



Patterns of Heterochromatin Transitions Linked to Changes in the Expression of *Plasmodium falciparum* Clonally Variant Genes

 Lucas Michel-Todó,^a Cristina Bancells,^a Núria Casas-Vila,^a Núria Rovira-Graells,^a Carles Hernández-Ferrer,^{b,c} Juan R. González,^{b,c}
 Alfred Cortés^{a,d}

^aISGlobal, Hospital Clínic, Universitat de Barcelona, Barcelona, Catalonia, Spain

^bISGlobal, Barcelona, Catalonia, Spain

^cCentro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

^dICREA, Barcelona, Catalonia, Spain


ABSTRACT The survival of malaria parasites in the changing human blood environment largely depends on their ability to alter gene expression by epigenetic mechanisms. The active state of *Plasmodium falciparum* clonally variant genes (CVGs) is associated with euchromatin characterized by the histone mark H3K9ac, whereas the silenced state is characterized by H3K9me3-based heterochromatin. Expression switches are linked to euchromatin-heterochromatin transitions, but these transitions have not been characterized for the majority of CVGs. To define the heterochromatin distribution patterns associated with the alternative transcriptional states of CVGs, we compared H3K9me3 occupancy at a genome-wide level among several parasite subclones of the same genetic background that differed in the transcriptional state of many CVGs. We found that *de novo* heterochromatin formation or the complete disruption of a heterochromatin domain is a relatively rare event, and for the majority of CVGs, expression switches can be explained by the expansion or retraction of heterochromatin domains. We identified different modalities of heterochromatin changes linked to transcriptional differences, but despite this complexity, heterochromatin distribution patterns generally enable the prediction of the transcriptional state of specific CVGs. We also found that in some subclones, several *var* genes were simultaneously in an active state. Furthermore, the heterochromatin levels in the putative regulatory region of the *gdv1* antisense noncoding RNA, a regulator of sexual commitment, varied between parasite lines with different sexual conversion rates.

IMPORTANCE The malaria parasite *P. falciparum* is responsible for more than half a million deaths every year. *P. falciparum* clonally variant genes (CVGs) mediate fundamental host-parasite interactions and play a key role in parasite adaptation to fluctuations in the conditions of the human host. The expression of CVGs is regulated at the epigenetic level by changes in the distribution of a type of chromatin called heterochromatin. Here, we describe at a genome-wide level the changes in the heterochromatin distribution associated with the different transcriptional states of CVGs. Our results also reveal a likely role for heterochromatin at a particular locus in determining the parasite investment in transmission to mosquitoes. Additionally, this data set will enable the prediction of the transcriptional state of CVGs from epigenomic data, which is important for the study of parasite adaptation to the conditions of the host in natural malaria infections.

KEYWORDS ChIP-seq, H3K9me3, *Plasmodium falciparum*, chromatin, *gdv1*, clonally variant gene expression, epigenetics, heterochromatin, malaria, transcriptomics

The apicomplexan parasite *Plasmodium falciparum*, the causative agent of malaria, has a complex life cycle divided between female *Anopheles* mosquitoes and humans. Human infection starts with the injection of parasites at the sporozoite stage during an

Editor Laura A. Kirkman, Weill Cornell Medicine

Ad Hoc Peer Reviewers Francesca Florini, Weill Cornell Medicine;  Jun Miao, University of South Florida

Copyright © 2022 Michel-Todó et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Alfred Cortés, alfred.cortes@isglobal.org.

The authors declare no conflict of interest.

Received 4 August 2022

Accepted 22 November 2022

Anopheles blood meal. After multiplication in the liver, thousands of parasites at the merozoite stage are released into the bloodstream. Merozoites invade red blood cells (RBCs) and start continuous rounds of the asexual intraerythrocytic development cycle (IDC), in which parasites develop through the intraerythrocytic ring, trophozoite, and schizont stages and then burst and release up to 32 new merozoites. Repeated rounds of the ~48-h IDC are responsible for all of the clinical manifestations of malaria. At each round, some parasites differentiate into sexual forms called gametocytes, which, after developing for ~10 days, can infect an *Anopheles* mosquito during a new blood meal. In the mosquito vector, male and female gametocytes mate and form zygotes that, after meiosis and several asexual replication steps, generate new sporozoites (1).

The complex life cycle of *P. falciparum* requires adaptation to the different conditions of each niche, which is in large part regulated at the transcriptional level. Therefore, the parasite has a different transcriptome at each stage of development (2–4). In addition, the parasite must adapt to fluctuations in the conditions of each niche. In the human blood, where infections can last for several months, parasites are exposed to changes in the availability of nutrients, in temperature (associated with fever episodes), in the presence of toxic compounds such as antimalarial drugs, and in host immune responses, among others (5–8). Rapid adaptation to fluctuations in the conditions of the human blood can be achieved by directed transcriptional responses, which involve sensing followed by changes in gene expression that protect against the challenge. However, malaria parasites can produce directed transcriptional responses to only a limited number of conditions (9–12). An alternative adaptive mechanism to survive under fluctuating conditions is based on epigenetic variation. In *P. falciparum*, epigenetic variation is linked to clonally variant genes (CVGs), which can be found in different transcriptional states (active or silenced) in parasites living in the same environment, with the same genome sequence, and at the same stage of the IDC. The transcriptional state of CVGs is transmitted from one round of the IDC to the next, but low-frequency switches between the active and silenced states occur, ensuring the constant generation of epigenetic diversity within parasite populations (13–17). As a consequence of this expression dynamics, different individual parasites within an isogenic population have different combinations of active and silenced CVGs. Transcriptional heterogeneity in parasite populations results in phenotypic heterogeneity, which provides the grounds for the continuous natural selection of epigenetic variants best adapted to the dynamic conditions of the environment. This constitutes a bet-hedging adaptive strategy (18).

Essentially all CVGs are involved in host-parasite interactions (13–17). The best-characterized family of CVGs is the *var* gene family, consisting of ~60 genes that encode *P. falciparum* erythrocyte membrane protein 1 (PfEMP1), a major virulence factor. PfEMP1 is exposed at the surface of infected RBCs and mediates antigenic variation and adhesion to endothelial cell receptors, which results in parasite sequestration within different host tissues. Individual parasites generally express a single member of the *var* gene family in a mutually exclusive manner, which is important for immune evasion (19–21). Other gene families linked to processes such as antigenic variation (e.g., *rifin* and *stevor*) (21), RBC remodeling (e.g., *pfmc-2tm* and *phist*) (22, 23), RBC invasion (e.g., *eba* and *Rh*) (24), lipid metabolism (e.g., *acs* and *acbp*) (25), or solute transport and nutrient acquisition (*clag*, with *clag3.1* and *clag3.2* displaying mutually exclusive expression) (26–29), among others, also show clonally variant expression (13, 18, 19, 26, 30–33). A peculiar CVG is the *pfap2-g* gene, which encodes a transcription factor of the ApiAP2 family that drives the conversion of asexual forms into nonreplicative sexual gametocytes (34–36). This gene is silenced in asexually growing parasites and, upon activation, triggers a transcriptional cascade that results in sexual conversion (37, 38). The proportion of parasites that activate *pfap2-g* at each round of the IDC determines the proportion of parasites that convert into sexual forms, i.e., the sexual conversion rate.

The epigenetic regulation of the expression of *P. falciparum* CVGs is based on reversible chromatin transitions between euchromatin (active) and facultative heterochromatin (silenced) states (13–17). The heterochromatic state is associated with the

histone posttranslational modification histone H3 trimethylation of lysine 9 (H3K9me3) and the associated heterochromatin protein 1 (PfHP1) (39–42). Furthermore, several ApiAP2 DNA binding proteins such as PfAP2-HC are also associated with heterochromatic regions (43, 44). The active state of CVGs is characterized by the absence of H3K9me3 and PfHP1 as well as the acetylation of H3K9 (H3K9ac) (45–48). The chromatin at CVG loci is considered bistable because, once established, both the euchromatic and the heterochromatic states are maintained across the full IDC and can be transmitted to the next generations. Infrequent transitions between the euchromatic and heterochromatic states at these loci underlie the expression switches of CVGs (15, 16).

In higher eukaryotes, H3K9me3 is typically associated with constitutive heterochromatin localized in subtelomeric and pericentromeric repetitive regions, but recent research has demonstrated that it also participates in cell differentiation processes (49, 50). In *P. falciparum*, H3K9me3 and PfHP1 are absent from pericentromeric regions, but they occupy subtelomeric repeats in addition to CVG loci. CVGs are located mainly in subtelomeric regions and a few chromosome internal clusters, and there are also a few stand-alone CVG loci in internal chromosome regions (40, 42, 44, 51, 52). In total, over 500 genes (~10% of the genome) carry heterochromatin marks and are considered CVGs. In contrast to the restricted distribution of H3K9me3, the histone mark associated with the active state of CVGs, H3K9ac, is widespread across the *P. falciparum* genome and marks constitutively euchromatic genes in addition to active CVGs. In non-CVGs, H3K9ac levels correlate with transcript levels and the temporal patterns of gene expression across the IDC (53, 54).

Previous studies showed that the overall distribution of heterochromatin at CVGs is stable across different stages of the *P. falciparum* IDC (40, 45–47, 52) but changes during transmission stages (52, 55–57). Indeed, in contrast to the relatively stable expression patterns of CVGs during multiple rounds of the IDC, the expression patterns of CVGs are reset during transmission stages (58–60). The overall distribution of heterochromatin is similar across *Plasmodium* spp., suggesting that it plays similar roles across the genus (52). A recent study showed that different *P. falciparum* lines have a similar overall distribution of heterochromatin, with differences at specific loci possibly reflecting different transcriptional states of the genes at these loci among the parasite lines analyzed, although gene expression was not assessed (52). However, the parasite lines used in that study had not been recently subcloned. Consequently, they were likely transcriptionally heterogeneous such that for many CVGs, the population is a mixture of parasites that have the gene in an active state and parasites that have it in a silenced state, as previously demonstrated for other parasite lines under long-term culture (18).

Here, we combined heterochromatin profiling with transcriptomic analysis to investigate in detail the chromatin differences between the active and silenced states of CVGs at a genome-wide level. For this, we compared the distribution of H3K9me3 between subcloned lines of the same genetic background that differ in the expression of dozens of CVGs. The use of subclones ensures that most genes are in a homogeneously active or silenced state within each population. We also compared the genome-wide distribution of heterochromatin between parasite lines that produce different numbers of gametocytes, with the aim of identifying epigenetic differences linked to sexual conversion rates.

RESULTS

Overview of the study. We analyzed the genome-wide distribution of H3K9me3 using chromatin immunoprecipitation followed by sequencing (ChIP-seq) and the transcriptome across the full IDC using two-channel microarrays for five different subclones of the 3D7 genetic background: 1.2B, 10G, A7, E5, and B11 (Fig. 1). The transcriptome analysis of the 1.2B and 10G subclones was previously reported (18). The analysis of 5 subclones enables 10 possible pairwise comparisons to identify genes that are active in one subclone and silenced in another. The selection of the subclones was based on

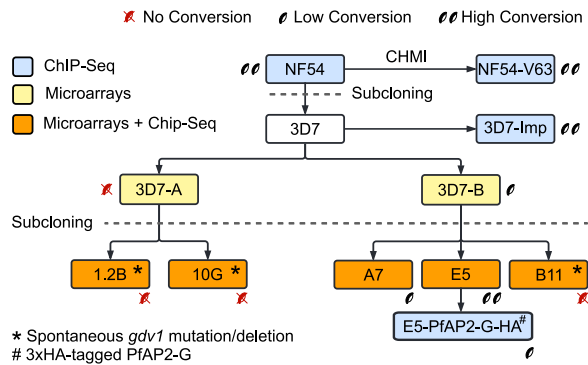


FIG 1 Schematic of the parasite lines used in this study. 3D7-A and 3D7-B are stocks of the clonal 3D7 line that were maintained in different laboratories for some years. Blue indicates parasite lines that were analyzed by ChIP-seq, yellow indicates parasite lines for which new or previously published transcriptomic data (using microarrays) were generated, and orange indicates parasite lines for which both ChIP-seq and transcriptomic data were generated. Parasite lines with a spontaneous deletion or a premature stop codon in *gdv1* (*) or with a genetic edition in the *pfap2-g* locus (#) are indicated. Parasite lines are semiquantitatively classified as high-gametocyte producers (two gametocytes, >10% sexual conversion rate), low-gametocyte producers (one gametocyte, <3% sexual conversion rate), or non-gametocyte producers (red, crossed gametocytes). NF54-V63 was obtained after human infection with the NF54 line in a controlled human malaria infection (CHMI) trial.

maximizing the number of differentially expressed genes among them and covering a broad range of sexual conversion rates. Subclones 1.2B and 10G were derived from the 3D7-A stock of 3D7, whereas A7, E5, and B11 were derived from the 3D7-B stock. The 3D7-A and 3D7-B stocks were maintained in different laboratories for several years and have phenotypic differences in growth rates, sexual conversion rates, density-dependent growth inhibition, the use of invasion pathways, and cytoadherence (26, 34, 61–64). The basal sexual conversion rate of the subclones ranges from zero (10G, 1.2B, and B11) to intermediate (<3%) (A7) or high (~15%) (E5) (34). To preserve transcriptional homogeneity, subclones were maintained in culture for the shortest time needed to obtain sufficient RNA or chromatin for the analyses (~4 to 5 weeks between limiting dilution to obtain the subclones and epigenomic or transcriptomic analysis).

Furthermore, we analyzed H3K9me3 occupancy in additional parasite lines with different basal sexual conversion rates: an E5-derived transgenic parasite line in which PfAP2-G was 3× hemagglutinin (3×HA) tagged (E5-PfAP2-G-HA) and the sexual conversion rate was lower (<3%) than that in the parental E5 line (34, 65); a stock of 3D7 obtained from Imperial College (3D7-Imp) with a high basal sexual conversion rate (66) (conversion rate in Albumax medium, 18.3% [range, 16.6 to 20%] [$n = 2$]); and the NF54 isolate from which the 3D7 clone was obtained, also a high-gametocyte producer (66) (conversion rate in Albumax medium, 21.8% [range, 18.8 to 24.8%] [$n = 2$]). In some of the analyses, we also included published ChIP-seq data for the NF54-V63 line that was recovered after infecting a volunteer with the NF54 line in a controlled human malaria infection (CHMI) trial (58). We also analyzed the distribution of H3K9ac in several of the parasite lines.

Transcriptomic analysis of 3D7 subclones. To identify genes that were transcriptionally active in one subclone and silenced in another, we analyzed the transcriptomes of tightly synchronized cultures (5-h age window) of the A7, E5, and B11 lines at six time points along the IDC. The analysis pipeline to identify genes with large expression differences between the subclones, similar to the method used in our previous study including 10G, 1.2B, and 3D7-B (18), was based on the statistically estimated parasite age (67), the average fold change (AFC) in transcript levels (between subclones) at four overlapping time intervals of the IDC, and the maximum (among the time intervals) average fold change (mAFC) parameter (18, 58). After applying several filters (see Materials and Methods), we generated for each pair of subclones a high-confidence list of differentially expressed genes with an mAFC of >4 [$\log_2(\text{mAFC})$ of >2].

