcript levels during gametocyte development revealed a progressive decrease after stage I (Supplementary Fig. 9c).

PfAP2-G expression and sexual commitment can start before the schizont stage.

We previously described that PfAP2-G-HA is localized in the nucleus of some parasites in cultures synchronized to the ring, trophozoite, or schizont stage[10]. However, if sexually-committed rings and trophozoites exist, in IFA experiments they would be morphologically indistinguishable from sexual rings and stage I gametocytes, respectively. Using Pfs16 protein as a gametocyte marker[26] that is absent from sexually-committed parasites, we found that PfAP2-G-HA is expressed in the nucleus from presumably committed trophozoites to stage I gametocytes (Fig. 5a), although some stages could not be unambiguously identified by this approach (see below). The distribution of PfAP2-G within the nucleus shows very limited overlap with heterochromatin, as expected from the euchromatic location of the majority of its predicted targets[10,13], but in many parasites the PfAP2-G signal appeared to concentrate in the nuclear periphery (Supplementary Fig. 11a-b). At later stages of sexual development (stage II-V gametocytes) PfAP2-G-HA signal is near background levels over the whole parasite and does not show nuclear localization (Fig. 5b). Western blot analysis confirmed a decrease of PfAP2-G-HA protein levels with gametocyte development (Supplementary Fig. 11c).

We observed Pfs16 signal in some parasites without detectable pigment (Fig. 5a, sexual ring), consistent with the observation that expression of this marker starts before the previously reported 30-40 hpi[26] (Supplementary Fig. 4). These results suggest that single-nucleated, hemozoin-containing parasites that are positive for PfAP2-G but negative for Pfs16 are sexually-committed trophozoites rather than very early stage I gametocytes that don't express Pfs16 at detectable levels yet. To confirm that committed forms preceding the schizont stage occur, we analyzed tightly synchronized E5-HA-DD cultures by IFA with Shld added at times that result in different use of the SCC route (10-15 or 20-25 hpi) (Fig. 5c). IFA performed at 40-45 hpi of the same cycle of Shld addition unambiguously distinguished sexually-committed parasites (all multinucleated schizonts, PfAP2-G-positive/ Pfs16-negative) from stage I gametocytes (single-nucleated, Pfs16-

positive). At 30-35 hpi, the majority of PfAP2-G-positive/ Pfs16-negative parasites were mononucleated, suggesting that they are sexually-committed trophozoites. If they were stage I gametocytes in which Pfs16 is not detectable yet, many of them would become Pfs16-positive at 40-45 hpi, but this was not observed: the distribution of parasite types between the two IFA time points indicates that the majority of PfAP2-G-positive/ Pfs16-negative parasites at 30-35 hpi actually are committed trophozoites that later develop to committed schizonts. Many Pfs16-positive stage I gametocytes did not express nuclear PfAP2-G-HA and the proportion of nuclear PfAP2-G-HA-positive stage I gametocytes decreased between 30-35 and 40-45 hpi (Fig. 5d), indicating that PfAP2-G-HA is present in the nucleus of early but not older stage I gametocytes.

The SCC route implies that *de novo* expression of PfAP2-G can start before nuclear replication. Indeed, we observed PfAP2-G-eYFP-positive rings among the progeny of FACS-sorted PfAP2-G-eYFP-negative schizonts of the E5-eYFP line analyzed by IFA 20 h after sorting (Fig. 5e). In this set of experiments we again determined the proportion of new Pfs16-positive parasites (gametocytes) formed in the first cycle after reinvasion, which roughly corresponded to the proportion of PfAP2-G-positive rings (Fig. 5e). This result suggests that activation of PfAP2-G in rings typically results in SCC, although it remains possible that some PfAP2-G-positive rings develop as replicating sexually-committed forms (those would be sexually-committed rings). The idea that PfAP2-G expression can be activated before the committed schizont stage was also supported by IFA analysis over consecutive stages, which revealed a higher proportion of PfAP2-G-positive parasites in cultures at the ring or trophozoite stages than in schizonts of the previous cycle (Supplementary Fig. 12).

PfAP2-G is not essential from gametocyte stage I onwards.

To identify at which stages PfAP2-G is needed for productive sexual conversion via the NCC and the SCC routes, Shld was removed from tightly synchronized E5-HA-DD cultures at different times. Under conditions promoting the NCC route, PfAP2-G stabilization was required until at least the ring stage of the second cycle (sexual ring stage) for high gametocyte production (Fig. 5f). In contrast, when conversion was restricted to the SCC route, more than half of the gametocytes could still form when Shld was removed as early as 41-46 hpi of the first cycle (Fig. 5g). These results indicate that

for both conversion routes PfAP2-G is no longer essential at the gametocyte stage I or even earlier.

DISCUSSION

Based on the results described here and published data, we propose an extended model for the early steps of sexual differentiation in *P. falciparum*. After the chromatin at the *pfap2-g* locus adopts a permissive state[10,19,20], either spontaneously[18,35] or induced by external cues[1-6,36,37], *pfap2-g* is transcribed mainly at the ring and late schizont stages. High expression of the gene requires a positive feedback loop involving the PfAP2-G protein. In some parasites, *pfap2-g* expression starts at the early ring stage and PfAP2-G levels reach a threshold sufficient to trigger the sexual transcriptional program before the onset of nuclear division, resulting in expression of early gametocyte markers and inhibition of the replicative asexual program. In such parasites, sexual conversion initiates within the same cycle of commitment as marked by PfAP2-G expression (SCC route) (Fig. 5h). In contrast, in other parasites in which PfAP2-G expression is activated, the protein levels necessary to trigger conversion are not reached early enough for SCC. We predict the existence of a checkpoint after which conversion in the same cycle is no longer possible. These parasites complete the replicative cycle as sexually-committed forms, which can include at least sexually-committed trophozoites and schizonts. After reinvasion, the new rings (sexual rings) have the *pfap2-g* locus in an active chromatin conformation inherited from the previous cycle by epigenetic mechanisms[18], and already carry nuclear PfAP2-G protein ready to drive its own expression[10,13]. Together, these conditions guarantee that high PfAP2-G levels are reached very early during the ring stage, resulting in sexual conversion one cycle after commitment (NCC route) (Fig. 5h). The observation that PfAP2-G is absent from the nucleus and dispensable from early stages of gametocyte development is consistent with a role for this protein largely restricted to triggering the sexual transcriptional program, which later involves additional ApiAP2 transcriptional regulators[13,38,39].

The single evidence for the idea that an additional round of multiplication after sexual commitment is absolutely required[1,3,4,6,9] came from experiments that suggested that all merozoites from the same schizont generate only pure sexual or pure asexual clusters in plaque assays[9]. However, using an improved protocol for plaque assays with tightly synchronized cultures (see Methods for details on the modifications relative to previous

protocols that can explain the different results obtained), we found that many schizonts produce only asexual or only sexual plaques, as previously observed[8,9], but a substantial amount of schizonts produce mixed plaques that correspond to sexual conversion events via the SCC route. The SCC route implies that PfAP2-G expression can start before nuclear replication, which was demonstrated by *de novo* PfAP2-G expression in rings arising from FACS-sorted PfAP2-G-negative schizonts, and is also consistent with the temporal dynamics of basal *pfap2-g* expression. However, PfAP2-G expression may also be activated at different times, as shown by a recent study on induced sexual conversion demonstrating activation in late stages[37]. We cannot rule out the possibility that some changes at the chromatin or transcriptional level may already occur at the *pfap2-g* locus or at other loci involved in sexual commitment (e.g. *gdv1*[22]) in individual nuclei of the PfAP2-G-negative schizonts that produce parasites that later convert via de SCC route. Future research should characterize the events that result in *pfap2-g* activation, identify the precise stages at which they can occur, and test the possibility that the SCC and NCC routes may reflect alternative mechanisms of *pfap2-g* activation, e.g. by determining the requirement of GDV1[22] and the role of lysophosphatiddylcholine[37] in both conversion routes. In any case, the SCC route is clearly distinct from the canonical NCC route, as it doesn't involve PfAP2-G-expressing schizonts or a homogeneous sexual or asexual fate for all parasites originated from the same schizont. The idea that the asexual or sexual fate of a new ring is not always irreversibly determined from the previous cycle is also demonstrated by an accompanying manuscript on sexual conversion in *P. berghei*[40].

New studies should establish the relative importance of sexual conversion via the SCC or the NCC routes in different parasite lines, including parasites of a genetic background other than 3D7. The use of the SCC and NCC routes during natural malaria infections also remains to be determined, but it is reasonable to speculate that the existence of two alternative routes, one resulting in prompt conversion and the other amplifying the number of sexual forms, may confer an evolutionary advantage to the parasites.

Altogether, our results revise the current model for the formation of malaria transmission stages and provide a necessary framework for the design and interpretation of studies aimed at characterizing the initial molecular events involved. Of note, gametocytes are a priority target for public health interventions aiming to reduce transmission and eventually eliminate malaria[41,42].

METHODS

Parasites.

The 3D7 subclone E5 (derived from the 3D7-B stock[43]), the E5-PfAP2-G–HAx3 transgenic line expressing 3xHA-tagged endogenous PfAP2-G (clone 9A, here named E5-HA), the ligand-regulatable transgenic line E5-PfAP2-G–ddFKBP expressing PfAP2-G fused to a 3xHA-tag and the FKBP-derived destabilizing domain (clone C2, here named E5-HA-DD), the 3D7 stock at Imperial College (3D7-Imp.) and the gametocyte-deficient 3D7-subclone F12 have been previously described and characterized[10,30,44,45]. The generation of the E5-eYFP line is described below. These parasite lines typically show sexual conversion rates of 7-15% (E5 and 3D7-Imp.), 1-5% (E5-HA and E5-eYFP), 20-50% (E5-HA-DD + Shld), and 0% (F12 and E5-HA-DD without Shld). Parasites were cultured in B+ erythrocytes (3% hematocrit) under standard conditions with media containing Albumax II and no human serum. Erythrocytes were obtained from the "Banc de Sang i Teixits" (Barcelona) after ethical approval by the Hospital Clínic (Barcelona) ethics committee (comitè ètic d'investigació clínica, CEIC). Cultures were regularly synchronized by sorbitol lysis, which eliminates erythrocytes infected with late asexual stages (trophozoites and schizonts). In some experiments, cultures were tightly synchronized to a defined 5 h age window by purification of parasites at the schizont stage using Percoll gradients (63% Percoll) followed by sorbitol lysis 5 h later. For the purification of mature parasites for transcriptional analysis in some experiments we used Percoll-sorbitol gradients (80-60-40% Percoll in the presence of 4% sorbitol) or magnetic separation using a VarioMACS magnetic separator and CS columns (Miltenyi Biotec). The PKG inhibitor ML10[31] was used at a 80 nM concentration to completely inhibit merozoite egress. It was added at the trophozoite stage (25-30 hpi) and maintained until 45-50 hpi, when essentially all schizonts were fully mature. To stabilize PfAP2-G in the E5-HA-DD line, 0.5 µM AquaShield-1 (Shld, Cheminpharma) was added to the cultures.

Sexual conversion rates were measured by treating synchronized cultures at the ring-stage (5-10% parasitemia, day 0) with 50 mM GlcNAc (Sigma) for 5-7 days to eliminate asexual parasites, approximately as described[10]. After treatment, asexual parasites arrest at the trophozoite stage and later die, dispensing the need to dilute the cultures. The sexual conversion rate was calculated as the gametocytemia at day 5-7 relative to the initial rings parasitemia at day 0. Parasitemia and gametocytemia were measured by light

microscopy analysis of Giemsa-stained smears. We measured spontaneous sexual conversion. Hence, we didn't intentionally stress the cultures with high parasitemia or used any other common method for environmental induction of gametocyte formation[36,46,47].

Generation of the E5-eYFP line and FACS-sorting experiments.

To generate the E5-eYFP line expressing endogenous PfAP2-G as a fusion protein with eYFP, E5 cultures were co-transfected with 60 µg of plasmid pDC2-Cas9-U6-*hdhfr-ap2g* and 12 µg of linearized plasmid pHR*ap2g*-eYFP and selected with 10 nM WR99210 for 4 days as previously described[48]. The plasmid pDC2-Cas9-U6-*hdhfr-ap2g*, derived from pDC2-Cas9-U6-h*dhfr*[49], encodes a single guide RNA that recognizes a sequence located -29 to -10 bp upstream from the *pfap2-g* stop codon. The plasmid pHR*ap2g*-eYFP, derived from pL6-eGFP-yFCU[50], contains a homology region corresponding to the sequence immediately upstream of the Cas9 cleavage site (HR1, positions -576 to -8 bp from the *pfap2-g* stop codon) and a homology region corresponding to a sequence downstream of the gene (HR2, positions +168 to +647 bp from the *pfap2-g* stop codon). HR1 and HR2 flank an in frame recodonized version of the sequence between the end of HR1 and the stop codon followed by the *eyfp* coding sequence and the *hsp90 3'* regulatory region. Details about the generation of the plasmids and Southern blot validation of the E5-eYFP line are provided in Supplementary Fig. 2.

E5-eYFP cultures were tightly synchronized to a 5 h age window and at 38-43 hpi schizonts were Percoll-purified before FACS-sorting eYFP-negative parasites in a FACSAria SORP (BD Biosciences; 5 laser, 18 parameters) cell sorter. Cultures of the parental line E5 were processed in parallel to provide a control of eYFP-negative parasites. Parasites were maintained in regular parasite culture media at 37ºC all the time. All measurements were made using the 488 nm-laser at 100 MW. The cell sorting conditions were as follows: flow chip diameter, 100 µm; sheath solution, ISOTON (Beckman Coulter); sheath pressure, 20 PSI; sort mode, purity. The erythrocytes population was identified and gated on SSC-A vs FSC-A plots. To avoid the sorting of cell doublets or cell aggregates, single cells were selected on FSC-H vs FSC-A plots. E5-eYFP-negative parasites were selected from plots combining the fluorescence channels 488-525/50-505LP-A (PfAP2-G-eYFP) and 488-582/15-556LP, for a better visualization of the positive eYFP signal (Supplementary Fig. 3). Sorted eYFP-negative parasites were

collected into a tube with culture media and 2-5 µl of packed erythrocytes at 37ºC and placed back in culture. In each sorting experiment we obtained 2-4x10^6 PfAP2-G-negative parasites. The proportion of eYFP-positive schizonts was determined by flow cytometry and by IFA with antibodies against eYFP within 3 h of sorting (only pigmented parasites were counted). The proportion of eYFP-positive rings (among non-pigmented parasites) was determined by IFA 20 h after sorting, using parental wild type E5 cultures processed in parallel as a negative control and unsorted E5-eYFP cultures as a positive control. The proportion of Pfs16-positive parasites (among pigmented parasites) was determined by IFA within 3 h of sorting and also 48 h later when the majority of asexual parasites were again at the schizont stage. Flow cytometry data was collected using BD FASCDiVa v.6.1.3 software and analyzed using FlowJo 10.2.

Immunofluorescence assays (IFAs).

IFAs were performed on air-dried smears fixed for 10 min with 1% formaldehyde (Electron Microscopy Sciences) and permeabilized for 5 min with 0.1% Triton X-100 in PBS. After incubation with primary and secondary antibodies, nuclei were stained with DAPI (5 µg/ml; Applichem Lifescience). The primary antibodies used were rat anti-HA (1:100; Roche #11867423001; we used this antibody for all experiments except for experiments in Supplementary Fig. 12 that were performed with the rabbit anti-HA antibody), rabbit anti-HA (1:100; Life technologies #71-5500, lot #936618A1), rabbit anti-GFP (1:1,000; Invitrogen #A11122, lot #1828014; this antibody also reacts with eYFP), mouse anti-Pfs16[51] (1:400-1:2,000, 32F717:B02, a gift from Robert Sauerwein, Radboud University), mouse anti-Pfg27[23] (1:2,000, 4B2, a gift from Richard Carter, University of Edinburgh), rabbit anti-H3K4me3 (1:10,000; Merck-Millipore #04-745, lot #DAM1606783) and rabbit anti-H3K9me3 (1:1,000; Merck-Millipore #07-442, lot #DAM1810831). To validate the specificity of the primary antibodies, we determined their signal in parasites at stages in which the protein is not expressed, and in the case of antibodies against an artificial 3xHA or eYFP tag we compared the signal between transgenic parasites carrying the tag and wild type parasites at the same stage that do not carry the tag. Additional characterization of the temporal dynamics and specificity of the signal with the anti-Pfs16 and anti-Pfg27 antibodies is presented in Supplementary Fig. 4. The secondary antibodies used were goat-anti-rat IgG conjugated with Alexa Fluor 488 (1:1,000, Thermo Fisher #A11006), goat-anti-mouse IgG - Alexa Fluor 594 (1:1,000, Thermo Fisher #A11007), goat-anti-mouse IgG - Alexa Fluor 488 (1:1,000, Thermo

Fisher #A11029), donkey-anti-mouse IgG - Alexa Fluor 546 (1:1,000, Thermo Fisher #A10036), and goat-anti-rabbit IgG - Alexa Fluor 594 (1:1,000, Thermo Fisher #A11012). When a single antigen was detected, we used secondary antibodies coupled with Alexa Fluor 488. For co-staining of more than one antigen, we used the antibody coupled with Alexa Fluor 488 to detect the HA or eYFP tags (marking PfAP2-G) and antibodies coupled with Alexa Fluor 594 or 546 to detect the other antigen. After mounting with Vectashield medium (Palex Medical), preparations were visualized under an Olympus IX51 epifluorescence microscope. For quantitative determinations we always counted >100 cells (in each replicate experiment), or the number of cells indicated. Images were acquired with an Olympus DP72 camera using CellSens Standard 1.11 software, and processed (including pseudo-coloring) using ImageJ.

Plaque assays.

Plaque assays were performed following a previously described approach[9], but with several important modifications. The main changes, described in detail below, were: i) overlaid schizont cultures were incubated on the erythrocyte monolayer for only 5 h, rather than 18-22h. This resulted in tight synchronization of the parasites growing in the monolayers; ii) plaques were analyzed at ~40-45 hpi, before another round of reinvasion started; iii) to identify gametocytes by IFA we used antibodies against Pfs16, which is an earlier gametocyte marker than Pfg27[26] (validated in Supplementary Fig. 4); iv) we identified asexual parasites by the presence of multiple nuclei (in our assay using tightly synchronized cultures, at the time of plaque analysis asexual parasites are at the schizont stage) rather than using IFA with an anti-MSP2 antibody that recognizes an antigen expressed only during a narrow age window at the very end of the asexual cycle[9]. We analyzed the number of hemozoin crystals to distinguish multinucleated parasites from multiply infected erythrocytes; v) in addition to shaking, we avoided high parasitemia in the overlaid culture to reduce the number of multiply infected erythrocytes.

