scientific reports

Check for updates

OPEN

Genetic factors regulating *Plasmodium falciparum* gametocytogenesis identified by phenotypic screens

Camilla V. Pires[1,7], Jyotsna Chawla[1,2,7], Lauriane Sollelis[4], Jenna Oberstaller[1], Min Zhang[1], Chengqi Wang[1], Justin Gibbons[1], Julian C. Rayner[3], Thomas D. Otto[5], Matthias Marti[4] & John H. Adams[1,6]✉

Successful transmission of *Plasmodium falciparum* from one person to another relies on the complete intraerythrocytic development of non-pathogenic sexual gametocytes infectious for anopheline mosquitoes. Understanding the genetic factors that regulate gametocyte development is vital for identifying transmission-blocking targets in the malaria parasite life cycle. Toward this end, we conducted a forward genetic study to characterize the development of gametocytes from sexual commitment to mature stage V. We described a new analysis pipeline for the *piggyBac* transposon-based mutagenesis phenotypic screen to identify genes that influence both early and late gametocyte stages. We classified individual mutants that increased or decreased parasite abundance as the hypoproducer and hyperproducer phenotypes, respectively, revealing distinctive temporal genetic factors early and late in the sexual development cycle. The study identifies that disruption in factors involved in transcription, protein trafficking and DNA repair are associated with decreasing gametocyte production, while modifications in phosphatase activity are linked to hyperproduction of gametocytes. Our study provides an optimized approach on genotype–phenotype evaluation, offering a new resource for understanding potential targets for therapeutic intervention strategies to disrupt transmission.

Malaria kills more than 600,000 people each year and there are more than 200 million cases of clinical disease each year[1]. One of the crucial threats to future advancements in malaria control is the emergence and dissemination of drug resistance in the highly virulent *P. falciparum* parasite[2]. Drug resistance propagation involves the asexual replicative phases of the life cycle occurring in the bloodstream. However, resistance can only disseminate in endemic populations through transmission of *Plasmodium*sexual stages by anopheline mosquitoes[3].

Malaria parasites are transmitted exclusively through nonpathogenic sexual stages known as gametocytes, and in *P. falciparum,*these stages complete development within ~ 14 days[3]. In previous studies, we identified a human serum lipid, lysophosphatidylcholine (LysoPC), as a suppressor of parasite commitment or conversion to the sexual stage in *P. falciparum* cells[4,5]. Under permissive conditions, conversion to the gametocyte development pathway initiates with the activation of a specific transcription factor, AP2-G, which serves as a pivotal switch driving development into the sexual pathway through metabolic processes distinct from the asexual blood-stage cycle[6,7]. While AP2-G is a critical switch in activating sexual development, the multitude of parasite genetic factors enabling the complex developmental processes that lead to viable gametocytes capable of mosquito transmission are poorly understood. In a human host and in in vitro cultures, *P. falciparum* gametocyte development can be classified into five well-defined morphological stages, enumerated as I to V[8–10]. The characteristic elongated shape of these gametocytes emerges around day 2 of maturation as a crescent-shaped stage II gametocyte and continues to elongate as it matures (by day 14). In vivo, *P. falciparum*in its early stages is mostly confined to

[1]Center for Global Health and Inter-Disciplinary Research, College of Public Health, University of South Florida, Tampa, FL, USA. [2]Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, USA. [3]Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. [4]Institute of Parasitology Zurich, VetSuisse Faculty, University of Zurich, Zurich, Switzerland. [5]Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK. [6]Department of Clinical Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [7]These authors contributed equally: Camilla V. Pires and Jyotsna Chawla. ✉email: ja2@usf.edu

the extravascular niche of the bone marrow[11], and only mature stage V gametocytes present in the peripheral circulation can infect mosquitoes[12]. Unraveling the genetic factors governing sexual commitment and differentiation is crucial for prioritizing potential high-value targets for preventing transmission.

Understanding gene function in gametocytogenesis and development has relied mainly on targeted reverse genetics methods and transcriptional profiling[4,13–15]. Recently, we fine-tuned phenotype screening of *P. falciparum piggyBac* single-insertion mutants to identify late-stage gametocyte phenotypes for hypoproducer and hyperproducer mutants[16]. Building on this approach, we expanded the study to a large-scale *piggyBac* library of ~ 500 *pB*-mutants[17] and introduced an early timepoint during gametocyte development to broaden the spectrum of genetic regulatory factors functionally annotated by this experimental approach. In this study, we identified phenotypes associated with decreased or increased gametocytogenesis as hypoproducer and hyperproducer phenotype mutants, respectively, in the early and late stages. By screening the expanded library of mutants, a wider range of genes and pathways implicated in the regulation of gametocyte development could be identified. Our findings demonstrate that disruptions in elements responsible for transcription, protein transport, and DNA repair correlate with reduced gametocyte production. Conversely, alterations in phosphatase activity are associated with increased gametocyte production.

Results

Development of a large-scale gametocyte screen

A library of ~ 500 *piggyBac* isogenic mutants, referred to herein as the "Half-K library"[17]randomly generated in a previous study[18] was used for a large-scale gametocyte screen. This collection of mutants represents ~ 11% of the parasite genome, and each mutant carries a single insertion in the intergenic, exonic or untranslated (UTR) regions of genes (Supplementary Tables S1, S2). Gametocyte production was induced in in vitro cultures, and genomic DNA was harvested on day 3 (early gametocyte timepoint) and day 14 (late gametocyte timepoint) to identify genes functionally important in early and late gametocyte development, respectively (Fig. 1).

Quantitative insertion-site sequencing (QIseq) analysis (Supplementary Tables S1, S2)[16,19]was used to identify the sexual development phenotypes based on quantifying fold-change differences in the relative abundance of each mutant at early (day 3) and late (day 14) timepoints versus the start of gametocyte culture (day 0, induction day). We used the DEseq2 package for normalization, fold-change calculation and significance assessment (see Method) as recommended in previous genetic screen studies[20,21]. Mutant parasites with significantly decreased relative abundance were classified as hypoproducer mutants ("hypo") (Log_2fold-change < -0.75 and adjusted p value [padj] < 0.05), while those with significantly increased abundance were classified as hyperproducer mutants ("hyper") (Log_2fold-change > 0.75 and padj < 0.05). The mutants that did not fall into either category were defined as neutral (Fig. 1, Supplementary Table S3).

Early-stage gametocytes

At the early gametocyte endpoint (day 3 post-induction), we identified 29 and 21 genes linked to the hypoproducer and hyperproducer mutants, respectively (Supplementary Table S3, Fig. 2A). Notably, these genes show peak expression in Stage V gametocytes (and not during early stage development) (Fig. 2B[22],Supplementary Table S4). Genes linked to hyperproducer mutants are also significantly upregulated in sexual gametocyte stages II compared to schizont and late trophozoite asexual stages, while genes linked to the hypoproducer mutants have no significant difference in transcript abundance between any sexual and asexual stages.