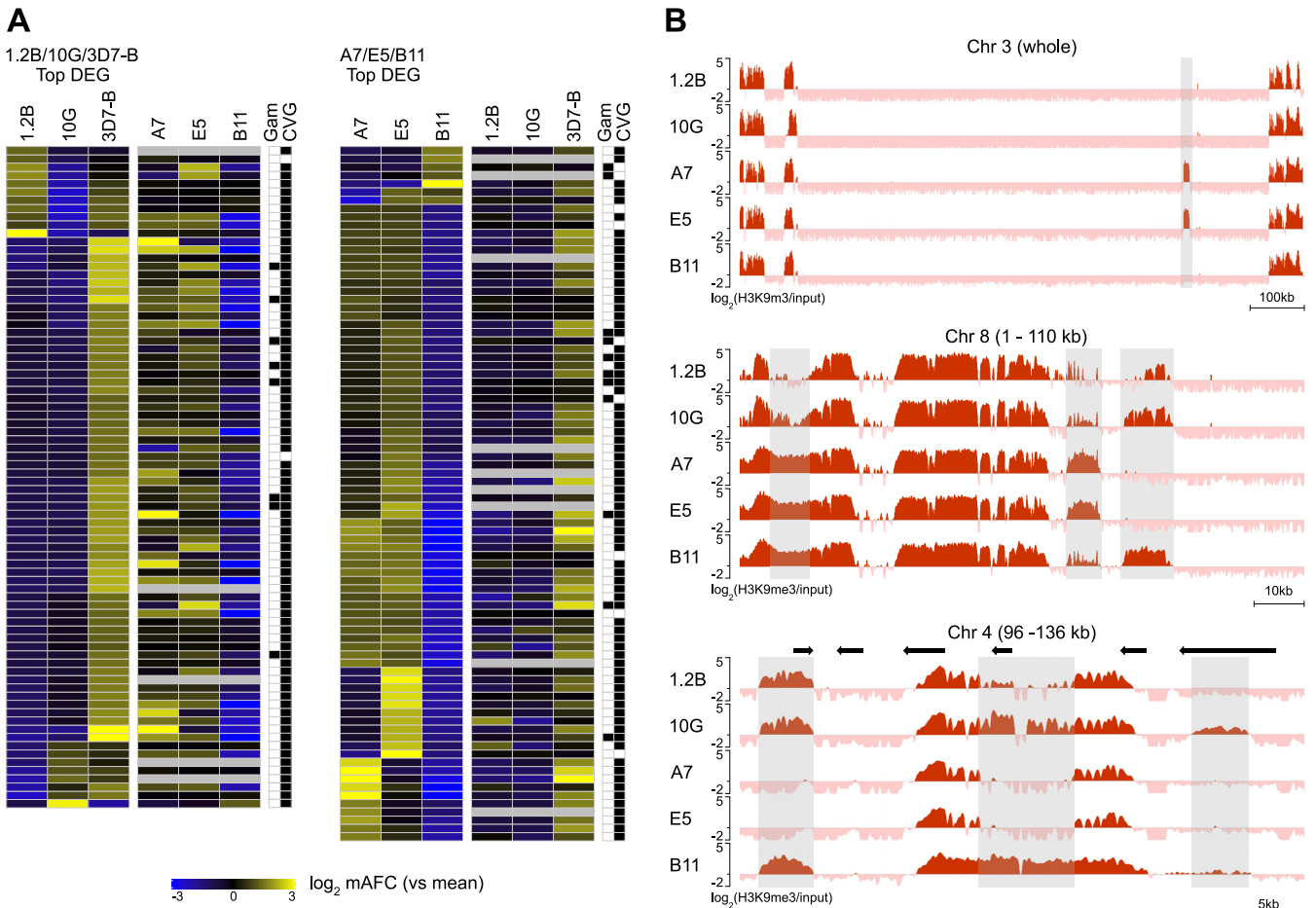


FIG 2 Overview of the transcriptomic and epigenomic analyses of the subclones. (A) Expression patterns of the top differentially expressed genes (DEG). (Left) Genes with differential expression (mAFC of >4) among the 1.2B, 10G, and 3D7-B lines (previously published data [18]). Values for the expression differences in the same genes among the A7, E5, and B11 lines (newly generated data) are shown. All values are the mean-centered log₂ of the mAFC in expression [log₂(mAFC)] among the three parasite lines. (Right) Analogous analysis for the top differentially expressed genes in the comparison of A7, E5, and B11. The last two columns indicate whether a gene is a known gametocyte marker or has been previously classified as a CVG. The values presented in this panel are provided in Data Set S2 in the supplemental material. (B) H3K9me3 distribution in a representative chromosome or chromosome regions in the five subclones at different magnification levels. Gray shades indicate regions with major differences in H3K9me3 coverage among subclones. Values are log₂(H3K9me3/input).

Between 19 and 74 genes were differentially expressed in the pairwise comparisons between 10G and 1.2B or among A7, E5, and B11 (Fig. 2A; see also Data Sets S1 and S2 in the supplemental material). Since 3D7-A subclones (1.2B and 10G) and 3D7-B subclones (A7, E5, and B11) were analyzed using microarrays with slightly different designs (18, 35, 68) and with different preparations of the reference pool, a direct comparison between subclones of the two different groups was not possible. However, we combined data for A7, E5, and B11 with data for the direct comparison of 1.2B and 10G against 3D7-B (18) to identify genes that were likely differentially expressed between 10G or 1.2B and A7, E5, or B11 (see Materials and Methods). In total, 121 genes were differentially expressed with high confidence in at least one of the possible pairwise comparisons, of which 111 were genes previously classified as CVGs based on the presence of heterochromatin marks or variant expression in previous studies (see the list of CVGs in reference 58). A few of the differentially expressed genes are gametocyte markers, and their different transcript levels may reflect the different sexual conversion rates among the subclones (Fig. 2A).

A large fraction of the differentially expressed genes showed lower expression levels in B11 than in A7 and E5. Unexpectedly, many of these genes were also expressed at lower levels in 1.2B and 10G than in 3D7-B, suggesting that the CVG expression patterns in B11 were relatively similar to those in 1.2B and 10G (Fig. 2A). Analysis of

genetic polymorphisms in these parasite lines using the ChIP-seq input data revealed several large deletions in B11, including an ~20-kb deletion affecting the *gdv1* locus (69, 70) (Fig. S1 and Data Set S3). Of note, 10G and 1.2B also have a defect in *gdv1*, in this case a mutation that results in a premature stop codon (Q578X) (Data Set S3) (35).

Heterochromatin distribution in the 3D7 subclones and general association with gene expression. ChIP-seq analysis of H3K9me3 revealed similar global distribution of heterochromatin among the five subclones, although differences were apparent at specific loci (see Fig. 2B for representative examples). The mean pairwise correlation among parasite lines was 0.88 (range, 0.83 to 0.95) using input-normalized coverage values (Fig. S2). We performed individual peak calling using the MACS2 tool (Data Set S4) and differential peak calling for all possible pairwise comparisons using custom scripts. There was an average of 142 (range, 68 to 214) differential peaks in pairwise comparisons, which map to an average of 62 (range, 34 to 99) genes (Data Set S5). Differential peaks involving changes in H3K9me3 coverage at the upstream region (1,000 bp) or the first half of the coding sequence (CDS) of a gene were generally associated with transcriptional differences such that the parasite line with higher coverage had lower transcript levels. However, when H3K9me3 coverage was different in only the second half of the coding sequence or the downstream region of a gene, typically, there were no associated transcriptional differences (see Fig. 3A for the comparison of 10G versus 1.2B and Fig. S3A to C for the other direct pairwise comparisons; Data Set S2). In addition to differences in the heterochromatin occupancy observed at chromosome regions containing genes, there were also differences at the telomere-associated repetitive elements (TAREs) (71, 72) located between the telomeres and subtelomeric genes (Fig. S1).

To further assess the relationship between differences in heterochromatin distribution and transcriptional differences, we first calculated for all differentially expressed genes in any of the pairwise comparisons the level of correlation between the H3K9me3 coverage at different landmark regions of the genes and transcript levels. The H3K9me3 coverage was calculated for fixed-length regions at different positions relative to the start or stop codons and also at the 5' untranslated region (UTR), the coding sequence, and the 3' UTR (according to the genome annotation in PlasmoDB v52). The strongest negative correlation between H3K9me3 coverage and transcript levels occurred at the 5' UTR and in the fixed-length -500 bp to ATG regions, but strong correlations (Pearson's $|r|$ value of >0.6) were observed throughout the upstream and coding sequences (Fig. S4). In genes with large transcriptional differences (mAFC of >4) among the subclones, there was a strong negative association between transcript levels and H3K9me3 coverage at the region from -1000 to +500 bp such that the subclone that expressed a gene at higher levels tended to have less heterochromatin (Fig. 3B, Fig. S3D to F, and Data Set S2).

Changes in heterochromatin distribution at CVG loci between the active and silenced states. To characterize in more detail the changes in the heterochromatin distribution that drive the transition between the silenced and active states of CVGs, we analyzed for each gene the H3K9me3 coverage across 23 bins encompassing the neighboring genes, the upstream region (until the first upstream gene), the CDS, and the downstream region (until the first downstream gene). For the 121 genes in the high-confidence list of differentially expressed genes, we compared the heterochromatin distribution between the two subclones with the maximum expression differences (Fig. 4). The subclone with the highest expression level was used to represent the active state of the gene, and the subclone with the lowest expression level was used to represent the silenced state. For genes in which the largest expression difference occurred between the 1.2B or 10G subclone and 3D7-B, with higher expression levels in 3D7-B than in 10G or 1.2B, we used the 3D7-B subclone (A7, E5, or B11) with the highest expression level as the subclone representative of the active state (and vice versa; i.e., for genes with lower expression levels in 3D7-B than in 10G or 1.2B, the 3D7-B subclone with the lowest expression level was used as the representative of the silenced state). The H3K9me3 coverages in the active and silenced states and the coverage difference

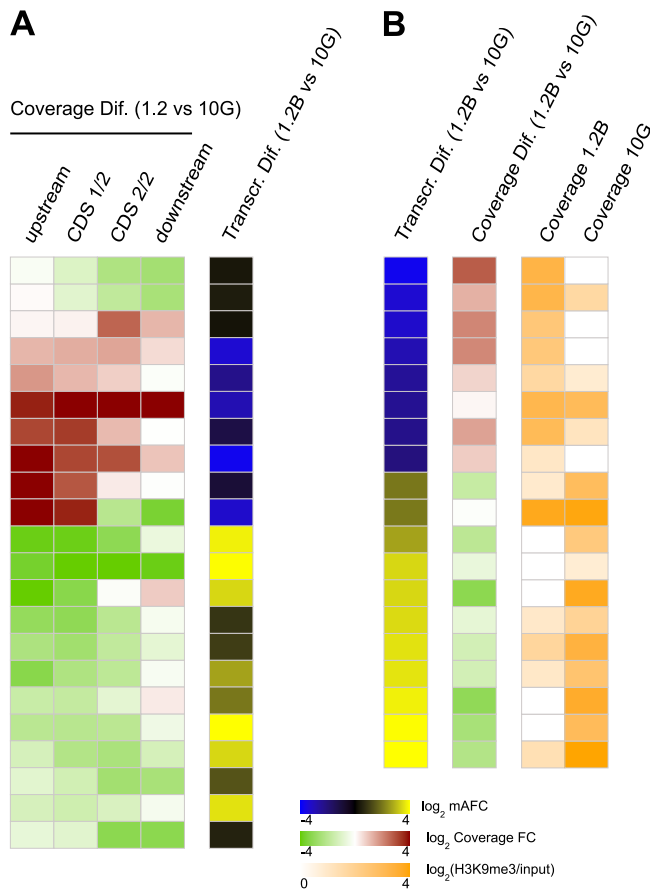


FIG 3 Association between H3K9me3 coverage differences and transcriptional differences. Results for subclones 1.2B and 10G are shown (see Fig. S3 in the supplemental material for the other pairwise comparisons). (A) Transcriptional differences in genes overlapping (from -1000 to +500 bp relative to the ATG start codon) an H3K9me3 differential peak between 1.2B and 10G. The \log_2 of the input-normalized coverage fold change (FC) in 1.2B relative to 10G is shown for the upstream region (1,000 bp before the ATG codon), the first or second half of the coding sequence (CDS 1/2 or CDS 2/2, respectively), and the downstream region (1,000 bp after the stop codon). Transcriptional differences are expressed as the \log_2 of the mAFC (1.2B versus 10G). (B) Heterochromatin levels in genes differentially expressed between 1.2B and 10G (mAFC of >4). Transcriptional differences are shown as described above for panel A. Heterochromatin coverage (\log_2 of the input-normalized coverage) and coverage differences between 10G and 1.2B (\log_2 of the input-normalized coverage fold change) are shown for the region from -1000 to +500 bp relative to the ATG codon. The values presented in this figure are provided in Data Set S2.

between the two were used to group the 121 differentially expressed genes into seven distinct clusters using the *k*-means algorithm (Fig. 4 and Data Set S2).

Cluster 1 (22 genes) corresponds to genes that are largely devoid of heterochromatin in the upstream and coding regions in their active state but fully heterochromatic in these regions and also in the downstream region in their silenced state. We called this the “global transition” cluster (Fig. 4 and Fig. 5A to C). Genes in cluster 2 (14 genes) have heterochromatin patterns that are relatively similar to those of the genes in cluster 1, but in the silenced state, the downstream region remains euchromatic (“mainly 5’ transition” cluster) (Fig. 4 and Fig. 5D and E).

Clusters 3, 4, and 5 (20, 13, and 29 genes, respectively) correspond to genes in which the full locus (clusters 3 and 5) or the upstream region and the beginning of the coding region (cluster 4) are marked by H3K9me3 in the silenced state, and in the active state, there is only a localized reduction of the coverage of this mark in the upstream region (Fig. 4 and Fig. 5F and G). We called these the “localized 5’ transition” clusters. In a few genes from these clusters, changes in gene expression were not accompanied by noticeable differences in heterochromatin levels. We cannot exclude

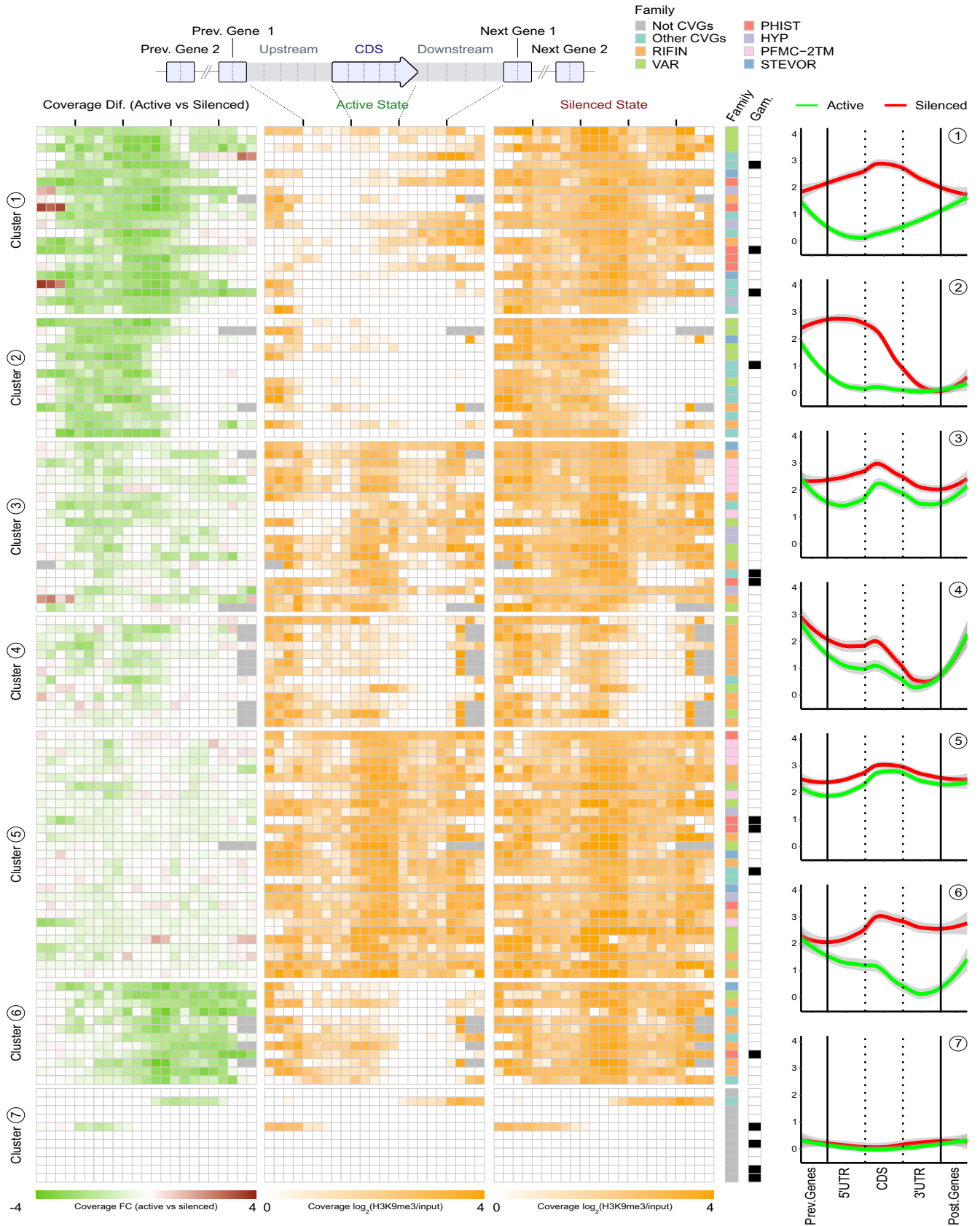


FIG 4 Distribution of heterochromatin in the active and silenced states of CVGs. The H3K9me3 coverage in the top differentially expressed genes among 1.2B, 10G, A7, E5, and B11 in the active compared to the silenced state is shown. For each gene, the subclone with the highest expression level was (Continued on next page)

the possibility that, even in recently subcloned lines, some of these genes may be expressed in only a small fraction of the parasites in the subclone in which they are tagged as active.