We consider that using tightly synchronized cultures and not restricting the identification of asexual parasites to a narrow age window provide especially important advantages. Using less synchronized cultures and a late schizont antigen to detect asexual parasites, some mixed plaques may be incorrectly classified as pure sexual plaques if the asexual parasites have not started expressing the marker (MSP2) yet, or entered a new replicative cycle, becoming mononucleated, MSP2-negative rings (in previous studies,

plaques containing parasites not stained with either the sexual or the asexual marker were included in the analysis and classified according only to parasites positive for one of the antigens[9]). In such cases, mixed plaques may be classified as pure sexual plaques because expression of gametocyte markers is maintained for several days. A low proportion of multiply-infected erythrocytes in the overlaid culture, resulting in a low background level of mixed colonies, is also critically important.

To prepare erythrocyte monolayers, UV-sterilized coverslips were first incubated in 35 mm Petri dishes with 10 µg/ml concanavalin A (Sigma) in PBS for 30 min at 37ºC. After washing twice with RPMI-HEPES (washing medium), a 1% erythrocytes suspension in the same medium was added and incubated for 2 h at 37ºC. To remove unbound erythrocytes, the dishes were carefully agitated and the media was removed by aspiration. After two additional washes with washing medium, dishes were maintained at 37ºC until parasite cultures were added.

Sorbitol-synchronized parasite cultures were grown under shaking conditions (130 rpm) and at low parasitemia to minimize the number of multiply-infected erythrocytes. The proportion of multiply-infected erythrocytes was determined at the ring stage by microscopy analysis of Giemsa-stained smears and validated by flow cytometry performed approximately as described[52], using a FACScalibur flow cytometer (Becton Dickinson) and SYTO 11 to stain nucleic acids (Supplementary Fig. 5a). When the majority of parasites reached the schizont stage, late forms were purified using Percoll gradients. Purified schizonts were diluted to 2-8x10^4 mature schizonts/ml in parasite culture media and 1.5 ml of the solution was placed on the erythrocyte monolayer. After 5 h incubation under regular culture conditions, dishes were carefully agitated and washed twice with complete parasite culture medium to remove the remaining schizonts that had not burst and reinvaded. This results in tight synchronization of the culture growing in the erythrocyte monolayer to a 5 h age window.

These assays were performed with the E5, 3D7-Imp. and E5-HA-DD parasite lines. E5-HA-DD cultures were tested under two different conditions: adding Shld approximately 18 h before Percoll purification (~30 hpi) to favor sexual conversion via the NCC route, or adding Shld just after washing out schizonts from the erythrocyte monolayer (0-5 hpi) to favor sexual conversion via the SCC route.

IFA analysis was performed 39-40 h after removing unbound parasites (parasite age ~40-45 hpi). To verify adequate progression of parasite development in the monolayers (e.g. the vast majority of asexual parasites had evolved to the schizont stage), we stained with Giemsa one of several replicate dishes. We selected for analysis preparations that had been seeded at a concentration of schizonts that resulted in widely spaced plaques (average of less than 1 plaque per field using a 63x objective). Coverslips containing the monolayers were air dried for ~40 min, delimitated with a hydrophobic pen (Electron Microscopy Sciences), fixed and permeabilized for IFA as described above. We used a mouse anti-Pfs16 monoclonal antibody[51] (1:400 dilution) and a goat-anti-mouse IgG - Alexa Fluor 488 secondary antibody at 1:1,000. Preparations were visualized on a Leica AF6000 or a Zeiss Cell Observer HS fluorescence microscope. Images were first acquired from >100 plaques and then from gametocyte-containing plaques until we obtained images from >100 plaques containing ≥1 Pfs16-positive parasite. Images were analyzed in a computer screen using LAS-AF Lite or Zen 2012 software.

We limited the analysis to clearly separated plaques containing at least two well-developed parasites (schizont or gametocyte stages) and only a single hemozoin-containing residual body from the parental schizont. The latter criterion was used to exclude plaques derived from multiply-infected erythrocytes in the overlaid schizont culture. Hence, in addition to having a very low abundance of multiply-infected erythrocytes in the starting culture (Supplementary Tables 1-2 and Supplementary Fig. 5a), the few plaques arising from multiply-infected erythrocytes were excluded during image analysis. At the time of analysis (~40-45 hpi), asexual parasites are at the schizont stage. Single nucleated parasites without hemozoin or Pfs16 signal, which were observed in some of the plaques, are merozoites that did not develop in the immobilized erythrocyte monolayer and were not scored. Presence of merozoites that failed to invade and absence of ring stage parasites was confirmed by Giemsa-stained smears prepared in parallel with samples for IFA analysis. A parasite was scored as asexual if it had multiple nuclei, a single hemozoin pigment, and was negative for Pfs16; and it was scored as sexual if it was positive for Pfs16. We didn't score a parasite as asexual if based on the number of hemozoin pigment signals and nuclei it may correspond to an erythrocyte infected with multiple single-nucleated parasites rather than a multinucleated parasite. Of note, very low abundance of multiply-infected erythrocytes is critically

important in the overlaid schizonts culture, but multiple infection in the immobilized erythrocytes cannot be avoided and doesn't interfere with the identification of pure or mixed plaques.

The expected proportion of mixed plaques attributable to multiply-infected erythrocytes in the overlaid culture was calculated according to the method described by Bruce *et al.*[9]. In brief, the calculation was based on the percentage of multiply-infected erythrocytes in the overlaid culture and the probability of a double infection producing a mixed plaque. We also calculated the expected proportion of mixed plaques attributable to multiply-infected erythrocytes among plaques containing ≥1 Pfs16-positive parasite. For this calculation, the percentage of infected erythrocytes in the overlaid culture that were multiply-infected was multiplied by the probability of a schizont in the overlaid culture being asexual (according to the experimentally determined sexual conversion rate).

Transcriptional analysis.

RNA for reverse transcription-quantitative PCR (RT-qPCR) was obtained using the Trizol method, DNAse treated, and column purified using a procedure optimized for low amounts of RNA[53]. Reverse transcription was performed with AMV reverse transcriptase (Promega) using a mixture of random primers and oligo (dT). cDNAs were analyzed by qPCR with the Power SYBR Green Master Mix (Applied Biosystems), using the standard curve method approximately as described[54]. Transcript levels were normalized against *ubiquitin-conjugating enzyme (uce*; PF3D7_0812600) or *seryl tRNA synthetase* (*serrs*; PF3D7_0717700), which show relatively stable expression across blood stages (www.plasmodb.org) and are commonly used to normalize gene expression[17,55]. Primers used for RT-qPCR analysis are described in Supplementary Table 4.

Gamete egress assay.

For these experiments we used the E5-HA-DD line and the parental E5 line. E5-HA-DD cultures were tested under two different conditions: adding Shld approximately at 30 hpi and GlcNAc at the next cycle (ring stage) to obtain gametocytes that mainly converted via the NCC route, or adding Shld simultaneously with GlcNAc at the very early ring stage to obtain gametocytes that converted via the SCC route. In the E5 line GlcNAc was added at the ring stage. Gamete egress assays were performed essentially as described[56] at day 9-15 after GlcNAc addition, when the majority of gametocytes were mature (stage V).

In brief, mature gametocyte cultures were incubated in complete parasite culture medium with 5 µg/ml WGA-Oregon Green 488 conjugate (Invitrogen #W6748) and 2 µg/ml Hoechst 33342 (Sigma) for 30 min at 37ºC. Cultures were washed once with washing medium and gametogenesis was induced by adding complete culture medium with 20 µM xanthurenic acid (XA; Sigma) and incubating at room temperature for 30 min. As a control of non-activated gametocytes, cultures were maintained in parallel at 37ºC without adding XA. Activated gametocyte preparations were diluted in PBS with 0.75% BSA + 20 µM XA and placed in a chambered cover glass (Thermo Scientific #155411) for live observation on a fluorescence microscope (Olympus IX51). Images were acquired using CellSens Standard 1.11 software and processed with Image J software.

Western blot.

To obtain total parasite protein extracts, infected erythrocytes were lysed with cold 0.1% saponin and washed with cold PBS containing protease inhibitors (Roche #11873580001). Isolated parasites were directly resuspended in SDS-PAGE loading buffer, boiled for 5 min at 95ºC and stored at -80ºC. Before performing SDS-PAGE, β-mercaphotetanol was added to a final concentration of 4% and samples boiled again for 5 min at 95ºC. Protein extracts were separated on 8% SDS-polyacrylamide gels and transferred to nitrocellulose membranes following standard procedures. We used a rat anti-HA monoclonal antibody (Roche #11867423001) at 1:200 and as a secondary antibody an HRP conjugated goat-anti-rat IgG (Thermo Fisher #A10549) at 1:1,000. To control for equal loading of parasite material between different samples, membranes were stripped with Restore stripping buffer (Thermo Scientific) according to the manufacturer's instructions and re-probed with an antibody against PfHSP70 (1:10,000; StressMarq Biosciences #SPC-186C, lot #1007) and an HRP conjugated goat-anti-rabbit IgG (Sigma #A6154) secondary antibody at 1:5,000.

Re-analysis of single-cell transcriptomic data.

Single-cell analysis was carried out using the Seurat R package as previously described[13]. Single-cell transcript counts were normalized to 10,000 transcripts per cell. Clustering resolution was chosen such that visually distinct groups of more than 100 cells were assigned to individual clusters.

Data availability.

The single-cell RNA sequencing data analyzed in this study has been deposited to the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) with the study accession code SRP116718. The authors declare that all other relevant data generated or analyzed during this study are included in this published article or its supplementary information files. Materials and protocols are available from the corresponding author on reasonable request.

Code availability.

The scripts used for analysis and figure generation of single-cell RNA-seq data are available on https://github.com/KafsackLab/.

REFERENCES

1 Alano, P. *Plasmodium falciparum* gametocytes: still many secrets of a hidden life. *Mol Microbiol* **66**, 291-302 (2007).

2 Baker, D. A. Malaria gametocytogenesis. *Mol Biochem Parasitol* **172**, 57-65 (2010).

3 Bousema, T. & Drakeley, C. Epidemiology and infectivity of *Plasmodium falciparum* and *Plasmodium vivax* gametocytes in relation to malaria control and elimination. *Clin Microbiol Rev* **24**, 377-410 (2011).

4 Josling, G. A. & Llinas, M. Sexual development in *Plasmodium* parasites: knowing when it's time to commit. *Nat Rev Microbiol* **13**, 573-587 (2015).

5 Meibalan, E. & Marti, M. Biology of Malaria Transmission. *Cold Spring Harb Perspect Med* **7** (2017).

6 Dixon, M. W., Thompson, J., Gardiner, D. L. & Trenholme, K. R. Sex in *Plasmodium*: a sign of commitment. *Trends Parasitol* **24**, 168-175 (2008).

7 Carter, R. & Miller, L. H. Evidence for environmental modulation of gametocytogenesis in *Plasmodium falciparum* in continuous culture. *Bull World Health Organ* **57 Suppl 1**, 37-52 (1979).

8 Inselburg, J. Gametocyte formation by the progeny of single *Plasmodium falciparum* schizonts. *J Parasitol* **69**, 584-591 (1983).

9 Bruce, M. C., Alano, P., Duthie, S. & Carter, R. Commitment of the malaria parasite *Plasmodium falciparum* to sexual and asexual development. *Parasitology* **100 Pt 2**, 191-200 (1990).

10 Kafsack, B. F. *et al.* A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507**, 248-252 (2014).

11 Sinha, A. *et al.* A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature* **507**, 253-257 (2014).

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

13 Poran, A. *et al.* Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature* **551**, 95-99 (2017).

14 Flueck, C. *et al. Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog* **5**, e1000569 (2009).
15 Lopez-Rubio, J. J., 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* **5**, 179-190 (2009).
16 Fraschka, S. A. *et al.* Comparative Heterochromatin Profiling Reveals Conserved and Unique Epigenome Signatures Linked to Adaptation and Development of Malaria Parasites. *Cell Host Microbe* **23**, 407-420 (2018).
17 Rovira-Graells, N. *et al.* Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res* **22**, 925-938 (2012).
18 Cortes, A. & Deitsch, K. W. Malaria Epigenetics. *Cold Spring Harb Perspect Med* **7** (2017).
19 Brancucci, N. M. *et al.* Heterochromatin protein 1 secures survival and transmission of malaria parasites. *Cell Host Microbe* **16**, 165-176 (2014).
20 Coleman, B. I. *et al.* A *Plasmodium falciparum* histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host Microbe* **16**, 177-186 (2014).
21 Eksi, S. *et al. Plasmodium falciparum* gametocyte development 1 (Pfgdv1) and gametocytogenesis early gene identification and commitment to sexual development. *PLoS Pathog* **8**, e1002964 (2012).
22 Filarsky, M. *et al.* GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* **359**, 1259-1263 (2018).
23 Carter, R. *et al. Plasmodium falciparum*: an abundant stage-specific protein expressed during early gametocyte development. *Exp Parasitol* **69**, 140-149 (1989).
24 Gupta, S. K., Schulman, S. & Vanderberg, J. P. Stage-dependent toxicity of N-acetyl-glucosamine to *Plasmodium falciparum*. *J Protozool* **32**, 91-95 (1985).
25 Ponnudurai, T., Lensen, A. H., Meis, J. F. & Meuwissen, J. H. Synchronization of *Plasmodium falciparum* gametocytes using an automated suspension culture system. *Parasitology* **93 (Pt 2)**, 263-274 (1986).
26 Bruce, M. C., Carter, R. N., Nakamura, K., Aikawa, M. & Carter, R. Cellular location and temporal expression of the *Plasmodium falciparum* sexual stage antigen Pfs16. *Mol Biochem Parasitol* **65**, 11-22 (1994).
27 Silvestrini, F. *et al.* Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* **9**, 1437-1448 (2010).
28 Eksi, S. *et al.* Identification of a subtelomeric gene family expressed during the asexual-sexual stage transition in *Plasmodium falciparum*. *Mol Biochem Parasitol* **143**, 90-99 (2005).
29 Lopez-Barragan, M. J. *et al.* Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* **12**, 587 (2011).
30 Alano, P. *et al. Plasmodium falciparum*: parasites defective in early stages of gametocytogenesis. *Exp Parasitol* **81**, 227-235 (1995).
31 Baker, D. A. *et al.* A potent series targeting the malarial cGMP-dependent protein kinase clears infection and blocks transmission. *Nat Commun* **8**, 430 (2017).
32 Tiburcio, M. *et al.* Specific expression and export of the *Plasmodium falciparum* Gametocyte EXported Protein-5 marks the gametocyte ring stage. *Malar J* **14**, 334 (2015).

33 Painter, H. J., Carrasquilla, M. & Llinas, M. Capturing in vivo RNA transcriptional dynamics from the malaria parasite *Plasmodium falciparum*. *Genome Res* **27**, 1074-1086 (2017).
34 Otto, T. D. *et al.* New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. *Mol Microbiol* **76**, 12-24 (2010).
35 Waters, A. P. Epigenetic Roulette in Blood Stream *Plasmodium*: Gambling on Sex. *PLoS Pathog* **12**, e1005353 (2016).
36 Fivelman, Q. L. *et al.* Improved synchronous production of *Plasmodium falciparum* gametocytes in vitro. *Mol Biochem Parasitol* **154**, 119-123 (2007).
37 Brancucci, N. M. B. *et al.* Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite *Plasmodium falciparum*. *Cell* **171**, 1532-1544 e1515 (2017).
38 Modrzynska, K. *et al.* A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the *Plasmodium* Life Cycle. *Cell Host Microbe* **21**, 11-22 (2017).
39 Yuda, M., Iwanaga, S., Kaneko, I. & Kato, T. Global transcriptional repression: An initial and essential step for *Plasmodium* sexual development. *Proc Natl Acad Sci USA* **112**, 12824-12829 (2015).
40 Kent, R. S. *et al.* Inducible developmental reprogramming redefines commitment to sexual development in the malaria parasite *Plasmodium berghei*. *Nat Microbiol*, doi: 10.1038/s41564-41018-40223-41566, in press (2018). This issue?
41 Alonso, P. L. *et al.* A research agenda to underpin malaria eradication. *PLoS Med* **8**, e1000406 (2011).
42 Sinden, R. E. Developing transmission-blocking strategies for malaria control. *PLoS Pathog* **13**, e1006336 (2017).

REFERENCES THAT APPEAR ONLY IN THE METHODS

43 Cortés, A., Benet, A., Cooke, B. M., Barnwell, J. W. & Reeder, J. C. Ability of *Plasmodium falciparum* to invade Southeast Asian ovalocytes varies between parasite lines. *Blood* **104**, 2961-2966 (2004).
44 Delves, M. J. *et al.* Male and female *Plasmodium falciparum* mature gametocytes show different responses to antimalarial drugs. *Antimicrob Agents Chemother* **57**, 3268-3274 (2013).
45 Delves, M. J. *et al.* Routine in vitro culture of *P. falciparum* gametocytes to evaluate novel transmission-blocking interventions. *Nat Protoc* **11**, 1668-1680 (2016).
46 Roncales, M., Vidal-Mas, J., Leroy, D. & Herreros, E. Comparison and Optimization of Different Methods for the In Vitro Production of *Plasmodium falciparum* Gametocytes. *J Parasitol Res* **2012**, 927148 (2012).
47 Brancucci, N. M., Goldowitz, I., Buchholz, K., Werling, K. & Marti, M. An assay to probe *Plasmodium falciparum* growth, transmission stage formation and early gametocyte development. *Nat Protoc* **10**, 1131-1142 (2015).
48 Knuepfer, E., Napiorkowska, M., van Ooij, C. & Holder, A. A. Generating conditional gene knockouts in *Plasmodium* - a toolkit to produce stable DiCre recombinase-expressing parasite lines using CRISPR/Cas9. *Sci Rep* **7**, 3881 (2017).
49 Lim, M. Y. *et al.* UDP-galactose and acetyl-CoA transporters as *Plasmodium* multidrug resistance genes. *Nat Microbiol* **1**, 16166 (2016).
50 Ghorbal, M. *et al.* Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat Biotechnol* **32**, 819-821 (2014).