GO enrichment analysis (Fig. 2C, Supplementary Table S5) revealed that genes of early gametocyte hypoproducer mutants are enriched for processes involved with gene regulation, such as GO:0003700 'DNA-binding transcription factor activity' and GO:0006355 'regulation of transcription, DNA-templated'. This included three AP2 genes: PF3D7_1115500 (ApiAP2), PF3D7_1139300 (AP2-G5) and PF3D7_1456000 (AP2-HC). PF3D7_1139300 (AP2-G5) was documented to inhibit the transcriptional activity of the *pfap2*-g gene by binding to both its upstream region and exonic gene regions, thereby maintaining the local heterochromatin structure and preventing sexual commitment[23].

Late-stage gametocytes

At the late gametocyte endpoint (day 14 post-induction), we identified 18 and 15 genes linked to mutants of the hypoproducer and hyperproducer phenotype, respectively (Supplementary Table S3, Fig. 3A). Genotype–phenotype associations using a *piggyBac*-based screen with a late-stage gametocyte endpoint have previously been piloted using a smaller library of different composition than the one used here[16]. Importantly, this pilot screen identified *piggyBac* mutants of similar genotype–phenotype associations, but with insertions in different locations of the same gene (Supplementary Table S7). Examples for late-stage gametocyte hyperproducer phenotypes, that were previously identified as hyperproducers[16] included PF3D7_1231800 (asparagine-rich protein) and PF3D7_0615900 (protein phosphatase) (Fig. 3A). The genes of hypoproducer mutants (*e.g.* PF3D7_1213800 [proline-tRNA ligase PRS] and PF3D7_0419500) and hyperproducer mutants (*e.g.,* kinesin-15, PF3D7_1245600) encode for proteins that are enriched in stage V gametocytes according to a proteomic study[24] (Supplementary Table S9; Fig. 3A).

Genes linked to hypoproducer and hyperproducer mutants in late stages show peak expression in late stage gametocytes (Fig. 3B)[22]. Late-stage genes linked to hypoproducer mutants were enriched for cellular components GO:0005764 'lysosome' and GO:0016021 'integral component of membrane' suggesting a link to the characteristic morphological changes essential to gametocyte development. In contrast, late-stage genes linked to hyperproducer mutants are enriched for 'cellular lipid metabolism' GO:0044255 (Fig. 3C).

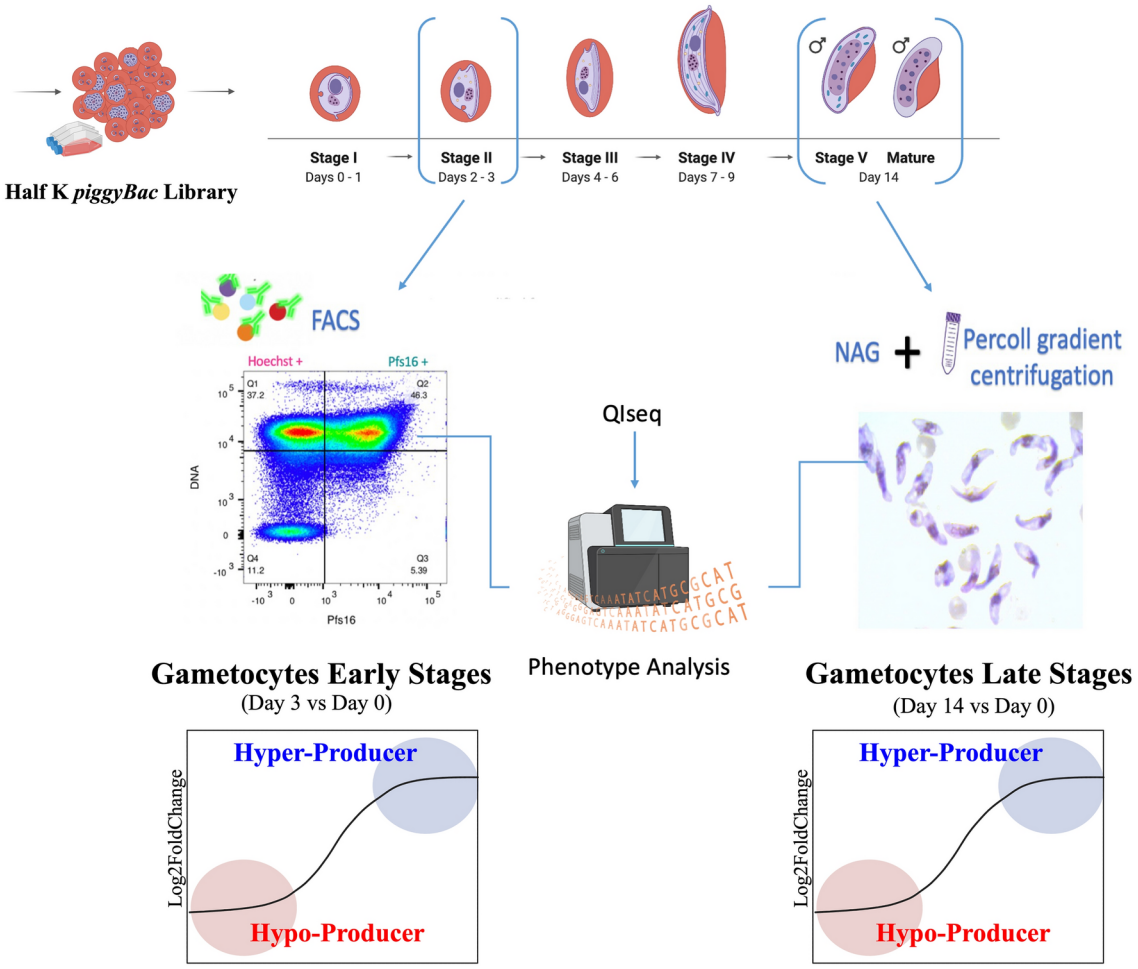

Fig. 1. Approach for the large-scale gametocyte phenotype screen, using *piggyBac* Half-K library. Methodology for enrichment of gametocyte populations. Early gametocytes were isolated with FACS selection by antibody staining for Pfs16. Gametocytes cultures at day 3 were first enriched by Percoll® (see methods) and incubated with anti-Pfs16 primary antibody and fluorescent secondary antibody followed by incubation with Hoechst. DNA from day 3 gametocytes (upper right quadrant) positive for both Hoechst and Pfs16 was harvested and sequenced. Stage V gametocytes (> 90% mature gametocytes) were isolated using Percoll® purification (see Methods). Parasite genomic DNA was harvested from day 0, day 3 and day 14. QIseq quantified each *piggyBac* mutant in the library by sequencing from the 5′ ends of a *piggyBac* insertion-site. Ranked *piggyBac* mutants based on gametocyte phenotypes in the library. QIseq results rely on counts of insertion sites for each mutant, which are normalized to calculate the relative abundances at day 3 and day 0 (Gametocytes Early stages), and day 14 and day 0 (Gametocytes Early stages). Gametocyte fold changes (day3 / day 0 and day 14 / day 0) were calculated for each mutant as a measure of their ability to commit into early gametocytes stages and to differentiate into mature gametocytes.