Cluster 6 (12 genes) includes genes in which heterochromatin is present in the upstream, coding, and downstream regions in the silenced state, whereas in the active state, heterochromatin is depleted mainly in the second half of the coding region and the downstream region (“mainly 3’ transition” cluster) (Fig. 4 and Fig. 5H and I). This pattern suggests that heterochromatin in the downstream region of these genes may play a role in their transcriptional regulation, although we cannot exclude the possibility that their expression was actually driven by differences in the upstream region. Indeed, differences in the upstream region between the active and silenced states occurred in the majority of cluster 6 genes, albeit these were less pronounced than the differences in the downstream region.

Finally, cluster 7 (11 genes) corresponds to genes largely devoid of heterochromatin in both the active and silenced states. Therefore, the transcriptional variation of these genes is not controlled by euchromatin-heterochromatin transitions. Of note, the vast majority of the genes in this cluster had not been previously classified as CVGs.

Except for the non-CVGs from cluster 7, in the majority of the differentially expressed genes, there was heterochromatin at some part of the locus or at the neighboring genes, even in the active state there was heterochromatin at some part of the locus or at neighboring genes. Furthermore, the transition from the active to the silenced state was typically associated with an increase in heterochromatin levels in a continuous region spanning from nearby heterochromatic regions to the regulatory regions of the gene. This is consistent with the silencing and activation of CVGs commonly involving the spreading and retraction of heterochromatin, respectively, according to an accordion-like mechanism. In several cases, this involved changes in the position of the outer limit of the heterochromatin domain (Fig. 5A, D, and H to J), whereas in other cases, the transition from the silenced to the active state was associated with reduced heterochromatin levels in a region surrounded by heterochromatin on both sides. The latter is consistent with an opening at an internal point within the heterochromatin domain followed by heterochromatin retraction (Fig. 5B, F, and G). We defined four main transition patterns between the active and silenced states of CVGs (Fig. 5L) and qualitatively assigned each of the 109 classifiable differentially expressed genes in clusters 1 to 6 to the pattern with which they best matched (Fig. 5L and Data Set S5). Only 5 differentially expressed genes showed transitions involving heterochromatin *de novo* formation or complete removal, whereas 65 genes showed transitions consistent with heterochromatin expansion/retraction, either affecting the external limits of a heterochromatin domain [“expansion/retraction (limits)”] (42 genes) or involving an opening within a heterochromatin domain that was continuous in the silenced state [“expansion/retraction (internal)”] (23 genes). In 39 genes, the transition between the active and silenced states involved small changes affecting heterochromatin levels in only a very small region (“localized closing/opening”) (Fig. 5L). However, it is important to note that the four general patterns likely represent a mechanistic continuum rather than discrete, fully independent mechanisms, and for some genes, the classification was ambiguous. Furthermore, in some genes, transitions followed a different pattern depending on the state of the neighboring genes.

FIG 4 Legend (Continued)

selected as the representative for the active state, and the subclone with the lowest expression level was selected as the representative for the silenced state. The distribution of heterochromatin in each subclone (\log_2 of the input-normalized coverage) and the coverage difference (\log_2 of the fold change [FC] in coverage in the active versus the silenced state) are shown for 23 bins spanning the coding sequence (CDS) (divided into 5 bins), the upstream and downstream regions (up to the next gene) (5 bins each), and the coding sequences of the two upstream and downstream neighboring genes (2 bins each). Genes were clustered according to their H3K9me3 coverage and coverage differences (excluding the bins from the neighboring genes) using *k*-means clustering ($k = 7$, guided by silhouette-and-elbow plots). Within each cluster, genes are ordered by fold changes in transcript levels between the active and silenced states. The two columns at the right of the heat map indicate the gene family and whether a gene is a known gametocyte marker. The plots on the right are locally estimated scatterplot smoothing (LOESS) regression plots (shades are 95% confidence intervals) showing the heterochromatin distribution (\log_2 of the input-normalized H3K9me3 coverage) in the active and silenced states for each cluster. The details of the genes included in this figure are provided in Data Set S2 in the supplemental material.

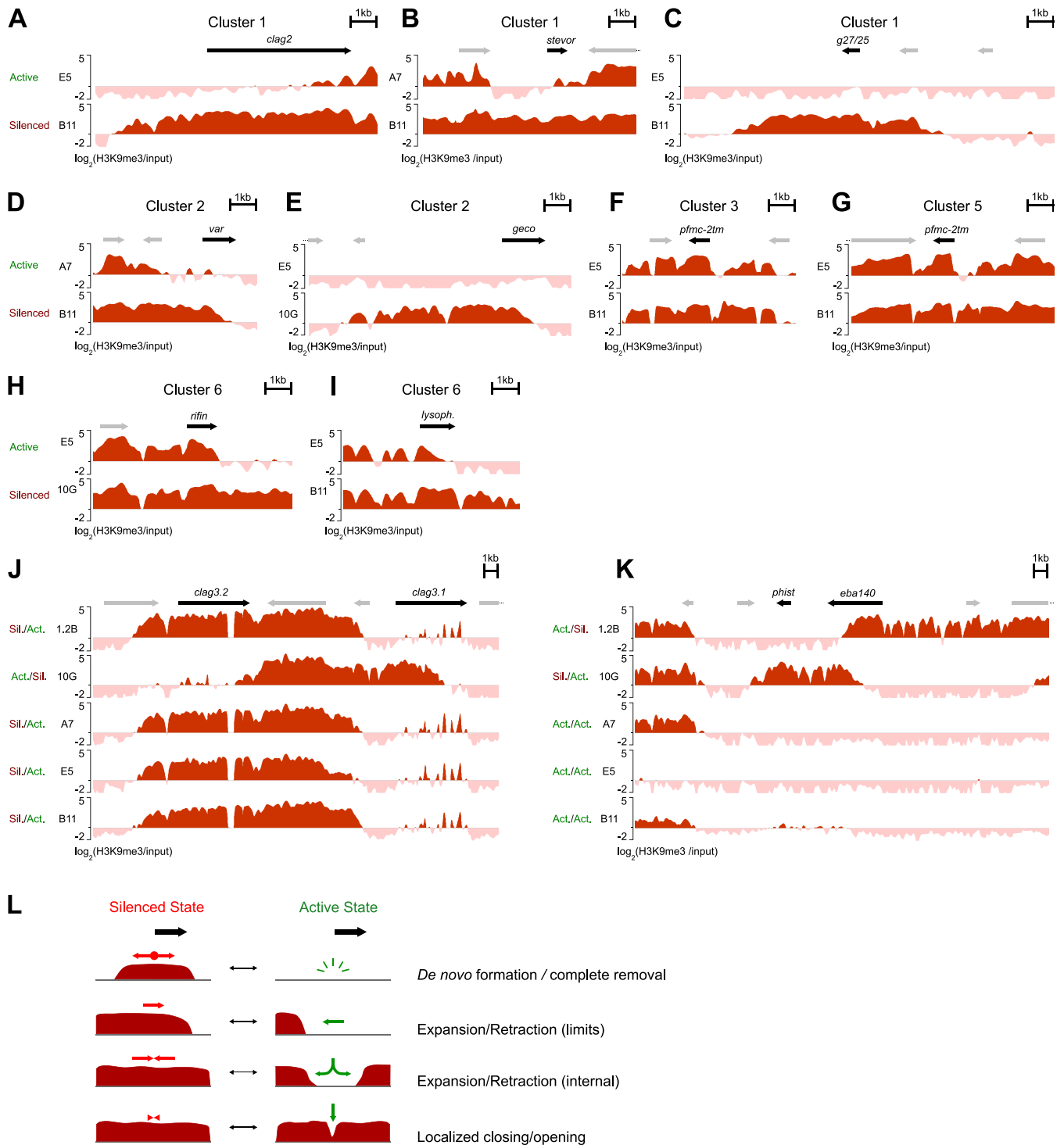


FIG 5 Representative examples of the different modes of heterochromatin transitions between the active and silenced states of CVGs. (A to I) Representative examples of heterochromatin distribution [$\log_2(\text{H3K9me3}/\text{input})$] associated with the active or silenced state of CVGs from the different clusters shown in Fig. 4. The genes are representative of “global transition” (cluster 1 [A to C]), “mainly 5’ transition” (cluster 2 [D and E]), “localized 5’ transition” (cluster 3 [F] and cluster 5 [G]), or “mainly 3’ transition” (cluster 6 [H and I]). Panels C and E are representative of transitions that involve the *de novo* formation/complete removal of a heterochromatin domain, whereas the other panels are representative of transitions that can be explained by heterochromatin expansion/retraction. (J and K) Representative examples of heterochromatin distribution in subclones that have a gene in a different transcriptional state and subclones that have it in the same transcriptional state. In all panels, the active or silenced state corresponds to the genes displayed as black arrows (with names). Other annotated genes are shown as gray arrows. The identifiers of the genes displayed as black arrows are PF3D7_0220800 (A), PF3D7_0300900 (B), PF3D7_1302100 (C), PF3D7_0713300 (D), PF3D7_1253000 (E), PF3D7_1039700 (F), PF3D7_0114100 (G), PF3D7_1480000 (H), PF3D7_0936700 (I), PF3D7_0302200 and PF3D7_0302500 (J), and PF3D7_1301500 and PF3D7_1301600 (K). (L) Schematic representation of the four main modalities of heterochromatin transitions between the active and silenced states of CVGs: the *de novo* formation of heterochromatin or the complete disruption of a heterochromatin domain (*De novo* formation/complete removal); the expansion/

(Continued on next page)

While for most genes activation or silencing could be explained by heterochromatin expansion or retraction across a continuous region or small localized transitions, there were 13 H3K9me3 differential peaks in which the *de novo* formation or the complete removal of a heterochromatin domain actually appeared to occur (Fig. 5C and E and Data Set S5). This involved heterochromatin islands isolated from other heterochromatic regions in which, in the active state, the region occupied by these islands and the adjacent regions (at least 5 kb on each side) were fully euchromatic. Unexpectedly, the majority of the 20 genes embedded in these regions were expressed at very low levels both in subclones in which they were euchromatic and in subclones in which they were heterochromatic, and only 5 of them showed differential expression among the 5 subclones (i.e., they were included in the high-confidence list of 121 differentially expressed genes). Indeed, most of these 20 genes were previously reported to be predominantly expressed in gametocytes or mosquito stages rather than during the IDC (Data Set S5). Together, these results suggest that variant gene expression linked to the adaptation of asexual blood-stage parasites to fluctuating conditions during the IDC is regulated mainly by heterochromatin expansion and retraction, whereas the *de novo* formation or the complete removal of heterochromatin is largely restricted to heterochromatin islands containing genes expressed at other stages of development.

The heterochromatin distribution associated with the active or silenced state of a specific CVG is generally conserved. For most CVGs, the analysis included more than one subclone in which the gene was either active or silenced. To compare the distribution of heterochromatin between subclones in which the gene was in the same transcriptional state, we first classified all genes in the five subclones as CVGs or non-CVGs and as active or silenced/inactive based on our transcriptomic data (Cy5 signal and expression differences between subclones) and publicly available transcriptome sequencing (RNA-seq) data sets (see Materials and Methods). Genes for which the active or silenced state could not be unambiguously determined were classified as “undetermined” (Data Set S6). In the majority of CVGs, the distribution of heterochromatin was almost identical among subclones in which the gene was in the same transcriptional state (Fig. 5J, Fig. S5, and Data Set S2). In the few genes in which prominent heterochromatin differences were observed between subclones that had the gene in the same state, differences typically occurred toward the end of the coding sequence or downstream region and could be explained by a different transcriptional state of the neighboring genes. As an example, in the active state, *eba-140* had heterochromatin in the second half of the coding sequence and downstream region only when the downstream *phist* gene was silenced (Fig. 5K).

Heterochromatin differences between the active and silenced states in different families of CVGs. None of the clusters described in Fig. 4 were strictly linked to a specific CVG family, but different families were unevenly represented among the clusters. Different scenarios were observed for different gene families (Fig. S6). In the *rif*, *stevor*, *pfmc-2tm*, and *var* families, the majority of genes that did not show transcriptional differences among the five subclones analyzed were fully heterochromatic, whereas in the *clag*, *acs*, *hyp*, and *phist* families, nondifferentially expressed genes were typically euchromatic. In some gene families, the majority of differentially expressed genes showed similar heterochromatin transition patterns between the active and silenced states (e.g., localized 5' transition in the *pfmc-2tm* family), but in other families, the transition patterns were rather heterogeneous.

In specific genes for which we had previously characterized the heterochromatin

FIG 5 Legend (Continued)

retraction of heterochromatin at the boundaries of a heterochromatin domain [Expansion/Retraction (limits)]; the opening of heterochromatin within a heterochromatin domain, followed by heterochromatin retraction, or the expansion resulting in a continuous heterochromatin domain from two separate neighboring domains [Expansion/Retraction (internal)]; or the opening/closing within a heterochromatin domain with minimal or no expansion/retraction (Localized closing/opening). Changes associated with silencing and activation are indicated by red and green arrows, respectively.

distribution associated with the active or silenced state in the 10G and 1.2B subclones using chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR) (45), the results were fully consistent. In the mutually exclusively expressed *clag3.1* and *clag3.2* genes (26, 45, 73), a heterochromatin domain centered at the *var* pseudogene located between the two genes expanded toward one gene or the other (Fig. 5J). In *eba-140*, the active state was associated with the absence of heterochromatin in the upstream region and the beginning of the coding sequence, whereas heterochromatin in the downstream region did not influence the expression of the gene and reflected the silencing of the neighboring *phist* gene, as described above (Fig. 5K).

Multiple *var* genes are simultaneously active in some subclones. Genes of the *var* family typically show mutually exclusive expression such that individual parasites express only one gene of the family at a time and keep all of the others repressed (19–21). To obtain a semiquantitative estimate of the expression patterns of *var* genes in the five 3D7 subclones, we used the normalized Cy5 signal (i.e., not relative to the reference pool) (Fig. 6A), as we did in previous studies (18, 58). Subclones 1.2B and B11 showed the expected pattern of mutually exclusive *var* gene expression, with a single *var* gene being predominantly expressed, whereas no expressed *var* gene was identified in 10G. Unexpectedly, subclones A7 and E5 expressed several *var* genes (four and six, respectively) at relatively high levels.

Two different scenarios can explain this observation: (i) authentic simultaneous expression of multiple *var* genes in individual parasites or (ii) transcriptional heterogeneity such that despite the use of recently subcloned lines, different subsets of parasites express different *var* genes. To distinguish between these two possibilities, we used the H3K9me3 ChIP-seq data. If each *var* gene was active in only a subset of the parasites, a partial reduction of heterochromatin levels would be expected, but heterochromatin was almost completely depleted at the regulatory regions of all but one (PF3D7_0413100 in A7) of the highly expressed *var* loci (Fig. 6B). This result suggests that multiple *var* genes were actually simultaneously active in the A7 and E5 subclones. The majority of the *var* genes that were expressed at high levels in one or more of the subclones and silenced in the others showed major differences between the active and silenced states (cluster 1, 2, or 6 in Fig. 4), although two showed more localized differences (cluster 3) (Data Set S2). While an unambiguous demonstration of the simultaneous expression of multiple *var* genes in the same cell requires analyses at the single-cell level, such as RNA fluorescence *in situ* hybridization (RNA-FISH) or single-cell RNA-seq (scRNA-seq), and we did not determine if the multiple *var* genes in an active chromatin state produce full-length mRNA and PfEMP1 protein, our results are consistent with previous reports indicating that the mutually exclusive expression of *var* genes may not be strict (74, 75).