51 Moelans, I. I. M. D. Pfs16, a potential vaccine candidate against the human malaria parasite Plasmodium falciparum. *Thesis University of Nijmegen*, ISBN 90-9007799-9007795 (1995).
52 Rovira-Graells, N., Aguilera-Simon, S., Tinto-Font, E. & Cortes, A. New Assays to Characterise Growth-Related Phenotypes of *Plasmodium falciparum* Reveal Variation in Density-Dependent Growth Inhibition between Parasite Lines. *PLoS ONE* **11**, e0165358 (2016).
53 Mira-Martinez, S. *et al.* Expression of the *Plasmodium falciparum* Clonally Variant *clag3* Genes in Human Infections. *J Infect Dis* **215**, 938-945 (2017).
54 Crowley, V. M., Rovira-Graells, N., de Pouplana, L. R. & Cortés, A. Heterochromatin formation in bistable chromatin domains controls the epigenetic repression of clonally variant *Plasmodium falciparum* genes linked to erythrocyte invasion. *Mol Microbiol* **80**, 391-406 (2011).
55 Aguilar, R. *et al.* Molecular evidence for the localization of *Plasmodium falciparum* immature gametocytes in bone marrow. *Blood* **123**, 959-966 (2014).
56 Suárez-Cortés, P., Silvestrini, F. & Alano, P. A fast, non-invasive, quantitative staining protocol provides insights in *Plasmodium falciparum* gamete egress and in the role of osmiophilic bodies. *Malar J* **13**, 389 (2014).

CORRESPONDENCE

Correspondence and requests for materials should be addressed to A.C. (alfred.cortes@isglobal.org).

ACKNOWLEDGMENTS

We are grateful to Pietro Alano (Istituto Superiore di Sanità, Italy) for the F12 line, to Michael J. Delves (Imperial College, UK) for the 3D7 line from Imperial College (3D7-Imp.), to Richard Carter (University of Edinburgh, UK) for the anti-Pfg27 monoclonal antibody, to Robert W. Sauerwein (Radboud University, The Netherlands) for the anti-Pfs16 monoclonal antibody, to Marcus Lee (Wellcome Sanger Institute, UK) for plasmid pDC2-Cas9-U6-hdhfr, to José-Juan López-Rubio (University of Montpellier, France) for plasmid pL6-eGFP-yFCU, to Manuel Llinás (Pennsylvannia State University, USA) for providing the E5-HA-DD line and a plasmid containing the *eyfp* gene followed by a *P. falciparum* terminator, and to Simon Osborne (LifeArc, UK) and David Baker (LSHTM, UK) for providing the compound ML10 and advice on its use. We are also grateful to Júlia Romero Ortolà and Carla Sànchez Guirado for assistance with the generation of the plasmids, and to Sara Pagans (Universitat de Girona, Spain) for critical reading of the manuscript. We are indebted to the Flow Cytometry core facility of the IDIBAPS for technical help. This work was supported by grants from the Spanish Ministry of Economy and Competitiveness (MINECO)/ Agencia Estatal de Investigación (AEI) [SAF2013-43601-R and SAF2016-76190-R to A.C.], co-funded by the European Regional

Development Fund (ERDF, European Union); and the Secretary for Universities and Research under the Department of Economy and Knowledge of the Catalan Government [2014 SGR 485 to A.C.]. C.B. was supported by postdoctoral fellowship 2011-BP-B 00060 from the same Secretary for Universities and Research (Catalan Government). O.L.-B. is supported by a FPU fellowship from the Spanish Ministry of Education, Culture and Sports (FPU014/02456). ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya. Single-cell experiments were supported by WCM internal startup funds (B.F.C.K.) and the NSF CAREER award (DBI-10549646 to O.E.), LLS SCOR (7006-13 & 7012016 to O.E.), Hirschl Trust Award (O.E.), Starr Cancer Consortium (I6-A618 to O.E.), and NIH (1R01CA194547 to O.E.). A.P. and C.N. were supported by WCM graduate fellowships.

AUTHOR CONTRIBUTIONS

C.B. and A.C. conceived the project. C.B., O.L.-B. and A.C. designed and interpreted the experiments. C.B., O.L.-B., N.R.-G. and A.C. performed the experiments. A.P. and B.F.C.K. analyzed single-cell RNA-seq data. A.P., C.N., O.E. and B.F.C.K. contributed resources or data. C.B. and A.C. wrote the article, with major input from O.L.-B. and B.F.C.K.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1. Gametocytes can form at the same cycle of PfAP2-G activation. a, Proposed nomenclature for the initial steps and stages of sexual differentiation. Sexual conversion is marked by the onset of gametocyte-specific expression of proteins absent from any replicating blood stages (asexual or sexually-committed). Sexual commitment is defined as a cell state that deterministically results in sexual conversion at a later point. Currently, sexually-committed forms are morphologically undistinguishable from their asexual counterparts. Sexual rings, which have also been named as ring gametocytes or gametorings, are still morphologically indistinguishable from asexual rings, but they mature into stage I gametocytes. Expression of known protein markers for the different stages is shown. Vertical lines indicate reinvasion. **b,** E5-HA-DD cultures with PfAP2-G

stabilized at the ring stage (+Shld) form stage I gametocytes before reinvasion. Gametocytes were detected as mononucleated parasites positive for the early gametocyte marker Pfg27. Similar results were obtained in four independent experiments. Scale bar, 5 µm. **c,** Gametocytes develop from E5-HA-DD cultures treated simultaneously with Shld and GlcNAc at the ring stage. Similar results were obtained in five independent experiments. Scale bar, 5 µm. **d-e,** Sexual conversion (proportion of parasites that differentiate into gametocytes, see Methods) in E5-HA-DD cultures with Shld added at different h post-invasion (hpi). GlcNAc was added at the ring stage of the next cycle to determine total sexual conversion (d) or simultaneously with Shld to measure SCC (e). Individual data points, mean and SEM from three independent biological replicates (except for 10-15 hpi, *n*=2) are shown. *p* values were calculated using one-way ANOVA. **f,** FACS-sorted PfAP2-G-eYFP-negative schizonts of the E5-eYFP line produce gametocytes within the first cycle after reinvasion. Bar charts show the proportion of PfAP2-G-eYFP-positive schizonts in unsorted controls and sorted samples, and the proportion of Pfs16-positives among pigmented parasites (within 3 h of sorting and 48 h later), as determined by IFA. In each experiment, >1,000 parasites of each sample were counted. Values are the average of four independent experiments, with S.E.M (except for Pfs16 “sorted”, *n*=3). *p* values were calculated using a two-tailed unpaired t-test with equal variance. Scale bar, 5 µm.

Figure 2. Plaque assays reveal that some individual schizonts generate mixed sexual and asexual plaques. a, Schematic of the possible types of plaques originated from schizonts expressing or not expressing PfAP2-G (expression indicated by red nuclei). Early gametocytes expressing the Pfs16 marker are indicated in green. Mixed plaques can arise from asexual schizonts if some rings activate PfAP2-G expression and convert within the same cycle (SCC route). **b,** Representative IFA images of pure asexual, mixed, and pure sexual plaques. Asexual parasites are multinucleated schizonts, Pfs16-negative and contain a large hemozoin pigment, whereas gametocytes are mono-nucleated, Pfs16-positive and have small hemozoin pigment granules. Arrowheads indicate free hemozoin pigment (residual body) from the parental schizont that originated the plaque. Single nucleated parasites without hemozoin pigment or Pfs16 signal are merozoites that failed to develop. Images are representative of four independent experiments, each including at least three different samples. Scale bar, 5 µm. **c,** Distribution of plaque types in the wild type line E5, E5-HA-DD treated with Shld at

0-5 hpi (favoring the SCC route), and E5-HA-DD treated at ~30 hpi of the previous cycle (favoring the NCC route). At least 100 plaques of each culture were counted in each experiment. **d,** Distribution of pure and mixed plaques only among plaques containing ≥1 Pfs16-positive parasite. At least 100 plaques of each culture containing ≥1 sexual parasite were counted in each experiment. In panels c and d, values are the average of three independent biological replicates (red dots), with S.E.M. The distribution of plaque types in panels c and d was significantly different between E5, E5-HA-DD-SCC and E5-HA-DD-NCC (p=0.000 using a two-tailed Fischer's exact test).

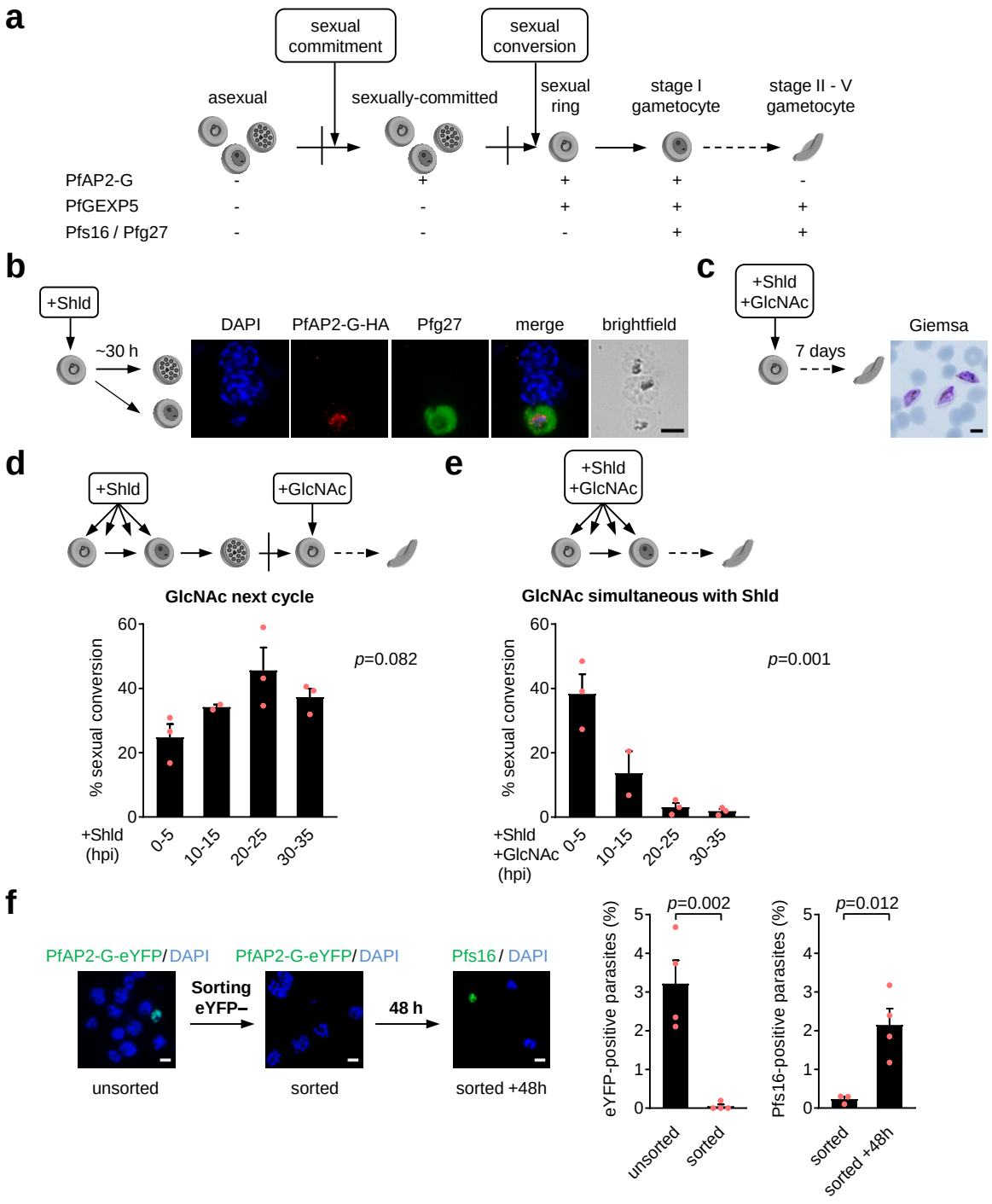
Figure 3. Single-cell RNA-seq identification and characterization of same cycle conversion (SCC) gametocytes. a, Cluster analysis of previously published single-cell transcriptomics data from parasites of the E5-HA-DD line treated with Shld at ~4-16 h post-invasion (hpi) and isolated at ~30, ~36 or ~42 hpi of the same cycle (cycle 1) or at ~42 hpi of the next cycle (cycle 2, stage I gametocytes). See Supplementary Figure 6 legend for a detailed definition of tSNE axes. The plot at the right shows the proportion of parasites from cycle 1 and cycle 2 in selected clusters (normalized per number of cells in each sample) from the analysis of 7,472 cells (5,736 from cycle 1 and 1,736 from cycle 2). The cycle 1 versus cycle 2 composition of cluster 13, cluster 14, and the combined remaining clusters were all non-randomly distributed (all $p<1x10^{-15}$, two-sided Exact Poisson test). **b,** Distribution of cells from cycle 1 (Shld-treated and untreated) and cycle 2 between the different clusters from the analysis of 10,509 cells. Cycle 1 values are normalized to account for the number of cells analyzed at each time point. **c,** Relative abundance of cycle 1 Shld-treated (bright shading) or untreated (pale shading) cells within each cluster. Abundance is normalized to account for the number of cells collected at each treatment condition. Numbers at the right show the relative enrichment for treated cells within each cluster (RE) and the number of cells in each cluster (n). Error bars indicate 95% confidence interval based on two-sided Exact Poisson test. **d,** Average expression of cluster 13 cycle 2 gametocyte markers (also see Supplementary Table 3 and Supplementary Fig. 6-7) shown for Shld-treated cycle 1 cells outside of cluster 13 (5,552 cells) or within cluster 13 (184 cells), and for Shld-treated cycle 2 cluster 13 cells (1,003 cells). Expression is normalized to 10,000 transcripts per cell. Mean and estimated 95% confidence intervals are shown. **e,** Same analysis for additional gametocyte markers.

Figure 4. Temporal dynamics of *pfap2-g* transcript levels. a, Reverse transcription-quantitative PCR (RT-qPCR) time-course analysis of *pfap2-g* expression in tightly synchronized cultures of the non-transgenic parasites lines F12 and E5. Parasite age is expressed in h post-invasion (hpi). **b,** Expression of *pfap2-g* in cultures in which 80 nM ML10 was added at 25-30 hpi, and in control cultures without the inhibitor. At 45-50 hpi ML10-treated cultures contained only mature schizonts, whereas control cultures already contained abundant rings. ML10 was washed out of treated cultures and RNA collected 2 h later, when substantial re-invasion had occurred ("wash +2 h"). Images of Giemsa-stained smears of the different preparations (representative of three independent experiments) are shown. Scale bar, 5 µm. **c,** Time-course analysis of *pfap2-g* expression in cultures of the E5-HA-DD line maintained in the absence of Shld (-Shld) or with Shld added at 0-5 hpi. Re-synchronization to a 5 h age window was performed between cycles 1 and 2. In cycle 1, *pfap2-g* expression was significantly higher in the presence of Shld compared to the -Shld condition at all time points analyzed (p=0.002, 0.002 and 0.027 at 10-15, 20-25 and 30-35 hpi, respectively, using a two-sided t-test with equal variance). In all panels, transcript levels were normalized against *ubiquitin-conjugating enzyme (uce).* Values are the average of two (c) or three (a-b) independent biological replicates (red dots). Error bars in panels a-b are S.E.M.

Figure 5. PfAP2-G expression dynamics. a, IFA analysis of PfAP2-G-HA and the early gametocyte marker Pfs16 in E5-HA cultures. Images are representative of six independent experiments. Scale bar, 2 µm. **b,** IFA analysis of PfAP2-G-HA and the gametocyte marker Pfg27 in E5-HA GlcNAc-treated gametocyte cultures. Images are representative of three independent experiments. Scale bar, 2 µm. **c,** IFA analysis of PfAP2-G-HA and the gametocyte marker Pfs16 at different h post-invasion (hpi) in Shld-treated E5-HA-DD cultures. **d,** Proportion of stage I gametocytes (Pfs16-positive) positive for PfAP2-G-HA nuclear signal. In panels c and d, values are the average of two independent biological replicates (red dots). **e**, IFA analysis of PfAP2-G-eYFP expression in rings arising from FACS-sorted eYFP-negative schizonts of the E5-eYFP line. IFA was performed 3 h ("sorted") and 20 h ("sorted +20h", ≤20 hpi rings) after sorting. Pfs16-expression and p values were determined as in Fig. 1f. Values are the average of four independent experiments (red dots), with S.E.M. **f,** Sexual conversion in E5-HA-DD cultures treated with Shld at 20-25 hpi and with GlcNAc at 5-10 hpi of the following cycle. Shld was removed at the times indicated by the horizontal bars. Gametocytemia was

determined at day 7 after GlcNAc addition by analysis of Giemsa-stained smears. “Relative sexual conversion” is the level of sexual conversion (%) relative to the condition in which Shld is present all the time. **g,** Same as in panel f, but GlcNAc was added simultaneously with Shld at 0-5 hpi to obtain only gametocytes formed by the SCC route. For consistency with panel f, age is expressed as if parasites continued the replicative cycle. In panels f and g, values are the average of three independent biological replicates, with S.E.M. **h,** A new model of the *P. falciparum* life cycle in the human blood. Red circles indicate nuclear PfAP2-G expression. When PfAP2-G expression is activated early during the ring stage, gametocyte differentiation proceeds without additional replication (same cycle conversion, SCC). In contrast, when PfAP2-G is activated later, parasites go through one additional round of replication as committed forms before differentiating into gametocytes (next cycle conversion, NCC).

