Published online 21 July 2015

Nucleic Acids Research, 2015, Vol. 43, No. 17 ***8243–8257***
doi: 10.1093/nar/gkv730

Deciphering the principles that govern mutually exclusive expression of *Plasmodium falciparum clag3* genes

Núria Rovira-Graells[1,2], Valerie M. Crowley[2,†], Cristina Bancells[1,†], Sofía Mira-Martínez[1], Lluís Ribas de Pouplana[2,3] and Alfred Cortés[1,2,3,*]

[1]ISGlobal, Barcelona Ctr. Int. Health Res. (CRESIB), Hospital Clínic—Universitat de Barcelona, 08036 Barcelona, Catalonia, Spain, [2]Institute for Research in Biomedicine (IRB), 08028 Barcelona, Catalonia, Spain and [3]Catalan Institution for Research and Advanced Studies (ICREA), 08010 Barcelona, Catalonia, Spain

Received March 6, 2015; Revised June 22, 2015; Accepted July 7, 2015

ABSTRACT

The product of the *Plasmodium falciparum* genes *clag3.1* and *clag3.2* plays a fundamental role in malaria parasite biology by determining solute transport into infected erythrocytes. Expression of the two *clag3* genes is mutually exclusive, such that a single parasite expresses only one of the two genes at a time. Here we investigated the properties and mechanisms of *clag3* mutual exclusion using transgenic parasite lines with extra copies of *clag3* promoters located either in stable episomes or integrated in the parasite genome. We found that the additional *clag3* promoters in these transgenic lines are silenced by default, but under strong selective pressure parasites with more than one *clag3* promoter simultaneously active are observed, demonstrating that *clag3* mutual exclusion is strongly favored but it is not strict. We show that silencing of *clag3* genes is associated with the repressive histone mark H3K9me3 even in parasites with unusual *clag3* expression patterns, and we provide direct evidence for heterochromatin spreading in *P. falciparum*. We also found that expression of a neighbor ncRNA correlates with *clag3.1* expression. Altogether, our results reveal a scenario where fitness costs and non-deterministic molecular processes that favor mutual exclusion shape the expression patterns of this important gene family.

INTRODUCTION

Regulation of gene expression plays a central role in the biology of the apicomplexan parasite *Plasmodium falciparum*, which is responsible for the most severe forms of human malaria. Progression along the complex life cycle of malaria parasites, which involves multiple well-differentiated stages, is mainly controlled at the transcriptional level (1,2). Dynamic transcriptional regulation along the life cycle involves transcription factor cascades (3) and also changes at the chromatin level (4,5), although in the context of life cycle progression the latter probably play an effector role and do not transmit epigenetic information from one generation to the next (6). Another important process in *P. falciparum* biology controlled at the transcriptional level is clonally variant gene expression. Genes under clonally variant expression can be found in either an active or a silenced state in genetically identical parasites at the same stage of the life cycle. Recent research has established that disparate *P. falciparum* gene families involved in different aspects of host-parasite interactions show clonally variant expression (7). This type of expression is believed to play an important role in parasite survival by allowing the adaptation of parasite populations to changing environments by bet-hedging adaptive strategies (7). While the main role of some large *P. falciparum* clonally variant gene families is antigenic variation and immune evasion, other clonally variant genes confer functional variation (7–10) or control developmental decisions (11).

Clonally variant gene expression is a truly epigenetic phenomenon, involving stochastic choices and transmission of non-genetic information from one generation to the next (6). The molecular basis for clonally variant expression is not completely understood (8,12), but post-translational modifications in histone H3 lysine 9 (H3K9) play a key role in determining and transmitting the expression state of clonally variant genes, similar to stochastic gene expression processes in higher eukaryotes (13). Acetylation at this position (H3K9ac) is associated with the active state of *P. falciparum* clonally variant genes, whereas tri-methylation

*To whom correspondence should be addressed. Tel: +34 932275400; Fax: +34 933129410; Email: alfred.cortes@cresib.cat
†These authors contributed equally to the paper as second authors.
Present address: Valerie M. Crowley, Toronto General Research Institute (TGRI), University Health Network (UHN), Toronto, Ontario, Canada.

(H3K9me3) is associated with their silenced state (7,14–19). Clonally variant genes are located in bistable chromatin domains, such that once established both the euchromatin (active) and the heterochromatin (silenced) states are stable and clonally inherited (16). Since the epigenetic transmission of chromatin states is less faithful than DNA replication, spontaneous transitions between the two chromatin states occur, albeit at low frequency, resulting in switches between the active and repressed transcriptional states.

The regulation of two of the *P. falciparum* gene families under clonally variant expression involves an additional layer of complexity: mutually exclusive expression. This phenomenon implies that an individual cell only expresses one gene of the family at a time. This type of regulation has been studied for gene families such as olfactory receptor genes and protocadherins in higher eukaryotes (20), and for genes linked to antigenic variation in protozoan parasites such as *Trypanosoma brucei* or *Giardia lamblia* (21), among others. Different gene families use disparate molecular mechanisms to achieve mutual exclusion. The two gene families under mutually exclusive expression in *P. falciparum* are *var* and cytoadherence-linked asexual gene 3 (*clag3*) (8). *var* genes, a large family of about 60 genes per genome, encode the *P. falciparum* exported membrane protein 1 (PfEMP-1). This protein is exported to the infected erythrocyte surface, where it participates in two processes linked to malaria virulence: antigenic variation and cytoadherence (22).

In contrast to the large *var* family, there are only two *clag3* genes, *clag3.1* and *clag3.2* (PF3D7_0302500 and PF3D7_0302200, respectively, previous IDs PFC0120w and PFC0110w), which show ∼95% sequence identity and are located only 10 kb apart from each other. These genes are part of the five-member *clag* family, which also includes the more distantly related genes *clag2*, *clag8* and *clag9* (23). Mutually exclusive expression only affects the two *clag3* genes, such that an individual parasite expresses either *clag3.1* or *clag3.2*, whereas *clag2* undergoes independent clonally variant expression and *clag8* and *clag9* appear to be expressed in all parasites (24). Mutually exclusive expression of *clag3* genes has now been observed by several different laboratories in parasite lines of different genetic backgrounds (16,17,24–26), and also in parasites in which a recombination event resulted in the presence of three *clag3* genes (26) or in transgenic parasites where the *clag3.2* ORF was disrupted by insertion of a selectable marker (17). Similar to other clonally variant genes, silencing of one of the *clag3* genes is mediated by H3K9me3-based heterochromatin, whereas activation is associated with an increase in the alternative modification at this position, H3K9ac (16,17). Additionally, silencing of *clag3* genes is associated with a different position for specific nucleosomes and in the case of *clag3.2* also reduced accessibility to restriction enzymes, which is consistent with a heterochromatic conformation (16).

The proteins encoded by *clag3* genes, CLAG3s (also known as RhopH1/CLAG3), play a fundamental role in parasite biology by determining solute transport at the infected erythrocyte membrane (25). Ions, nutrients, and also certain toxic compounds with antimalarial activity require CLAG3 for their transport into infected erythrocytes (9,10,25–27). However, a role for CLAG proteins in cytoadherence or erythrocyte invasion had been previously proposed. Whether CLAGs actually participate in these processes in addition to their established role in solute transport remains to be determined (23). In any case, the temporal expression pattern of this gene family is intriguing: all *clag* genes are expressed late in the asexual cycle (schizont stage), but CLAG3 proteins only reach the infected erythrocyte membrane and participate in solute transport after schizont bursting and reinvasion, about 20 h post-invasion in the following cycle of asexual growth (25). Another unresolved question is why parasites express only one *clag3* gene at a time. Recent investigations have revealed functional differences in the transport of some solutes between parasites expressing CLAG3.1 or CLAG3.2, which is suggestive of a role for mutually exclusive expression of these genes in restricting the entrance of specific solutes (9). An alternative possible role for mutual exclusion of *clag3* genes is immune evasion: although this gene family is apparently too small to mediate effective antigenic variation, variant expression and genetic polymorphism may play synergistic roles to escape immune responses (28).

While the products of both *var* and *clag3* genes play fundamental roles in parasite biology, only mutually exclusive expression of *var* has been extensively studied (8,12,29). Mutually exclusive expression of *var* genes is controlled by two genetic elements, the *var* upstream region including the main promoter and the *var* intron including a second promoter (30–32). Small genetic elements that mediate promoter-intron pairing and the mutual exclusion element at the *var* upstream region are necessary for *var* mutually exclusive expression (33,34). Furthermore, several chromatin regulators have been identified that play specific roles in *var* mutual exclusion, including sirtuins, SET domain methyltranferases, and the histone variants H2A.Z and H2B.Z (reviewed in ref. (12)).

We have previously characterized the *clag3* promoters and found that in transient transfection assays episomal *clag3* promoter activity is independent of the endogenous *clag3* gene being expressed (16). However, in these assays only a small fraction of the parasites carry the transfected plasmid (transfection efficiency $\sim 10^{-6}$), making it impossible to determine how the episomal *clag3* promoter affects expression of endogenous *clag3* genes. To unravel the 'rules' that govern mutually exclusive expression of *clag3* genes, here we used transgenic parasite lines with additional *clag3* promoters driving the expression of reporter genes or selectable markers, either in stable episomes or integrated next to the endogenous *clag3* loci. We also investigated the molecular mechanisms that control the expression of these genes.