Early and late-stage gametocyte phenotypes overlap

Most of the genetic factors controlling early and late gametocyte development appeared to be distinct, reflecting the different metabolic processes required at each stage[13,25]. Nonetheless, we identified mutants that consistently displayed significant mutant phenotypes throughout both early and late gametocyte stages (Table 1). In the population of the half K *piggyBac* mutant library, 7 disrupted genes were linked to the hypoproducer mutant phenotypes in early and late gametocytes. These include PF3D7_1250800 (rhp16), a DNA repair protein gene, and genes linked to trafficking proteins (PF3D7_0202200, EMP1-trafficking protein [PTP2], PF3D7_0207600, serine repeat antigen 5 [SERA 5] and PF3D7_1011200 [BET1-like protein]). Meanwhile, the 9 genes shared among mutants for early and late stage hyperproducer mutant phenotypes included genes producing host cell surface binding proteins, such as PHISTs, phosphatase activity genes (PF3D7_0615900 and PF3D7_1249300 [PPM4]), and a gene coding for a protein linked to fatty acid metabolism ACS11 (PF3D7_1238800) (Table 1).

Validation of gametocyte *piggyBac* phenotypes

To validate the hypoproducer and hyperproducer mutant phenotypes and evaluate the efficacy of our screens, we selected mutants linked to different phenotypes; these were two hypoproducers, two hyperproducers and one neutral phenotype (Fig. 4A). Additionally, we included the wild-type NF54 line in the phenotype validation

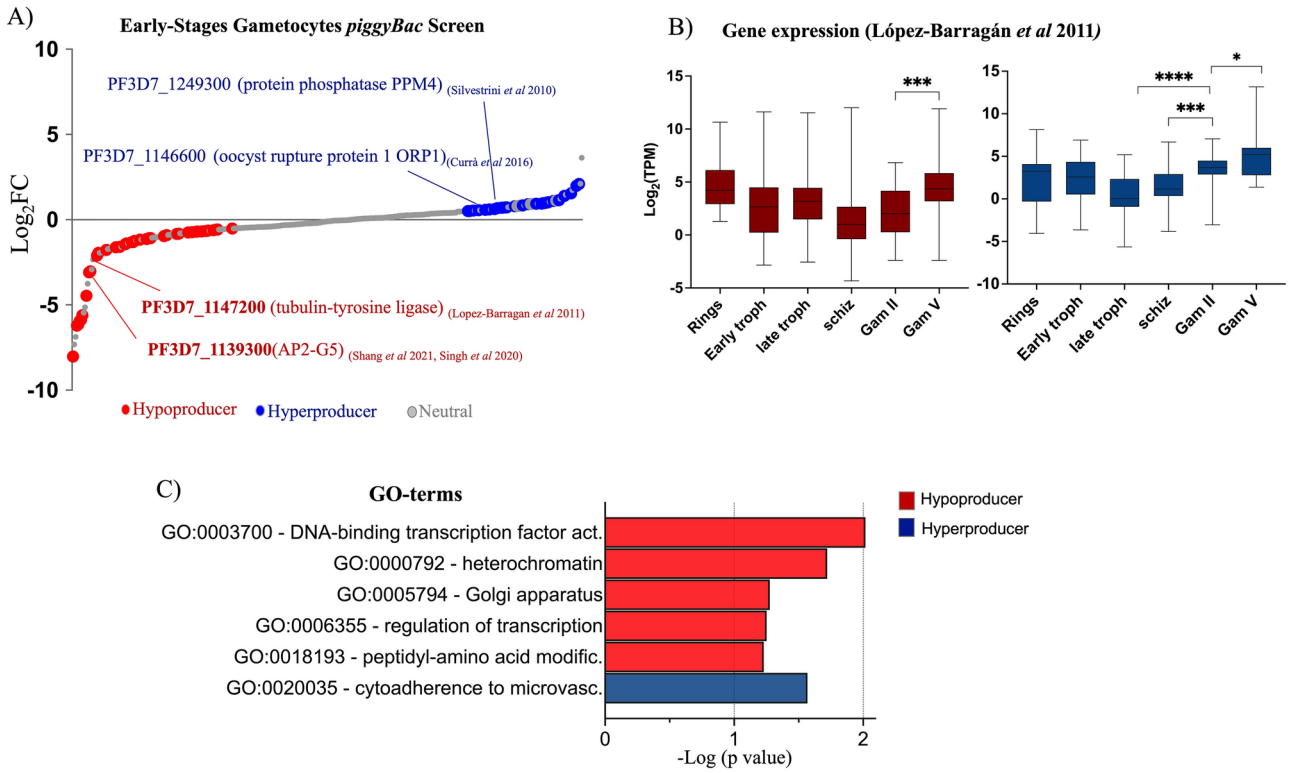

Fig. 2. Genetic factors associated with early-stage gametocytes identified in the *piggyBac* phenotypic screen. (**A**) Relative differentiation of each *piggyBac* mutant in the library was determined by ranking mutants from low to high. The genes in the bottom indicated in red are significant and inferred as linked to gametocyte hypoproducers (Log_2fold-change < -0.75 and adjusted $p < 0.05$). The top ranked genes indicated in blue are significant and inferred as kinked to gametocyte hyperproducers (Log_2fold-change < -0.75 and adjusted $p < 0.05$). Hits highlighted include genes that have been previously associated with sexual development. The entire *piggyBac* screen dataset is provided in Table S3. **B**) The mean number of transcripts per kilobase per million per kilobase per million (TPM) of all the genes linked to hypoproducer and hyperproducer mutant phenotypes were determined from published transcriptome sequencing RNAseq data from Lopez-Barragan et al. (Table S4) (mean and maximum and minimum; $^{*}p < 0.01$ $^{**}p < 0.001$, $^{***}p < 0.0001$ ANOVA follow by Turkey's test). (**B**) Functional enrichment of significant gene ontology (GO) terms for early gametocyte hypo and hyperproducer *piggyBac* mutant's *vs* all other mutants in the library. The entire GO-dataset is provided in Table S5.

experiments as a comparator to the parental strain's baseline phenotype. Note that this parental line is does not carry an integrated transposon and therefore is absent from the *piggyBac* pool libraries[17–19,26]. The analysis used a different gametocyte induction protocol known for its higher efficiency in sexual commitment, which is also referred to as the minimum fatty acids (mFA) induction protocol[13,22,25,27]. As previously suggested[16], hypoproducer mutants might exhibit reduced sexual commitment and/or defects during gametocyte development before reaching stage V. Early timepoints from our current screen allowed us to infer genes involved in sexual commitment (Fig. 3). The lack of intermediate time points in our library screens (Fig. 1) was addressed with this additional experimental approach (Fig. 4), also. Besides evaluating and validating the *piggyBac* mutant capacity to convert to sexual stages (at Day 5), we tracked gametocyte development by measuring the relative gametocyte abundance at days 8, 10, and 14 (see Methods). The mutants included in this approach were successfully cloned out and identified from the Half-k *piggyBac* mutant library (Fig. 4A): PF3D7_0522900 (zinc finger), PF3D7_1351800 (conserved *Plasmodium* protein, unknown function), PF3D7_1250800 (DNA repair protein rhp16), and PF3D7_0830600 (*Plasmodium* exported protein, PHISTc). We also included a clone from Chawla et al.[16] that showed similar phenotypes but had insertions in different locations of the same gene, the hyperproducer PF3D7_0615900 (phosphatase protein) (Supplementary Table S7).