Fig. 1.


a
sexual commitment
sexual conversion
asexual
sexually-committed
sexual ring
stage I gametocyte
stage II - V gametocyte
PfAP2-G - + + -
PfGEXP5 - - + + +
Pfs16 / Pfg27 - - - + +
b
+Shld
~30 h
DAPI
PfAP2-G-HA
Pfg27
merge
brightfield
c
+Shld +GlcNAc
7 days
Giemsa
d
+Shld
+GlcNAc
GlcNAc next cycle
% sexual conversion
p=0.082
+Shld (hpi)
0-5
10-15
20-25
30-35
e
+Shld +GlcNAc
GlcNAc simultaneous with Shld
% sexual conversion
p=0.001
+Shld +GlcNAc (hpi)
f
PfAP2-G-eYFP/DAPI
Sorting eYFP–
PfAP2-G-eYFP/DAPI
48 h
Pfs16/DAPI
unsorted
sorted
sorted +48h
eYFP-positive parasites (%)
p=0.002
Pfs16-positive parasites (%)
p=0.012

990

Fig. 2.

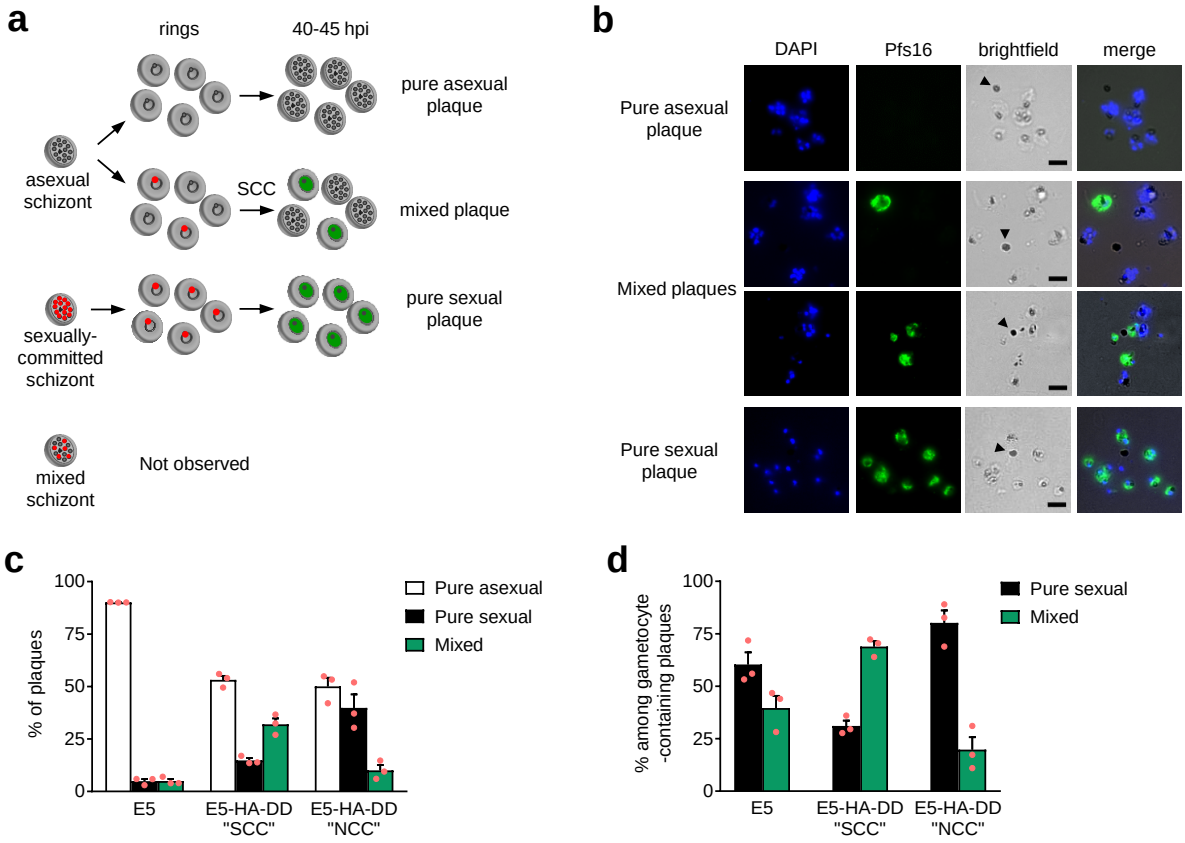

a
rings
40-45 hpi
pure asexual plaque
asexual schizont
SCC
mixed plaque
pure sexual plaque
sexually-committed schizont
mixed schizont
Not observed
b
DAPI
Pfs16
brightfield
merge
Pure asexual plaque
Mixed plaques
Pure sexual plaque
c
% of plaques
100
75
50
25
0
E5
E5-HA-DD "SCC"
E5-HA-DD "NCC"
Pure asexual
Pure sexual
Mixed
d
% among gametocyte -containing plaques
100
75
50
25
0
E5
E5-HA-DD "SCC"
E5-HA-DD "NCC"
Pure sexual
Mixed

Fig. 3.

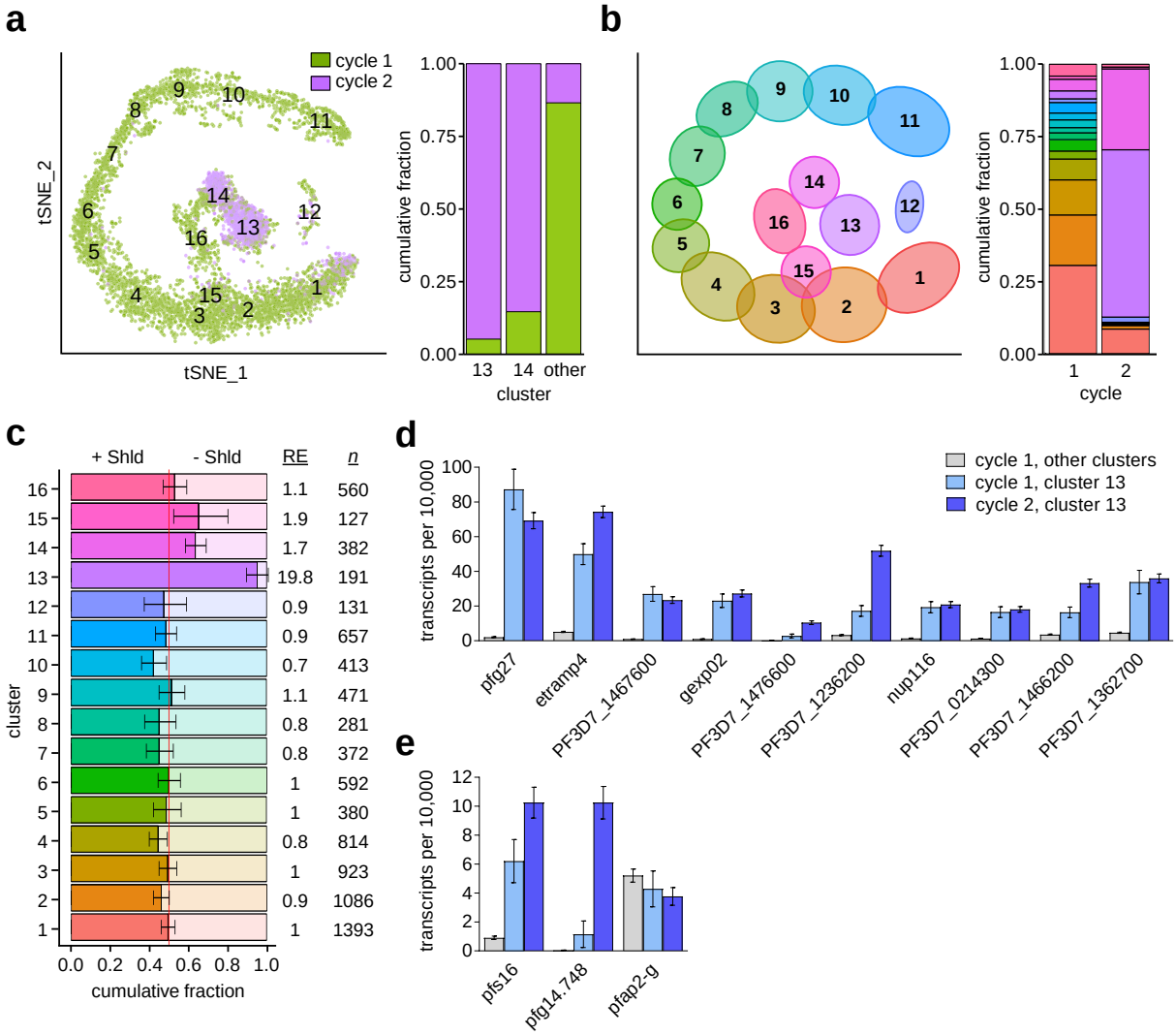

a
cycle 1
cycle 2
tSNE_2
tSNE_1
cumulative fraction
cluster
13
14
other
b
cycle
1
2
c
+ Shld
- Shld
RE
n
cluster
cumulative fraction
d
cycle 1, other clusters
cycle 1, cluster 13
cycle 2, cluster 13
transcripts per 10,000
pfg27
etramp4
PF3D7_1467600
gexp02
PF3D7_1476600
PF3D7_1236200
nup116
PF3D7_0214300
PF3D7_1466200
PF3D7_1362700
e
pfs16
pfg14.748
pfap2-g

Fig. 4.

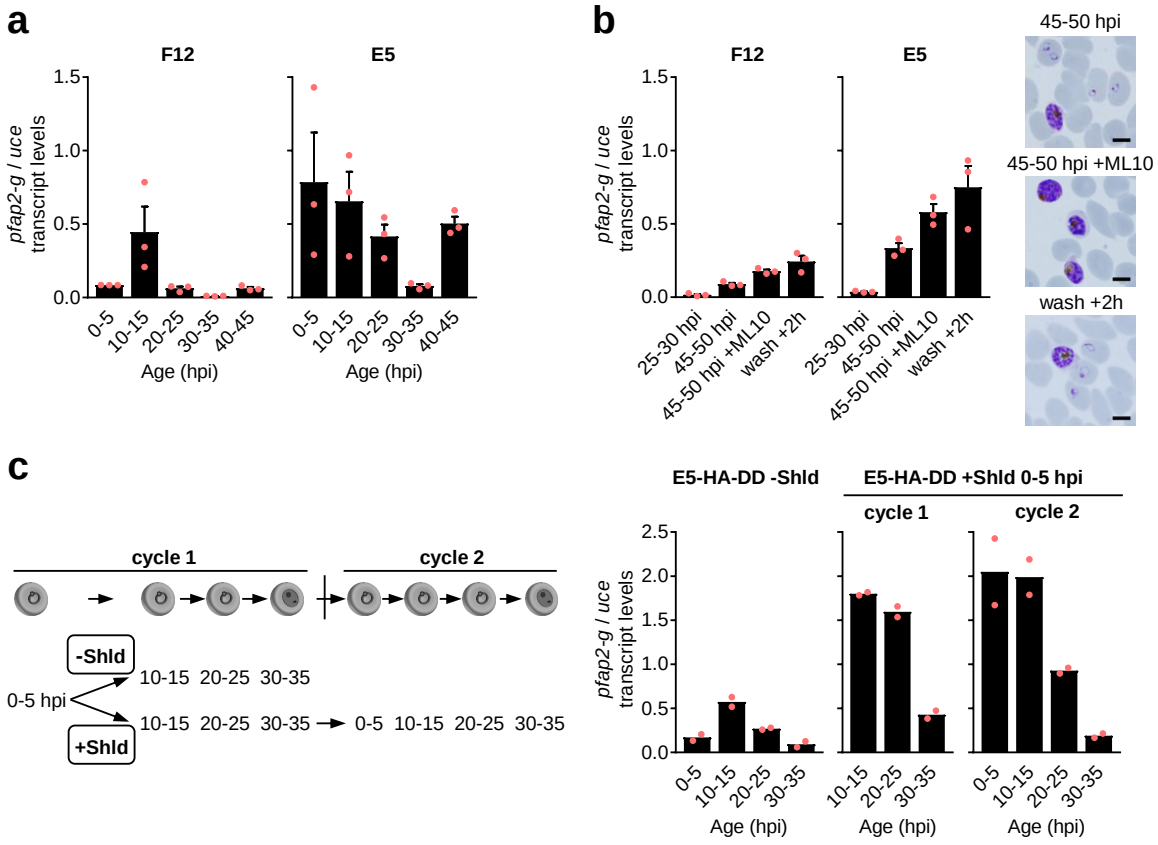
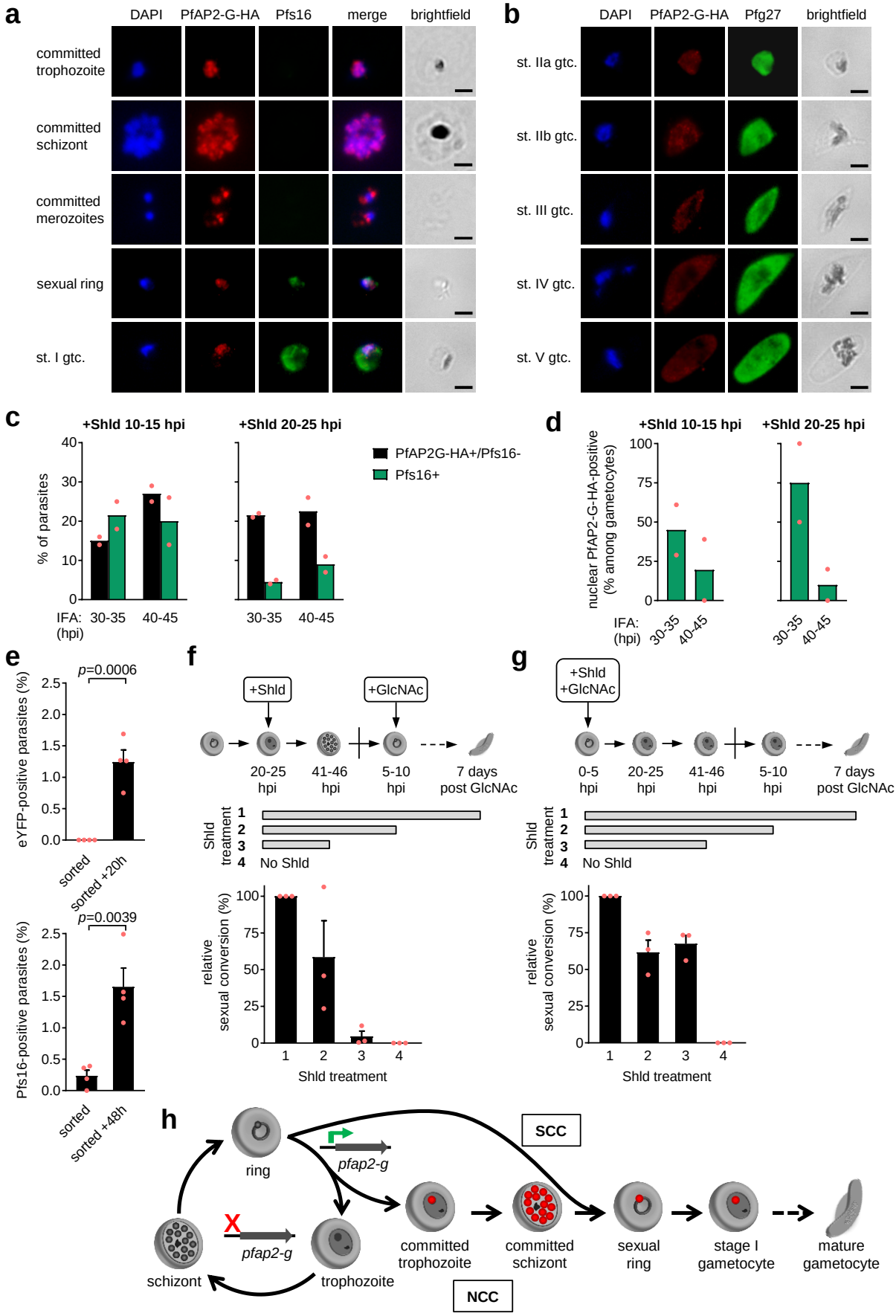

a
F12
E5
pfap2-g / uce transcript levels
1.5
1.0
0.5
0.0
0-5
10-15
20-25
30-35
40-45
Age (hpi)
b
F12
E5
pfap2-g / uce transcript levels
25-30 hpi
45-50 hpi
45-50 hpi +ML10
wash +2h
45-50 hpi
45-50 hpi +ML10
wash +2h
c
cycle 1
cycle 2
-Shld
+Shld
0-5 hpi
10-15 20-25 30-35
10-15 20-25 30-35 → 0-5 10-15 20-25 30-35
E5-HA-DD -Shld
E5-HA-DD +Shld 0-5 hpi
cycle 1
cycle 2
pfap2-g / uce transcript levels
2.5
2.0
1.5
1.0
0.5
0.0
Age (hpi)

Fig. 5.


a
DAPI
PfAP2-G-HA
Pfs16
merge
brightfield
committed trophozoite
committed schizont
committed merozoites
sexual ring
st. I gtc.
b
DAPI
PfAP2-G-HA
Pfg27
brightfield
st. IIa gtc.
st. IIb gtc.
st. III gtc.
st. IV gtc.
st. V gtc.
c
+Shld 10-15 hpi
+Shld 20-25 hpi
PfAP2G-HA+/Pfs16-
Pfs16+
% of parasites
IFA: (hpi)
30-35
40-45
d
+Shld 10-15 hpi
+Shld 20-25 hpi
nuclear PfAP2-G-HA-positive (% among gametocytes)
IFA: (hpi)
e
p=0.0006
eYFP-positive parasites (%)
sorted
sorted +20h
p=0.0039
Pfs16-positive parasites (%)
sorted
sorted +48h
f
+Shld
+GlcNAc
20-25 hpi
41-46 hpi
5-10 hpi
7 days post GlcNAc
Shld treatment
1
2
3
4
No Shld
relative sexual conversion (%)
Shld treatment
g
+Shld +GlcNAc
0-5 hpi
20-25 hpi
41-46 hpi
5-10 hpi
7 days post GlcNAc
Shld treatment
No Shld
relative sexual conversion (%)
Shld treatment
h
ring
pfap2-g
SCC
X
pfap2-g
schizont
trophozoite
committed trophozoite
committed schizont
sexual ring
stage I gametocyte
mature gametocyte
NCC

SUPPLEMENTARY INFORMATION

Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*

Cristina Bancells, Oriol Llorà-Batlle, Asaf Poran, Christopher Nötzel, Núria Rovira-Graells, Olivier Elemento, Björn F.C. Kafsack and Alfred Cortés

Supplementary Table 1. Observed frequency of mixed plaques and expected frequency if they originated only from multiply-infected erythrocytes in the overlaid culture. Plaque assays with the E5 and E5-HA-DD lines.