MATERIALS AND METHODS

Plasmids

Plasmids 3.2–1371-LH and 3.2-LH-bsdR (previously named 3.2–1371-LH-bsdR) have been described before (9,16). Plasmids 3.2-LH-bsd and ama1-LH-bsd were generated by adding the blasticidin S (BSD) deaminase resistance cassette from plasmid pHBupsC^{R} (32) to plasmids 3.2–1371-LH or ama1–1570-LH (16), respectively. This was

achieved by inserting a PstI-SmaI fragment of pHBupsC^R, treated with T4 DNA polymerase to generate blunt ends, into a SmaI site in the 3.2–1371-LH or ama1–1570-LH plasmids. Plasmid pHBc3.2 was derived from pHBupsC^R (32) by replacing the upsC promoter and rep20 sequences with the *clag3.2* upstream region (full 5′ intergenic region). The *clag3.2* promoter was amplified by polymerase chain reaction (PCR) with primers Prom3.2fPst and Prom3.2rBam (Supplementary Table S1), digested with PstI and BamHI, and inserted into the pHBupsC^R plasmid digested with the same enzymes.

Parasites

The 3D7-A stock of the clonal parasite line 3D7 and 3D7-A subclones 10G and 1.2B have been previously described and characterized (24,35,36). Parasites were cultured in B+ erythrocytes under standard conditions with media containing Albumax II and no human serum. Cultures were typically synchronized by sorbitol lysis, which eliminates erythrocytes infected with parasites at the trophozoite or schizont stages. Synchronization of the parasite line 10G-0.6–2, which is resistant to sorbitol lysis (9), was performed using either magnetic columns to purify erythrocytes infected with parasites at mature stages (experiments to analyze expression of a ncRNA) or by treatment with L-Proline (to synchronize 10G-0.6–2 and control 10G cultures for histone modifications analysis). In brief, L-Proline lysis was performed using a procedure similar to sorbitol lysis, but treating cultures for 7 min at 37°C with a L-Proline solution (seven pellet volumes) instead of sorbitol. The concentration of L-Proline used ranged from 280 mM in initial experiments to 150 mM in later experiments to minimize the effects of the treatment on the viability of the cultures. For selected experiments, parasites were tightly synchronized to a defined age window by purification of parasites at the schizont stage (using Percoll gradient) followed by sorbitol lysis 5 or 7 h later. Transfection was performed by electroporation according to standard procedures (37). To select for parasites stably carrying the plasmids, transfected cultures were selected with BSD at 2.5 μg/ml and to select for parasites that integrated the plasmid we additionally performed two off/on drug cycles. Other drug selection experiments used BSD at the same concentration or 10 nM WR99210 unless otherwise stated. Subcloning was performed by limiting dilution. In the case of parasites with an integrated plasmid (pHBc3.2), subcloning was performed in the absence of any drug. Luciferase assays on transiently or stably transfected parasites were performed as previously described (16).

Transcriptional analysis

Transcriptional analysis by reverse-transcriptase quantitative PCR (RT-qPCR) was performed on synchronized cultures at the schizont stage unless otherwise stated. RNA was obtained using the Trizol method, DNAse treated, column purified and reverse transcribed as previously described (24). Reverse transcription was performed with a mixture of random primers and oligo (dT). cDNAs were analyzed by qPCR with the Power SYBR Green Master Mix (Applied Biosystems), using the standard curve method. In brief, each plate included a standard curve consisting of five serial 1:10 dilutions of gDNA for each primer pair. All plates also included–RT controls for each sample and no template controls for each primer pair. All samples and controls were analyzed in triplicate. Results for test samples were obtained by interpolating the Ct value in the standard curve and normalized against expression values for control genes with a similar temporal expression pattern. This approach corrects for differences in parasite age between RNA preparations from different cultures. Unless otherwise stated, *rhoph2* (ID PF3D7_0929400) was used to normalize the expression of genes under the control of schizont-specific promoters (such as *clag3* promoters) and seryl tRNA synthetase (*serrs*, ID PF3D7_0717700) was used to normalize the expression of genes under the control of promoters expressed at similar levels throughout the asexual blood cycle, or to investigate the stage of expression. The gDNA used for the standard curve was typically from the same parasite line under transcriptional analysis. In the case of stably transfected parasite lines (episomal or integrated plasmids), using this approach implies that transcript levels indicate the activity of one copy of the transgene (average of expression of all copies). When the same gDNA standard curve was used to analyze transgenic parasite lines with different plasmid copy number, results for the genes in the plasmid were adjusted by plasmid copy number relative to the parasite line used for the standard curve. The primers used for RT-qPCR analysis are described in Supplementary Table S1.

Genetic analysis

gDNA for qPCR or Southern blot analysis was prepared by sodium dodecyl sulphate (SDS) lysis followed by double phenol–chloroform extraction and ethanol precipitation. Integrated or episomal plasmid copy number relative to the gDNA used for the standard curve (gDNA from one of the transgenic lines analyzed) was determined by qPCR using the standard curve method as in the transcriptional analysis. Southern blot analysis was performed as previously described (16). The primers used to amplify the probes for Southern blot are described in Supplementary Table S1.

Chromatin immunoprecipitation (ChIP)

Chromatin was obtained from saponin-extracted parasites at the late trophozoite or schizont stage, cross-linked with formaldehyde and sonicated as previously described (11,16). For the experiments in Supplementary Figure S5 we used the chromatin immunoprecipitation (ChIP) procedure previously described, with Magna ChIP^{TM} Protein G magnetic beads (Merck-Millipore) and magnetic separation (11) instead of agarose beads and centrifugation (16). For these experiments we used antibodies from Millipore (#07–442 and #07–352) (16) that have been widely used for ChIP in malaria research, including ChIP–chip experiments (19). For the experiments presented in Figure 4, we immunoprecipitated chromatin (prepared as described above) using the MAGnify^{TM} Chromatin Immunoprecipitation System (Life Technologies) and antibodies against H3K9me3 and H3K9ac from Diagenode (pAb-193–050 and pAb-004-050, respectively). This new procedure described here results

in highly improved signal to noise ratios. In all cases, immunoprecipitates were analyzed by qPCR using the standard curve method, with the primers detailed in Supplementary Table S1.

RESULTS

Activity of a *clag3* promoter controlling a reporter gene in stable episomes

To study the expression of a *clag3* promoter in stable episomes, we first designed the 3.2-LH-bsdR plasmid (9) (Figure 1A), which contains the *clag3.2* promoter region (full 5′ intergenic region) controlling the expression of a luciferase (*luc*) reporter gene and a BSD resistance cassette consisting of the BSD deaminase gene (*bsd*) under the control of the constitutive *hsp90* (ID PF3D7_0708400) promoter. This plasmid also contains the subtelomeric repeat sequence rep20 that improves plasmid segregation (38). We first used a transient transfection assay to compare luciferase expression between parasites transfected with this plasmid or an equivalent plasmid without the BSD resistance cassette and rep20 element (plasmid 3.2–1371-LH, Figure 1A) (16). The plasmid 3.2-LH-bsdR showed lower luciferase activity (Figure 1B), indicating that the *clag3.2* promoter is partially silenced in this plasmid. This is not explained by the presence of the rep20 element because we have previously demonstrated that addition of this element increases luciferase activity in 3.2–1371-LH (16). Hence, partial repression of the *clag3* promoter in plasmid 3.2-LH-bsdR is likely to be a consequence of the presence of a second promoter (the constitutive *hsp90* promoter). This situation is reminiscent of observations with *var* promoters, which require the intron or an alternative paired promoter for transcriptional silencing (30,39,40). However, silencing in 3.2-LH-bsdR may be a consequence of the specific plasmid architecture, as suggested by experiments presented below.

Next we transfected two subclones of the parasite line 3D7-A, termed 10G and 1.2B, with the plasmid 3.2-LH-bsdR and applied BSD selection to obtain populations stably carrying the plasmid. 10G and 1.2B predominantly express *clag3.2* or *clag3.1*, respectively (16,24). Consistent with previous observations from experiments using transient transfection (16), luciferase assays revealed similar *clag3.2* promoter activity between the two stably transfected lines (Figure 1C), confirming that expression of episomal *clag3* promoters is independent of which endogenous *clag3* is active. This result suggests that expression of one or the other *clag3* gene does not depend on differentially-expressed *trans*-acting factors. However, luciferase assays do not inform about the activity of the episomal *clag3.2* promoter relative to the activity of endogenous *clag3* promoters. Hence, we performed a transcriptional analysis of the stably transfected 10G line using RT-qPCR. We found that expression of the episomal *luc* gene was much lower than expression of endogenous *clag3* genes (Figure 1D), indicating that the *clag3.2* promoter in the stable episome is predominantly silenced. We obtained subclones of 10G stably transfected with plasmid 3.2-LH-bsdR and analyzed their *luc* expression, but these experiments did not conclusively determine whether the episomal *clag3.2* promoter undergoes clonally variant expression or is homogeneously expressed at low levels in this transgenic parasites population (Supplementary Figure S1).