First, to validate the screens, we compared the *piggyBac* clones among themselves, given that the library screen phenotypes rely on the ranked relative abundance of each mutant (see method). We observed a consistent tendency of phenotypes across the library screens (Fig. 4A) and all five clones (Fig. 4B). Specifically, the two hypoproducer mutants have significantly lower gametocyte conversion rates (GCR%) compared to at least one of the two hyperproducer mutants. This difference persists from day 5 to day 14. Second, we compared the *piggyBac* clones with wild-type NF54. The two hypoproducers showed lower GCR% compared to the wild-type line; particularly, the hypoproducer PF3D7_1250800 (DNA repair protein rhp16) showed significant lower GCR% compared to NF54, from day 5 to day 14 post-induction (Fig. 4B). All clones, including the wild-type NF54,

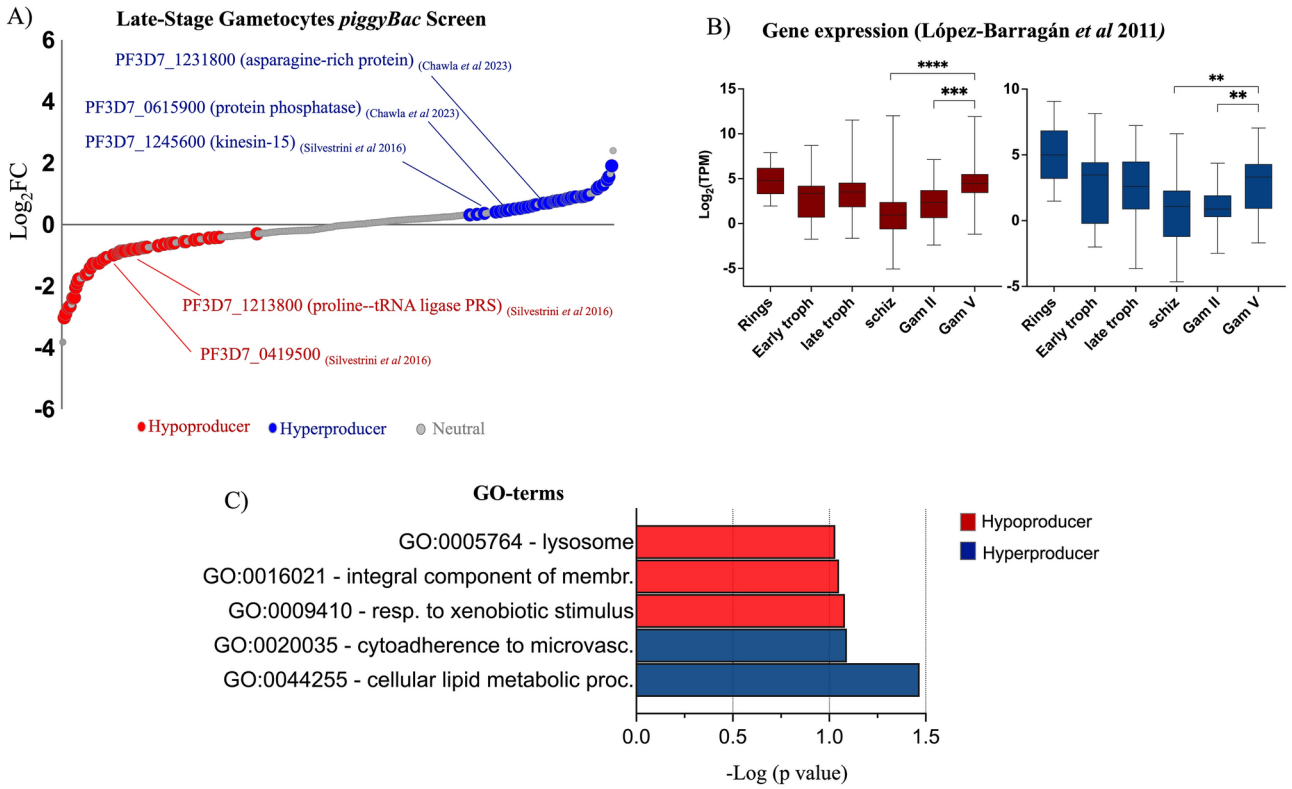

Fig. 3. Genetic factors associated with late-stage gametocytes identified in the *piggyBac* phenotypic screen. (**A**) Relative differentiation of each *piggyBac* mutant in the library was determined by ranking mutants from low to high. The genes in the bottom indicated in red are significant and inferred as gametocyte hypoproducers (Log_2fold-change < -0.75 and adjusted $p < 0.05$). The top ranked genes indicated in blue are significant and inferred as linked to gametocyte hyperproducers (Log_2fold-change > 0.75 and adjusted $p < 0.05$). Hits highlighted include genes that have been previously associated with sexual development. The entire *piggyBac* screen dataset is provided in Table S3. (**B**) The mean number of transcripts per kilobase per million (TPM) of all the genes linked to hypoproducer and hyperproducer phenotype mutants were determined from published transcriptome sequencing RNAseq data from Lopez-Barragan et al.[22] (Table S4) (mean and maximum and minimum; *$p < 0.01$ **$p < 0.001$, ***$p < 0.0001$, ****$p < 0.00001$, ANOVA follow by Turkey's test). (**C**) Functional enrichment of significant gene ontology (GO) terms for early gametocyte hypoproducer and hyperproducer *piggyBac* mutant's *vs* all other mutants in the library. The entire GO-dataset is provided in Table S6.

	geneID	insertion-site	Product description
Hypoproducer in Early and Late Stages	PF3D7_0202200	chr2:112097	EMP1-trafficking protein—PTP2
	PF3D7_0207600	chr2:307896	serine repeat antigen 5—SERA 5
	PF3D7_1011200	chr10:437715	BET1-like protein, putative
	PF3D7_1022600	chr10:947928	kelch protein K10
	PF3D7_1250800	chr12:2081187	DNA repair protein rhp16, putative—rhp6
	PF3D7_1351200	chr13:2042606	conserved Plasmodium protein, unknown function
	PF3D7_1410300	chr14:413547	WD repeat-containing protein, putative
Hyperproducer in Early and Late Stages	PF3D7_0530600	chr5:1250257	XAP-5 DNA binding protein, putative
	PF3D7_0615900	chr6:664011	protein phosphatase, putative
	PF3D7_0830600	chr8:1305128	Plasmodium exported protein (PHISTc)
	PF3D7_0921200	chr9:870606	conserved Plasmodium membrane protein
	PF3D7_1000700	chr10:54859	Plasmodium exported protein (PHISTa), unknown function, pseudogene
	PF3D7_1238800	chr12:1613083	acyl-CoA synthetase—ACS11
	PF3D7_1249300	chr12:2015676	protein phosphatase PPM4, putative
	PF3D7_1458500	chr14:2404280	spindle assembly abnormal protein 4, putative—SAS 4

Table 1. Genes linked to overlapping early and late-stage gametocyte mutant phenotypes.