	Exp. 2			Exp. 3		
	E5	E5-HA-DD "SCC"	E5-HA-DD "NCC"	E5	E5-HA-DD "SCC"	E5-HA-DD "NCC"
Multiply-infected erythrocytes in the overlaid culture (%)	1.6	2.9	2.9	3.5	3.4	3.4
Observed mixed plaques (%)	**4.0**	**36.6**	**14.7**	**7.0**	**27.0**	**6.0**
Number of plaques analyzed (*n*)	101	101	102	100	100	100
Expected mixed plaques (%)	0.3	1.3	1.4	0.6	1.6	1.7
p value	0.000	0.000	0.000	0.000	0.000	0.000
Observed mixed plaques among Pfs16-positive plaques (%)	**46.7**	**72.3**	**31.1**	**44.0**	**64.0**	**11.0**
Number of plaques analyzed (*n*)	92	101	103	100	100	100
Expected mixed plaques among Pfs16-positive plaques (%)	1.5	2.5	2.0	3.4	2.9	1.7
p value	0.000	0.000	0.000	0.000	0.000	0.000

The proportion of mixed plaques observed is much higher than expected if they originated only from multiply-infected erythrocytes in the overlaid culture. Observed proportions of mixed plaques were obtained by counting ~100 plaques (top) or ~100 gametocyte-containing plaques (bottom) in each sample of each experiment (*n* is the number of plaques included in the analysis after excluding plaques that could not be unambiguously classified during the image analysis in a computer screen). The proportion of multiply-infected erythrocytes in the overlaid schizont cultures was determined by light microscopy analysis of Giemsa-stained smears when parasites were at the ring stage. Flow cytometry analysis gave similar results, but only microscopy values are shown and were used for all calculations. For two independent experiments, the expected proportion of mixed plaques arising from multiply-infected erythrocytes was calculated according to the method described by Bruce *et al.*[1], and the expected proportion of mixed plaques among plaques containing ≥1 Pfs16-positive parasite was determined as described in the Methods section (an accurate determination of the proportion of multiply-infected erythrocytes was not available for another independent biological replicate). *p* values were obtained using a two-tailed Fisher's exact test. In all cases, the proportion of mixed plaques observed was significantly higher than expected from multiply-infected erythrocytes only. Additionally, most plaques arising from multiply-infected

erythrocytes in the overlaid cultures were excluded during the image analysis step because they had more than one hemozoin-containing residual body (see Methods).

Supplementary Table 2. Observed frequency of mixed plaques and expected frequency if they originated only from multiply-infected erythrocytes in the overlaid culture. Plaque assays with the 3D7-Imp. line.

	Exp. 1	**Exp. 2**
Multiply-infected erythrocytes in the overlaid culture (%)	7.4	2.9
Observed mixed plaques (%)	**12.0**	**17.9**
Number of plaques analyzed (*n*)	100	106
Expected mixed plaques (%)	2.9	1.2
p value	0.000	0.000
Observed mixed plaques among Pfs16-positive plaques (%)	**33.0**	**40.6**
Number of plaques analyzed (*n*)	100	101
Expected mixed plaques among Pfs16-positive plaques (%)	6.0	2.5
p value	0.000	0.000

Comparison between the proportion of mixed plaques observed and the proportion expected if they originated only from multiply-infected erythrocytes in the overlaid culture, for experiments with the parasite line 3D7-Imp. The analysis was performed as in Supplementary Table 1. *p* values were obtained using a two-tailed Fisher's exact test. The number of mixed plaques observed was significantly higher than expected from multiply-infected erythrocytes only.

Supplementary Table 3. Single-cell RNA-seq analysis of genes differentially expressed in cluster 13.

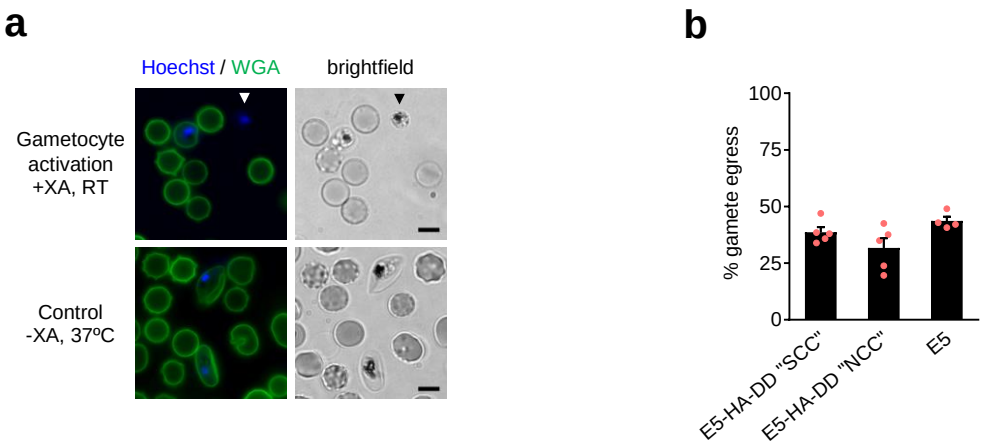
GeneID	Gene Name	Description	% expression in Cycle 2 Cluster 13	% expression outside of cluster 13	% difference in expression	expression fold-change	p.value	Max. Expression blood stage	Peptides detected in T / GC I-II / GC V)
PF3D7_1302100	pfg27	gamete antigen 27/25 (G27/25)	79.5%	10.7%	68.8%	16.6	0	GC II	0/31/29
PF3D7_0423700	etramp4	early transcribed membrane protein 4 (ETRAMP4)	92.3%	30.0%	62.3%	10.7	0	GC V	0/5/2
PF3D7_1467600		conserved Plasmodium protein unknown function	57.8%	7.6%	50.2%	9.4	0	GC II	0/19/0
PF3D7_1102500	gexp02	Plasmodium exported protein (PHISTb) unknown function (GEXP02)	64.5%	9.2%	55.3%	8.5	0	late T, GC V	7/35/24
PF3D7_1476600		Plasmodium exported protein unknown function	38.3%	2.3%	36.0%	8.5	2.07E-259	GC V	no data
PF3D7_1236200		conserved Plasmodium protein unknown function	79.6%	17.7%	61.9%	8.5	0	late T, GC V	no data
PF3D7_1473700	nup116	nucleoporin NUP116/NSP116 putative (NUP116)	55.6%	9.5%	46.1%	6.8	5.14E-268	late T	0/30/7
PF3D7_0214300		conserved Plasmodium protein unknown function	50.3%	10.8%	39.5%	5.4	1.45E-200	GC V	0/1/24
PF3D7_1466200		conserved Plasmodium protein unknown function	68.2%	22.1%	46.1%	4.8	9.84E-253	GC V	0/5/1
PF3D7_1362700		conserved Plasmodium protein unknown function	68.7%	27.0%	41.7%	4.7	3.78E-244	GC V	0/12/0

Differentially expressed genes between cycle 2 cluster 13 cells (1,003 cells) and cycle 1 and cycle 2 cells outside of cluster 13 (6,285 cells). The criteria for different expression was a minimum 4.5-fold change in average expression and 35% difference in the fraction of cells with detectable transcript. *p* values were calculated using a likelihood-ratio test for single-cell gene expression[2], not adjusted for multiple comparisons. The last two columns show expression of the genes in previous transcriptomic[3] or proteomic[4] analysis of asexual stages and gametocytes. “T” refers to trophozoites, “GC” refers to gametocytes.

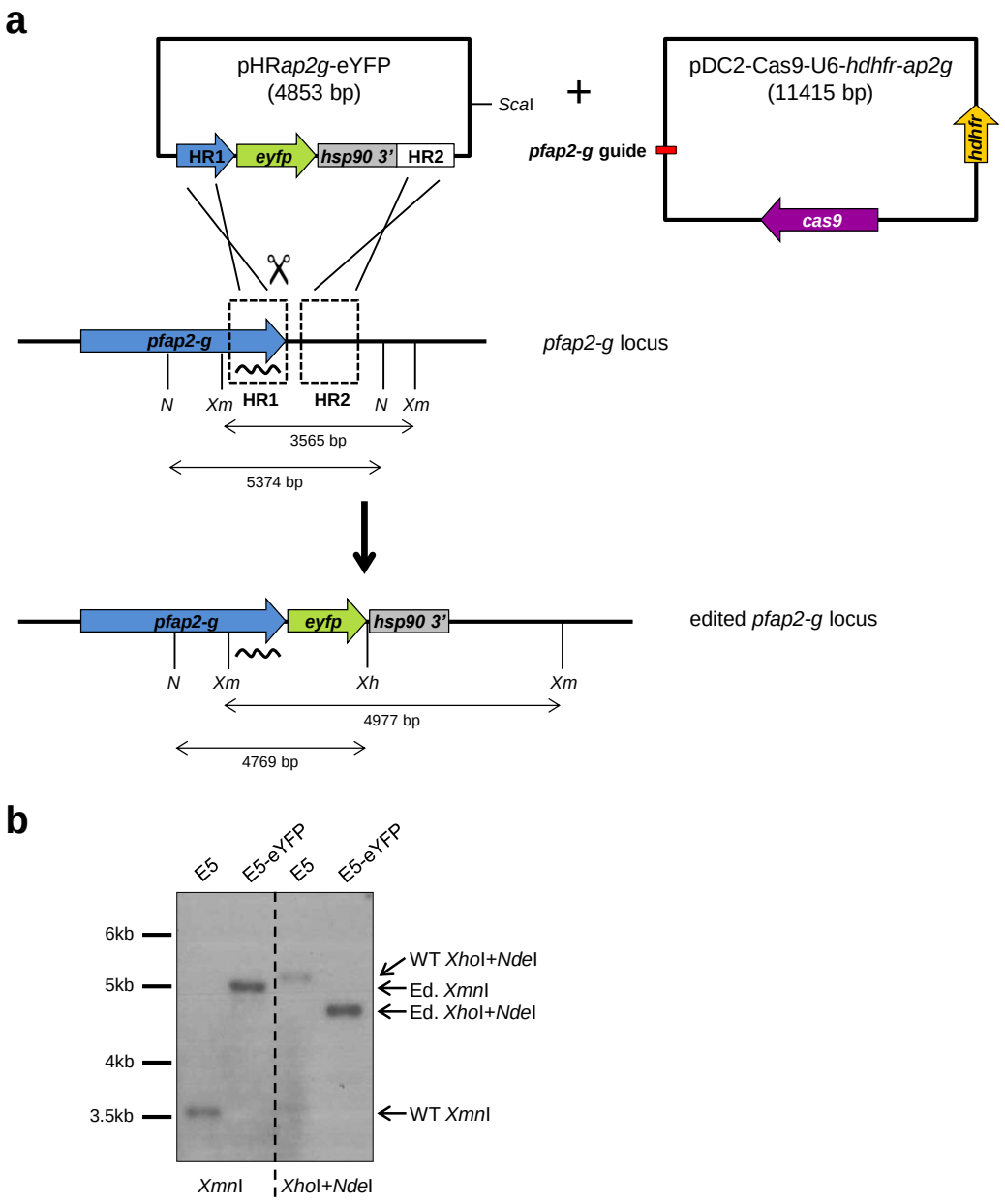
Supplementary Table 4. Primers used in this study.

Primer Name	Sequence 5' – 3'
Primers used for real time PCR	
PF3D7_1222600_F1 (*pfap2-g*) *	AACAACGTTCATTCAATAAATAAGG
PF3D7_1222600_R1 (*pfap2-g*) *	ATGTTAATGTTCCCAAACAACCG
PF3D7_0717700_F (*serrs*) [5]	AAGTAGCAGGTCATCGTGGTT
PF3D7_0717700_R (*serrs*) [5]	TTCGGCACATTCTTCCATAA
PF3D7_0812600_F (*uce*) [# 6]	GGTGTTAGTGGCTCACCAATAGGA
PF3D7_0812600_R (*uce*) [# 6]	GTACCACCTTCCCATGGAGTA
PF3D7_0406200_F (*pfs16*) [7]	TCAGGTGCCTCTCTTCATGCT
PF3D7_0406200_R (*pfs16*) [7]	GCTGAGTTTCTAAAGGCATTTTGTC
PF3D7_1477300_F (*pfg14.744*) [7]	GATGTACCGAAGTATGAGAATGATT
PF3D7_1477300_R (*pfg14.744*) [7]	TGGATAACGGCAAGGATATTTCTT
PF3D7_1302100_F (*pfg27*) [7]	GAAGCGTATCATGAACGACAAGA
PF3D7_1302100_R (*pfg27*) [7]	CTTATTCTTGCTGCTGCGTC
PF3D7_0936600_F (*pfgexp5*) *	GGATGCAAGTCCGAGGTAG
PF3D7_0936600_R (*pfgexp5*) *	TATGTGTACATGATTCCATTGG
Primers used to generate the E5-eYFP transgenic line	
pfap2g_HR1_F	ttgtta**actagt**GAACTTCAAAATGTGAATAATG
pfap2g_HR1_R	ccacaa**cttaag**gATgTTtCTGTTGTTTCCCCCTTTGTG
pfap2g_HR2_F	ttgtca**gaattcatcgat**CATCTTGTGACATTTTTAAATA
pfap2g_HR2_R	gaacaa**ccatgg**CACATATACATGTATGATTACA
pfap2g_gRNA_IF_F	taagtatataatattACCACAAAGGGGGAAACAACgttttagagctagaa
pfap2g_gRNA_IF_R	ttctagctctaaaacGTTGTTTCCCCCTTTGTGGTaatattatatactta
eyfp_IF_F	aacaacagaaacatc**agatct**gcagcagcagcagcagcaGTGAGCAAGGGCGAGGAGC
hsp90 3'_IF_R	aaatgtcacaagatg**gaattc**TATTTGATGAATTAACTACACTTA

All the primers for quantitative (real-time) PCR were used at 200 nM except those indicated with the symbols * (500 nM) and # (250 nM). A reference is provided for primers that had been described before. Restriction sites are indicated in bold and are underlined. Upper case indicates the part of the primer hybridizing with its target sequence.

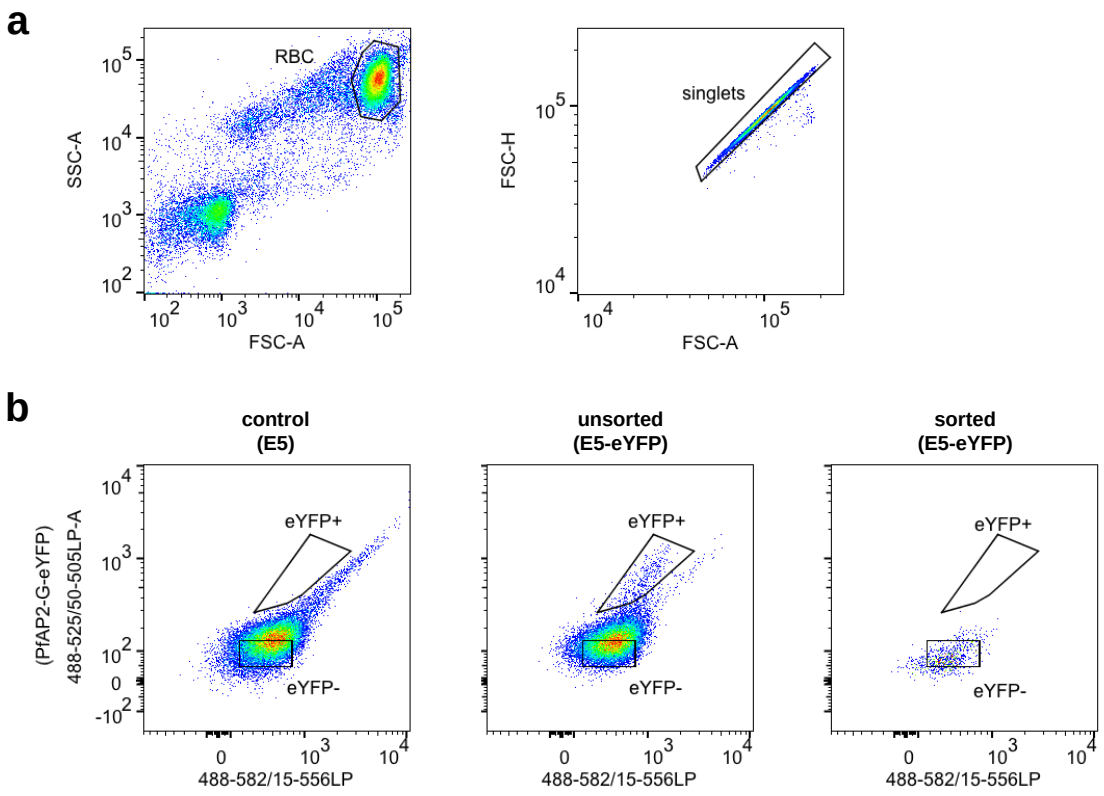


Supplementary Figure 1. Gametocytes formed by either the SCC or the NCC routes have similar gamete egress efficiency. a, Representative images of gametocyte activation and egress assays. Hoechst (blue) stains parasite DNA, whereas WGA-Oregon Green 488 (green) marks erythrocyte membranes. +XA, RT refers to xanthurenic acid addition and low temperature (room temperature) used to activate the gametocytes, whereas -XA, 37ºC refers to control conditions. Activation of gametocytes was evidenced by rounding up (observed in all samples), whereas egress was detected by absence of WGA surface staining (arrowhead). Under control conditions, gametocytes were elongated and did not egress. Images are representative of five independent experiments, with at least two independent samples each. Scale bar, 5 µm. **b,** Quantification of gamete egress efficiency for E5-HA-DD cultures under conditions that result in sexual conversion either via the SCC route or mainly via the NCC route (see Methods), and for the wild type parasite line E5. Parasites were scored as egressed gametes if they were round-shaped and free of WGA signal. 50 to 150 gametocytes were counted for each experiment. Values are the average of five (E5-HA-DD) or four (E5) independent biological replicates, with S.E.M. Red dots are individual data points. Differences in egress efficiency were not significant between any of the parasite lines or conditions (using a two-sided t-test with equal variance).