To determine whether the low transcriptional activity of the episomal *clag3.2* promoter is attributable to an intrinsic ability to nucleate heterochromatin, we conducted a new series of experiments using the plasmids 3.2-LH-bsd (identical to 3.2-LH-bsdR except for the absence of the rep20 element) and ama1-LH-bsd (Figure 1A). In the latter plasmid, the *clag3.2* upstream region is replaced by the upstream region (1570 bp) of *ama1*, which is not prone to heterochromatin formation or clonally variant expression and is expressed at the schizont stage in all parasites (7,19). Stable transfection of 10G and 1.2B with these plasmids followed by transcriptional analysis revealed that both promoters driving *luc* expression (*clag3.2* and *ama1*) are largely silenced in the episome compared with their chromosomal counterparts, which indicates that silencing is independent of the presence of a clonally variant promoter (Figure 1E). Hence, low expression of the promoters in these plasmids is mainly attributable to the episomal status or the specific plasmid architecture and does not seem to depend on mutual exclusion of *clag3* promoters or the ability to nucleate heterochromatin. While this result was unexpected, low expression from an episomal promoter compared to a chromosomal copy of the same promoter has been previously observed by others (41). It is difficult to estimate if this occurs commonly in malaria episomal expression systems because in the majority of published articles transcript levels are not compared between a transgene and the endogenous gene controlled by the same promoter.

Total expression of endogenous *clag3* genes was similar between parasite lines stably transfected with the 3.2-LH-bsdR plasmid and untransfected 10G and 1.2B parasites. Likewise, total endogenous *clag3* expression was also similar between parasites transfected with minus rep20 plasmids containing either a *clag3.2* promoter or an *ama1* promoter (Figure 1F). However, stably transfected 10G cultures switched to different extents from predominant *clag3.2* expression to a mixed population of parasites expressing one or the other *clag3* gene, independently of the presence or absence of the episomal *clag3.2* promoter. Parasites that express CLAG3.1 limit the transport of BSD across the infected erythrocyte membrane and in the presence of this drug have a survival advantage over those that express CLAG3.2 (9), which explains the progressive elimination of CLAG3.2-expressing parasites in BSD selected cultures. In any case, the low level of transcriptional activity of the episomal promoters precludes determining from these experiments whether a fully active episomal *clag3* promoter would affect the expression of endogenous *clag3* genes.

Activation of an episomal *clag3* promoter does not result in silencing of endogenous *clag3* genes

To determine whether an active episomal *clag3* promoter infiltrates the *clag3* mutual exclusion program, resulting in simultaneous silencing of the two endogenous *clag3* genes, we designed a new plasmid where a *clag3.2* promoter controls the expression of the selectable marker *hdhfr*, which confers resistance to the drug WR99210 (plasmid pHBc3.2, Fig-

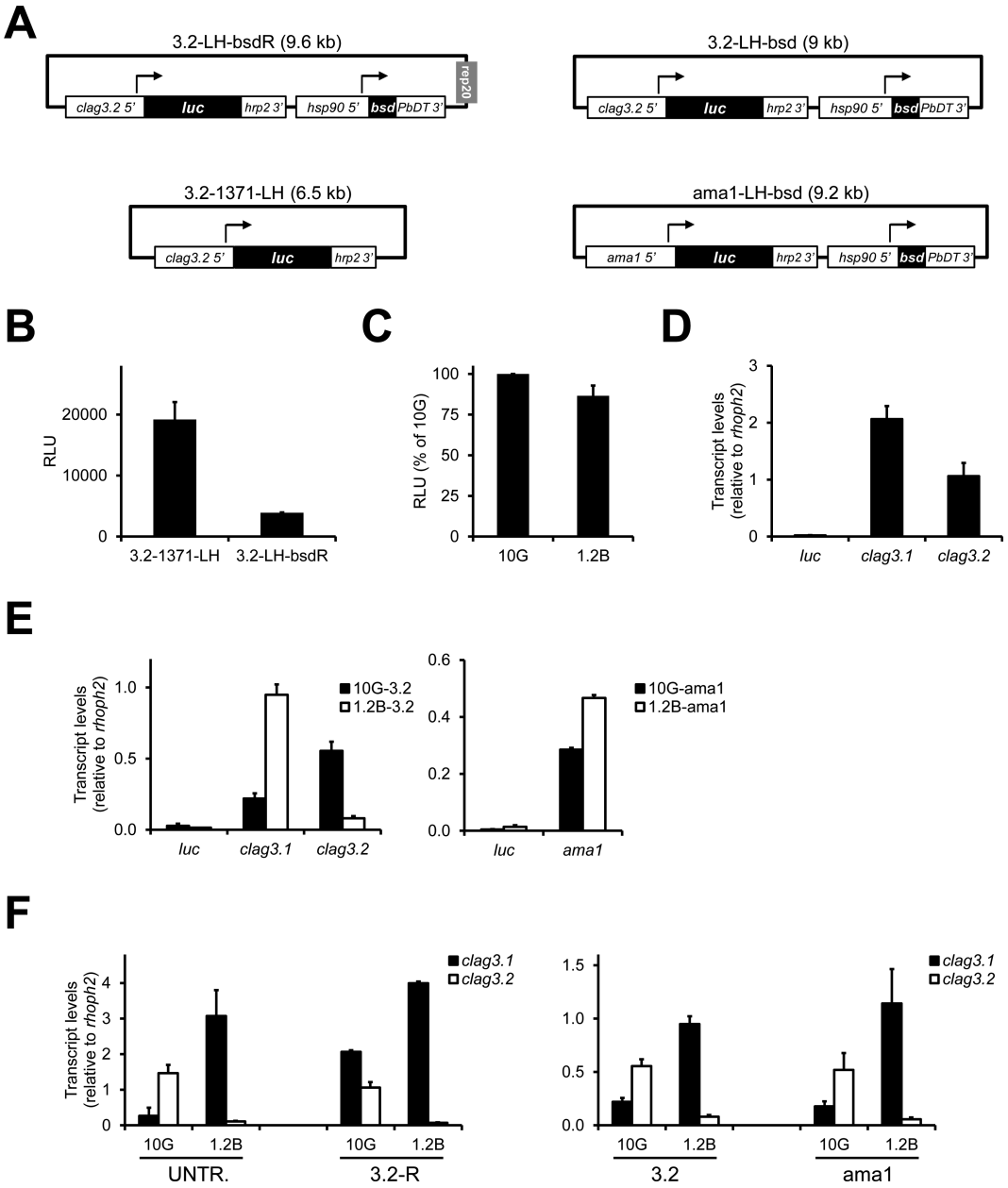

Figure 1. Characterization of *clag3* promoters driving expression of a luciferase gene in stable episomes. (**A**) Schematic of the plasmids used in this study with a *clag3* promoter controlling a *luc* gene (not to scale). Note that the *hsp90* (ID PF3D7_0708400) promoter is referred to as *hsp86* promoter in many previous publications. (**B**) Transient transfection luciferase assays of 3D7-A with plasmids 3.2–1371-LH and 3.2-LH-bsdR. Luciferase activity was measured in parasites at the schizont stage. Results are normalized to 1% parasitemia and expressed in relative light units (RLU). Values are the average of two independent biological replicates each performed in duplicate plates, with SEM. (**C**) Luciferase activity in 10G and 1.2B parasite lines stably transfected with plasmid 3.2-LH-bsdR. Cultures were synchronized to a 0–5 h age window and harvested for luciferase assays 40 h later (40–45 h post invasion). Results are normalized to 1% parasitemia and expressed as RLU in 1.2B relative to 10G. Values are the average of two independent biological replicates each performed in duplicate plates, with SEM. (**D**) RT-qPCR transcriptional analysis (schizont stage) of the 10G line stably transfected with plasmid 3.2-LH-bsdR. Transcript levels were normalized against *rhoph2*. Since we used a qPCR standard curve consisting of gDNA of the same parasite line as the test samples, expression levels for the episomal gene reflect expression per copy of the gene (average of all copies, see 'Materials and Methods' section). Values are the average of three independent biological replicates, with SEM. (**E**) Transcriptional analysis of 10G and 1.2B cultures (schizont stage) stably transfected with plasmids 3.2-LH-bsd or ama1-LH-bsd ('3.2' and 'ama1', respectively). Values are the average of two independent biological replicates, with SEM. (**F**) Expression of endogenous *clag3* genes in 10G and 1.2B cultures transfected with plasmid 3.2-LH-bsdR ('3.2-R') or minus rep20 plasmids '3.2' and 'ama1'(as in panel E). Expression of *clag3* genes in untransfected 10G and 1.2B cultures ('UNTR.') is shown for comparison. Each series of experiments is shown in a separate bar chart. Values are the average of two or three independent biological replicates, with SEM.

ure 2A). This plasmid also contains a BSD resistance cassette under the control of the constitutive *hsp90* promoter. 10G cultures were transfected with pHBc3.2 and selected with BSD to obtain a population of parasites stably carrying the plasmid. We then maintained these cultures for 3 weeks either in the presence of BSD alone or in the presence of BSD plus 10 nM WR99210 to select for parasites in which the *hdhfr* gene was active. These cultures were initially sensitive to WR99210, consistent with a silenced state of the episomal *clag3.2* promoter, but a resistant population was apparent after about 1 week of WR99210 selection. On a technical note, this result demonstrates that the *hdhfr* gene under the control of a promoter expressed only at the schizont stage can efficiently mediate drug resistance.