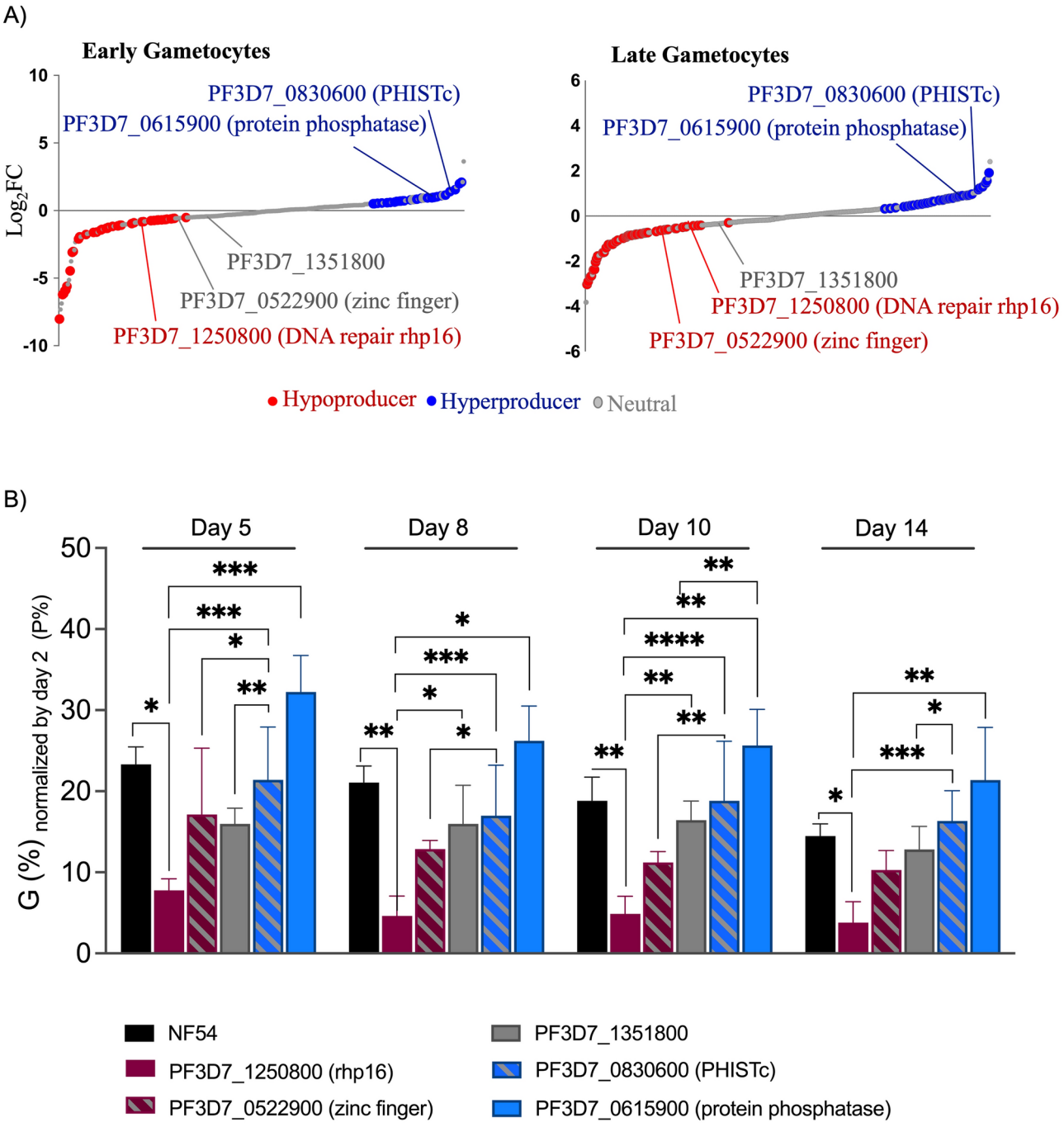

Fig. 4. Gametocyte Development of Individual *piggyBac* Mutant Clones. (**A**) Five *piggyBac* mutant clones highlighted in the ranked plots from the early and late-stage gametocyte screens from the Half-k library (red for hypoproducer phenotype, blue for hyperproducer phenotype, and gray for neutral phenotype) were selected for phenotype validation. (**B**) Gametocyte development of these *piggyBac* mutants highlighted in (**A**) and NF54 were monitored over a period of 14 days. Gametocyte abundance was estimated by Giemsa-stained thin blood smears on Days 5, 8, 10 and 14 post-induction (Supplementary Table S9). To allow comparison among different lines, Gametocytemia (G %) for Days 5, 8, 10, and 14 were normalized by the parasitemia (P %) of Day 2 post-induction. At day 5, G% is commonly named gametocytes conversion rate (GCR%). The assay was performed with 3–5 biological replicates per parasite line (means with SD [error bars] are shown). Comparisons among parasite lines were analyzed using one-way ANOVA followed by Tukey's test (*$p < 0.01$, **$p < 0.001$, ***$p < 0.0001$, ****$p < 0.00001$).

hypoproducer and hyperproducer mutants (but not the neutral clone PF3D7_1351800) exhibited a decreasing GCR % from day 5 to day 14 post-induction (Supplementary Figure S1).

Discussion

We optimized and employed a large-scale phenotypic screen of *P. falciparum piggyBac* mutants to characterize the genetic factors that affect the gametocyte development. We previously used random *piggyBac*-transposon insertional mutagenesis to uncover genes related to *P. falciparum* asexual blood stage growth, as well as their survival under heat-shock[26]and upon exposure to antimalarial drugs[17]. The improved *piggyBac* gametocyte screen demonstrates a scalability to complex mixed mutant populations from our first phenotypic screen[16], to reveal specific genetic factors that impact both early and late gametocyte development. These results identified gene disruptions linked to gametocyte phenotypes evident either by suppressing or enhancing gametocyte development, defined herein as hypoproducer and hyperproducer mutant phenotypes, respectively. The genotype–phenotype screen approach included early and late time-points for in vitro gametocytogenesis thereby enabling us to differentiate phenotypes for some *piggyBac* mutants from gametocyte conversion/early development vs maturation/late development mutant while other mutations exhibited overlapping phenotypes for both. The study approach provides a more robust and accurate phenotypic analysis than the previously published gametocyte screens[16] and it can be readily taken to genome scale to infer the functional roles of genes that regulate and/or contribute to gametocyte development.