Differences in heterochromatin distribution between parasite lines with different sexual conversion rates. The proportion of parasites that convert into gametocytes at each round of the IDC is called the sexual conversion rate. It is well established that different parasite lines, and even different subclones of the same genetic background, have different rates of basal (noninduced) sexual conversion that are inherited across multiple generations of asexual blood growth (34, 76). Sexual conversion depends on the activation of *pfap2-g*, which is normally silenced by heterochromatin in asexually growing blood-stage parasites. This gene is a peculiar CVG because, unlike other CVGs, only the silenced state can be transmitted from one generation of the IDC to the next: parasites in which the gene is activated convert into nonreplicating gametocytes and abandon the IDC. Therefore, the stable differences in the conversion rates between parasite lines imply the transmission of the probability of activation of *pfap2-g* rather than the transmission of an active or silenced state. To identify potential epigenetic determinants of the probability of activation of *pfap2-g*, we compared the distribution of heterochromatin among the five 3D7 subclones and additional parasite lines with a broad range of sexual conversion rates (0 to >10%) (Fig. 1A). All parasite lines were maintained in standard culture medium containing Albumax II and no human serum, with the exception of NF54-V63 (previously published data), which was analyzed in medium containing human serum (58).

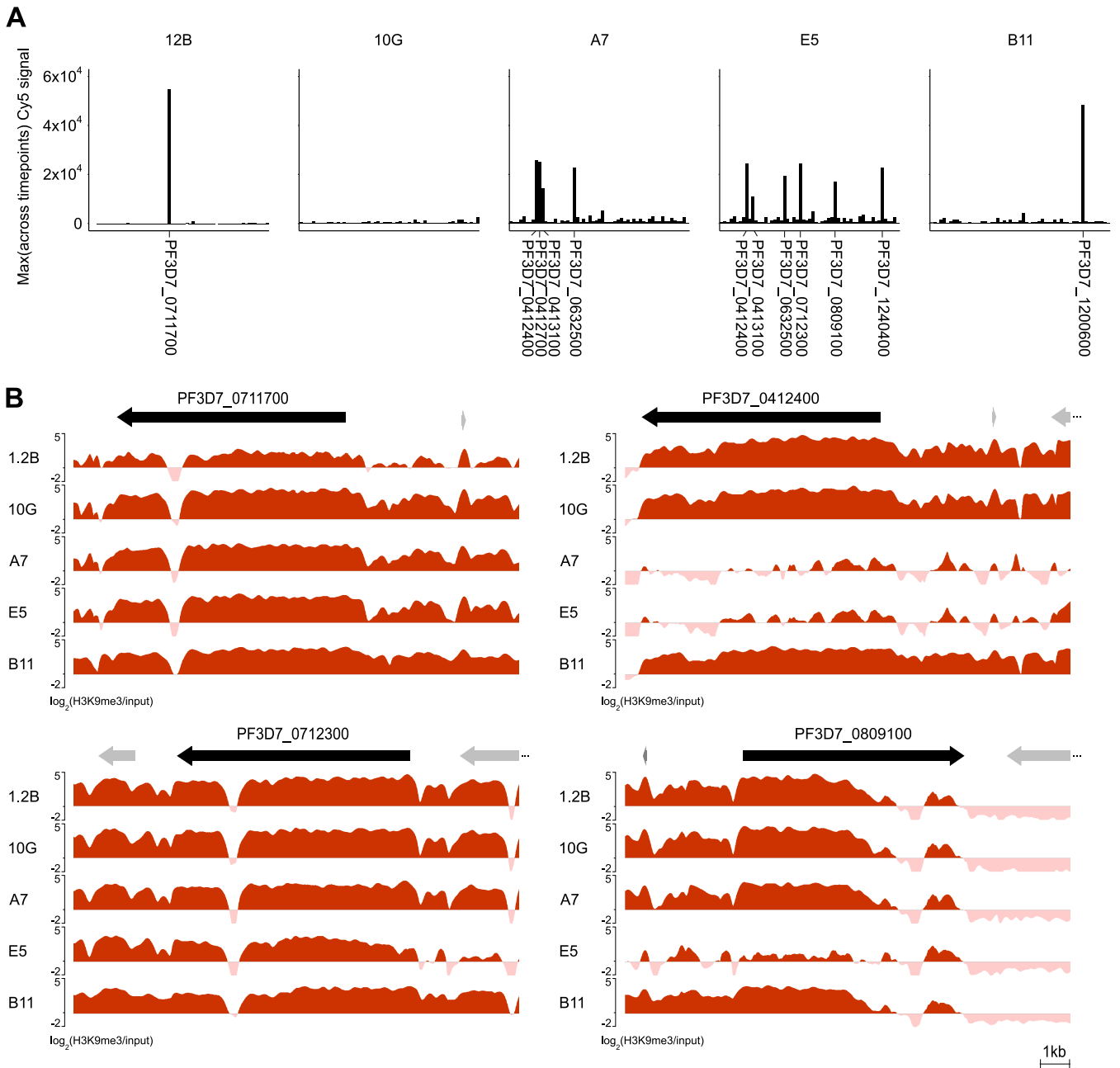


FIG 6 Expression patterns and heterochromatin distribution at *var* gene loci. (A) Expression of *var* genes in the five subclones. Values are the normalized sample signal (Cy5 channel), which provides a semiquantitative estimation of the absolute expression levels. The gene identifier is provided only for *var* genes expressed at high levels (normalized Cy5 signal intensity of >10,000). (B) H3K9me3 distribution at representative *var* genes that are highly expressed in some of the subclones. Black arrows (with identifiers) are the differentially expressed *var* genes; other annotated genes are shown as gray arrows. Values are $\log_2(\text{H3K9me3}/\text{input})$.

The heterochromatin occupancy at the *pfap2-g* locus was similar among all of the parasite lines, with similar 5' and 3' boundaries of the H3K9me3-enriched region (Fig. 7). Therefore, a different extension of the heterochromatin domain at the *pfap2-g* locus does not appear to underlie the different probabilities of *pfap2-g* activation between parasite lines. There were subtle differences between parasite lines in the H3K9me3 coverage at the 5' and 3' ends of the H3K9me3 domain, but these may be explained by the different proportions of parasites that were sexually committed (and therefore had *pfap2-g* in an active state) rather than authentic epigenetic differences transmitted across multiple rounds of the IDC.

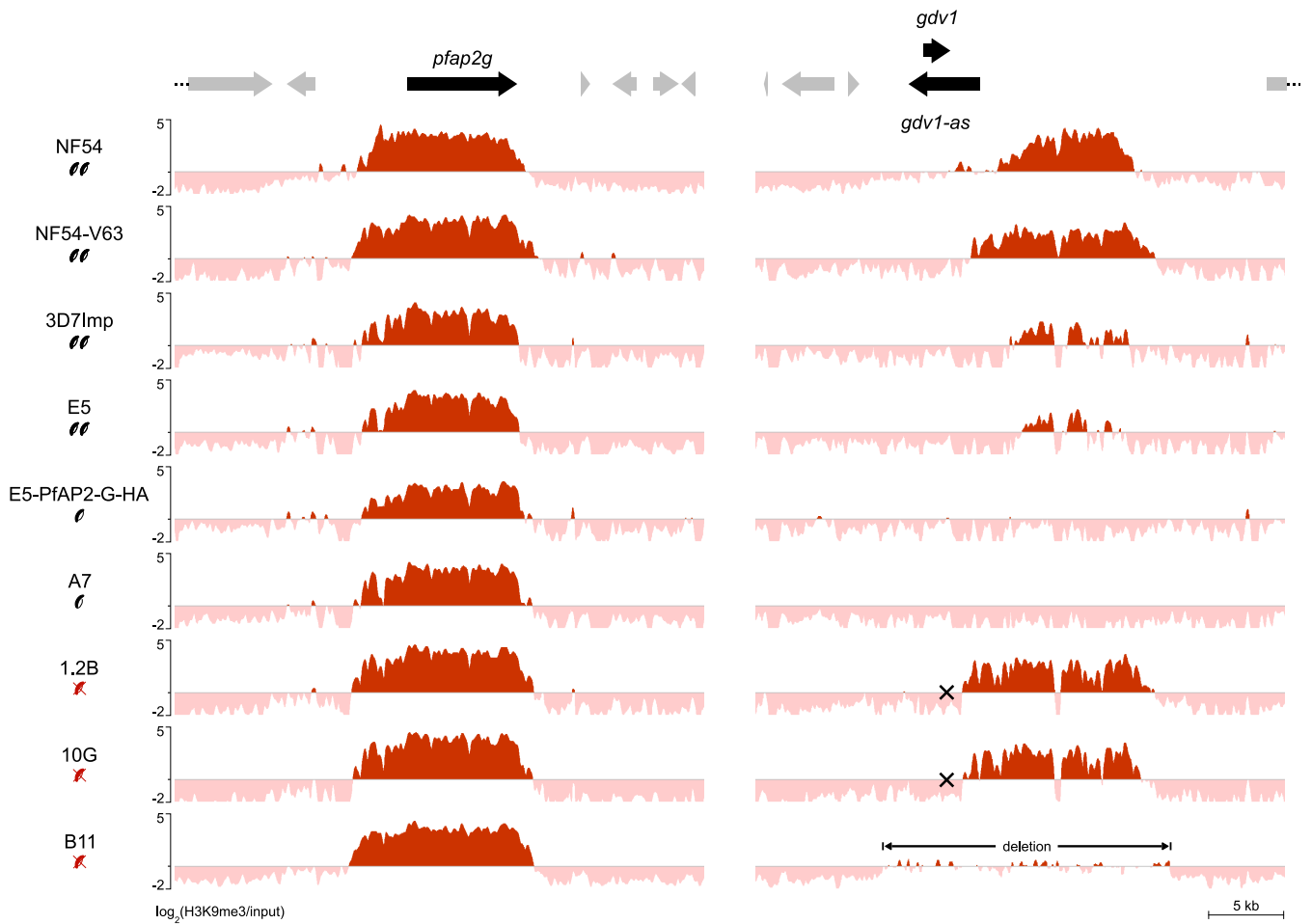


FIG 7 Distribution of heterochromatin at the *pfap2-g* and *gdv1* loci encoding regulators of sexual conversion. The distribution of H3K9me3 at the *ap2-g* and *gdv1* loci in different parasite lines is shown. Values are $\log_2(\text{H3K9me3}/\text{input})$. Parasite lines are semiquantitatively classified as high-gametocyte producers (two gametocytes, >10% sexual conversion rate), low-gametocyte producers (one gametocyte, <3% sexual conversion rate), or non-gametocyte producers (red, crossed gametocytes), as described in the legend of Fig. 1. Mutations that produce a premature stop codon in *gdv1* (crosses) or a deletion including the full *gdv1* locus are indicated. Other annotated genes are indicated as gray arrows.

In contrast, we observed major differences between the parasite lines in the heterochromatin distribution at the locus encoding GDV1, an upstream positive regulator of *pfap2-g* (69, 70, 77) (Fig. 7). While *gdv1* is typically not considered a CVG or a heterochromatic gene, in some of the parasite lines, we observed a prominent H3K9me3 signal downstream of the coding sequence. This heterochromatin domain overlaps the putative promoter of the long noncoding antisense RNA (*gdv1-as*) that was shown previously to operate as a negative regulator of GDV1 (69) and other previously reported long noncoding RNAs located in this large intergenic region (78). Of note, parasite lines with high (>10%) sexual conversion rates (NF54, NF54-V63, 3D7-Imp, and E5) (34, 58, 66) tended to have prominent heterochromatin at this position, whereas heterochromatin was absent from parasite lines with low (<3%) conversion rates (A7 and E5-PfAP2-G-HA, the latter derived from E5 but with a much lower conversion rate after genome editing and subcloning) (34, 65). The exceptions were the parasite lines 1.2B and 10G, which had prominent heterochromatin at the *gdv1* locus despite not producing gametocytes. However, these subclones have a nonsense mutation that results in a premature stop codon in *gdv1* (Data Set S2) (35) that likely makes the protein nonfunctional and uncouples the regulation of the *gdv1* locus from sexual conversion. Another study also showed that the premature truncation of GDV1 makes the protein nonfunctional (79). The other non-gametocyte-producer line in our study, B11, has a large deletion that spans the *gdv1* coding sequence and the downstream

region (Fig. S1). Together, these results suggest that heterochromatin at the *gdv1* locus plays a role in the regulation of sexual conversion rates, possibly by affecting the expression of *gdv1-as*.

Changes in H3K9ac associated with the transcriptional state of CVGs. We also characterized the distribution of H3K9ac in the 10G, 1.2B, and B11 subclones by ChIP-seq as an exploratory analysis to determine if, in combination with H3K9me3, analysis of this mark may help to identify specific signatures for active and silenced CVGs. H3K9ac is typically associated with the active state of CVGs (45–48) and is also abundant at constitutively euchromatic genes, mainly in intergenic regions (53, 54). There was high overall similarity in the distribution of H3K9ac between subclones (mean pairwise correlation using input-normalized values of 0.92 [range, 0.90 to 0.96]) (Fig. S2A), and the distribution of H3K9ac was largely nonoverlapping with that of H3K9me3, as previously described (51, 54) (Fig. 8A and B). Active CVGs typically had H3K9ac in the region depleted of H3K9me3 (Fig. 8B). We also generated H3K9ac ChIP-seq data for the parasite lines 3D7-Imp, NF54, and E5-PfAP2-G-HA (available in the GEO database), but these data were not further analyzed because transcriptomic analysis was not performed for these parasite lines.

To compare the distribution of H3K9me3 and H3K9ac between the active or silenced CVGs and the active or inactive non-CVGs, we used the classifications of the genes in these different categories for each subclone (Data Set S6). We compared the coverages of H3K9ac and H3K9me3 in the upstream (1,000 bp), coding, and downstream (1,000 bp) regions between genes in the different categories (Fig. 8C). As expected, H3K9me3 was absent from non-CVGs and more abundant in silenced CVGs than in active CVGs. In contrast, the H3K9ac coverage was higher in non-CVGs than in CVGs. Among the CVGs, the H3K9ac coverage was higher in active CVGs than in silenced CVGs, especially in the upstream and coding regions. In these regions, the H3K9ac coverage in active CVGs was almost as high as that in non-CVGs. We also compared the ratios of H3K9ac to H3K9me3, which revealed a pattern similar to that for H3K9ac levels.

To determine if genes in the different categories showed distinct H3K9 signatures, we generated two-dimensional scatterplots with H3K9me3 and H3K9ac coverages at the upstream, coding, and downstream regions (Fig. 8D). Both non-CVGs and silenced CVGs formed compact clusters, whereas the active CVG cluster was more scattered and overlapped the two other clusters. Thus, the global levels of H3K9 epigenetic marks alone cannot be used to fully discriminate between non-CVGs and active CVGs or between active and silenced CVGs. Instead, algorithms that consider the distribution of H3K9 marks associated with the active or silenced state of each specific CVG will be needed to predict the transcriptional state of a CVG from epigenomic data.