The two genes in the episome, *hdhfr* and *bsd*, were expressed at low levels in the absence of WR99210, but transcript levels increased dramatically for both genes after selection with the drug (Figure 2B). However, the increase in transcript levels was paralleled by an increase in plasmid copy number. When transcript levels were normalized by the relative plasmid copy number, it was apparent that transcriptional activation was only about three-fold (Figure 2C). Hence, WR99210 pressure selected for parasites with both higher plasmid copy number and plasmids that were transcriptionally more active. These results indicate that the full plasmid, including the constitutive promoter driving *bsd* expression, is by default partly silenced. Importantly, total levels of endogenous *clag3* transcripts did not decrease in WR99210-selected parasites, demonstrating that activation of an episomal *clag3.2* promoter does not preclude expression of the endogenous *clag3* genes (Figure 2D). Instead, total endogenous *clag3* transcript levels were slightly higher in the presence of WR99210, and the switch from *clag3.2* to *clag3.1* expression was less pronounced in the presence of this drug. The explanation for this observation is that the increased plasmid copy number and the transcriptional activation of the full plasmid associated with WR99210 selection leads to a large increase in *bsd* transcript levels, which confers efficient BSD resistance and limits the selective advantage of parasites with *clag3* expression patterns that restrict BSD transport. Altogether, these results provide the first evidence, to our knowledge, that more than one *clag3* promoter can be simultaneously active.

A chromosomal *clag3* promoter driving the expression of a selectable marker is silenced by default but mutually exclusive expression is relieved under strong selective pressure

To characterize an additional copy of a *clag3* promoter in a chromosomal context, integrated next to the endogenous *clag3* loci, we transfected 3D7-A parasites with the plasmid pHBc3.2 (Figure 2A) and used on/off BSD cycles to select for parasites that had integrated the plasmid. Southern blot analysis revealed integration of multiple copies of the plasmid by single homologous recombination with the endogenous *clag3.2* upstream region (Figure 3A and Supplementary Figure S2). In spite of several attempts, we failed to obtain parasites with a single copy of the plasmid integrated. Transcriptional analysis of the transgenic line before subcloning revealed that the *hdhfr* gene controlled by the extra copies of the *clag3.2* promoter was silenced (Figure 3B), which is consistent with mutually exclusive expression of *clag3* promoters. While the parental 3D7-A population contains a similar proportion of parasites expressing *clag3.1* or *clag3.2*, the majority of parasites in the transgenic line expressed *clag3.1* (Figure 3B), as expected after having used BSD for selection (9). Nonetheless, we were still able to identify subclones of this transgenic line (Supplementary Figure S2) that predominantly express *clag3.1* (G9 subclone) or *clag3.2* (C5 subclone) (Figure 3C).

When C5 and G9 were maintained in the absence of any drug for 3 weeks, the additional *clag3.2* promoters driving *hdhfr* expression were silenced in both subclones (Figure 3C). Consistently, these subclones were initially highly sensitive to WR99210, but a population of resistant parasites emerged after a few cycles of drug selection. C5 and G9 cultures selected for 3 weeks with WR99210 expressed *hdhfr* at high levels, indicating that in the selected parasite populations at least some of the copies of the additional *clag3.2* promoters are in an active state (Figure 3C). Note that the *clag3.2* promoter is weaker than the *clag3.1* promoter, such that in a transcriptionally homogeneous parasite line expressing *clag3.2*, total *clag3* transcript levels are lower than in a line expressing *clag3.1* (Figures 1F and 3C), as previously observed (9,16). Importantly, we ruled out changes in gene copy number as a mechanism for the increased *hdhfr* transcript levels observed after WR99210 selection (Supplementary Figure S3). Thus, in contrast to experiments with episomal plasmids (Figure 2), transcript level changes associated with drug selection in parasites with chromosomal extra copies of a *clag3* promoter are exclusively attributable to dynamic epigenetic changes. Of note, the transgenic promoters displayed correct temporal expression patterns across the asexual blood cycle (Supplementary Figure S4), confirming their physiological functionality.

The activation of the additional *clag3.2* promoters did not lead to silencing of the endogenous *clag3* genes in either of the subclones (Figure 3C). This result unambiguously demonstrates that under selective pressure more than one *clag3* promoter can be simultaneously expressed. The *bsd* gene, driven by a constitutive promoter not susceptible to clonally variant expression or heterochromatin formation (7,19,42) and located between *hdhfr* and endogenous *clag3.2*, was silenced in the absence of drugs in G9, where both neighbor genes are silenced. In contrast, *bsd* was expressed at much higher levels in C5, where *hdhfr* is silenced but endogenous *clag3.2* is active (Figure 3C). The low level of *bsd* expression in the unsubcloned transgenic population (Figure 3B) reflects the predominance of *clag3.1* expressing parasites (G9-like) before subcloning. After WR99210 selection *bsd* transcript levels increased >10-fold in G9, but they were still lower than in C5 (Figure 3C). However, *bsd* expression driven by the *hsp90* promoter was always lower than expression of endogenous *hsp90* (Figure 3B), suggesting that only one or a few of the integrated copies of the *hsp90* promoter driving *bsd* expression are active in the C5 or WR99210-selected G9 lines. Note that in our analysis expression levels reflect expression per copy of the transgene (average of all copies, see 'Materials and Methods' section). Silencing of a constitutive promoter by the influence of the neighboring epigenetically-regulated genes is clearly suggestive of spreading of repressive chromatin structures (see

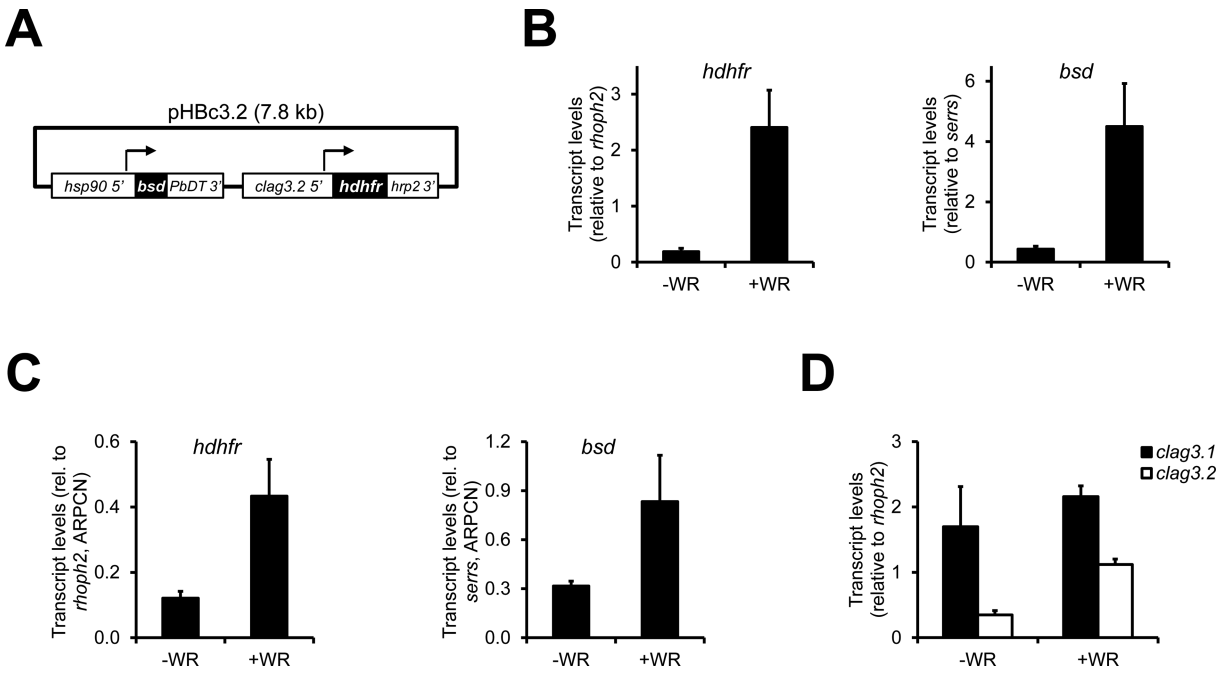

Figure 2. Characterization of *clag3* promoters driving expression of a selectable marker in stable episomes. (**A**) Schematic of the plasmid pHBc3.2. (**B**) Transcriptional analysis of 10G cultures (schizont stage) stably transfected with plasmid pHBc3.2 maintained in parallel for 3 weeks without or with WR99210 (WR). All cultures were maintained in the presence of BSD. Transcript levels were normalized against *rhoph2* for genes controlled by *clag3* promoters or against *serrs* for *bsd*, which is controlled by a promoter expressed at all asexual stages. (**C**) Transcript levels in panel B adjusted by relative plasmid copy number (ARPCN). (**D**) Expression of endogenous *clag3* genes in the same cultures. In all panels, values are the average of four independent biological replicates (except in panel C, $n = 3$ because relative plasmid copy number was not available for one of the experiments) from two independent WR99210 selections, with SEM.

below). In fact, *hdhfr* activation in the presence of WR99210 in G9 was accompanied not only by activation of the *bsd* gene but also by a small increase in transcription of endogenous *clag3.2* (Figure 3C) consistently observed in three independent selection experiments.

Next we maintained WR99210-selected cultures for three additional weeks either in the presence ('+' bars) or absence ('+/−' bars) of this drug. Parasites never exposed to WR99210 were also grown and analyzed in parallel ('−' bars, Figure 3D). Removal of the drug resulted in reversal of the changes associated with drug selection (Figure 3D). Of note, *hdhfr* was almost completely silenced in both C5 and G9 by 3 weeks after removal of the drug, indicating that simultaneous expression of multiple *clag3* promoters is only maintained under strong selective pressure and mutual exclusion is restored after removing the pressure. Last, treatment of C5 and G9 with BSD revealed that both parasite lines are relatively insensitive to this drug, showing that even the low levels of *bsd* expression in unselected G9 are sufficient to enable growth in the presence of BSD (albeit at reduced rates). Consequently, no major transcriptional changes were observed after 3 weeks of BSD selection, apart from partial activation of the *bsd* gene in G9 (Figure 3E). Interestingly, the moderate activation of *bsd* in G9 was accompanied by a small but consistent increase in *hdhfr* transcript levels (Figure 3E), confirming the inter-related chromatin-based epigenetic regulation of the two neighbor loci.