Fifteen *piggyBac* mutants exhibited consistent phenotypes across both early and late stages of gametocyte development (Table 1). However, mutants showing significant phenotypes only at one stage, appear to reflect the specific processes (e.g., transcriptional regulation, metabolism, host cell remodeling) during individual steps of gametocyte maturation[13,22,25,27]. For instance, a mutation in the *ap2-g5* gene, previously associated with sexual commitment[23], showed a hyperproducer phenotype in early but not late gametocyte (Supplementary Table S3 and S7)[16]. It is important to highlight that more significant phenotypes were identified in the early stages compared to the late stages. This observation is reflection of the relative abundance of gametocytes that defines the phenotypes in the *piggyBac* pool library screens. Two factors may affect this observation: 1) gametocyte conversion rates are more likely to be detected in early stages, as the rates of hypoproducer and hyperproducer phenotypes generally decrease over the course of development (Fig. 4B); and 2) defects may occur during gametocyte development before reaching stage V due to gene disruptions in the *piggyBac* mutants. For instance, as illustrated by Chawla et al. (2023), a *piggyBac* mutant with a disruption in the gene encoding the ER membrane protein complex subunit 3 (EMC3, PF3D7_1360200) exhibited atypical/pyknotic forms of gametocytes, arresting development at stage III[16]. It is important to emphasize that phenotype identification in this study and in the previously published[16,17,26] is based on the relative abundance (DNA abundance) of *piggyBac* mutants remaining in the population after the screen. This is measured by the fold change readouts (normalized QISeq read counts, see Methods) at Day 3 (early stage) or Day 14 (late stage) relative to Day 0. If one or more mutants exhibit a weak phenotype in the early stages and their abundance decreases during intermediate stages, this will affect the relative abundance of the remaining mutants in the later stages. Notably, we found a strong correlation between the normalized read counts at early and late stages (Supplementary Figure S2), suggesting that mutants with lower read counts (abundance) in early stages also tend to have lower read counts in late stages and vice versa.

Further analysis of individual gene functions, using additional intermediate time points, will be necessary to elucidate more precise temporal patterns of gene essentiality during gametocytogenesis. So far, gametocyte rates at intermediate time points have been addressed through individual gametocyte development studies (Figs. 4and)[16]. Despite using different induction protocols, the characterized phenotypes were consistent with the library screens (Fig. 4). Notably, DNA repair gene (PF3D7_1250800, rhp6) exhibited a consistent hypoproducer phenotype across early, late, and intermediate time points, both against hyperproducer mutants and wild-type NF54.

Our current study underscores the potential of forward genetic screens to identify genes involved in gametocyte production and development. It also provides an excellent starting point for future investigations of candidate genes towards the development of innovative transmission-blocking agents for *P. falciparum*.

Methods

Plasmodium falciparum culture and maintenance

The parasites were cultured at 5% hematocrit (O + erythrocytes from the Interstate Blood Bank, Memphis, TN) in complete medium containing 10% human AB serum (heat inactivated) and 2.5% sodium bicarbonate (using 7.5% stock solution) in RPMI 1640 medium (KD Medical) supplemented with 50 µg/mL hypoxanthine and 25 mM HEPES. The culture flasks were grown in an incubator at 37 °C and manually gassed with mixed gas (90% N_2, 5% CO_2 and 5% O_2).

Large-scale gametocyte screen

The "Half K" *piggyBac*mutant library was generated from our whole-genome random mutagenesis saturation project[18]. The gametocyte screens were performed with 2 biological replicates on early stage and late-stage gametocytes screens. Briefly, 1 ml of the mixed pool population was thawed in non-vented T25 flasks (5 ml, 5% hematocrit, manual gassing, 37 °C), and after reaching 1–2% parasitemia, the parasites were scaled up to 20-mL cultures in T75 flasks. Gametocytogenesis was induced by increasing the parasitemia and addition of spent media[28]. Genomic DNA was harvested from the induced flasks using a Qiagen DNA Extraction Kit (QIAamp, cat no. 51104) and sequenced by QIseq[19] as day 0. Gametocyte flasks were set with a starting parasitaemia of 0.5%−0.8% and maintained with daily medium change.

Early gametocyte enrichment on day 3 was accomplished by FACS using Pfs16, an early sexual-stage marker, expressing in the periphery of the immature gametocytes[29–32]. In brief, 200 ml of each gametocyte culture was first enriched using Percoll® gradient centrifugation[33]. A 100 µl aliquot of live, gametocyte-enriched parasite-infected RBCs was incubated with polyclonal anti-*P. falciparum* Pfs16 (diluted 1:200) for 30 min on ice. Next, Alexa Fluor 488-conjugated goat anti-rabbit IgG (diluted 1:25) was added and allowed to incubate for 30 min on ice. To stain the nuclei, Hoechst dye was added to the samples, which were incubated for 20 min at 37 °C. Afterwards, the samples were washed 3 times with FACS buffer made from phosphate-buffered saline supplemented with 1% bovine serum albumin (BSA). Flow cytometry was carried out in a high-speed BD FACSAria IIu sorter, where compensation to exclude background fluorescence on the vertical and horizontal axes was applied to samples single labeled with AF488, Hoechst and uninfected erythrocytes. Approximately 1 million cells (+ Pfs16-AF488, + Hoechst 3342) were sorted from each replicate (900,000 cells for T1_GamEarly_DP_1_1, T1_GamEarly_DP_2_1). Polyclonal anti-*P. falciparum* Pfs16 (antiserum, rabbit) was obtained from MR4 BEI resources (MRA-1276). The genomic DNA harvested from this population was sequenced by QIseq[19] as day 3.

Late gametocyte enrichment was achieved by treating cultures with 50 mM N-acetylglucosamine (NAG) from days 4–9 to eliminate any asexual parasites[34]. When the majority of the parasite population was stage V gametocytes on day 14, the mature stages were isolated by Percoll® gradient centrifugation and washed thrice with incomplete medium[33]. Subsequently, genomic DNA was harvested from this population and sequenced by QIseq[19] as day 14.

Phenotype identification

Previously described QIseq methodology was used to quantify insertion sites for each *piggyBac* mutant in the library screened[19](Supplementary Table S1). The R package DESeq2[20], which has been recommended for analysis of small size libraries[20,21], was used to normalize the original read counts (Supplementary Table S2) per insertion site, calculate fold changes for the early (day 3/day 0) and late (day 14/day 0) timepoints and assess the significance. DESeq2 uses Wald test with a significance cut-off of $p < 0.05$ from the subset of genes that pass an independent filtering step and p adjusted for multiple testing using the Benjamini and Hochberg[20]. Mutants that performed poorly, as determined by decreased relative abundance, were classified as hypoproducers (Log2FC < −0.75 and padj < 0.05), while those mutants that had an increased abundance (Log2FC > 0.75 and padj < 0.05), were classified as hyperproducers. Unclassified mutants were categorized as neutral (Supplementary Table S3).

Gene ontology (GO) enrichment

All GO-enrichment analyses were performed by testing GO-terms mapped to the gametocyte phenotypic categories of interest against a background of GO-terms mapped to all other genes using our R package pfGO[35] (v 1.1). Early gametocyte phenotypic categories of interest were hypoproducer, hyperproducer, and neutral (Fig. 2); late gametocyte phenotypic categories of interest were hypoproducer, hyperproducer, and neutral (Fig. 3). The GO-term database was created from the latest curated *P. falciparum*ontology available at the time of analysis from PlasmoDB[36] and enrichment was assessed via a weighted Fisher/elim-hybrid $p <= 0.05$ (v. 57). The fraction of genes represents the number of significant genes annotated to a given GO-term in each of the categories divided by the total number of genes annotated to that GO-term included in the analysis for all categories (background-set). (Supplementary Tables S5, S6).