DISCUSSION

Changes in the expression of CVGs underlie several important processes in malarial host-parasite interactions, including antigenic variation, alteration of infected RBC cytoadherence tropism and solute permeability, the use of alternative invasion pathways, and sexual conversion (13, 14, 17). Transitions between the heterochromatic and euchromatic states at CVG loci drive expression switches, but the heterochromatin distribution associated with the active and silenced states has been previously characterized for only a few genes (45–48). Here, we systematically characterized the heterochromatin distribution associated with the active and silenced states of over a hundred CVGs, which revealed diverse modalities of heterochromatin changes associated with transcriptional switches. Of note, all of the characterized expression switches and heterochromatin changes occurred during the IDC because all of the subclones were derived from the clonal 3D7 line and had not gone through mosquito passage since the 3D7 line was established. In the majority of the 121 genes for which we compared the distribution of heterochromatin between the active and silenced states, the transitions can be explained by heterochromatin retraction and expansion. Heterochromatin spreading was previously demonstrated to occur in *P. falciparum* (73). Therefore, we propose that CVG expression

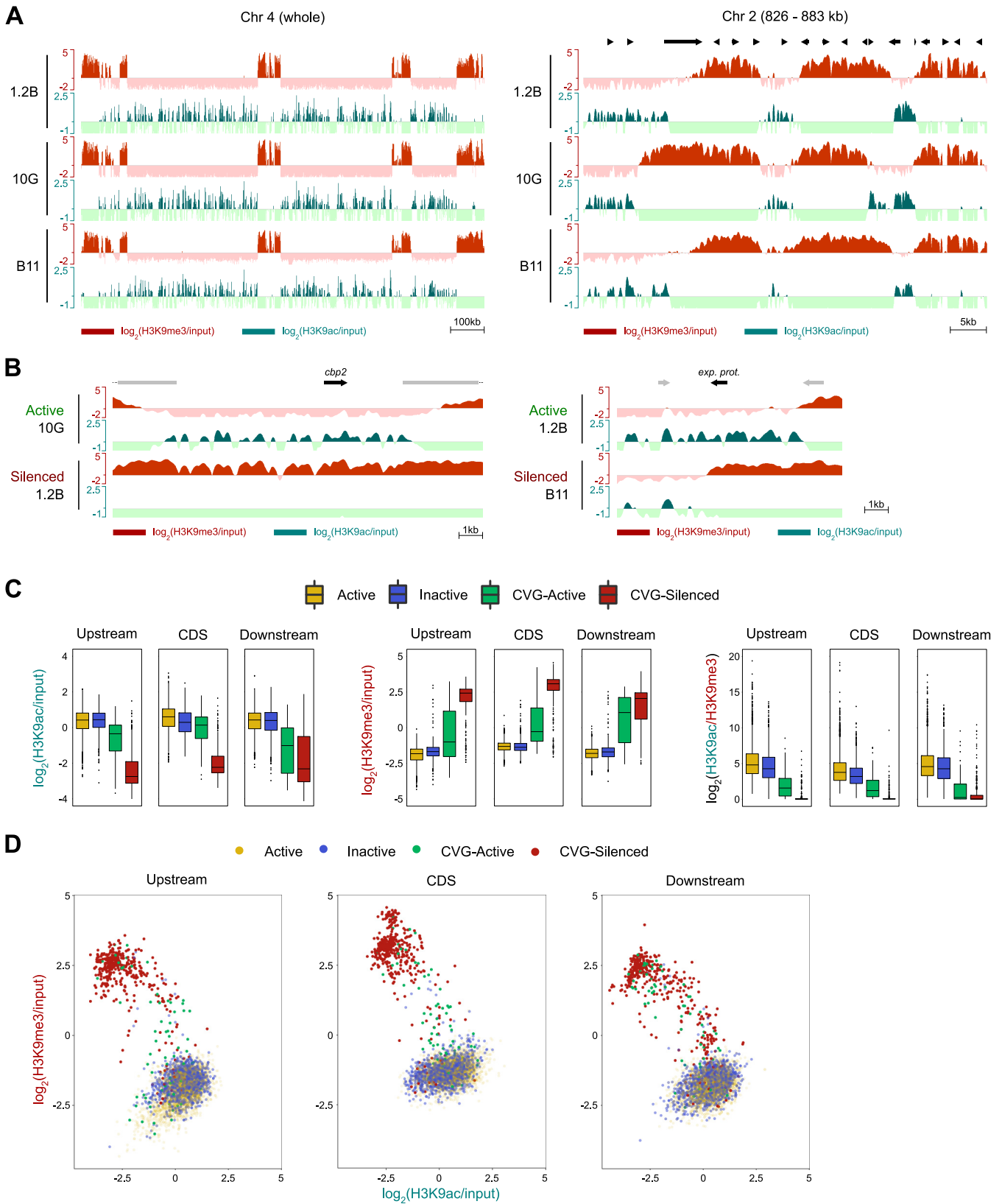


FIG 8 Distribution of H3K9ac and H3K9me3 in CVGs and non-CVGs. (A) General distribution of H3K9ac and comparison with the distribution of H3K9me3 in a representative chromosome or chromosomal region. (B) Representative examples of changes in H3K9me3 and H3K9ac associated with the active or silenced states of CVGs. The identifiers of the genes displayed as black arrows are PF3D7_1301700 (left) and PF3D7_0532600 (right). Other annotated genes are indicated as gray arrows. (C) Mean coverage of H3K9ac and H3K9me3 and H3K9ac/H3K9me3 ratios at the upstream region (1,000 bp), the coding (Continued on next page)

during the IDC is typically regulated by an accordion-like mechanism rather than by the *de novo* formation or the complete removal of heterochromatin domains (Fig. 5L). The latter mechanism appears to be common only for genes that are expressed at other stages of the life cycle (i.e., transmission stages). The significance of heterochromatin at these genes for expression during transmission stages is not known.

A key aspect of the design of our study was the use of subclones that had been cultured for a limited number of generations after subcloning, which ensures that they were largely homogeneous for the transcriptional state of the majority of the CVGs. In parasite lines that have not been recently subcloned, different individual parasites have different combinations of expressed and silenced CVGs (18), and this heterogeneity precludes the detailed characterization of the active and silenced states of the genes. Other important aspects of our study design were comparing subclones of the same genetic background and combining genome-wide epigenomic analysis at high resolution (using small chromatin fragments) with transcriptomic analysis across the full IDC of the same parasite lines.

Because the *P. falciparum* genome is compact, with relatively small intergenic regions, the distribution of heterochromatin in one CVG locus was often influenced by the state of neighboring CVGs. This likely contributes to the complexity and variety of patterns of heterochromatin differences observed between the active and silenced states. For example, in a gene that has a heterochromatic downstream neighbor, the end of the coding sequence is typically heterochromatic even in the active state. Despite this, CVGs are regulated as individual units such that a CVG can be activated while its neighboring CVGs remain silenced. This suggests the existence of boundary elements, such as barrier insulators (80, 81), that prevent the spreading of heterochromatin beyond certain sequence elements. Based on the observation that there are positions beyond which heterochromatin never spreads, neither in the parasite lines analyzed here nor in parasite lines of different genetic backgrounds (52), we speculate that there may be “strong” barrier elements that strictly prevent spreading into euchromatic areas containing essential non-CVGs. Additionally, there are specific positions where heterochromatin spreading commonly stops in different subclones that have a CVG in the active state but not in subclones in which the gene is silenced (heterochromatin domain limits are generally conserved among subclones with a CVG in the same state). This may reflect the presence of “weak” barrier elements that enable the stable activation of a CVG amid heterochromatic regions but can eventually be overcome, resulting in a switch toward the silenced state.

The phenotype of a parasite depends not only on its genome sequence but also on how it uses its CVGs. In natural human malaria infections, thousands of full parasite genomes have been sequenced (82, 83), but characterizing transcriptomes across the full IDC is far more complex (16). The main difficulty in studying *P. falciparum* gene expression in human infections is that only ring-stage parasites and mature gametocytes are found in the circulation, whereas all other blood stages are sequestered in different organs or tissues and therefore are not available in peripheral blood samples. Consequently, the expression of genes such as *var* or *pfap2-g*, which are active in ring stages, can be directly characterized in human blood samples (20, 84–89), but the expression of genes expressed at other stages of the IDC cannot. In studies using genome-wide transcriptome analysis of patient samples, only the expression of genes expressed in rings is captured (16, 84, 90, 91). Some full-transcriptome studies (58, 92, 93) and studies focused on specific CVGs such as *clag* (60) or invasion genes (31, 94, 95), all expressed in late stages of the IDC, used *ex vivo* parasite cultures to obtain samples for transcriptional analysis at different stages of the IDC. However, this approach is

FIG 8 Legend (Continued)

sequence (CDS), and the downstream region (1,000 bp) of genes classified as being (non-CVG) active, (non-CVG) inactive, CVG active, or CVG silenced. All genes from the 1.2B, 10G, and B11 subclones assigned to one of these categories were included in the analysis. Boxes are interquartile ranges (with medians), and whiskers indicate the 5th and 95th percentiles. (D) Scatterplot of H3K9ac versus H3K9me3 coverage for genes classified as being active, inactive, CVG active, or CVG silenced.

time-consuming and requires culture facilities at the place of sample collection, which compromises its feasibility. Together, these difficulties limit our capacity to investigate CVG expression under different clinical presentations, in hosts of different ages, or under different epidemiological conditions.

The distribution of heterochromatin does not change during the full IDC (45–47, 52), implying that H3K9me3 or PfHP1 ChIP-seq analysis performed at a single time point of the IDC informs about the distribution of heterochromatin for CVGs expressed at any asexual blood stage. Since we found that the distribution of heterochromatin is generally conserved among different parasite lines that have a CVG in the same transcriptional state, it is reasonable to expect that it should be possible to predict the transcriptional state of a CVG across the full IDC from the heterochromatin distribution at a single stage (e.g., the ring stage available in peripheral blood samples). However, the influence of the state of neighboring CVGs needs to be taken into consideration because for some genes, this influence results in more than one possible heterochromatin distribution associated with the same transcriptional state of a gene. We envisage that algorithms using machine learning or other approaches can be developed to accurately predict in human infections the expression state of essentially all CVGs from H3K9me3 ChIP-seq analysis of circulating rings. While developing such classification algorithms is clearly beyond the aim of our study, the data set presented here could be used as a seed for this type of algorithm, which later on would be improved with further data sets as they become available. We are aware that chromatin extraction and ChIP-seq are technically more demanding than RNA extraction and microarray or RNA-seq analysis, but this would clearly be compensated for by avoiding the need for *ex vivo* culture and sampling at multiple times to capture the full IDC. Furthermore, new technical developments such as cleavage under targets and tagmentation (CUT&Tag) (96), which uses streamlined protocols and requires a small amount of sample, may facilitate heterochromatin profiling directly from human infections. In addition to H3K9me3, many other histone posttranslational modifications have been described in *P. falciparum* (54, 97–99), but the majority are a consequence, rather than a cause, of the transcriptional state of a gene, or they fluctuate as the parasite progresses along the IDC. In contrast, H3K9me3-based heterochromatin actually determines the state of CVGs, and its distribution is stably maintained across the full IDC over multiple generations, making it a truly epigenetic mark that carries heritable information (100). Therefore, profiling H3K9me3 (or PfHP1) may be sufficient to predict the transcriptional state of most CVGs. Additionally, profiling H3K9ac may also be informative, especially when a CVG is active in only a small fraction of the parasites.

In addition to the global analysis of heterochromatin changes at CVGs, we paid special attention to specific genes with important known functions in host-parasite interactions. We found that in some subclones, several *var* genes were simultaneously in an active chromatin state (and expressed at high levels) in the majority of parasites in the population. These results suggest that the mutually exclusive expression of *var* genes may not be strict, consistent with previous studies that reported the expression of multiple *var* genes in a single parasite after selection for binding to multiple receptors (74) or the expression of several *var* genes in some recently subcloned lines (75). As an exploratory analysis, we also investigated epigenomic traits that may correlate with sexual conversion rates, a phenotype for which the molecular basis of variation has not been established. This resulted in the identification of a heterochromatin domain at the *gdv1* locus as a candidate regulator of sexual conversion rates. Heterochromatin was present at the putative promoter of the *gdv1-as* long noncoding RNA in parasite lines with high sexual conversion rates or lacking functional GDV1, but not in parasite lines with low sexual conversion rates and an intact *gdv1* gene. This result suggests that heterochromatin at this locus may prevent the expression of *gdv1-as*, which may enable the expression of GDV1 (69) and the subsequent activation of the master regulator of sexual conversion, *pfap2-g*. However, further studies will be needed to test this hypothesis, including analyses under defined conditions that induce or repress sexual conversion (9, 101) and manipulation of the locus using transfection approaches.

MATERIALS AND METHODS

Parasites. The 3D7-A and 3D7-B stocks of the *P. falciparum* clonal line 3D7; the 3D7-A subclones 10G and 1.2B; the 3D7-B subclones A7, E5, and B11; and the transgenic E5-PfAP2-G-HA line (9A subclone) were previously described and characterized (18, 26, 34, 61, 62). The 3D7-Imp line (66, 102) is the stock of 3D7 (103) maintained at Imperial College. The NF54 line (103) was from the stock maintained at Sanaria (104), and the NF54-V63 line was obtained after infecting a volunteer with Sanaria cryopreserved NF54 sporozoites in a CHMI trial (58, 105). Cultures for all of the new experiments presented in this study (and for the previous transcriptomic analysis of 1.2B, 10G, and 3D7-B [18]) were maintained at 3% hematocrit with B⁺ RBCs in standard parasite culture medium with Albumax II and no human serum in an atmosphere containing 5% CO₂ and 3% O₂, balanced with N₂. However, the NF54-V63 line (for which ChIP-seq data generated as part of a previous study are included in some analyses) was cultured in medium containing human serum (58). Sexual conversion rates were determined using *N*-acetyl-D-glucosamine treatment and light microscopy, as previously described (65).

Transcriptomic experiments. Samples for transcriptomic analysis of the A7, E5, and B11 lines were obtained from tightly synchronized cultures (5-h age window) obtained by Percoll purification of RBCs infected with mature stages followed by 5% sorbitol lysis 5 h later. RNA was extracted from samples collected at 10 to 15, 20 to 25, 30 to 35, 37 to 42, 40 to 45, and 43 to 48 h postinvasion (hpi) using the TRIzol method, as previously described (18, 106). After reverse transcription and labeling, Cy5-labeled samples were hybridized against a Cy3-labeled reference pool, as previously described (68). The reference pool consisted of a mixture of equal amounts of cDNA from rings, trophozoites, and schizonts from 3D7-A and 3D7-B, similar to the reference pool used for the analysis of 1.2B, 10G, and 3D7-B but from a different biological preparation (18). Samples were hybridized on previously described custom Agilent microarrays (AMADID no. 085763) (35) modified from the AMADID no. 037237 design (68).

ChIP-seq experiments. Samples for ChIP-seq analysis were obtained from sorbitol-synchronized cultures at the late trophozoite/early schizont stage. Since the heterochromatin distribution is stable during the IDC, we did not use cultures tightly synchronized to a 5-h age window for these experiments. Chromatin extraction and immunoprecipitation were performed essentially as previously described (35, 58). In brief, after saponin lysis and formaldehyde cross-linking, chromatin was extracted using the MAGnify chromatin immunoprecipitation system (Life Technologies). Chromatin was sonicated using a Covaris M220 sonicator to obtain fragments of ~150 to 200 bp. For immunoprecipitation, we used 4 μ g of chromatin and 8 μ g of antibodies against H3K9me3 (catalog no. C15410193; Diagenode) or H3K9ac (catalog no. C15410004; Diagenode). For library preparation, we used 4 ng of input or immunoprecipitated DNA, the NEBNext multiplex oligonucleotides for Illumina (New England BioLabs [NEB]), 8 to 10 cycles of amplification with the Kapa HiFi PCR kit (Kapa Biosystems), and AMPure XP beads for purification steps (35, 107). Sequencing was performed using an Illumina HiSeq 2500 system, obtaining 8 million to 39 million 125-bp paired-end reads per sample.

Microarray data analysis. For the 1.2B, 10G, and 3D7-B lines, we used previously published processed expression values (18). For the A7, E5, and B11 subclones, microarray data were analyzed similarly, but normalized Cy5 and Cy3 values were obtained directly using Agilent Feature Extraction software. The rest of the analysis was performed using mainly the Bioconductor and the Tidyverse suite libraries in an R environment (v.3.6.3), essentially as previously described (58). In brief, for each channel, the background signal intensity was calculated as the median signal of the 100 lowest-signal probes in each array. Probes with a signal for both channels below three times the background were excluded from further analysis. For each probe, the expression value was defined as the log₂(Cy5/Cy3) signal. Probes were collapsed into genes using median polish. For each array, the estimated parasite age in hours postinvasion was calculated using a previously described maximum likelihood method (67). For each gene and subclone, expression plots were generated with gene expression values on the y axis and estimated hours postinvasion on the x axis. These plots were used to calculate the average fold change (AFC) for each gene over four overlapping time intervals of half the estimated parasite age difference (in hours postinvasion) between the first and last analysis time points. The AFC was calculated for all possible pairwise comparisons among A7, E5, and B11 according to a previously described pipeline (18). The maximum AFC (mAFC) was defined as the AFC at the time interval in which it had the highest absolute value.