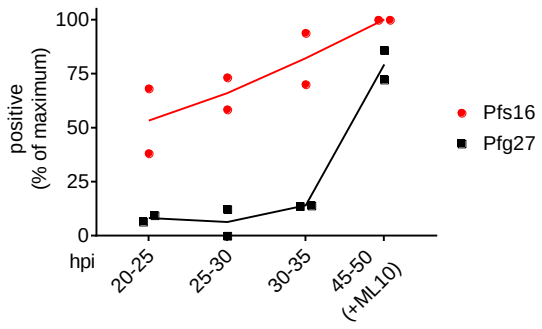


Supplementary Figure 2. Generation of parasites expressing endogenous PfAP2-G fused to eYFP using CRISPR-Cas9 technology. a. Schematic (not to scale) of *pfap2-g* tagging with eYFP using the CRISPR-Cas9 system. The plasmid pDC2-Cas9-U6-*hdhfr-ap2g* was generated by inserting a 20 bp *pfap2-g* guide obtained by annealing complementary oligonucleotides pfap2g_gRNA_IF_F and pfap2g_gRNA_IF_R (all primers are described in Supplementary Table 4) into the *Bbs*I digested plasmid pDC2-Cas9-U6-h*dhfr* [8] using the In-Fusion system (Takara). The scissors indicate the Cas9 cleavage site. The *pfap2-g* homology regions in plasmid pHR*ap2g*-eYFP were PCR-amplified from 3D7 genomic DNA using primer sets pfap2g_HR1_F and pfap2g_HR1_R (HR1) and pfap2g_HR2_F and pfap2g_HR2_R (HR2). Primer pfap2g_HR1_R included a recodonized sequence encoding

the amino acids between the cleavage site and the *pfap2-g* stop codon. The GFP homology regions from plasmid pL6-eGFP-yFCU[9] were sequentially replaced by *pfap2-g* HR1 (using *Spe*I-*Afl*II *sites*) and *pfap2-g* HR2 (*Nco*I-*Eco*RI sites). The resulting plasmid was digested with *Afl*II and *Eco*RI to replace the *hdhfr* gene by a fragment containing a *eyfp-hsp90 3'* sequence, obtained by PCR-amplification from plasmid pDC2-*ap2g*-eYFP (a gift from Manuel Llinás, The Pennsylvania State University, USA) using primers eyfp_IF_F and hsp90 3'_IF_R. The fragment was cloned in frame using the In-Fusion system. Finally, the U6 promoter-terminator guide cassette was excised by sequentially digesting with *Nco*I and *Aat*II, end-blunting using T4 DNA polymerase, and religation. Plasmid pHR*ap2g*-eYFP was linearized with *Sca*I before transfection. The positions of *Xmn*I (*Xm*), *Xho*I (*Xh*) and *Nde*I (*N*) restriction sites used for Southern blot analysis are indicated. The undulating line indicates the position of the probe used for Southern blot, which was PCR amplified using primers pfap2g_HR1_F and pfap2g_HR1_R. **b.** Southern blot analysis of E5 parasites transfected with plasmids pHR*ap2g*-eYFP and pDC2-Cas9-U6-*hdhfr-ap2g* and selected with WR99210 (E5-eYFP line). Parental E5 parasites were also analyzed. "Ed. *Xmn*I*"* and "Ed. *Xho*I+*Nde*I" indicate the position of the bands expected for the correctly edited *pfap2-g* locus in genomic DNA digested with the respective enzymes (4,977 and 4,769 bp, respectively). "WT *Xmn*I*"* and "WT *Xho*I+*Nde*I" indicate the position of the bands expected for the intact *pfap2-g* locus (3,565 and 5,374 bp, respectively). This analysis revealed that the *pfap2-g* locus was edited in essentially all E5-eYFP parasites, as bands corresponding to the wild type locus were not observed. A single Southern blot experiment was performed. Edition of the *pfap2-g* locus was also confirmed by PCR (data not shown).

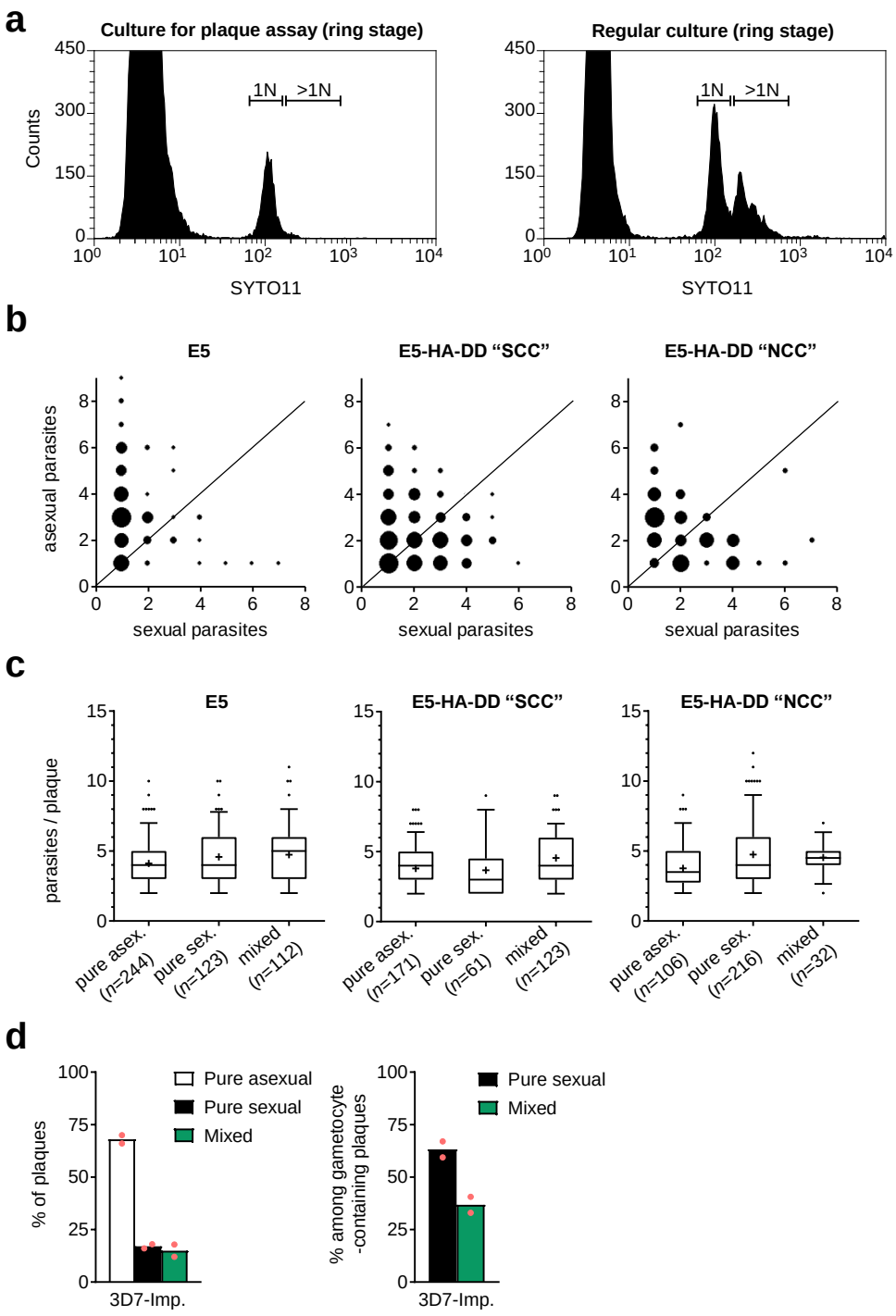


Supplementary Figure 3. FACS-sorting of PfAP2-G-negative schizonts. a, Representative images of the initial gates used to sort PfAP2-G-negative parasites. First, the erythrocytes (RBC) population was identified and gated on SSC-A vs FSC-A plots. Next, singlets were selected on FSC-A vs FSC-H plots. **b,** Representative images of 488-525/50-505LP-A (PfAP2-G-eYFP) vs 488-582/15-556LP (autofluorescence) plots. The plot on the left (control) corresponds to an E5 culture maintained and processed in parallel with E5-eYFP cultures, whereas the plot in the middle (unsorted) is an E5-eYFP culture before sorting and the plot on the right (sorted) is the same E5-eYFP culture after sorting eYFP-negative parasites. The gates used to identify eYFP-positive parasites, which are absent from E5 cultures, and the gate used to sort eYFP-negative parasites, are shown. The latter was defined conservatively to collect only parasites with very low eYFP signal. In the culture shown in this representative example, 1.8% of parasites were scored as eYFP-positive and 51.5% fell within the eYFP-negative gate. In all experiments, none of the sorted eYFP-negative parasites were eYFP-positive (0.0%). The plots shown in panels a and b are representative of eight independent biological replicates (independent sorting experiments). We attempted sorting of eYFP-positive parasites but the yield was very low and the purity never approached 100% (typically 60-70%); sorted eYFP-positive parasites were not used for any of the experiments presented in this article.



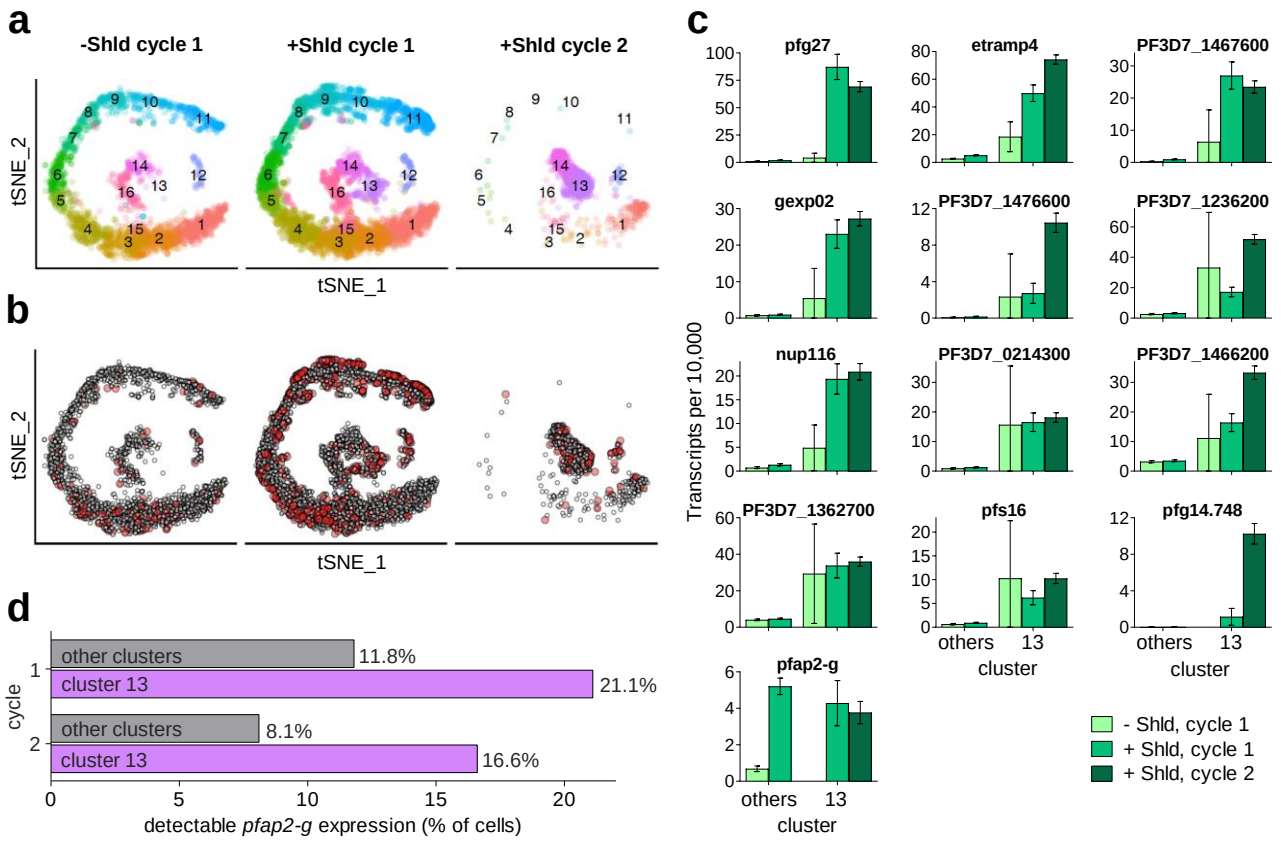
Supplementary Figure 4. Validation of the temporal expression dynamics of the early gametocyte markers Pfs16 and Pfg27. Time-course analysis of the expression of Pfs16 and Pfg27 in E5 cultures synchronized to a 5 h age window. The proportion of positive parasites was determined by IFA with antibodies against Pfs16 and Pfg27 at the times indicated in h post-invasion (hpi). 45-50 hpi samples were treated with ML10 to prevent reinvasion. Values are expressed relative to the highest proportion of positives in each experiment (i.e., Pfs16 at the latest time point). Individual data points and average (line) of two independent biological replicates are shown. For each time point and experiment, >1,000 parasites were scored as positive or negative for each marker. Double staining of the two markers was not possible because both antibodies available were raised in mice. Gametocytes formed at previous cycles are expected to be positive for the two markers from the first time point. They occur at low density because cultures are regularly diluted (asexual parasites multiply every 48 h whereas gametocytes are non-replicative). The temporal pattern observed indicates that the onset of Pfs16 expression in new gametocytes typically occurs between 20-35 hpi or even earlier, whereas Pfg27 expression starts at 35-45 hpi. However, the similar proportion of positive parasites between both markers at the latest time point indicates that sufficiently-mature gametocytes express Pfs16 and Pfg27 and that both are appropriate gametocyte markers, as previously reported[10,11]. While both antibodies clearly mark gametocytes, anti-Pfs16 antibodies show a better signal-to-background ratio, which together with earlier expression makes Pfs16 a superior marker for experiments such as plaque assays. The specificity of anti-Pfs16 antibodies was further demonstrated by experiments with the E5-HA-DD line cultured in the absence of Shld. No single Pfs16-positive parasite was ever observed in these PfAP2-G-deficient cultures (0 Pfs16-positives out of >1,000 parasites scored, and we never observed a Pfs16-positive parasite in such cultures in multiple experiments that were not analyzed quantitatively). In contrast, anti-

Pfg27 antibodies stain a small subset of schizonts (<0.5%) with a signal of similar intensity to gametocytes.



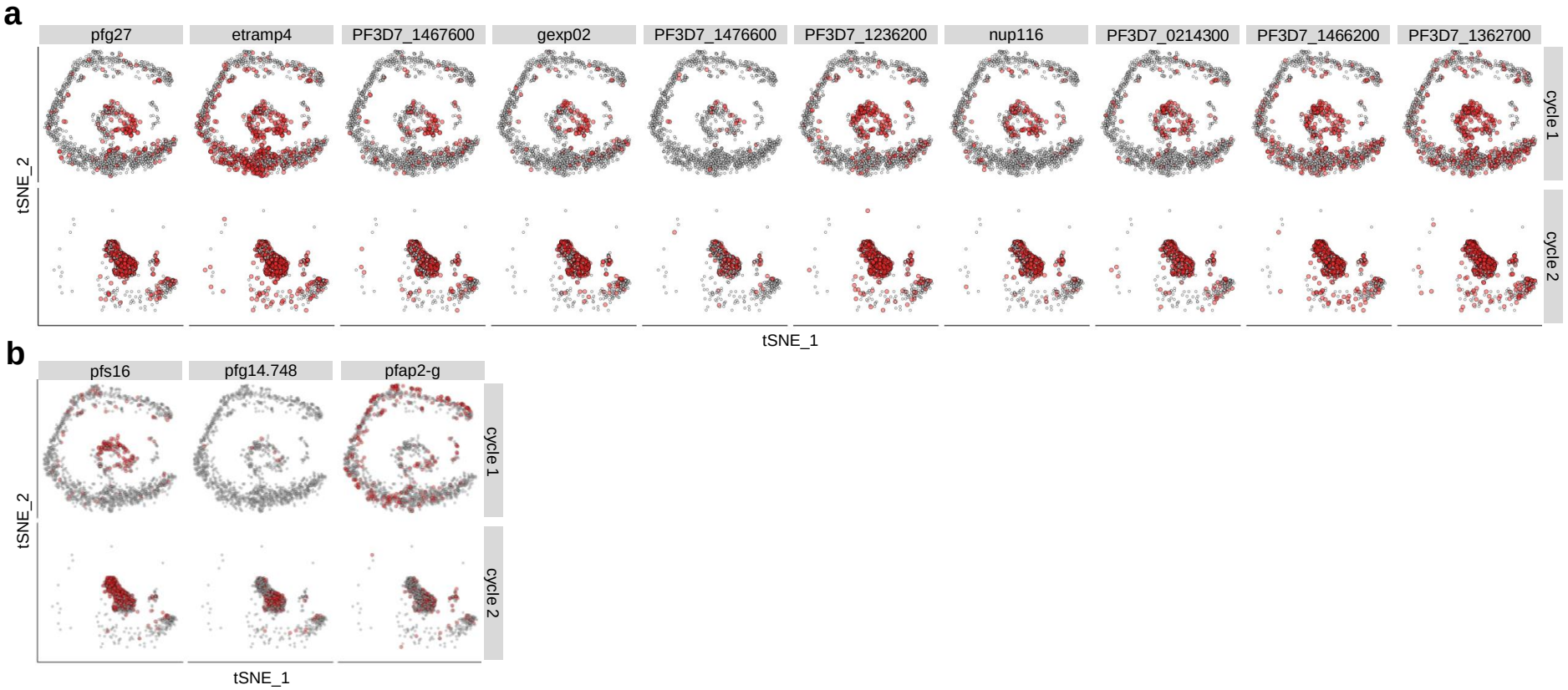
Supplementary Figure 5. The vast majority of mixed plaques do not originate from multiply-infected erythrocytes. a, Representative example of a flow cytometry analysis of synchronized cultures at the ring stage used to generate the overlaid schizont cultures for plaque assays. Cultures at the same stage prepared under regular conditions are shown for comparison. 1N indicates the position of the peak corresponding to single-infected

erythrocytes, whereas >1N are multiply-infected erythrocytes. Data analysis was performed using Summit v4.3 software. The plots are representative of two independent experiments (with two independent samples each). The comparison of the proportion of mixed plaques observed with the proportion expected if they originated only from multiply-infected erythrocytes in the overlaid cultures is presented in Supplementary Tables 1 and 2. The flow cytometry quantification of 1N and >1N cells was not used for any of the calculations shown in these Supplementary Tables. **b**, Distribution of sexual and asexual parasites in mixed plaques. Each dot represents a number of mixed plaques proportional to the size of the dot. The distribution observed roughly reflects the sexual conversion rate of each parasite line: the E5 line shows a clear predominance of asexual parasites in its mixed plaques, whereas the E5-HA-DD line shows a similar abundance of sexual and asexual parasites. This is consistent with the lower sexual conversion rate of E5 compared to E5-HA-DD (Fig. 2c in the main article). If mixed plaques originated mainly from multiply-infected erythrocytes, the distribution of sexual and asexual parasites within mixed plaques would be similar for all parasite lines. Plots include data from three independent biological replicates. **c**, Number of parasites in pure and mixed plaques. The number of parasites is similar in the different types of plaques: including all experiments with the different parasite lines there were an average of 3.94 parasites/plaque in pure asexual plaques, 4.53 in pure sexual plaques, and 4.63 in mixed plaques. Although the difference between pure asexual and other types of plaques was statistically significant (p=0.000 using a two-tailed unpaired t-test with unequal variance), this difference is of low magnitude and unlikely to have any biological significance. *n* indicates the number of plaques included in the analysis (plaques are from three independent experiments). Boxes show median and interquartile range, whereas whiskers are the 5-95 percentiles, "+" indicates the mean, and dots are outliers. **d**, Plaque assays performed with the parasite line 3D7-Imp. (the 3D7 stock at Imperial College)[12,13]. Experiments were performed and quantified as the experiments in Fig. 2c-d in the main article. Additional analysis of these experiments is presented in Supplementary Table 2. At least 100 plaques (left bar chart) or 100 plaques containing ≥1 Pfs16-positive sexual parasite (right bar chart) were counted in each experiment. Values are the average of two independent experiments, with individual data points (red dots).