Epigenetic silencing at the *clag3* region is invariably mediated by H3K9me3-based heterochromatin

We used ChIP analysis to directly determine whether silencing of the *bsd* gene in the G9 parasite line cultured in the absence of any drug is mediated by H3K9me3-based heterochromatin. The *bsd* gene driven by the constitutive *hsp90* promoter is highly enriched in H3K9me3 and has low levels of H3K9ac, similar to heterochromatin positive controls (a *var* gene and *clag3.2*, which is silenced in this parasite line) (Figure 4A). The ratio of acetylation to tri-methylation at H3K9, which provides a more sensitive measure of the euchromatic or heterochromatic state of a locus, clearly confirms that the *bsd* cassettes are in a heterochromatic state in G9 (Figure 4A). Considering that the *hsp90* promoter is not clonally variant and does not carry epigenetic marks of silencing in its endogenous location (7,19,42), this result directly demonstrates spreading of heterochromatin from neighbor loci (Figure 5). ChIP analysis of the G9 line selected with WR99210 showed a pattern of H3K9 modifications similar to the unselected line (data not shown), consistent with the idea that only some of the copies of the *bsd* cassette are activated upon selection (Figure 3B and C). Given that all integrated copies of the *bsd* cassette are identical, it is impossible to analyze by ChIP the individual copies separately.

We also used ChIP analysis to determine whether simultaneous silencing of the two *clag3* genes is mediated by H3K9me3-based heterochromatin (Figure 4B and Supplementary Figure S5). We and others have previously demon-

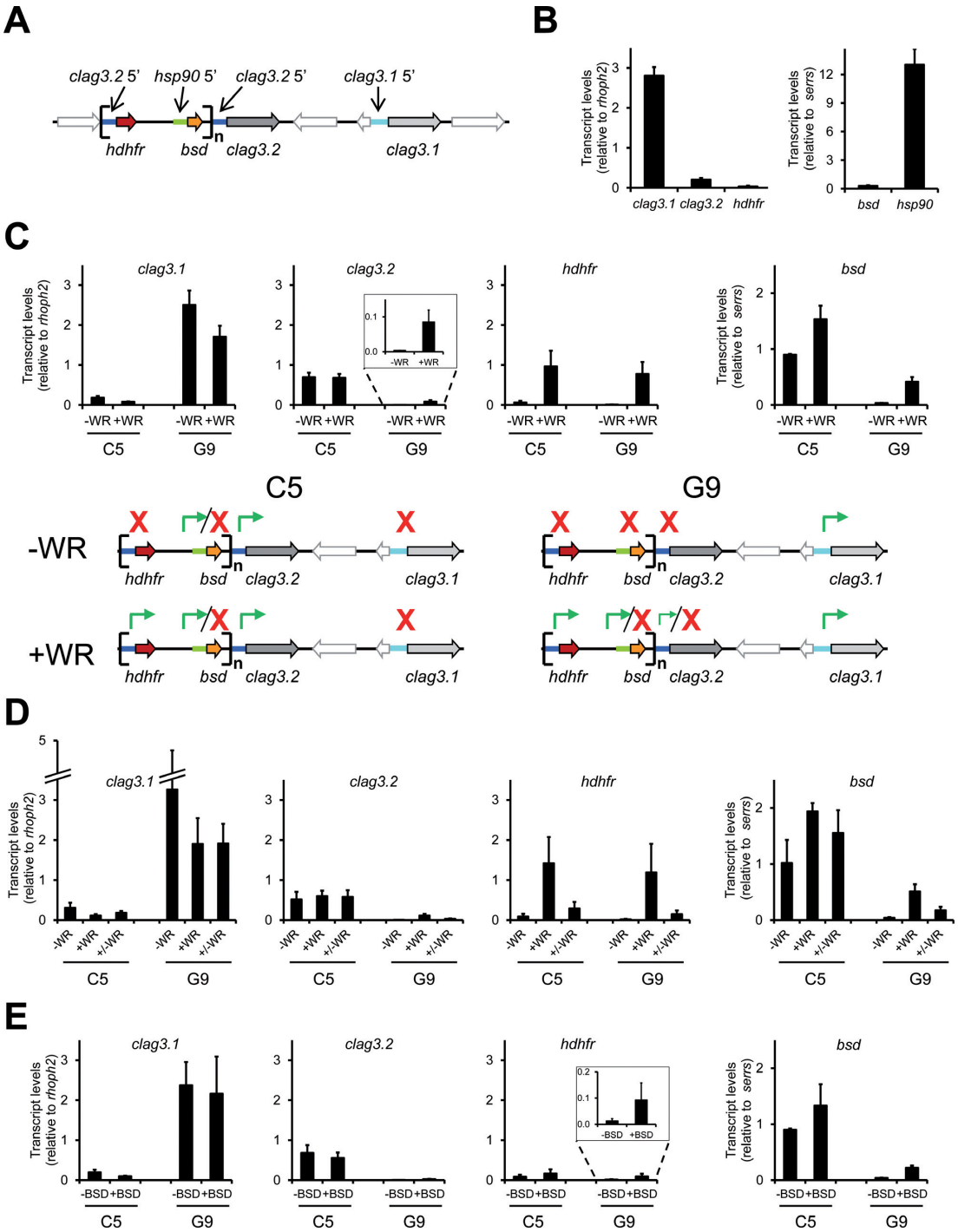

Figure 3. Characterization of parasite lines with additional chromosomal *clag3* promoters integrated in the vicinity of endogenous *clag3* loci. (**A**) Schematic (not to scale) of the *clag3* loci in transgenic parasites that have integrated the pHBc3.2 plasmid by single homologous recombination with the *clag3.2* upstream region. Genes or annotated ncRNAs are indicated by block arrows. The 5′ promoter regions of the genes analyzed are highlighted. The area in brackets is the integrated plasmid, which is repeated several times because multiple copies are integrated (Supplementary Figure S2). (**B**) Transcriptional analysis (schizont stage) of 3D7-A with integrated pHBc3.2 (unsubcloned population). Values are the average of two independent biological replicates, with SEM. (**C**) Transcriptional analysis of C5 and G9 subclones selected for 3 weeks with WR99210 or grown in parallel in the absence of drug. Values are the average of three independent biological replicates (independent drug selections), with SEM. The schematic indicates the active (green arrow) or silenced (red cross) state of the genes in C5 and G9 under the different conditions tested. Activation of only some of the integrated copies (*bsd* gene) or activation in only a small fraction of the parasites (endogenous *clag3.2* promoter) is indicated by green arrow/red cross. It is likely that active expression of *hdhfr* after WR99210 selection also occurs in only some of the copies of the integrated plasmid. (**D**) Transcriptional analysis of C5 and G9 transgenic subclones grown in parallel for 6 weeks with no drug, with WR99210 or 3 weeks with WR99210 followed by 3 weeks without drug (+/− WR). Values are the average of two independent biological replicates (independent drug selections), with SEM. (**E**) Transcriptional analysis of C5 and G9 transgenic subclones grown in parallel for 3 weeks in the presence or absence of BSD. Results are the average of two independent biological replicates (independent drug selections), with SEM.