From this point, data from 1.2B, 10G, 3D7-B, A7, E5, and B11 were analyzed in parallel, although comparisons were performed separately for 1.2B, 10G, and 3D7-B or for A7, E5, and B11 (because these two groups of samples were analyzed with slightly different microarray platforms and reference pools). Genes with an mAFC of >4 (in at least one of the pairwise comparisons among 1.2B, 10G, and 3D7-B or among A7, E5, and B11) were considered to have markedly different transcript levels. To identify differentially expressed genes with high confidence, we applied several filters, essentially as previously described (58). For each pairwise comparison, only genes that had an mAFC of >4 and passed all of the filters were included in the final list of differentially expressed genes. First, tRNAs and genes lacking a PlasmoDB identification were excluded. Second, genes with very low absolute expression values were excluded. For this, genes with a Cy5 signal within the lowest 15% (at all time points) in the parasite lines with an mAFC of >4 in the pairwise comparison were filtered out. Third, genes that showed large expression differences only at the time periods in which they were expressed at low levels were excluded. For this, genes that at the time interval of maximum expression did not have an AFC of >4 were excluded. Finally, genes (as annotated in PlasmoDB v52, including the 5' UTR and 3' UTR when available) that overlapped a duplication or a deletion (see below) in at least one of the parasite lines among which the expression difference was observed were also excluded. To classify genes as previously reported CVGs, we used a previously described list of CVGs (58), whereas for the identification of gametocyte markers (genes with higher expression levels in sexual than in asexual parasites), we used a

list based on previously reported transcriptomic studies of gametocytes (see Data Set S7 in the supplemental material).

For some analyses of differentially expressed genes, we focused on the pairwise comparison with the largest fold difference in expression to identify representative subclones for the silenced and active states. For this, for each gene, we selected the pairwise comparison that passed all filters and had the largest mAFc among the comparisons of 1.2B versus 10G, 1.2B versus 3D7-B, 10G versus 3D7-B, A7 versus E5, A7 versus B11, and E5 versus B11. When the largest mAFc occurred between 3D7-B (which is not a transcriptionally homogeneous subclone and for which we did not generate ChIP-seq data) and 1.2B or 10G, we first determined which 3D7-B subclone (A7, E5, or B11) had the maximum (when the gene was active in 3D7-B and silenced in 10G or 1.2B) or minimum (when the gene was silenced in 3D7-B and active in 10G or 1.2B) expression value and selected this subclone as the representative subclone for the active or silenced state of the gene.

ChIP-seq data analysis. ChIP-seq data analysis, including the generation of plots, was performed using custom Python (v.3.6.9) and R (v.3.6.3) scripts, the latter mainly using Tidyverse suite packages. The initial processing of the data was performed approximately as described previously (35). After a quality check using FastQC (v.0.11.9), reads were trimmed, and repetitive *k*-mers were removed using BBDDUK (v.36.99) with the parameters `ktrim = r`, `k = 22`, and `mink = 6`. Reads were aligned to the reference genome (*P. falciparum* 3D7 [PlasmoDB v28]) using Bowtie2 (v.2.3.0), and duplicate reads were excluded using the PICARD suite (v.4.1.9.0) with the `RemoveDuplicates` command. Peak calling was performed using MACS2 (v.2.2.7.1) with the parameters `-f BAMPE -B -g 2.41e7 --keep-dup all --fe-cutoff 1.5 --nomodel --extsize 150`. To calculate the normalized coverage of enrichment over the input (\log_2 transformed) at a 10-bp resolution, we used the DeepTools (v.3.5.0) `BamCompare` command with the parameters `--normalizeUsing RPKM -bs 10 --smoothLength 200 --pseudocount 10`. After setting negative coverage values to zero to prevent the detection of coverage differences in nonheterochromatic regions, the coverage for landmark regions (e.g., upstream, coding, or downstream sequences of genes) was calculated as the mean value for the 10-bp bins within the region using BedTools (v.2.26.0). Custom scripts were used to create BED files demarcating the regions of interest for each gene for each particular analysis.

To identify differential H3K9me3 peaks in all possible pairwise comparisons among subclones 10G, 1.2B, E5, A7, and B11, we used custom scripts. First, the input-normalized coverage was calculated for 100-bp bins in all regions overlapping a MACS2 peak in at least one of the parasite lines of the pairwise comparison. The distribution of coverage intensities was adjusted to a skewed normal distribution, and for each bin, we calculated a cumulative density function (CDF) value for each parasite line. Bins with a CDF difference of >0.3 were considered to have differential coverage. Next, bins with differential coverage separated by <500 bp were merged, and differential peaks of $<1,000$ bp after merging were filtered out.

Identification of genetic changes in the parasite lines used in this study. To identify genetic polymorphisms, we analyzed the sequences from the input samples of the ChIP-seq experiments. To identify single nucleotide polymorphisms (SNPs) and short indels, we used the variant-calling protocol in the GATK suite according to GATK best practices (Data Set S2). Mutations with a coverage of <20 reads were excluded. Since recently subcloned lines are genetically highly homogeneous, a very high prevalence is expected for real mutations; therefore, we retained only mutations (relative to the reference 3D7 genome in PlasmoDB v52) in which the frequency of the reference allele was ≤ 0.5 . Mutations were annotated using the ENSEMBL variant effect predictor (VEP). To define regions with relatively large (>500 -bp) duplications or deletions, we used custom scripts. We first calculated coverage values for the inputs of the ChIP-seq data using the DeepTools (v.3.5.0) `BamCoverage` command with the parameters `--normalizeUsing RPKM -bs 10 --smoothLength 200`, similar to the calculation of the coverage for input-normalized ChIP-seq values. For each sample, the mean coverage over the whole genome was calculated. Bins with a coverage value <0.05 times the mean coverage were marked as underrepresented, and bins with a coverage value >1.75 times the mean coverage were marked as overrepresented. Bins marked as underrepresented or overrepresented that were separated by <200 bp were joined, and the resulting fragments were discarded if they were <500 bp. Finally, to avoid marking as deletion regions of low coverage due to the underlying sequence, we removed all regions with putative deletions that were common to all of the subclones analyzed. The remaining fragments were marked as deletions or duplications.

Classification of genes according to their transcriptional state. In order to assign a transcriptional state to each gene in subclones 1.2B, 10G, A7, E5, and B11, we designed a tree-like classification system. The final possible categories were active, inactive, undetermined, CVG active, CVG silenced, and CVG undetermined. The classification system is illustrated in Fig. S7, and the state of each gene is provided in Data Set S6.

The main parameters used for the classification of genes were: (i) previous classification as a CVG on the basis of variant expression or the presence of heterochromatin marks, according to a previously published list of CVGs (58); (ii) differential expression among the five parasite subclones in the data set presented here, with an mAFc of >2 at the time of maximum expression; (iii) Cy5 (sample channel) value percentiles in our data set; and (iv) expression level percentiles in published RNA-seq data sets (53, 107–109). In brief, genes were first classified as CVGs or non-CVGs. Next, regardless of whether a gene was a CVG or not, genes that showed no expression differences among the subclones were classified as active or inactive/silenced according to their maximum expression percentiles, whereas genes with expression differences were classified as active or inactive/silenced according to their relative expression levels. Genes with intermediate expression levels, duplications, or deletions were classified as undetermined or CVG undetermined.

Availability of data. The new microarray and ChIP-seq data described in this article have been deposited in the GEO database with accession no. [GSE208561](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE208561). ChIP-seq data are also available for visualization

at the UCSC genome browser (https://genome.ucsc.edu/s/lucas_michel_todo/3D7A%20and%203D7B%20subclones%20H3K9me3%20and%20H3K9ac and https://genome.ucsc.edu/s/lucas_michel_todo/NF54_3D7imp_E5HA). All of the scripts used for the data analysis are available at GitHub (https://github.com/LucasMichelTodo/Heterochromatin_Transitions.git).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 2.4 MB.

SUPPLEMENTAL FILE 2, XLSX file, 1.1 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.2 MB.

SUPPLEMENTAL FILE 4, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 5, XLSX file, 0.6 MB.

SUPPLEMENTAL FILE 6, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 7, XLSX file, 0.4 MB.

SUPPLEMENTAL FILE 8, XLSX file, 0.02 MB.

ACKNOWLEDGMENTS

We are grateful to Mariona Bustamante (ISGlobal) for help in the acquisition of funding, support in setting up the project, and critical reading of the manuscript; S. Hoffman (Sanaria) for providing the NF54 line; and M. Delves (Imperial College) for providing the 3D7-Imp line.

This work was supported by grants from the Spanish Ministry of Science and Innovation (MCI)/Agencia Estatal de Investigación (AEI) (SAF2016-76190-R and PID2019-107232RBI00/AEI/10.13039/501100011033 to A.C.), cofunded by the European Regional Development Fund (ERDF) (European Union), the la Caixa Banking Foundation (HR18-00267 to A.C.), and the ISGlobal Alliance pilot project program (to C.B., J.R.G., and A.C.). L.M.-T. is supported by a fellowship from the Spanish Ministry of Economy and Competitiveness (BES-2017-081079), cofunded by the European Social Fund (ESF). Our research is part of ISGlobal's Program on the Molecular Mechanisms of Malaria, which is partially supported by the Fundación Ramón Areces. We acknowledge support from the Spanish Ministry of Science and Innovation through the Centro de Excelencia Severo Ochoa 2019–2023 Program (CEX2018-000806-S) and support from the Generalitat de Catalunya through the CERCA Program.

A.C. and C.B. designed the experiments. C.B. optimized and performed ChIP-seq experiments, and N.C.-V. and N.R.-G. performed transcriptomic experiments. L.M.-T. performed the bioinformatic analysis of the data and generated the figures. C.H.-F. and J.R.G. performed preliminary analyses of the initial data. L.M.-T. and A.C. interpreted the results and wrote the manuscript, with input from all authors. A.C. conceived the project. C.B., J.R.G., and A.C. obtained funding.

We declare that we have no competing interests.

REFERENCES

1. Venugopal K, Hentzschel F, Valkiūnas G, Marti M. 2020. *Plasmodium* asexual growth and sexual development in the haematopoietic niche of the host. *Nat Rev Microbiol* 18:177–189. <https://doi.org/10.1038/s41579-019-0306-2>.
2. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. 2003. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol* 1:E5. <https://doi.org/10.1371/journal.pbio.0000005>.
3. Howick VM, Russell AJC, Andrews T, Heaton H, Reid AJ, Natarajan K, Butungi H, Metcalf T, Verzier LH, Rayner JC, Berriman M, Herren JK, Billker O, Hemberg M, Talman AM, Lawniczak MKN. 2019. The Malaria Cell Atlas: single parasite transcriptomes across the complete *Plasmodium* life cycle. *Science* 365:eaaw2619. <https://doi.org/10.1126/science.aaw2619>.
4. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ, Winzeler EA. 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 301:1503–1508. <https://doi.org/10.1126/science.1087025>.
5. Cowman AF, Healer J, Marapana D, Marsh K. 2016. Malaria: biology and disease. *Cell* 167:610–624. <https://doi.org/10.1016/j.cell.2016.07.055>.
6. Kumar M, Skillman K, Duraisingh MT. 2021. Linking nutrient sensing and gene expression in *Plasmodium falciparum* blood-stage parasites. *Mol Microbiol* 115:891–900. <https://doi.org/10.1111/mmi.14652>.
7. Zuzarte-Luis V, Mota MM. 2018. Parasite sensing of host nutrients and environmental cues. *Cell Host Microbe* 23:749–758. <https://doi.org/10.1016/j.chom.2018.05.018>.
8. Oakley MS, Gerald N, McCutchan TF, Aravind L, Kumar S. 2011. Clinical and molecular aspects of malaria fever. *Trends Parasitol* 27:442–449. <https://doi.org/10.1016/j.pt.2011.06.004>.
9. Brancucci NMB, Gerdt JP, Wang C, De Niz M, Philip N, Adapa SR, Zhang M, Hitz E, Niederwieser I, Boltryk SD, Laffitte MC, Clark MA, Gruring C, Ravel D, Blancke Soares A, Demas A, Bopp S, Rubio-Ruiz B, Conejo-García A, Wirth DF, Gendaszewska-Darmach E, Duraisingh MT, Adams JH, Voss TS, Waters AP, Jiang RHY, Clardy J, Marti M. 2017. Lysophosphatidylcholine regulates sexual stage differentiation in the human malaria parasite *Plasmodium falciparum*. *Cell* 171:1532–1544.e15. <https://doi.org/10.1016/j.cell.2017.10.020>.
10. Mancio-Silva L, Slavic K, Grilo Ruivo MT, Grosso AR, Modrzynska KK, Vera IM, Sales-Dias J, Gomes AR, MacPherson CR, Crozet P, Adamo M, Baena-Gonzalez E, Tewari R, Llinas M, Billker O, Mota MM. 2017. Nutrient sensing modulates