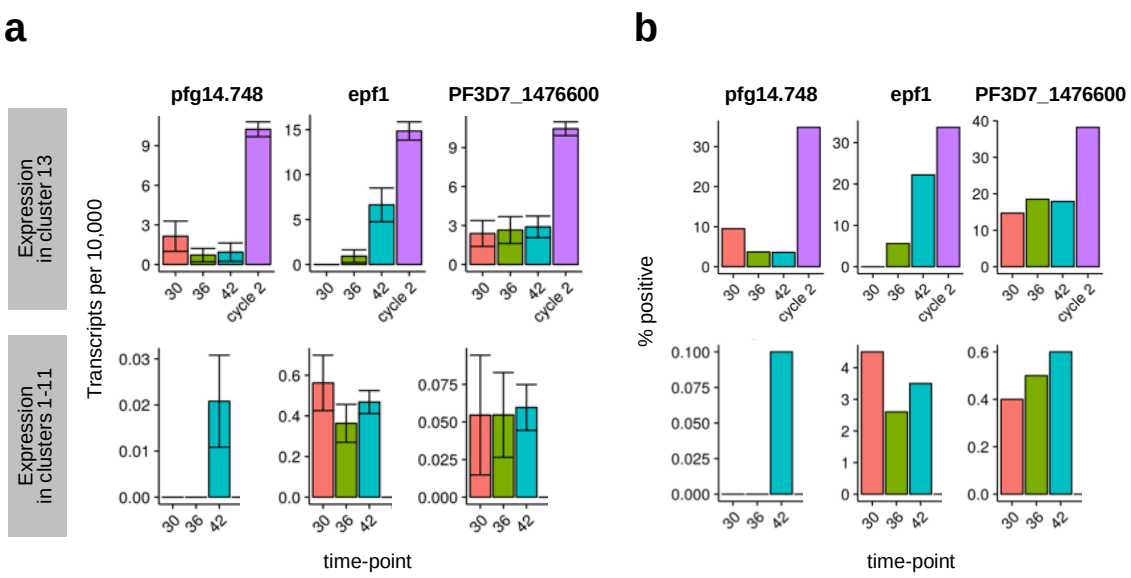


Supplementary Figure 6. Single-cell RNA-seq reveals a cluster of cells that corresponds to stage I gametocytes. a, Cluster analysis of single-cell RNA-seq data. Panels a and b show 2D representations of the single-cell transcriptomes using the tSNE algorithm. This non-linear dimensionality reduction algorithm aims to preserve similarity in the high-dimensional space and represent it as proximity in this 2D plot. The axes represent the two dimensions; however, they do not directly represent transcriptional features. Single-cell RNA-seq was performed on E5-HA-DD samples collected at ~30, ~36 or ~42 hpi of the same cycle of treatment with either solvent control (left panel, -Shld cycle 1, 3,037 cells) or Shld (middle panel, +Shld cycle 1, 5,736 cells), or samples collected at ~42 hpi of the next cycle (right panel, +Shld cycle 2, 1,736 cells). All samples were magnet-purified before the analysis to remove ring-stage parasites and retain only gametocytes and pigmented trophozoites/schizonts. Cycle 2 samples were treated with GlcNAc after reinvasion[14]. The majority of cells from cycle 2 (containing mainly stage I gametocytes after GlcNAc treatment and magnet purification) fall within clusters 13 or 14. Cluster 13 also contains cells from cycle 1 cultures treated with Shld, but virtually no cells from untreated cycle 1 cultures. **b,** Expression of *pfap2-g* (shown in red) in the different samples. The *pfap2-g* transcripts are detected in a significantly lower proportion of cells and only at earlier trophozoite time-points

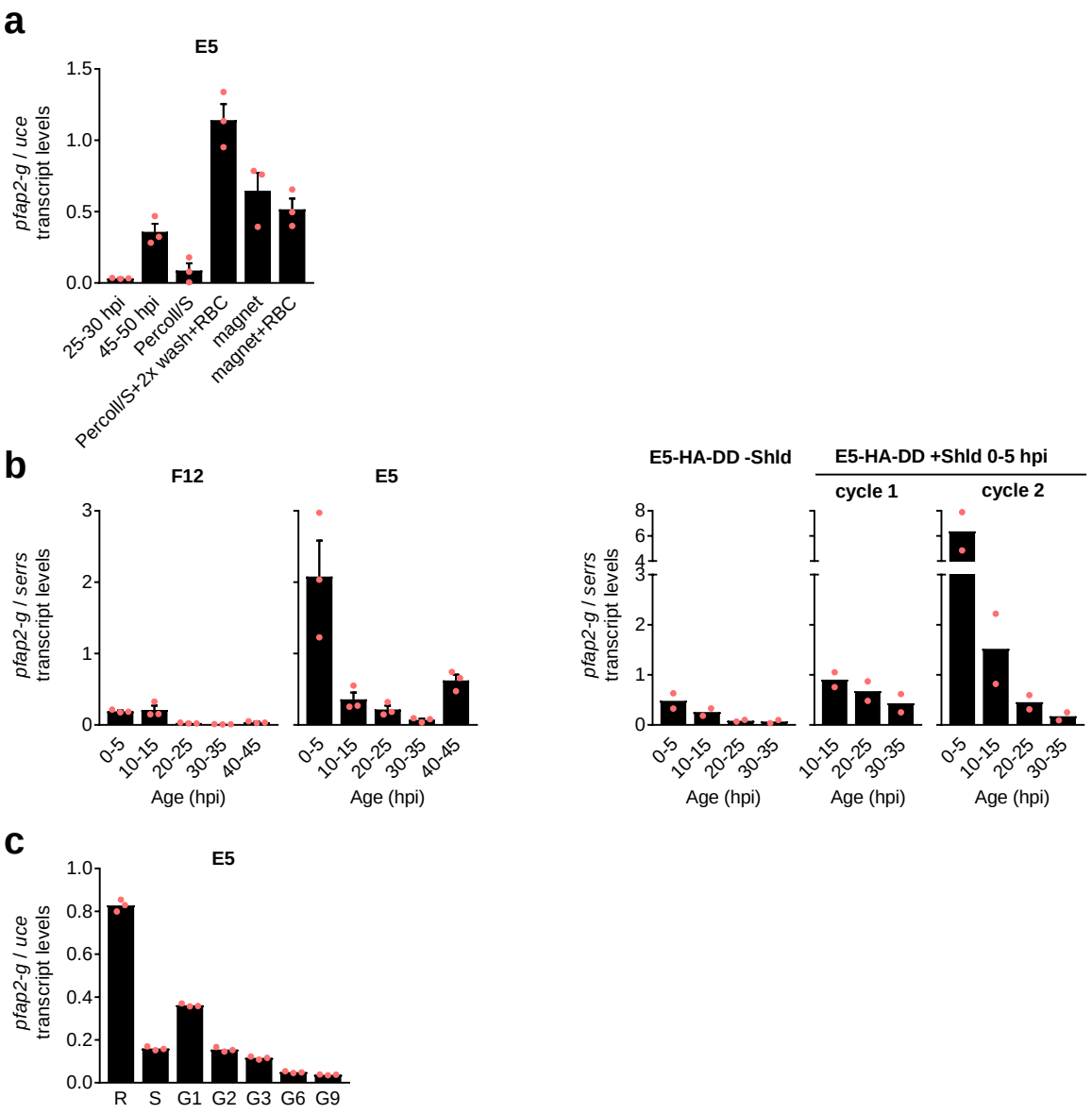
in untreated cultures. The same cells as in panel a are shown. **c,** Expression of genes most highly expressed in cluster 13 cycle 2 cells compared to non-cluster 13 cells (also see Supplementary Table 3 and Supplementary Fig. 7) and additional selected markers. Average expression in cluster 13 or other clusters is shown for untreated cycle 1 cells (others: 3,030 cells, cluster 13: 7 cells), Shld-treated cycle 1 cells (others: 5,552 cells, cluster 13: 184 cells) or Shld-treated cycle 2 cells (cluster 13: 1,003 cells). The very few untreated cells that fall within cluster 13 (7 cells) express some cluster 13 markers but have very low levels of transcripts for some established gametocytes genes such as *pfg27*, *pfg14.748* or *pfap2-g*. This can explain the absence of gametocytes in untreated cultures in spite of containing a few cells that fall within cluster 13. Expression is normalized to 10,000 transcripts per cell. Mean and 95% confidence intervals are shown. **d,** Percentage of Shld-treated cells with detectable *pfap2-g* transcripts from cycle 1 and from cycle 2 is shown for cluster 13 (stage I gametocytes; 184 cells for cycle 1, 1,003 cells for cycle 2) and for other clusters (5,552 cells for cycle 1, 733 cells for cycle 2).



Supplementary Figure 7. Single-cell RNA-seq analysis of genes differentially expressed in cluster 13 and other gametocyte markers. a, Expression of genes differentially expressed between cycle 2 cluster 13 cells and cycle 1 and cycle 2 cells outside of cluster 13. Differentially expressed genes are described in Supplementary Table 3. Expression is shown for Shld-treated cells in cycle 1 (top row) and cycle 2 (bottom row, composed mainly of stage I gametocytes). Equal numbers of randomly sampled cells (1,736 cells) are shown for cycle 1 and cycle 2. Cells with detectable transcripts are shown in red. See Supplementary Figure 6a legend for a detailed definition of tSNE axes. **b,** Same analysis for additional selected markers.

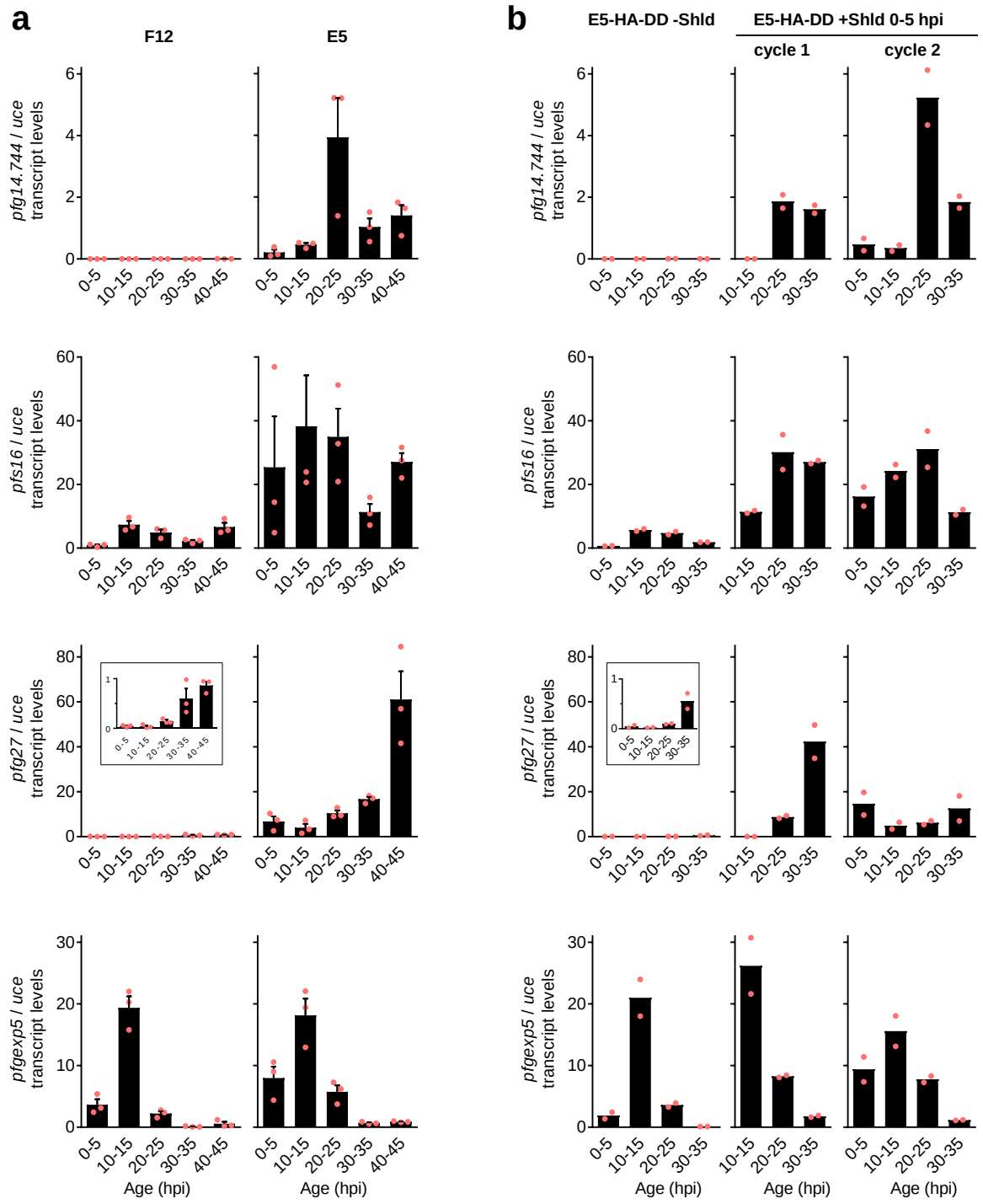


Supplementary Figure 8. Genes differentially expressed between cluster 13 cells from cycle 1 and cycle 2. The criteria to identify differentially expressed genes were a minimum 3-fold change in average expression and 20% difference in the fraction of cells with detectable transcripts. IDs: PF3D7_1477700 (*pfg14.748*); PF3D7_0114000 (*epf1*). Only genes showing higher expression in cycle 2 gametocytes are presented because the majority of genes showing higher expression in cycle 1 gametocytes are among the most abundantly transcribed genes in asexual parasites and their higher number of transcripts in cycle 1 gametocytes may be attributable to technical issues of single-cell RNA-seq. **a,** Average expression of differentially expressed genes in Shld-treated cells within cluster 13 from cycle 1 (time-points ~30 hpi: 41 cells; ~36 hpi: 47 cells; and ~42 hpi: 96 cells) and cycle 2 (1,003 cells) (top), or within clusters 1-11 cycle 1 (~30 hpi: 508 cells; ~36 hpi: 821 cells; and ~42 hpi: 3,356 cells) (bottom). Values are average number of transcripts per 10,000 transcripts, with S.E.M. For all genes, expression in cluster 13 was significantly higher in cycle 2 cells than in cycle 1 cells collected at either time point, using a two-sided t-test (*pfg14.748*: p=3x10^{-8}, 8x10^{-28} and 2x10^{-21} for 30, 36 and 42 h, respectively; *epf1*: p=5x10^{-44}, 2x10^{-25} and 2x10^{-4}; PF3D7_1476600: p=9x10^{-10}, 4x10^{-9}, 8x10^{-13}). **b,** Proportion of cells with detectable transcripts of the differentially-expressed genes.