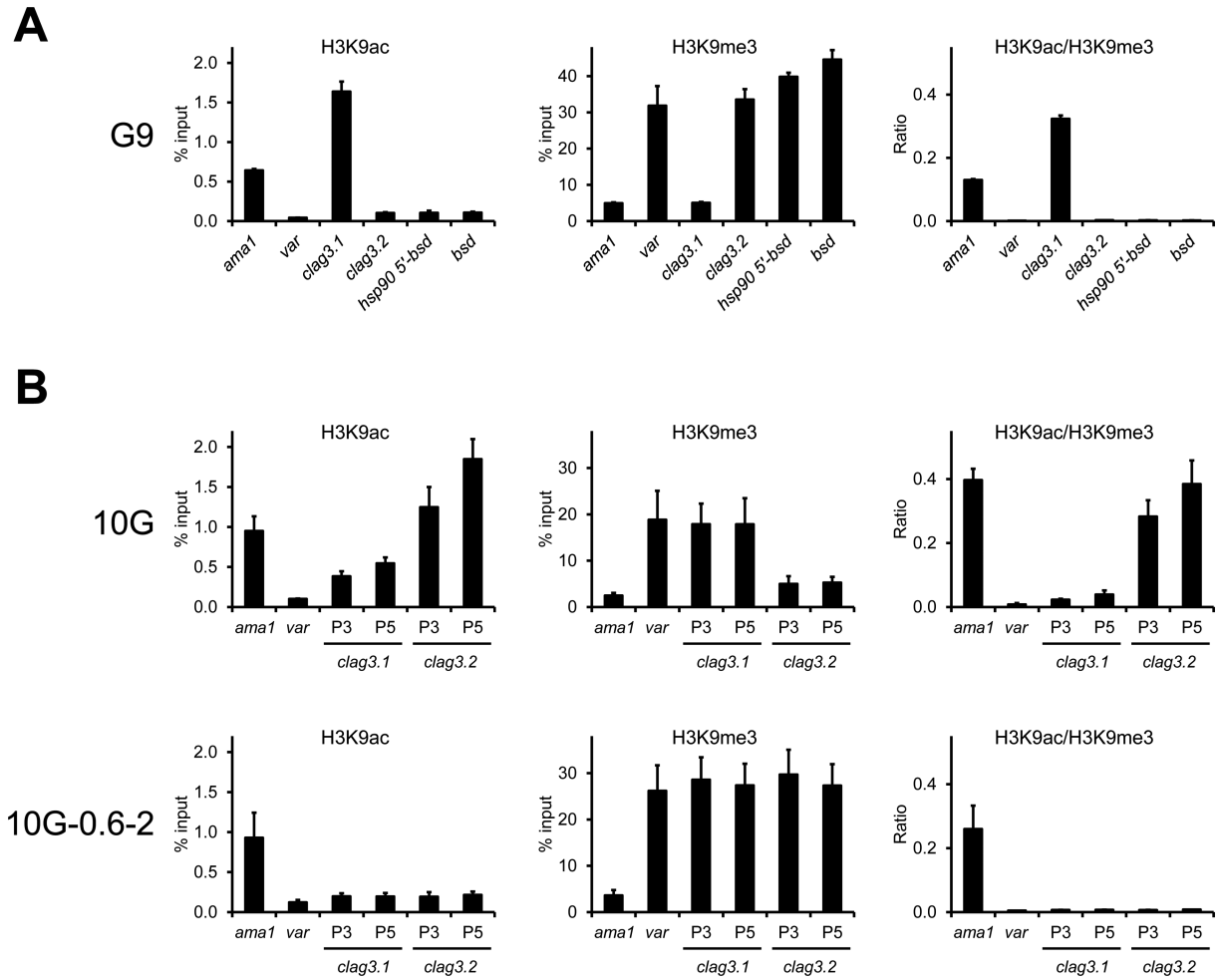

Figure 4. ChIP analysis of the *clag3* loci in transgenic and BSD-selected parasite lines. (**A**) ChIP analysis of H3K9me3 and H3K9ac in the transgenic G9 parasite line (late trophozoite stage) grown in the absence of drug. *ama1* and the *var* gene PF3D7_1240300 are euchromatin and heterochromatin controls, respectively. Primers at the ORF of *clag3.1* or *clag3.2* (near the 5′ end, primer pair 5) were used for the analysis of *clag3* genes. In this parasite line *clag3.1* is expressed and *clag3.2* is silenced; hence, these genes are additional controls for the euchromatic and heterochromatic states, respectively. The *bsd* loci were analyzed with a forward primer at the *hsp90* 5′ region and a reverse primer at the *bsd* ORF (*hsp90* 5′-*bsd*) and with primers at the *bsd* ORF (*bsd*). Results are expressed as percentage of immunoprecipitated chromatin relative to the input sample (% input). Results are also expressed as the ratio of % input for H3K9ac divided by % input for H3K9me3. High values of this ratio indicate a euchromatic state, as in the *ama1* control, whereas low values indicate an heterochromatic state, as in the *var* control. Values are the average of two independent biological replicates, with SEM. (**B**) ChIP analysis of the *clag3* loci in the parasite line 10G (*clag3.2* is expressed) and the BSD-selected line10G-0.6–2 (the two *clag3* genes are silenced) at the schizont stage. For each *clag3* gene, primer pair 3 (P3) is near the transcription start sites whereas primer pair 5 (P5) is near the 5′ end of the ORF. Values are the average of three independent biological replicates, with SEM. Both H3K9ac and H3K9me3 levels at the *clag3.2* locus were significantly different between 10G and 10G-0.6–2 ($P < 0.05$ using a *t*-test). Similar results were obtained using a different protocol and alternative antibodies widely used for ChIP analysis of malaria parasites (Supplementary Figure S5).

strated that toxic compounds that require CLAG3s to enter infected erythrocytes can select for parasites in which both *clag3* promoters are simultaneously silenced (9,10). Here we analyzed the 10G-0.6–2 line, in which both *clag3* genes are silenced after selection with a high concentration of BSD (9), and a 10G culture maintained in parallel without BSD as a control. In 10G-0.6–2 the *clag3.2* locus is enriched in H3K9me3 and shows low levels of H3K9ac, similar to the heterochromatic control (*var*), whereas in 10G, where *clag3.2* is active, the opposite pattern is observed, with H3K9me3 and H3K9ac levels similar to the euchromatic gene *ama1*. On the other hand, the levels of H3K9me3 and H3K9ac were similar between 10G and 10G-0.6–2 at the *clag3.1* locus, which is silenced in both lines (Figure 4B and Supplementary Figure S5). These results demonstrate that the same mechanism based on H3K9me3 mediates silencing of only one *clag3* gene in unselected parasite populations (16,17) and simultaneous silencing of the two paralogous genes in cultures under strong selective pressure. This is in contrast to the results for an analogous parasite line developed at a different laboratory, for which double *clag3* silencing was proposed to operate via a different epigenetic mechanism involving reduced levels of activating histone marks, as observed here for H3K9ac, but not the heterochromatin mark H3K9me3 (10).

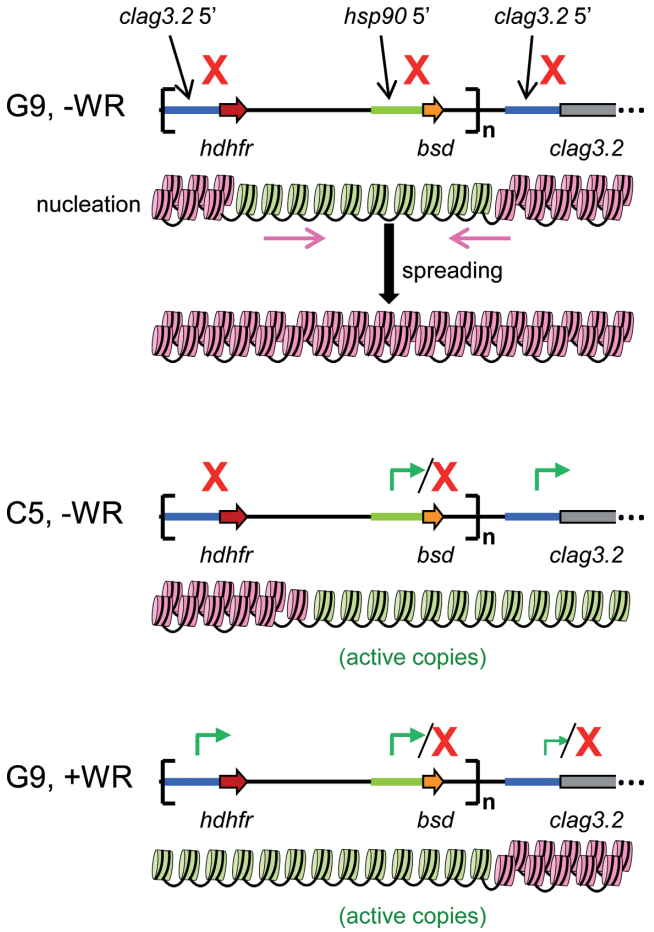

Figure 5. Hypothetical model for chromatin nucleation and spreading at *clag3* promoters. In the transgenic subclone G9 grown in the absence of any drug (−WR), where both the endogenous *clag3.2* promoter and the *clag3.2* promoter controlling *hdhfr* are silenced, all copies of the *hsp90* constitutive promoter controlling *bsd* expression are repressed. However, one or more copies of the *hsp90* promoter are active in the C5 subclone (even in the absence of any drug), likely including the copy that is closer to the endogenous *clag3.2* promoter (active in this subclone). Likewise, some copies of the *hsp90* promoter are active in G9 selected with WR99210 (+WR), in which at least some of the flanking *clag3.2* promoters controlling *hdhfr* expression are active. Considering that the *hsp90* promoter sequence does not have the ability to nucleate heterochromatin, and silencing of *clag3.2* promoters is mediated by H3K9me3-based heterochromatin, the most plausible interpretation for these observations is that *hsp90* silencing is mediated by heterochromatin spreading from the adjacent loci. In support of this view, we directly observed the H3K9me3 mark at the silenced *bsd* locus in G9 (−WR). The schematic shows the predicted chromatin organization in the different parasite lines. Densely-packed nucleosomes with pink tones represent heterochromatin, characterized by H3K9me3, whereas spaced green nucleosomes represent transcriptionally-permissive euchromatin. The pink arrows indicate the possible directions of heterochromatin spreading.

Expression of the ncRNA PF3D7_0302400 correlates with *clag3.1* expression

ncRNAs play important roles in the regulation of gene expression in many eukaryotes, including *P. falciparum* (43). Interestingly, the first published RNA-seq study in *P. falciparum* reported two ncRNAs in chromosome 3 and both are located between the two *clag3* genes (44). One of these two ncRNAs corresponds to a *var*-pseudogene, whereas the other, PF3D7_0302400 (previous ID PF03TR002, abbreviated TR2), is an authentic >1 kb long ncRNA transcribed from the opposite strand than *clag3* genes ((45) and unpublished data available at PlasmoDB: 'Sanger_FRT' dataset by Otto *et al.*) (Figure 6A). Since published data indicates that TR2 is predominantly expressed at the schizont stage (44), similar to *clag* genes, we initially focused our analysis of TR2 expression on this life cycle stage. In two independent comparisons, we found that schizonts from parasite clones expressing *clag3.1* show much higher TR2 transcript levels than isogenic clones that express *clag3.2* (Figure 6B). Although TR2 transcript levels are rather low, as occurs for many intergenic long ncRNAs also in other species (46), they are clearly within the limits of detection of RT-qPCR. Additional support to the idea that TR2 expression in schizonts correlates with *clag3.1* expression came from the analysis of non-transfected 10G parasites selected with increasing BSD concentrations, which results in altered *clag3* expression patterns (9). In parallel to the switch from predominant *clag3.2* expression to predominant *clag3.1* expression associated with selection with BSD at low concentrations, a dramatic increase in the levels of TR2 transcripts was observed (Figure 6C). However, in parasites selected with a high concentration of the drug (2 μg/ml), in which both *clag3* genes are simultaneously silenced, TR2 is also silenced (Figure 6C). This result indicates that TR2 expression correlates with *clag3.1* expression, rather than with *clag3.2* silencing. Furthermore, 10G parasites transfected with a plasmid containing a BSD resistance cassette and drug selected also increase TR2 expression (Figure 6C) in parallel to their switch from *clag3.2* to *clag3.1* expression (9).