- malaria parasite virulence. *Nature* 547:213–216. <https://doi.org/10.1038/nature23009>.
11. Tintó-Font E, Michel-Todó L, Russell TJ, Casas-Vila N, Conway DJ, Bozdech Z, Llinás M, Cortés A. 2021. A heat-shock response regulated by the PfAP2-HS transcription factor protects human malaria parasites from febrile temperatures. *Nat Microbiol* 6:1163–1174. <https://doi.org/10.1038/s41564-021-00940-w>.
 12. Tintó-Font E, Cortés A. 2022. Malaria parasites do respond to heat. *Trends Parasitol* 38:435–449. <https://doi.org/10.1016/j.pt.2022.02.009>.
 13. Cortés A, Deitsch KW. 2017. Malaria epigenetics. *Cold Spring Harb Perspect Med* 7:a025528. <https://doi.org/10.1101/cshperspect.a025528>.
 14. Duraisingh MT, Skillman KM. 2018. Epigenetic variation and regulation in malaria parasites. *Annu Rev Microbiol* 72:355–375. <https://doi.org/10.1146/annurev-micro-090817-062722>.
 15. Guizetti J, Scherf A. 2013. Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell Microbiol* 15: 718–726. <https://doi.org/10.1111/cmi.12115>.
 16. Llorà-Battle O, Tinto-Font E, Cortes A. 2019. Transcriptional variation in malaria parasites: why and how. *Brief Funct Genomics* 18:329–341. <https://doi.org/10.1093/bfpg/elz009>.
 17. Voss TS, Bozdech Z, Bartfai R. 2014. Epigenetic memory takes center stage in the survival strategy of malaria parasites. *Curr Opin Microbiol* 20:88–95. <https://doi.org/10.1016/j.mib.2014.05.007>.
 18. Rovira-Graells N, Gupta AP, Planet E, Crowley VM, Mok S, Ribas de Pouplana L, Preiser PR, Bozdech Z, Cortés A. 2012. Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res* 22:925–938. <https://doi.org/10.1101/gr.129692.111>.
 19. Scherf A, Lopez-Rubio JJ, Riviere L. 2008. Antigenic variation in *Plasmodium falciparum*. *Annu Rev Microbiol* 62:445–470. <https://doi.org/10.1146/annurev.micro.61.080706.093134>.
 20. Smith JD, Rowe JA, Higgins MK, Lavstsen T. 2013. Malaria's deadly grip: cytoadhesion of *Plasmodium falciparum*-infected erythrocytes. *Cell Microbiol* 15:1976–1983. <https://doi.org/10.1111/cmi.12183>.
 21. Wahlgren M, Goel S, Akhouri RR. 2017. Variant surface antigens of *Plasmodium falciparum* and their roles in severe malaria. *Nat Rev Microbiol* 15:479–491. <https://doi.org/10.1038/nrmicro.2017.47>.
 22. Sargeant TJ, Marti M, Caler E, Carlton JM, Simpson K, Speed TP, Cowman AF. 2006. Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol* 7:R12. <https://doi.org/10.1186/gb-2006-7-2-r12>.
 23. Templeton TJ. 2009. The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, *Plasmodium falciparum*. *Mol Biochem Parasitol* 166:109–116. <https://doi.org/10.1016/j.molbiopara.2009.04.003>.
 24. Wright GJ, Rayner JC. 2014. *Plasmodium falciparum* erythrocyte invasion: combining function with immune evasion. *PLoS Pathog* 10:e1003943. <https://doi.org/10.1371/journal.ppat.1003943>.
 25. Bethke LL, Zilversmit M, Nielsen K, Daily J, Volkman SK, Ndiaye D, Lozovsky ER, Hartl DL, Wirth DF. 2006. Duplication, gene conversion, and genetic diversity in the species-specific acyl-CoA synthetase gene family of *Plasmodium falciparum*. *Mol Biochem Parasitol* 150:10–24. <https://doi.org/10.1016/j.molbiopara.2006.06.004>.
 26. Cortés A, Carret C, Kaneko O, Yim Lim BYS, Ivens A, Holder AA. 2007. Epigenetic silencing of *Plasmodium falciparum* genes linked to erythrocyte invasion. *PLoS Pathog* 3:e107. <https://doi.org/10.1371/journal.ppat.0030107>.
 27. Nguitragool W, Bokhari AA, Pillai AD, Rayavara K, Sharma P, Turpin B, Aravind L, Desai SA. 2011. Malaria parasite *clag3* genes determine channel-mediated nutrient uptake by infected red blood cells. *Cell* 145:665–677. <https://doi.org/10.1016/j.cell.2011.05.002>.
 28. Pillai AD, Nguitragool W, Lyko B, Dolinta K, Butler MM, Nguyen ST, Peet NP, Bowlin TL, Desai SA. 2012. Solute restriction reveals an essential role for *clag3*-associated channels in malaria parasite nutrient acquisition. *Mol Pharmacol* 82:1104–1114. <https://doi.org/10.1124/mol.112.081224>.
 29. Mira-Martínez S, Rovira-Graells N, Crowley VM, Altenhofen LM, Llinás M, Cortés A. 2013. Epigenetic switches in *clag3* genes mediate blasticidin S resistance in malaria parasites. *Cell Microbiol* 15:1913–1923. <https://doi.org/10.1111/cmi.12162>.
 30. Lavazec C, Sanyal S, Templeton TJ. 2007. Expression switching in the *stevor* and *Pfmc-2TM* superfamilies in *Plasmodium falciparum*. *Mol Microbiol* 64:1621–1634. <https://doi.org/10.1111/j.1365-2958.2007.05767.x>.
 31. Cortés A. 2008. Switching *Plasmodium falciparum* genes on and off for erythrocyte invasion. *Trends Parasitol* 24:517–524. <https://doi.org/10.1016/j.pt.2008.08.005>.
 32. Duraisingh MT, Triglia T, Ralph SA, Rayner JC, Barnwell JW, McFadden GI, Cowman AF. 2003. Phenotypic variation of *Plasmodium falciparum* merozoite proteins directs receptor targeting for invasion of human erythrocytes. *EMBO J* 22:1047–1057. <https://doi.org/10.1093/emboj/cdg096>.
 33. Petter M, Haeggstrom M, Khattab A, Fernandez V, Klinkert MQ, Wahlgren M. 2007. Variant proteins of the *Plasmodium falciparum* RIFIN family show distinct subcellular localization and developmental expression patterns. *Mol Biochem Parasitol* 156:51–61. <https://doi.org/10.1016/j.molbiopara.2007.07.011>.
 34. Kaf sack BF, Rovira-Graells N, Clark TG, Bancells C, Crowley VM, Campino SG, Williams AE, Drought LG, Kwiatkowski DP, Baker DA, Cortes A, Llinas M. 2014. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* 507:248–252. <https://doi.org/10.1038/nature12920>.
 35. Llorà-Battle O, Michel-Todó L, Witmer K, Toda H, Fernández-Becerra C, Baum J, Cortés A. 2020. Conditional expression of PfAP2-G for controlled massive sexual conversion in *Plasmodium falciparum*. *Sci Adv* 6:eaz5057. <https://doi.org/10.1126/sciadv.aaz5057>.
 36. Poran A, Notzel C, Aly O, Mencia-Trinchant N, Harris CT, Guzman ML, Hassane DC, Elemento O, Kaf sack BFC. 2017. Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature* 551:95–99. <https://doi.org/10.1038/nature24280>.
 37. Josling GA, Williamson KC, Llinas M. 2018. Regulation of sexual commitment and gametocytogenesis in malaria parasites. *Annu Rev Microbiol* 72:501–519. <https://doi.org/10.1146/annurev-micro-090817-062712>.
 38. Josling GA, Russell TJ, Venezia J, Orchard L, van Biljon R, Painter HJ, Llinas M. 2020. Dissecting the role of PfAP2-G in malaria gametocytogenesis. *Nat Commun* 11:1503. <https://doi.org/10.1038/s41467-020-15026-0>.
 39. Brancucci NM, Bertschi NL, Zhu L, Niederwieser I, Chin WH, Wampfler R, Freymond C, Rottmann M, Felger I, Bozdech Z, Voss TS. 2014. Heterochromatin protein 1 secures survival and transmission of malaria parasites. *Cell Host Microbe* 16:165–176. <https://doi.org/10.1016/j.chom.2014.07.004>.
 40. Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, Alako BT, Ehlgren F, Ralph SA, Cowman AF, Bozdech Z, Stunnenberg HG, Voss TS. 2009. *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog* 5:e1000569. <https://doi.org/10.1371/journal.ppat.1000569>.
 41. Perez-Toledo K, Rojas-Meza AP, Mancio-Silva L, Hernandez-Cuevas NA, Delgado DM, Vargas M, Martinez-Calvillo S, Scherf A, Hernandez-Rivas R. 2009. *Plasmodium falciparum* heterochromatin protein 1 binds to trimethylated histone 3 lysine 9 and is linked to mutually exclusive expression of *var* genes. *Nucleic Acids Res* 37:2596–2606. <https://doi.org/10.1093/nar/gkp115>.
 42. Lopez-Rubio JJ, Mancio-Silva L, Scherf A. 2009. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* 5:179–190. <https://doi.org/10.1016/j.chom.2008.12.012>.
 43. Carrington E, Cooijmans RHM, Keller D, Toenhake CG, Bártfai R, Voss TS. 2021. The ApiAP2 factor PfAP2-HC is an integral component of heterochromatin in the malaria parasite *Plasmodium falciparum*. *iScience* 24: 102444. <https://doi.org/10.1016/j.isci.2021.102444>.
 44. Shang X, Wang C, Fan Y, Guo G, Wang F, Zhao Y, Sheng F, Tang J, He X, Yu X, Zhang M, Zhu G, Yin S, Mu J, Culleton R, Cao J, Jiang M, Zhang Q. 2022. Genome-wide landscape of ApiAP2 transcription factors reveals a heterochromatin-associated regulatory network during *Plasmodium falciparum* blood-stage development. *Nucleic Acids Res* 50:3413–3431. <https://doi.org/10.1093/nar/gkac176>.
 45. Crowley VM, Rovira-Graells N, de Pouplana LR, Cortés A. 2011. 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. <https://doi.org/10.1111/j.1365-2958.2011.07574.x>.
 46. Jiang L, Lopez-Barragan MJ, Jiang H, Mu J, Gaur D, Zhao K, Felsenfeld G, Miller LH. 2010. Epigenetic control of the variable expression of a *Plasmodium falciparum* receptor protein for erythrocyte invasion. *Proc Natl Acad Sci U S A* 107:2224–2229. <https://doi.org/10.1073/pnas.0913396107>.
 47. Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R, Scherf A. 2007. 5' flanking region of *var* genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol Microbiol* 66:1296–1305. <https://doi.org/10.1111/j.1365-2958.2007.06009.x>.
 48. Cabral FJ, Fotoran WL, Wunderlich G. 2012. Dynamic activation and repression of the *Plasmodium falciparum rif* gene family and their relation to

- chromatin modification. *PLoS One* 7:e29881. <https://doi.org/10.1371/journal.pone.0029881>.
49. Allshire RC, Madhani HD. 2018. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol* 19:229–244. <https://doi.org/10.1038/nrm.2017.119>.
 50. Nicetto D, Zaret KS. 2019. Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. *Curr Opin Genet Dev* 55:1–10. <https://doi.org/10.1016/j.gde.2019.04.013>.
 51. Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AM, Cohen AM, Janssen-Megens EM, van de Vegte-Bolmer M, Selzer RR, Iniguez AL, Green RD, Sauerwein RW, Jensen ON, Stunnenberg HG. 2009. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 106:9655–9660. <https://doi.org/10.1073/pnas.0902515106>.
 52. Fraschka SA, Filarsky M, Hoo R, Niederwieser I, Yam XY, Brancucci NMB, Mohring F, Mushunje AT, Huang X, Christensen PR, Nosten F, Bozdech Z, Russell B, Moon RW, Marti M, Preiser PR, Bartfai R, Voss TS. 2018. Comparative heterochromatin profiling reveals conserved and unique epigenome signatures linked to adaptation and development of malaria parasites. *Cell Host Microbe* 23:407–420.e8. <https://doi.org/10.1016/j.chom.2018.01.008>.
 53. Bartfai R, Hoeijmakers WA, Salcedo-Amaya AM, Smits AH, Janssen-Megens E, Kaan A, Treeck M, Gilberger TW, Francoijs KJ, Stunnenberg HG. 2010. H2A.Z demarcates intergenic regions of the *Plasmodium falciparum* epigenome that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog* 6:e1001223. <https://doi.org/10.1371/journal.ppat.1001223>.
 54. Karmodiya K, Pradhan SJ, Joshi B, Jangid R, Reddy PC, Galande S. 2015. A comprehensive epigenome map of *Plasmodium falciparum* reveals unique mechanisms of transcriptional regulation and identifies H3K36me2 as a global mark of gene suppression. *Epigenetics Chromatin* 8:32. <https://doi.org/10.1186/s13072-015-0029-1>.
 55. Zanghi G, Vembar SS, Baumgarten S, Ding S, Guizetti J, Bryant JM, Mattei D, Jensen ATR, Rénia L, Goh YS, Sauerwein R, Hermsen CC, Franetich JF, Bordessoulles M, Silvie O, Soulard V, Scatton O, Chen P, Mecheri S, Mazier D, Scherf A. 2018. A specific PfEMP1 is expressed in *P. falciparum* sporozoites and plays a role in hepatocyte infection. *Cell Rep* 22:2951–2963. <https://doi.org/10.1016/j.celrep.2018.02.075>.
 56. Gomez-Diaz E, Yerbanga RS, Lefevre T, Cohuet A, Rowley MJ, Ouedraogo JB, Corces VG. 2017. Epigenetic regulation of *Plasmodium falciparum* clonally variant gene expression during development in *Anopheles gambiae*. *Sci Rep* 7:40655. <https://doi.org/10.1038/srep40655>.
 57. Bunnik EM, Cook KB, Varoquaux N, Batugedara G, Prudhomme J, Cort A, Shi L, Andolina C, Ross LS, Brady D, Fidock DA, Nosten F, Tewari R, Sinnis P, Ay F, Vert J-P, Noble WS, Le Roch KG. 2018. Changes in genome organization of parasite-specific gene families during the *Plasmodium* transmission stages. *Nat Commun* 9:1910. <https://doi.org/10.1038/s41467-018-04295-5>.
 58. Pickford AK, Michel-Todó L, Dupuy F, Mayor A, Alonso PL, Lavazec C, Cortés A. 2021. Expression patterns of *Plasmodium falciparum* clonally variant genes at the onset of a blood infection in malaria-naïve humans. *mBio* 12:e01636-21. <https://doi.org/10.1128/mBio.01636-21>.
 59. Bachmann A, Petter M, Krumkamp R, Esen M, Held J, Scholz JA, Li T, Sim BK, Hoffman SL, Kremsner PG, Mordmüller B, Duffy MF, Tannich E. 2016. Mosquito passage dramatically changes *var* gene expression in controlled human *Plasmodium falciparum* infections. *PLoS Pathog* 12:e1005538. <https://doi.org/10.1371/journal.ppat.1005538>.
 60. Mira-Martínez S, van Schuppen E, Amambua-Ngwa A, Bottieau E, Affara M, Van Esbroeck M, Vlieghe E, Guetens P, Rovira-Graells N, Gómez-Pérez GP, Alonso PL, D'Alessandro U, Rosanas-Urgell A, Cortés A. 2017. Expression of the *Plasmodium falciparum* clonally variant *clag3* genes in human infections. *J Infect Dis* 215:938–945. <https://doi.org/10.1093/infdis/jix053>.
 61. Cortés A. 2005. A chimeric *Plasmodium falciparum* *Pfnbp2b/Pfnbp2a* gene originated during asexual growth. *Int J Parasitol* 35:125–130. <https://doi.org/10.1016/j.ijpara.2004.11.004>.
 62. Cortés A, Benet A, Cooke BM, Barnwell JW, Reeder JC. 2004. Ability of *Plasmodium falciparum* to invade Southeast Asian ovalocytes varies between parasite lines. *Blood* 104:2961–2966. <https://doi.org/10.1182/blood-2004-06-2136>.
 63. Cortés A, Mellombo M, Mgone CS, Beck HP, Reeder JC, Cooke BM. 2005. Adhesion of *Plasmodium falciparum*-infected red blood cells to CD36 under flow is enhanced by the cerebral malaria-protective trait Southeast Asian ovalocytosis. *Mol Biochem Parasitol* 142:252–257. <https://doi.org/10.1016/j.molbiopara.2005.03.016>.
 64. Rovira-Graells N, Aguilera-Simon S, Tinto-Font E, Cortes A. 2016. New assays to characterise growth-related phenotypes of *Plasmodium falciparum* reveal variation in density-dependent growth inhibition between parasite lines. *PLoS One* 11:e0165358. <https://doi.org/10.1371/journal.pone.0165358>.
 65. Bancellis C, Llorca-Batlle O, Poran A, Notzel C, Rovira-Graells N, Elemento O, Kafsack BFC, Cortes A. 2019. Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*. *Nat Microbiol* 4:144–154. <https://doi.org/10.1038/s41564-018-0291-7>.
 66. Delves MJ, Straschil U, Ruecker A, Miguel-Blanco C, Marques S, Dufour AC, Baum J, Sinden RE. 2016. Routine in vitro culture of *P. falciparum* gametocytes to evaluate novel transmission-blocking interventions. *Nat Protoc* 11:1668–1680. <https://doi.org/10.1038/nprot.2016.096>.
 67. Lemieux JE, Gomez-Escobar N, Feller A, Carret C, Amambua-Ngwa A, Pinches R, Day F, Kyes SA, Conway DJ, Holmes CC, Newbold CI. 2009. Statistical estimation of cell-cycle progression and lineage commitment in *Plasmodium falciparum* reveals a homogeneous pattern of transcription in ex vivo culture. *Proc Natl Acad Sci U S A* 106:7559–7564. <https://doi.org/10.1073/pnas.0811829106>.
 68. Painter HJ, Altenhofen LM, Kafsack BF, Llinas M. 2013. Whole-genome analysis of *Plasmodium* spp. utilizing a new Agilent Technologies DNA microarray platform. *Methods Mol Biol* 923:213–219. https://doi.org/10.1007/978-1-62703-026-7_14.
 69. Filarsky M, Fraschka SA, Niederwieser I, Brancucci NMB, Carrington E, Carrio E, Moes S, Jenoe P, Bartfai R, Voss TS. 2018. GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* 359:1259–1263. <https://doi.org/10.1126/science.aan6042>.
 70. Eksi S, Morahan BJ, Haile Y, Furuya T, Jiang H, Ali O, Xu H, Kiattibutr K, Suri A, Czesny B, Adeyemo A, Myers TG, Sattabongkot J, Su X, Williamson KC. 2012. *Plasmodium falciparum* gametocyte development 1 (*Pfgdv1*) and gametocytogenesis early gene identification and commitment to sexual development. *PLoS Pathog* 8:e1002964. <https://doi.org/10.1371/journal.ppat.1002964>.
 71. Figueiredo LM, Pirrit LA, Scherf A. 2000. Genomic organisation and chromatin structure of *Plasmodium falciparum* chromosome ends. *Mol Biochem Parasitol* 106:169–174. [https://doi.org/10.1016/s0166-6851\(99\)00199-1](https://doi.org/10.1016/s0166-6851(99)00199-1).
 72. Gardner MJ, Hall N, Funk E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan M-S, Nene V, Shallow SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DMA, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419:498–511. <https://doi.org/10.1038/nature01097>.
 73. Rovira-Graells N, Crowley VM, Bancellis C, Mira-Martínez S, Ribas de Pouplana L, Cortés A. 2015. Deciphering the principles that govern mutually exclusive expression of *Plasmodium falciparum* *clag3* genes. *Nucleic Acids Res* 43:8243–8257. <https://doi.org/10.1093/nar/gkv730>.
 74. Joergensen L, Bengtsson DC, Bengtsson A, Ronander E, Berger SS, Turner L, Dalgaard MB, Cham GK, Victor ME, Lavstsen T, Theander TG, Arnot DE, Jensen AT. 2010. Surface co-expression of two different PfEMP1 antigens on single *Plasmodium falciparum*-infected erythrocytes facilitates binding to ICAM1 and PECAM1. *PLoS Pathog* 6:e1001083. <https://doi.org/10.1371/journal.ppat.1001083>.
 75. Merrick CJ, Jiang RH, Skillman KM, Samarakoon U, Moore RM, Dzikowski R, Ferdig MT, Duraisingh MT. 2015. Functional analysis of sirtuin genes in multiple *Plasmodium falciparum* strains. *PLoS One* 10:e0118865. <https://doi.org/10.1371/journal.pone.0118865>.
 76. Stewart LB, Freville A, Voss TS, Baker DA, Awandare GA, Conway DJ. 2022. *Plasmodium falciparum* sexual commitment rate variation among clinical isolates and diverse laboratory-adapted lines. *Microbiol Spectr* <https://doi.org/10.1128/spectrum.02234-22>.
 77. Boltryk SD, Passecker A, Alder A, Carrington E, van de Vegte-Bolmer M, van Gemert GJ, van der Starre A, Beck HP, Sauerwein RW, Kooij TWA, Brancucci NMB, Proelochs NI, Gilberger TW, Voss TS. 2021. CRISPR/Cas9-engineered inducible gametocyte producer lines as a valuable tool for *Plasmodium falciparum* malaria transmission research. *Nat Commun* 12:4806. <https://doi.org/10.1038/s41467-021-24954-4>.
 78. Broadbent KM, Broadbent JC, Ribacke U, Wirth D, Rinn JL, Sabeti PC. 2015. Strand-specific RNA sequencing in *Plasmodium falciparum* malaria identifies developmentally regulated long non-coding RNA and circular RNA. *BMC Genomics* 16:454. <https://doi.org/10.1186/s12864-015-1603-4>.
 79. Tibúrcio M, Hitz E, Niederwieser I, Kelly G, Davies H, Doerig C, Billker O, Voss TS, Treeck M. 2021. A 39-amino-acid C-terminal truncation of GDV1 disrupts sexual commitment in *Plasmodium falciparum*. *mSphere* 6:e01093-20. <https://doi.org/10.1128/mSphere.01093-20>.