Supplementary Figure 9. Analysis of *pfap2-g* relative transcript levels at different stages of development. a, Reverse transcription-quantitative PCR (RT-qPCR) analysis of *pfap2-g* transcript levels in mature schizonts. RNA was prepared from tightly synchronized E5 cultures at 25-30 and 40-45 hpi. At the latter time point the culture contains abundant rings in addition to schizonts, which potentially may account for the *pfap2-g* transcripts. Schizonts were purified from 40-45 hpi cultures using Percoll-sorbitol 80-60-40% gradients (Percoll/S) or using magnetic columns (magnet) before Trizol extraction of RNA. Percoll-sorbitol gradients were used instead or regular Percoll purification (see Methods) to obtain increased purity (~99%), but similar results were obtained when using regular Percoll purification (not shown). We also analyzed samples in which we extracted RNA from schizonts purified in Percoll-sorbitol gradients, performed an additional wash (the regular protocol includes a single washing step after Percoll

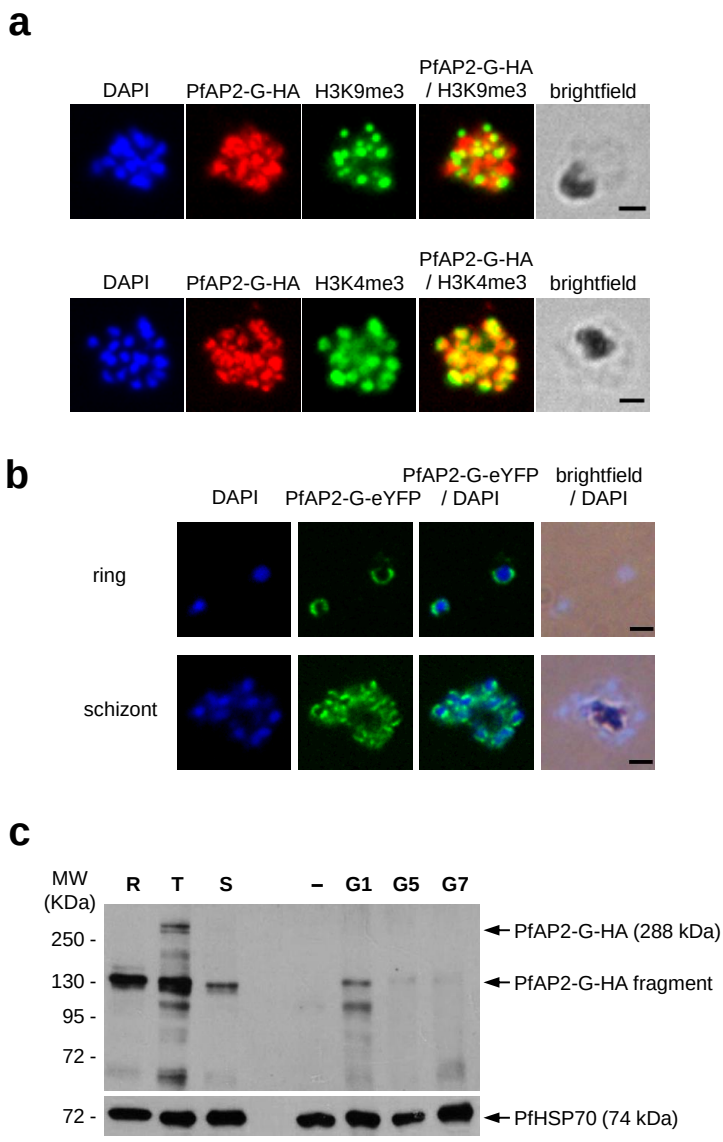
purification), and added uninfected erythrocytes immediately before Trizol extraction (Percoll/S+2x wash+RBC), and samples in which we magnet-purified schizonts and added uninfected erythrocytes immediately before Trizol extraction (magnet+RBC). Together with the experiments in Fig. 4b in the main article, these experiments show that *pfap2-g* can be abundantly expressed in mature schizonts. However, we measured very low transcript levels for this gene in late schizonts when Trizol extraction was performed directly on Percoll-purified schizont pellets, suggesting that left over Percoll selectively interferes with Trizol extraction of some transcripts. Thus, Trizol extraction directly from Percoll-purified schizonts can give artifactual results and should be avoided. Transcript levels were normalized against *ubiquitin-conjugating enzyme (uce).* Values are the average of three independent biological replicates (red dots), with S.E.M. **b,** Transcriptional time-course analysis of tightly synchronized cultures (5 h age window) of the parasite lines F12, E5 and E5-HA-DD (same samples as in Fig. 4a,c of the main article) normalized against *seryl tRNA synthetase* (*serrs*). While the overall temporal patterns of relative expression are similar to those observed when normalizing against *uce* (Fig. 4), apparent higher relative *pfap2-g* transcript levels are observed at 0-5 hpi. This is attributable to very low *serrs* expression at this time point (www.plasmodb.org). Values are the average of three (F12 and E5) or two (E5-HA-DD) independent biological replicates (red dots), with S.E.M. (F12 and E5 panels). **c,** *pfap2-g* transcript levels decrease with gametocyte maturation. RT-qPCR analysis of *pfap2-g* relative transcript levels in synchronized cultures of the E5 line at the late-ring stage (R), schizont stage (S, also containing ring-stages of the next cycle), and different stages of gametocyte development (G1, G2, G3, G6 and G9 correspond to days 1, 2, 3, 6 and 9 after GlcNAc addition to cultures at the ring stage, respectively). Sorbitol lysis was performed in the G1 and G2 samples to eliminate growth-arrested asexual parasites (although dead asexual parasites could not be completely removed). Transcript levels cannot be directly compared between rings/schizonts samples and gametocyte preparations because in the former only a subpopulation of the parasites is sexually committed and expresses *pfap2-g*, whereas in the latter essentially all parasites are sexual forms. Transcript levels were normalized against *ubiquitin-conjugating enzyme* (*uce*). Values are the average of three technical replicates (red dots).



Supplementary Figure 10. Transcriptional analysis of early gametocyte markers. a, Reverse transcription-quantitative PCR (RT-qPCR) time-course analysis of early gametocyte markers in tightly synchronized cultures (5 h age window) of the parasite lines F12 and E5. **b,** Analysis of the same markers in the transgenic parasite line E5-HA-DD maintained in the absence of Shld (-Shld) or with Shld added at 0-5 hpi. Samples are the same as in Fig. 4 of the main article. Insets show the same data with a different y-axis scale to highlight differential expression between time points. Transcript levels were normalized against *ubiquitin-conjugating enzyme (uce).* Values are the average of three

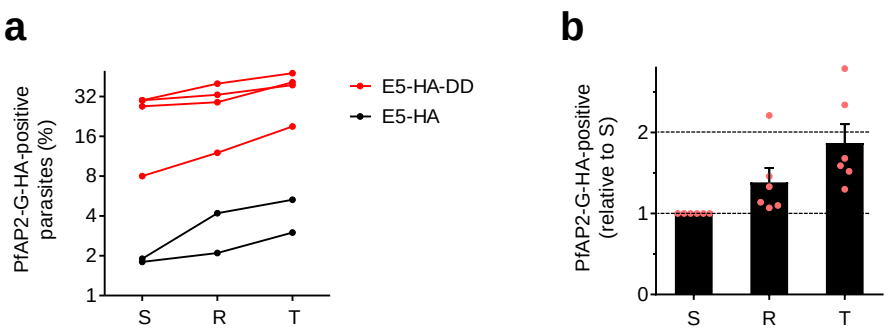
(panel a) or two (panel b) independent biological replicates (red dots), with S.E.M. (panel a).

pfs16[10] was the earliest gene to show an increase in transcript levels relative to untreated cultures when PfAP2-G was stabilized with Shld in the E5-HA-DD line. *pf14.744*[15] transcripts were the most specific of the markers tested, as transcripts for this gene were not detected in F12 and E5-HA-DD without Shld (however, at the protein level other markers such as Pfs16 are also highly specific. By IFA, we never observed Pfs16-positive parasites in the E5-HA-DD line in the absence of Shld). Transcripts of *pfs16* and *pfg27*[11] were detected in F12 and E5-HA-DD without Shld, but at much lower levels than in E5 and E5-HA-DD with Shld. Another recently described early gametocyte marker, *pfgexp5*[16], was expressed at high levels in the absence of functional PfAP2-G and gametocyte production (F12 and E5-HA-DD without Shld), which needs to be taken into account for the interpretation of the results of epidemiological studies or clinical trials using transcripts of this gene as a gametocyte marker[17]. However, this gene showed a marked increase in transcript levels at 0-5 hpi of cycle 2 after Shld addition.



Supplementary Figure 11. Subnuclear location of PfAP2-G and expression during gametocyte development. a, Representative IFA images of committed schizonts of the E5-HA line stained with antibodies against the HA tag and the H3K9me3 or H3K4me3 post-translational histone modifications, which mark heterochromatin and euchromatin, respectively. Experiments with anti-H3K9me3 and anti-H3K4me3 were independently repeated two and eight times, respectively, with similar results. PfAP2-G-HA does not preferentially localize within heterochromatin foci. Scale bar, 2 µm. **b**, In many parasites the PfAP2-G signal appears concentrated in the nuclear periphery. IFA images of ring and schizont stage parasites of the E5-eYFP line clearly showing this distribution of the PfAP2-G-eYFP signal are shown. IFA was performed with an anti-GFP antibody. Images are representative of four independent experiments. Scale bar, 2 µm. **c,** Western blot analysis of PfAP2-G-HA in E5-HA cultures synchronized to the ring (R), trophozoite (T) or schizont (S) stages, and in preparations of gametocytes collected at day 1 (G1), 5 (G5) or

7 (G7) after treatment of cultures at the ring stage with GlcNAc (day 0). Depending on the stage, we used between ~10^6 and ~10^7 parasites (the number of nuclei and total protein content varies between different stages). For the day 1 sample, sorbitol lysis was performed to eliminate asexual trophozoites, although contaminating arrested/ dead trophozoites could not be completely removed. Lane "–" is a negative control consisting of parasites of the wild type line E5 (no tag in PfAP2-G) at the trophozoite/schizont stage. Hybridization of the same membranes with antibodies against PfHSP70 was used as a loading control. The image is representative of two independent biological replicates, although variability was observed in the distribution of incomplete fragments of the protein.

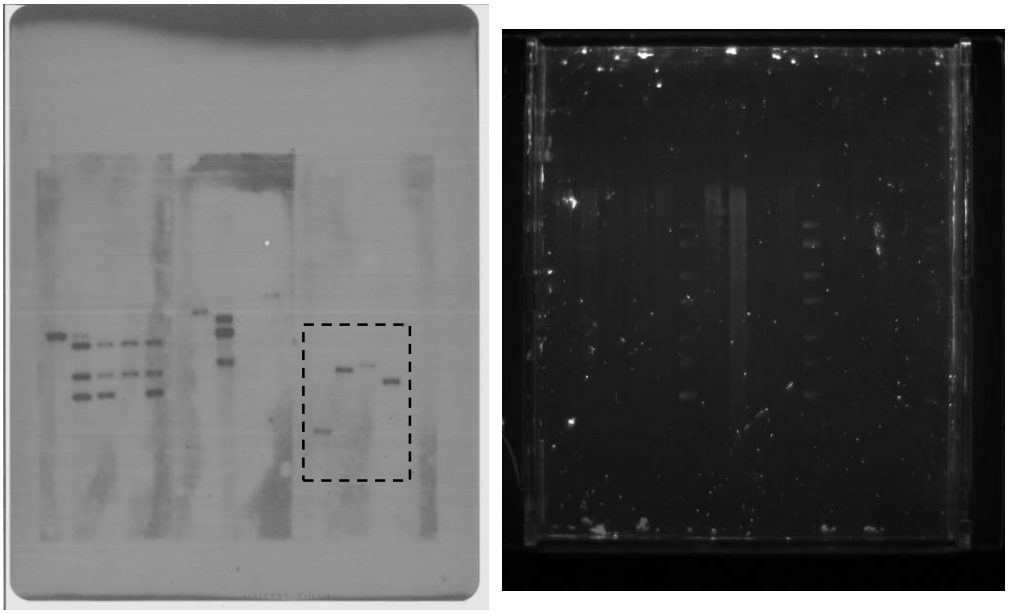


Supplementary Figure 12. Abundance of PfAP2-G-positive parasites in cultures synchronized to different stages of intraerythrocytic development. a, The proportion of PfAP2-G-HA-positive parasites was determined by IFA in synchronized cultures of the E5-HA and E5-HA-DD lines in consecutive schizont (S), ring (R) and trophozoite (T) stages. In the E5-HA-DD line, PfAP2-G was stabilized with Shld at the ring stage of the previous cycle. Each black or red line corresponds to a separate experiment. **b**, Proportion of PfAP2-G-positive parasites in synchronized cultures at the ring or trophozoite stage relative to the proportion of PfAP2-G-positive schizonts in the previous cycle. Values are the average of six independent experiments (red dots) using the E5-HA (*n*=2) and E5-HA-DD (*n*=4) lines, with S.E.M. The increase in the proportion of PfAP2-G-HA-positive parasites was significant from schizonts to trophozoites (p=0.013) but not from schizonts to rings (p=0.080) using a two-sided t-test with unequal variance.

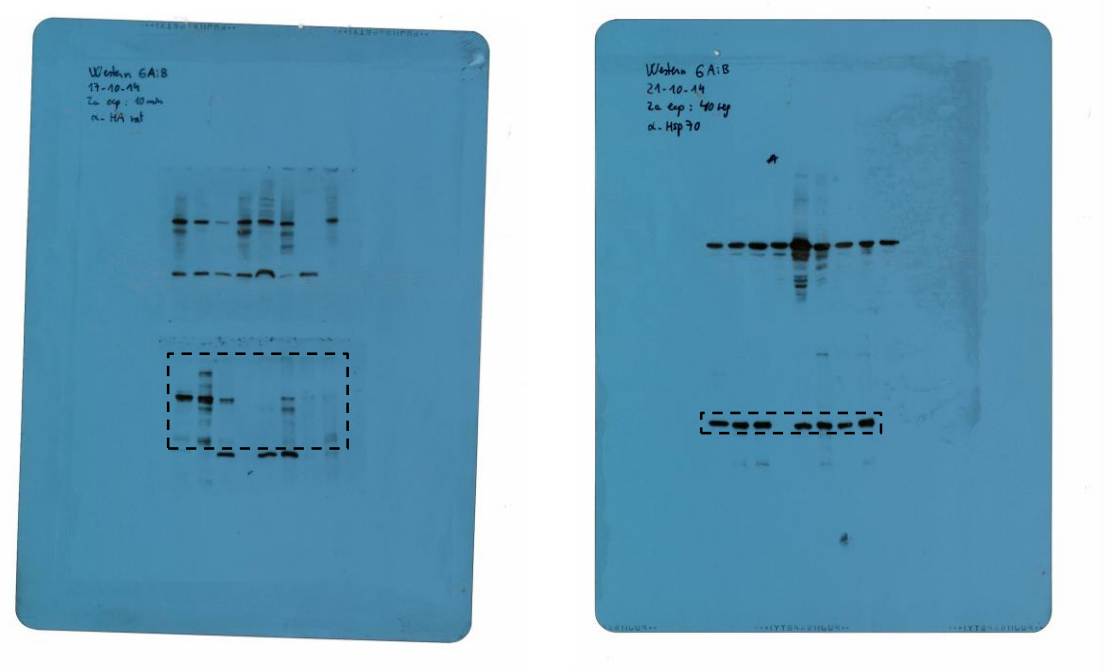
If the onset of PfAP2-G expression was always at the schizont stage, PfAP2-G-positive parasites in synchronized cultures at the schizont, ring or trophozoite stage would only include committed schizonts, sexual rings, and stage I gametocytes, respectively. Then, assuming that the average number of new parasites derived from a committed schizont is the same as for a non-committed schizont, the proportion of PfAP2-G-positive parasites in cultures at the ring or trophozoite stage would be expected to be the same as the proportion of positive parasites at the schizont stage of the previous cycle. Instead, if PfAP2-G activation (and sexual commitment) can start at the ring or trophozoite stage, a higher proportion of PfAP2-G-positive parasites would be expected in the ring and trophozoite cultures, as in addition to sexual rings or stage I gametocytes originated from sexually-committed schizonts of the previous cycle, they would include PfAP2-G-positive newly committed parasites (of which some would convert via the SCC route and some via

the NCC route). Hence, the higher proportion of PfAP2-G-positive parasites in ring and trophozoite cultures supports the idea that PfAP2-G expression and subsequent sexual commitment can start as early as the ring or trophozoite stage.

a



b



Supplementary Figure 13. Full length blots. a, Full length blot for Supplementary Fig. 2b. Left, autoradiography. Right, SYBR green-stained agarose gel used to determine the position of the molecular weight markers. **b,** Full length blots for Supplementary Fig. 11c. In the left image, the band excluded from the cropped image is non-specific, as it appears in the negative control lane (a wild type parasite line not containing the tag against which the antibody is directed, see Supplementary Fig. 11c).

Supplementary References

1 Bruce, M. C., Alano, P., Duthie, S. & Carter, R. Commitment of the malaria parasite *Plasmodium falciparum* to sexual and asexual development. *Parasitology* **100 Pt 2**, 191-200 (1990).

2 McDavid, A. *et al.* Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* **29**, 461-467 (2013).

3 Lopez-Barragan, M. J. *et al.* Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* **12**, 587 (2011).

4 Silvestrini, F. *et al.* Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* **9**, 1437-1448 (2010).

5 Salanti, A. *et al.* Selective upregulation of a single distinctly structured *var* gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. *Mol Microbiol* **49**, 179-191 (2003).

6 Aguilar, R. *et al.* Molecular evidence for the localization of *Plasmodium falciparum* immature gametocytes in bone marrow. *Blood* **123**, 959-966 (2014).

7 Rovira-Graells, N. *et al.* Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res* **22**, 925-938 (2012).

8 Lim, M. Y. *et al.* UDP-galactose and acetyl-CoA transporters as *Plasmodium* multidrug resistance genes. *Nat Microbiol* **1**, 16166 (2016).

9 Ghorbal, M. *et al.* Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat Biotechnol* **32**, 819-821 (2014).

10 Bruce, M. C., Carter, R. N., Nakamura, K., Aikawa, M. & Carter, R. Cellular location and temporal expression of the *Plasmodium falciparum* sexual stage antigen Pfs16. *Mol Biochem Parasitol* **65**, 11-22 (1994).

11 Carter, R. *et al. Plasmodium falciparum*: an abundant stage-specific protein expressed during early gametocyte development. *Exp Parasitol* **69**, 140-149 (1989).

12 Delves, M. J. *et al.* Routine in vitro culture of *P. falciparum* gametocytes to evaluate novel transmission-blocking interventions. *Nat Protoc* **11**, 1668-1680 (2016).

13 Delves, M. J. *et al.* Male and female *Plasmodium falciparum* mature gametocytes show different responses to antimalarial drugs. *Antimicrob Agents Chemother* **57**, 3268-3274 (2013).

14 Poran, A. *et al.* Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature* **551**, 95-99 (2017).

15 Eksi, S. *et al.* Identification of a subtelomeric gene family expressed during the asexual-sexual stage transition in *Plasmodium falciparum*. *Mol Biochem Parasitol* **143**, 90-99 (2005).

16 Tiburcio, M. *et al.* Specific expression and export of the *Plasmodium falciparum* Gametocyte EXported Protein-5 marks the gametocyte ring stage. *Malar J* **14**, 334 (2015).

17 Farid, R., Dixon, M. W., Tilley, L. & McCarthy, J. S. Initiation of gametocytogenesis at very low parasite density in *Plasmodium falciparum* infection. *J Infect Dis* **215**, 1167-1174 (2017).