Next we performed a time-course analysis of TR2 expression in the 3D7 parasite stocks 3D7-A and 3D7-B (35) and the 3D7-A subclones 10G and 1.2B (Figure 6D). Schizonts (40–45 h) of the 3D7-A line, which consist of a mixture of parasites expressing *clag3.1* or *clag3.2* (24), expresses TR2 at levels intermediate between 10G and 1.2B, as expected. On the other hand, the vast majority of parasites in the 3D7-B population express *clag3.1* (24), and correspondingly 3D7-B schizonts expresses TR2 at high levels similar to 1.2B. These experiments also revealed expression of TR2 mainly in schizonts in 3D7-A and subclones, consistent with previous data (44), but in 3D7-B high TR2 expression was also observed in 10–15 h rings (Figure 6D). Of note, the neighbor *var* pseudogene PF3D7_0302300 (Figure 6A) is expressed in rings at much higher levels in 3D7-B than in 3D7-A or its subclones, where it is completely silenced (7). These results suggest that TR2 expression in schizonts correlates with *clag3.1* expression, whereas expression of this ncRNA in rings correlates with expression of the neighbor *var* pseudogene.

DISCUSSION

Two different *P. falciparum* gene families show mutually exclusive expression: *var* and *clag3*. In both gene families, expression switches involve *in situ* silencing and activation regulated at the epigenetic level in the absence of DNA rearrangements (8,16,24,47). Here we investigated for the first time the basic properties of mutual exclusion in the *clag3*

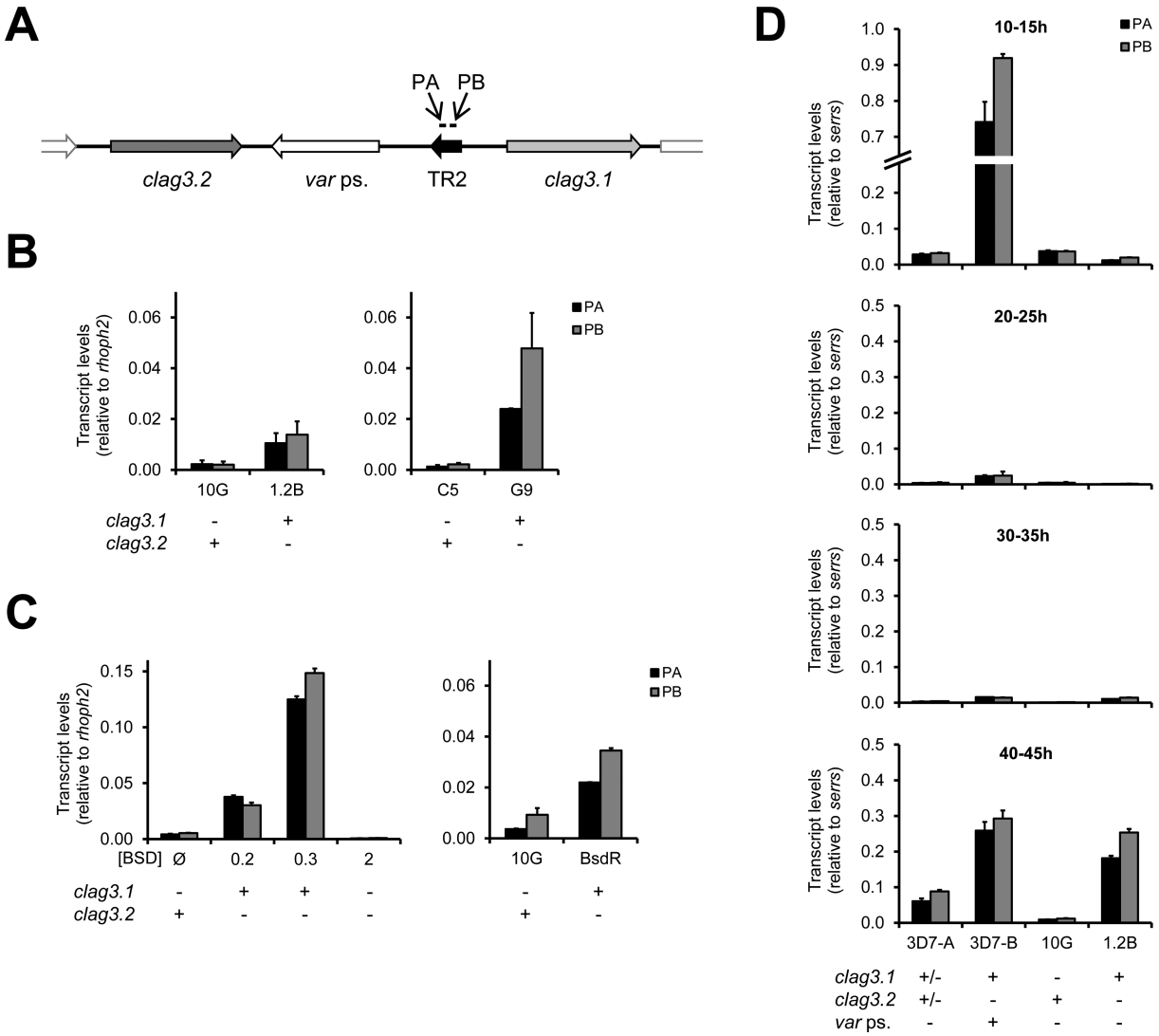

Figure 6. Transcriptional analysis of the ncRNA TR2 in parasite lines with different *clag3* expression patterns. (**A**) Schematic of the *clag3* loci with the *var* pseudogene PF3D7_0302300 (*var* ps.) and the ncRNA PF3D7_0302400 (TR2) between the two *clag3* genes. The position of the primers used for the transcriptional analysis of TR2, primer pairs A (PA) and B (PB), is indicated. (**B**) RT-qPCR comparison of TR2 transcript levels (relative to *rhoph2*) in schizonts of the isogenic wild-type subclones 10G and 1.2B (left), or the transgenic subclones C5 and G9 (right). The expression status of *clag3.1* and *clag3.2*, as determined here or previously (16,24), is indicated. Values are the average of three independent biological replicates, with SEM. TR2 transcript levels were significantly higher in parasite lines that express *clag3.1* than in those that express *clag3.2* ($P < 0.05$ using a paired *t*-test). (**C**) TR2 expression at the schizont stage in cultures of the 10G subclone grown in the absence of drugs (Ø) or selected with BSD at the concentrations indicated (in μg/ml). The bar chart on the right shows the comparison between untransfected 10G and a 10G culture transfected with plasmid BsdR (containing a *bsd* selectable marker under the control of a constitutive promoter) and maintained under BSD selection for 8 weeks (9). The *clag3* expression status indicated for each selected or control line was previously determined (9). Values are the average of reactions performed in triplicate, with SEM. (**D**) Time-course analysis of TR2 expression at different stages of the asexual life cycle (indicated in hours post-invasion). The expression status of *clag3* genes and the *var* pseudogene PF3D7_0302300 in each parasite line has been determined before (7,24). Values are the average of reactions performed in triplicate, with SEM, and representative of independent technical or biological replicates.

gene family and unexpectedly found that mutually exclusive expression of *clag3* genes is strongly favored but is not strict. We also gained insight into the molecular mechanisms controlling this type of expression.

First we interrogated the system by introducing an additional *clag3* promoter in a stable episome. To obtain parasites stably carrying the episome we used the *bsd* selectable marker and selection with BSD, a drug to which parasites can acquire resistance either via the *bsd* transgene or via changes in endogenous *clag3* expression (9,10). We found that three separate but inter-related mechanisms shaped transcript level patterns in this system: changes in the expression of endogenous *clag3* genes to restrict drug access, changes in episome copy number and transcriptional regulation of the genes in the episome. The relative use of the different mechanisms varied between the different transgenic lines and reflected dynamic selection of parasites, complicating the interpretation of the results. Furthermore, we found that some plasmids containing two expression cassettes are prone to transcriptional silencing independently

of the presence of a clonally variant promoter able to drive heterochromatin formation. In spite of these difficulties, we could unambiguously determine that active expression of an episomal *clag3* promoter (under selective pressure) does not preclude expression of endogenous *clag3* genes. This is in contrast to the *var* family, for which activation of episomal *var* regulatory regions can result in silencing of all endogenous *var* genes (31,32,48). Our results for episomal *clag3* promoters are more similar to the results of a recent study that analyzed expression in an episomal context of promoters from the clonally variant gene families *rif*, *stevor*, *pfmc-2tm* and *phist* (48). Active transcription from episomal promoters of these gene families did not repress the expression of endogenous members of the respective families, as expected for gene families that are clonally variant but are not subject to mutually exclusive expression. However, these promoters were reported to be active by default in an episomal context (48), in contrast to observations for episomal *clag3* promoters in our transgenic lines. Nonetheless, these results are not directly comparable to ours because in that study the transcriptional state of the episomal promoters was not defined based on comparison with the endogenous promoters.