80. Gaszner M, Felsenfeld G. 2006. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7:703–713. <https://doi.org/10.1038/nrg1925>.
81. Phillips-Cremins JE, Corces VG. 2013. Chromatin insulators: linking genome organization to cellular function. *Mol Cell* 50:461–474. <https://doi.org/10.1016/j.molcel.2013.04.018>.
82. Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, O'Brien J, Djimde A, Doumbo O, Zongo I, Ouedraogo J-B, Michon P, Mueller I, Siba P, Nzila A, Borrmann S, Kiara SM, Marsh K, Jiang H, Su X-Z, Amaratunga C, Fairhurst R, Socheat D, Nosten F, Imwong M, White NJ, Sanders M, Anastasi E, Alcock D, Drury E, Oyola S, Quail MA, Turner DJ, Ruano-Rubio V, Jyothi D, Amenga-Etego L, Hubbard C, Jeffreys A, Rowlands K, Sutherland C, Roper C, Mangano V, Modiano D, Tan JC, Ferdig MT, Amambua-Ngwa A, Conway DJ, Takala-Harrison S, Plowe CV, Rayner JC, et al. 2012. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* 487:375–379. <https://doi.org/10.1038/nature11174>.
83. MalariaGEN, Ahouidi A, Ali M, Almagro-Garcia J, Amambua-Ngwa A, Amaratunga C, Amato R, Amenga-Etego L, Andagalu B, Anderson TJC, Andrianaranjaka V, Apinjoh T, Ariani C, Ashley EA, Auburn S, Awandare GA, Ba H, Baraka V, Barry AE, Bejon P, Bertin GI, Boni MF, Borrmann S, Bousema T, Branch O, Bull PC, Busby GBJ, Chookajorn T, Chotivanich K, Claessens A, Conway D, Craig A, D'Alessandro U, Dama S, Day NPJ, Denis B, Dinkite M, Djimde A, Dolecek B, Dondorp AM, Drakeley C, Drury E, Duffy P, Echeverry DF, Ekwang TG, Erko B, Fairhurst RM, Faiz A, Fanello CA, Fukuda MM, Gamboa D, et al. 2021. An open dataset of *Plasmodium falciparum* genome variation in 7,000 worldwide samples. *Wellcome Open Res* 6:42. <https://doi.org/10.12688/wellcomeopenres.16168.2>.
84. Tonkin-Hill GQ, Trianty L, Noviyanti R, Nguyen HHT, Sebayang BF, Lampah DA, Marfurt J, Cobbold SA, Rambhatla JS, McConville MJ, Rogerson SJ, Brown GV, Day KP, Price RN, Anstey NM, Papenfuss AT, Duffy MF. 2018. The *Plasmodium falciparum* transcriptome in severe malaria reveals altered expression of genes involved in important processes including surface antigen-encoding *var* genes. *PLoS Biol* 16:e2004328. <https://doi.org/10.1371/journal.pbio.2004328>.
85. Prajapati SK, Ayanful-Torgby R, Pava Z, Barbeau MC, Acquah FK, Cudjoe E, Kakanev C, Amponsah JA, Obboh E, Ahmed AE, Abuaku BK, McCarthy JS, Amoah LE, Williamson KC. 2020. The transcriptome of circulating sexually committed *Plasmodium falciparum* ring stage parasites forecasts malaria transmission potential. *Nat Commun* 11:6159. <https://doi.org/10.1038/s41467-020-19988-z>.
86. Usui M, Prajapati SK, Ayanful-Torgby R, Acquah FK, Cudjoe E, Kakanev C, Amponsah JA, Obboh EK, Reddy DK, Barbeau MC, Simons LM, Czesny B, Raiculescu S, Olsen C, Abuaku BK, Amoah LE, Williamson KC. 2019. *Plasmodium falciparum* sexual differentiation in malaria patients is associated with host factors and GDV1-dependent genes. *Nat Commun* 10:2140. <https://doi.org/10.1038/s41467-019-10172-6>.
87. Kaestli M, Cortes A, Lagog M, Ott M, Beck HP. 2004. Longitudinal assessment of *Plasmodium falciparum* *var* gene transcription in naturally infected asymptomatic children in Papua New Guinea. *J Infect Dis* 189:1942–1951. <https://doi.org/10.1086/383250>.
88. Lavstsen T, Turner L, Saguti F, Magistrado P, Rask TS, Jespersen JS, Wang CW, Berger SS, Baraka V, Marquard AM, Seguin-Orlando A, Willerslev E, Gilbert MT, Lusingu J, Theander TG. 2012. *Plasmodium falciparum* erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children. *Proc Natl Acad Sci U S A* 109:E1791–E1800. <https://doi.org/10.1073/pnas.1120455109>.
89. Portugaliza HP, Natama HM, Guetens P, Rovira-Vallbona E, Somé AM, Millogo A, Ouédraogo DF, Valéa I, Sorgho H, Tinto H, van Hong N, Siteo A, Varo R, Bassat Q, Cortés A, Rosanas-Urgell A. 2022. *Plasmodium falciparum* sexual conversion rates can be affected by artemisinin-based treatment in naturally infected malaria patients. *EBioMedicine* 83:104198. <https://doi.org/10.1016/j.ebiom.2022.104198>.
90. Mok S, Ashley EA, Ferreira PE, Zhu L, Lin Z, Yeo T, Chotivanich K, Imwong M, Pukrittayakamee S, Dhorda M, Nguon C, Lim P, Amaratunga C, Suon S, Hien TT, Htut Y, Faiz JA, Onyamboko MA, Mayxay M, Newton PN, Tripura R, Woodrow CJ, Miotto O, Kwiatkowski DP, Nosten F, Day NPJ, Preiser PR, White NJ, Dondorp AM, Fairhurst RM, Bozdech Z. 2015. Drug resistance. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance. *Science* 347:431–435. <https://doi.org/10.1126/science.1260403>.
91. Daily JP, Scanzfeld D, Pochet N, Le Roch K, Plouffe D, Kamal M, Sarr O, Mboup S, Ndir O, Wypij D, Levesseur K, Thomas E, Tamayo P, Dong C, Zhou Y, Lander ES, Ndiaye D, Wirth D, Winzeler EA, Mesirov JP, Reggev A. 2007. Distinct physiological states of *Plasmodium falciparum* in malaria-infected patients. *Nature* 450:1091–1095. <https://doi.org/10.1038/nature06311>.
92. Tarr SJ, Diaz-Ingelmo O, Stewart LB, Hocking SE, Murray L, Duffy CW, Otto TD, Chappell L, Rayner JC, Awandare GA, Conway DJ. 2018. Schizont transcriptome variation among clinical isolates and laboratory-adapted clones of the malaria parasite *Plasmodium falciparum*. *BMC Genomics* 19:894. <https://doi.org/10.1186/s12864-018-5257-x>.
93. Mackinnon MJ, Li J, Mok S, Kortok MM, Marsh K, Preiser PR, Bozdech Z. 2009. Comparative transcriptional and genomic analysis of *Plasmodium falciparum* field isolates. *PLoS Pathog* 5:e1000644. <https://doi.org/10.1371/journal.ppat.1000644>.
94. Mensah-Brown HE, Amoako N, Abugri J, Stewart LB, Agongo G, Dickson EK, Ofori MF, Stoute JA, Conway DJ, Awandare GA. 2015. Analysis of erythrocyte invasion mechanisms of *Plasmodium falciparum* clinical isolates across 3 malaria-endemic areas in Ghana. *J Infect Dis* 212:1288–1297. <https://doi.org/10.1093/infdis/jiv207>.
95. Gomez-Escobar N, Amambua-Ngwa A, Walther M, Okebe J, Ebonyi A, Conway DJ. 2010. Erythrocyte invasion and merozoite ligand gene expression in severe and mild *Plasmodium falciparum* malaria. *J Infect Dis* 201:444–452. <https://doi.org/10.1086/649902>.
96. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. 2019. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* 10:1930. <https://doi.org/10.1038/s41467-019-09982-5>.
97. Gupta AP, Chin WH, Zhu L, Mok S, Luah YH, Lim EH, Bozdech Z. 2013. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite *Plasmodium falciparum*. *PLoS Pathog* 9:e1003170. <https://doi.org/10.1371/journal.ppat.1003170>.
98. Connacher J, von Grüning H, Birkholtz L. 2022. Histone modification landscapes as a roadmap for malaria parasite development. *Front Cell Dev Biol* 10:848797. <https://doi.org/10.3389/fcell.2022.848797>.
99. Tang J, Chisholm SA, Yeoh LM, Gilson PR, Papenfuss AT, Day KP, Petter M, Duffy MF. 2020. Histone modifications associated with gene expression and genome accessibility are dynamically enriched at *Plasmodium falciparum* regulatory sequences. *Epigenetics Chromatin* 13:50. <https://doi.org/10.1186/s13072-020-00365-5>.
100. Cortés A, Crowley VM, Vaquero A, Voss TS. 2012. A view on the role of epigenetics in the biology of malaria parasites. *PLoS Pathog* 8:e1002943. <https://doi.org/10.1371/journal.ppat.1002943>.
101. Portugaliza HP, Miyazaki S, Geurten FJ, Pell C, Rosanas-Urgell A, Janse CJ, Cortés A. 2020. Artemisinin exposure at the ring or trophozoite stage impacts *Plasmodium falciparum* sexual conversion differently. *Elife* 9:e60058. <https://doi.org/10.7554/eLife.60058>.
102. Delves MJ, Ruecker A, Straschil U, Lelievre J, Marques S, Lopez-Barragan MJ, Herreros E, Sinden RE. 2013. Male and female *Plasmodium falciparum* mature gametocytes show different responses to antimalarial drugs. *Antimicrob Agents Chemother* 57:3268–3274. <https://doi.org/10.1128/AAC.00325-13>.
103. Walliker D, Quakyi IA, Wellems TE, McCutchan TF, Szarfman A, London WT, Corcoran LM, Burkot TR, Carter R. 1987. Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science* 236:1661–1666. <https://doi.org/10.1126/science.3299700>.
104. Roestenberg M, Bijker EM, Sim BKL, Billingsley PF, James ER, Bastiaens GJH, Teirlinck AC, Scholzen A, Teelen K, Arens T, van der Ven A, Gunasekera A, Chakravarty S, Velmurugan S, Hermsen CC, Sauerwein RW, Hoffman SL. 2013. Controlled human malaria infections by intradermal injection of cryopreserved *Plasmodium falciparum* sporozoites. *Am J Trop Med Hyg* 88:5–13. <https://doi.org/10.4269/ajtmh.2012.12-0613>.
105. Gomez-Perez GP, Legarda A, Munoz J, Sim BK, Ballester MR, Dobano C, Moncunill G, Campo JJ, Cistero P, Jimenez A, Barrios D, Mordmuller B, Pardos J, Navarro M, Zita CJ, Nhamuave CA, Garcia-Basteiro AL, Sanz A, Aldea M, Manoj A, Gunasekera A, Billingsley PF, Aponte JJ, James ER, Guinovart C, Antonijano RM, Kreamsner PG, Hoffman SL, Alonso PL. 2015. Controlled human malaria infection by intramuscular and direct venous inoculation of cryopreserved *Plasmodium falciparum* sporozoites in malaria-naïve volunteers: effect of injection volume and dose on infectivity rates. *Malar J* 14:306. <https://doi.org/10.1186/s12936-015-0817-x>.
106. Casas-Vila N, Pickford AK, Portugaliza HP, Tintó-Font E, Cortés A. 2021. Transcriptional analysis of tightly synchronized *Plasmodium falciparum* intraerythrocytic stages by RT-qPCR. *Methods Mol Biol* 2369:165–185. https://doi.org/10.1007/978-1-0716-1681-9_10.

107. Kensch PR, Hoeijmakers WA, Toenhake CG, Bras M, Chappell L, Berriman M, Bartfai R. 2016. The nucleosome landscape of *Plasmodium falciparum* reveals chromatin architecture and dynamics of regulatory sequences. *Nucleic Acids Res* 44:2110–2124. <https://doi.org/10.1093/nar/gkv1214>.
108. Otto TD, Wilinski D, Assefa S, Keane TM, Sarry LR, Bohme U, Lemieux J, Barrell B, Pain A, Berriman M, Newbold C, Llinas M. 2010. New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. *Mol Microbiol* 76:12–24. <https://doi.org/10.1111/j.1365-2958.2009.07026.x>.
109. Toenhake CG, Frasnka SA, Vijayabaskar MS, Westhead DR, van Heeringen SJ, Bartfai R. 2018. Chromatin accessibility-based characterization of the gene regulatory network underlying *Plasmodium falciparum* blood-stage development. *Cell Host Microbe* 23:557–569.e9. <https://doi.org/10.1016/j.chom.2018.03.007>.