Validation of *piggyBac* gametocyte phenotypes

Isolation of piggyBac mutants from Half K library

Clones of *piggyBac* mutants were isolated from the half k library using limiting dilution cloning in 96-well plates and modification on the PCR screening to detect positive mutants[37]. For the qPCR, the Phusion Blood Direct PCR Kit (Thermo Scientific) was used, following the manufacturer's instructions, SYBR Green I nucleic acid stain (10,000 X) (Invitrogen) diluted 3X in water and then further diluted to 1X in the PCR master mix, and the microsatellite primers PE14D added with the sequences 5'-TGTAATGAATGATTCTAATACCAC-3' and 5'-TTGGACCATGCTTCACAG-3'. Potential positives were confirmed by Giemsa-stained thin blood smears were prepared by standard methods. Positive wells were subsequently expanded by transferring to 5, 10 and then 20 ml cultures for further DNA extraction and the creation of cryopreserved stocks. The insertion sites were confirmed using QISeq[19].

Individual piggyBac mutant gametocyte induction and development assay

The sexual conversion assay was adapted from Brancucci et al.[4,38]. Parasites were cultured in Albumax + choline medium, consisting of 0.5% Albumax II in RPMI 1640 medium (Invitrogen) supplemented with 50 mg/mL hypoxanthine (Sigma), 25 mM HEPES (Invitrogen), and 2 mM choline. The cultures were maintained in an incubator with a continuous flow of mixed gas (90% nitrogen, 5% CO2, and 5% O2). Parasite cultures were tightly synchronized using 5% sorbitol synchronization, MACS magnetic purification, and a second round of sorbitol synchronization. After synchronization, parasites 26–28 h old (late rings to early trophozoites) were induced to gametocytes by replacing the Albumax + choline medium for mFA (minimum fatty acid) medium (0.39% Bovine Serum Albumin fatty acid-free [Sigma-Aldrich], RPMI 1640 medium [Invitrogen], 50 mg/mL hypoxanthine [Sigma], 25 mM HEPES [Invitrogen], 30 µM palmitic acid [Sigma-Aldrich], and 30 µM oleic acid [Sigma-Aldrich]). After reinvasion at 24 h (Day 1), the medium was switched back to Albumax + choline. To prevent asexual reinvasion, 230 µg/mL heparin (Sigma) was added to the medium from Day 2 to Day 6 post-induction. The medium was replaced daily until Day 8, and then every other day until Day 14 post-induction. Gametocyte abundance (gametocytemia) was estimated by Giemsa-stained thin blood smears on Days 5, 8,

and 14 post-induction (Supplementary Table S9). To allow comparison among different lines, gametocytemia of Days 5, commonly named gametocytes conversion rate (GCR%), day 8, day 10, and day 14 post-induction were normalized by the parasitemia (sexual + asexual) of Day 2 post-induction. This experimental approach was performed with 3 technical replicates and 3—5 biological replicates. Comparisons among wild-type and mutant parasite clones were analyzed using one-way ANOVA followed by Tukey's test in GraphPad Prism version 10.2.3.

Whole genome sequence (WGS)

The WGS was performed to the hypoproducer clone PF3D7_1250800 to confirm that there is no genetic changes in the loci of the gametocyte-related genes *gdv1* (PF3D7_0935400, gametocyte development protein 1) and *ap2-g* (PF3D7_1222600, AP2 domain transcription factor AP2-G). Genomic DNA was extracted using the QIAGEN DNA extraction kit and purified according to the manufacturer's instructions. Library construction was carried out using the NEBNext Ultra II DNA Library Prep Kit (NEB #7645S/L). Sequencing was performed on the Illumina NextSeq 2000 system using the P1-300 cycles reagent kit at the USF Genomics Sequencing Core. Raw reads were aligned to the *Plamodium falciparum* 3D7 reference genome 47 downloaded from PlasmoDB (https://plasmoDB.org)[36], using BWA MEM version 0.7.17[39]. The resulting same files were sorted and coverted to bam files using samtools 1.15 [40]. The bam file was indexed using samtools. Read coverage was visualized using IGV 2.17. 4 [41].

Data availability

The raw QIseq data sets generated for this study were deposited in the European Nucleotide Archive. Access numbers and original QIseq read mapping data are provided in Supplementary Table S1. Whole genome sequence data generated for this study have been deposited to the NCBI Sequence Read Archive (SRA) database bioproject PRJNA893007 (https://www.ncbi.nlm.nih.gov/sra/?term = PRJNA893007), accession number SRR31046078 (https://www.ncbi.nlm.nih.gov/sra/?term = SRR31046078).

Received: 19 February 2024; Accepted: 3 December 2024
Published online: 28 December 2024

References

1. WHO. World malaria report 2021.
2. Menard, D. & Dondorp, A. Antimalarial drug resistance: A threat to malaria elimination. *Cold Spring Harb. Perspect. Med.* https://doi.org/10.1101/cshperspect.a025619 (2017).
3. Sinden, R. E. Targeting the parasite to suppress malaria transmission. *Adv. Parasitol.* **97**, 147–185. https://doi.org/10.1016/bs.apar.2016.09.004 (2017).
4. Brancucci, N. M. B. et al. Lysophosphatidylcholine regulates sexual stage differentiation in the human malaria parasite Plasmodium falciparum. *Cell.* **171**, 1532–1544. https://doi.org/10.1016/j.cell.2017.10.020 (2017).
5. Abdi, A. I. et al. Plasmodium falciparum adapts its investment into replication versus transmission according to the host environment. *Elife*. https://doi.org/10.7554/eLife.85140 (2023).
6. Kafsack, B. F. et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature*. **507**, 248–252. https://doi.org/10.1038/nature12920 (2014).
7. Sinha, A. et al. A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. *Nature*. **507**, 253–257. https://doi.org/10.1038/nature12970 (2014).
8. Hawking, F., Wilson, M. E. & Gammage, K. Evidence for cyclic development and short-lived maturity in the gametocytes of Plasmodium falciparum. *Trans. R. Soc. Trop. Med. Hyg.* **65**, 549–559. https://doi.org/10.1016/0035-9203(71)90036-8 (1971).
9. Meibalan, E. & Marti, M. Biology of malaria transmission. *Cold Spring Harb. Perspect. Med.* https://doi.org/10.1101/cshperspect.a025452 (2017).
10. Sinden, R. E. Gametocytogenesis of Plasmodium falciparum in vitro: an electron microscopic study. *Parasitology*. **84**, 1–11. https://doi.org/10.1017/s003118200005160x (1982).
11. Joice, R. et al. Plasmodium falciparum transmission stages accumulate in the human bone marrow. *Sci. Transl. Med.* https://doi.org/10.1126/scitranslmed.3008882 (2014).
12. Alano, P. Plasmodium falciparum gametocytes: still many secrets of a hidden life. *Mol. Microbiol.* **66**, 291–302. https://doi.org/10.1111/j.1365-2958.2007.05904.x (2007).
13. van Biljon, R. et al. Hierarchical transcriptional control regulates Plasmodium falciparum sexual differentiation. *BMC Genomics*. **20**, 920. https://doi.org/10.1186/s12864-019-6322-9 (2019).
14. Hitz, E., Balestra, A. C., Brochet, M. & Voss, T. S. PfMAP-2 is essential for male gametogenesis in the malaria parasite Plasmodium falciparum. *Sci. Rep.* **10**, 11930. https://doi.org/10.1038/s41598-020-68717-5 (2020).
15. Chawla, J., Oberstaller, J. & Adams, J. H. Targeting gametocytes of the Malaria Parasite Plasmodium falciparum in a functional genomics era: Next steps. *Pathogens*. **10**(3), 346. https://doi.org/10.3390/pathogens10030346 (2021).
16. Chawla, J. et al. Phenotypic screens identify genetic factors associated with gametocyte development in the human Malaria Parasite Plasmodium falciparum. *Microbiol. Spectr.* **11**, e0416422. https://doi.org/10.1128/spectrum.04164-22 (2023).
17. Pires, C. V. et al. Chemogenomic profiling of a Plasmodium falciparum transposon mutant library reveals shared effects of dihydroartemisinin and bortezomib on lipid metabolism and exported proteins. *Microbiol. Spectr.* **11**, e0501422. https://doi.org/10.1128/spectrum.05014-22 (2023).
18. Zhang, M. et al. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science*. https://doi.org/10.1126/science.aap7847 (2018).
19. Bronner, I. F. et al. Quantitative insertion-site sequencing (QIseq) for high throughput phenotyping of transposon mutants. *Genome Res.* **26**, 980–989. https://doi.org/10.1101/gr.200279.115 (2016).
20. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. https://doi.org/10.1186/s13059-014-0550-8 (2014).
21. Young, J. et al. A CRISPR platform for targeted in vivo screens identifies Toxoplasma gondii virulence factors in mice. *Nat. Commun.* **10**, 3963. https://doi.org/10.1038/s41467-019-11855-w (2019).
22. Lopez-Barragan, M. J. et al. Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. *BMC Genomics*. **12**, 587. https://doi.org/10.1186/1471-2164-12-587 (2011).
23. Shang, X. et al. A cascade of transcriptional repression determines sexual commitment and development in Plasmodium falciparum. *Nucleic Acids Res*. **49**, 9264–9279. https://doi.org/10.1093/nar/gkab683 (2021).

24. 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. https://doi.org/10.1074/mcp.M900479-MCP200 (2010).
25. Meerstein-Kessel, L. et al. Probabilistic data integration identifies reliable gametocyte-specific proteins and transcripts in malaria parasites. *Sci. Rep.* **8**, 410. https://doi.org/10.1038/s41598-017-18840-7 (2018).
26. Zhang, M. et al. The apicoplast link to fever-survival and artemisinin-resistance in the malaria parasite. *Nat. Commun.* **12**, 4563. https://doi.org/10.1038/s41467-021-24814-1 (2021).
27. Dogga, S. K. et al. A single cell atlas of sexual development in Plasmodium falciparum. *Science*. **384**, eadj4088. https://doi.org/10.1126/science.adj4088 (2024).
28. Saliba, K. S. & Jacobs-Lorena, M. Production of Plasmodium falciparum gametocytes in vitro. *Methods Mol. Biol.* **923**, 17–25. https://doi.org/10.1007/978-1-62703-026-7_2 (2013).
29. Baker, D. A., Daramola, O., McCrossan, M. V., Harmer, J. & Targett, G. A. Subcellular localization of Pfs16, a Plasmodium falciparum gametocyte antigen. *Parasitology*. **108**(Pt 2), 129–137. https://doi.org/10.1017/s0031182000068219 (1994).
30. 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. https://doi.org/10.1016/0166-6851(94)90111-2 (1994).
31. Lanfrancotti, A., Bertuccini, L., Silvestrini, F. & Alano, P. Plasmodium falciparum: mRNA co-expression and protein co-localisation of two gene products upregulated in early gametocytes. *Exp. Parasitol.* **116**, 497–503. https://doi.org/10.1016/j.exppara.2007.01.021 (2007).
32. Yahiya, S. et al. Plasmodium falciparum protein Pfs16 is a target for transmission-blocking antimalarial drug development. *BioRxiv.* **14**, 2021–2106. https://doi.org/10.1101/2021.06.14.448287 (2021).
33. Kariuki, M. M. et al. Plasmodium falciparum: purification of the various gametocyte developmental stages from in vitro-cultivated parasites. *Am. J. Trop. Med. Hyg.* **59**, 505–508. https://doi.org/10.4269/ajtmh.1998.59.505 (1998).
34. Gupta, S. K., Schulman, S. & Vanderberg, J. P. Stage-dependent toxicity of N-acetyl-glucosamine to Plasmodium falciparum. *J. Protozool.* **32**, 91–95. https://doi.org/10.1111/j.1550-7408.1985.tb03020.x (1985).
35. Oberstaller, J. pfGO: P. falciparum functional enrichment analysis tools. https://doi.org/10.5281/zenodo.5931520 (2021).
36. Amos, B. et al. VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center. *Nucleic Acids Res.* **50**, D898–D911. https://doi.org/10.1093/nar/gkab929 (2022).
37. Maher, S. P., Balu, B., Shoue, D. A., Weissenbach, M. E. & Adams, J. H. A highly sensitive, PCR-based method for the detection of Plasmodium falciparum clones in microtiter plates. *Malar. J.* **7**, 222. https://doi.org/10.1186/1475-2875-7-222 (2008).
38. 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. https://doi.org/10.1038/nprot.2015.072 (2015).
39. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595. https://doi.org/10.1093/bioinformatics/btp698 (2010).
40. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2079–2079. https://doi.org/10.1093/bioinformatics/btp352 (2009).
41. Robinson, J. T. et al. Integrative genomics viewer. *Nat Biotechnol* **29**, 24–26. https://doi.org/10.1038/nbt.1754 (2011).

Acknowledgements
We thank the Wellcome Sanger Institute (United Kingdom) for performing the QIseq. We thank the USF Genomics Sequence Core and Omics Hub for the sequence service and productive discussion. This work is supported by NIAID grants R01AI117017 and R01AI130171.

Author contributions
Conceptualization: CVP, JC, JHA; Experimental Design: JC, CVP, MZ, LS, JHA; Investigation: CVP, JC, MZ; Data Curation: CVP, JO, CW, JG, TDO; Data Analysis: CVP, JO, CW, LS, TDO; Visualization: CVP, JC; Writing (original draft): CVP; Writing (reviewing and editing): JC, CVP, JO, LS, JG, TDO, JCR, MM, JHA; Supervision: MZ, MM, JHA; Project Administration: JCR, TDO, MM, JHA; Funding acquisition: JCR, MM, JHA.

Funding
National Institute of Allergy and Infectious Diseases, R01AI117017

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-82133-z.

Correspondence and requests for materials should be addressed to J.H.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.