To overcome the problems associated with episomal constructs, we integrated a construct containing a *clag3* promoter driving the expression of a selectable marker in the parasite genome, next to the endogenous *clag3* loci. In this system the integrated plasmid is stable in the absence of selection so we could avoid the constant use of BSD and the effects of this drug on endogenous *clag3* expression. In this transgenic parasite line the additional chromosomal *clag3* promoters are silenced by default, a pattern consistent with mutually exclusive expression, but using drug pressure (WR99210) we could select for parasites in which they are activated. Importantly, activation of the additional *clag3* promoters did not affect the expression of endogenous *clag3* genes. Hence, again in contrast to analogous experiments with *var* genes (31,32), our results demonstrate that under selective pressure *clag3* promoters can break mutually exclusive expression, although this occurs only transiently because expression of a single *clag3* promoter is quickly recovered after removing the selective pressure. These findings indicate that mutually exclusive expression of *clag3* genes is not strict, such that parasites that simultaneously express more than one *clag3* promoter can exist, but mutual exclusion is strongly favored. We and others have previously demonstrated that parasites that keep the two *clag3* promoters simultaneously silenced can also exist but recover expression of a single *clag3* gene after removing selective pressure (9,10). Based on all of these observations, we propose a scenario where probabilistic molecular interactions that remain to be characterized favor expression of a single *clag3* gene, but other expression patterns are possible and occur at low frequency within parasite populations. Under strong selective pressure favoring survival of parasites expressing none or more than one *clag3* promoters, such parasites can be selected and become predominant in the population. However, the erythrocyte permeability phenotypes resulting from these expression patterns may have a fitness cost in the absence of selection, in addition to not being favored by regulatory mechanisms. While the characteristics of mutual exclusion appear to be different between *clag3* and *var* genes, it is possible that the scenario proposed for *clag3* may also apply for *var*. *var* mutual exclusion may be more strict because of the role of this gene family in immune evasion, but it is important to note that there are examples of wild-type parasites expressing more than one *var* gene simultaneously (49–51). Furthermore, absence of PfEMP-1 expression does not pose a fitness cost for the parasite under culture conditions, whereas absence of CLAG3 results in severely reduced growth (9). This could explain why drug pressure in transgenic parasites with a *var* promoter controlling a drug resistance gene selects for parasites in which all endogenous *var* genes are silenced whereas analogous experiments with a *clag3* promoter select for parasites expressing more than one *clag3* promoter simultaneously, although these different results may also be explained by a more strict regulation of mutual exclusion in *var*.

The *clag3* regulatory region used in our transgenic lines (the full *clag3.2* 5′ intergenic region) confers correct temporal regulation of expression (Supplementary Figure S4 and ref. (16)) and appears to be sufficient to mediate transcriptional silencing, indicating that *clag3* promoters play a predominant role in controlling the expression of this gene family. In our first set of experiments with *clag3* promoters in an episomal context we could not establish whether the default state of extra copies of *clag3* promoters is silenced, because some plasmids with two promoters appear to be largely silenced independently of the presence of a promoter with the ability to nucleate heterochromatin (Figure 1). This may be attributable to plasmid architecture and/or to episomal state. Evidence for a default silenced state for extra copies of *clag3* promoters comes from experiments using a plasmid with a different organization (pHBc3.2, with the *clag3.2* promoter located downstream of the *bsd* expression cassette), analogous to plasmids used to define the default silenced state of *var* promoters (32) and similar to plasmids used to analyze other clonally variant promoters that were reported to be active by default in these plasmids (48). When pHBc3.2 is integrated in the genome, *clag3.2* promoters have the ability to nucleate heterochromatin that spreads into neighbor genes (see below). These results support the idea that *clag3* promoters are sufficient to mediate transcriptional silencing, but it is important to note that even in the integrated construct each transgenic copy of the *clag3.2* promoter is paired with another promoter. In the case of *var* genes the intron (or an alternative paired promoter) is necessary for a *var* promoter to enter the mutually exclusive expression program, such that *var* promoters missing this genetic element are active by default (30,39,40,52). We cannot formally exclude the possibility that an analogous as yet unidentified regulatory element is necessary for the silencing of *clag3* genes, and that the paired promoters in our constructs mimic the role of this hypothetical element. We also considered the possibility that expression of a functional CLAG3 protein plays a role in the regulation of mutual exclusion, but this is highly unlikely because a previous study demonstrated that a *clag3.2* gene truncated at the beginning of its ORF still enters the mutual exclusion program and its activation precludes expression of *clag3.1* (17). This observation strongly indicates that mutually exclusive

expression of *clag3* genes is independent of protein production, similar to *var* genes (31,32).

The actual counting mechanism that determines that one and only one gene of a family under mutually exclusive expression is active in individual cells remains elusive even for extensively studied gene families such as *var* or olfactory receptors (8,20). However, some of the molecular factors that are necessary to maintain the silenced or the active state have been identified for *var* genes (8,12), while the factors involved in *clag3* regulation remain largely unexplored. Given that *var* and *clag3* are gene families of very different size and encode proteins with very different biological functions, it is possible that different mechanisms regulate mutual exclusion in the two families. Here we found that *clag3* epigenetic silencing relies on H3K9me3-based heterochromatin even in parasite lines where the two *clag3* genes are simultaneously silenced. This result suggests that *clag3* mutual exclusion depends on unknown molecular interactions that favor heterochromatin formation in all but one of the *clag3* promoters, but the mechanism operates in a non-strict manner that allows heterochromatin formation in alternative numbers of *clag3* promoters (e.g. all) in small selectable subpopulations of parasites. Additionally, we show that expression of the TR2 ncRNA, located between the two *clag3* genes, correlates with expression of *clag3.1*. Future studies will need to establish the direction of causality in the coordinated expression of TR2 and *clag3.1*, as TR2 expression may either be a cause or a consequence of *clag3.1* expression. In model eukaryotes, ncRNAs are involved in the regulation of many epigenetic processes and in *P. falciparum* antisense ncRNAs originating from *var* introns appear to be involved in the regulation of *var* expression (53). Furthermore, ncRNAs originating from subtelomeric repeats near the *var* loci have been identified and also proposed to participate in *var* regulation (54,55), although this hypothesis awaits experimental validation.

Our experiments with integrated constructs revealed that a constitutive promoter (*hsp90* upstream region controlling the *bsd* selectable marker) can be silenced by spreading of heterochromatin from the neighbor loci (Figure 5). While the ability of H3K9me3-based heterochromatin to spread into adjacent regions is well documented in other organisms (56), this is to our knowledge the first direct demonstration of this phenomenon in *P. falciparum*. Previous studies had shown silencing of a constitutive promoter located next to a *var* promoter in an episomal context, and increased transcription from the constitutive promoter upon activation of the neighbor *var* promoter (32). However, these results could be explained by reversible heterochromatin formation and spreading, as proposed by the authors, or alternatively by changes in plasmid copy number or by silencing related with episomal state. In contrast, we provide clear evidence for heterochromatin spreading in a stable chromosomal context, with isogenic parasite lines with different epigenetic states of neighbor genes providing appropriate controls. Furthermore, our ChIP analysis directly demonstrates that the heterochromatin mark H3K9me3 is present in the silenced *bsd* gene, as expected if silencing occurs by heterochromatin spreading. The genome of *P. falciparum* is organized in small transcriptional units corresponding to individual genes, such that euchromatin-heterochromatin transitions at a clonally variant locus do not affect the state of the neighbor genes (7,16,40,57). In the light of our results demonstrating that *P. falciparum* silenced chromatin has the potential to spread into neighbor regions, the independent regulation of endogenous neighbor clonally variant genes implies the existence of barrier elements or insulators (58) that limit the spreading of heterochromatin.

Altogether, our findings unravel some main features of *clag3* mutually exclusive expression. CLAG3 expression patterns are shaped by a combination of the effects of these proteins on parasite fitness and non-deterministic epigenetic mechanisms favoring expression of a single *clag3* gene in a non-strict manner. These mechanisms appear to operate by preventing heterochromatin-mediated silencing in only one of the *clag3* promoters. A stochastic mechanism that results in expression of a single *clag3* gene in the vast majority of parasites but allows alternative patterns in small selectable subpopulations of parasites can increase the adaptability of parasite populations and broaden the range of environmental challenges to which they can survive. Understanding mutually exclusive expression of *clag3* genes is important *per se* because CLAG3 proteins play a fundamental role in parasite biology, but the limited size of this gene family also makes it an attractive model to study conserved features of mutually exclusive gene expression, a largely enigmatic phenomenon repeatedly observed in evolution.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Till S. Voss (Swiss TPH) for providing the plasmid pHBupsC^{R} and to Jacobus Pharmaceuticals for providing the drug WR99210. We are also grateful to PlasmoDB for unpublished data deposited by T.D. Otto and C.I. Newbold (Sanger_FRT dataset).

FUNDING

Spanish Ministry of Science and Innovation (MICINN) [SAF2010-20111 to A.C.; BIO2009-09776 to LRdP]; Spanish Ministry of Economy and Competitiveness (MINECO) and European Regional Development Fund (ERDF, European Union) [SAF2013-43601-R to A.C.]; Secretary for Universities and Research under the Department of Economy and Knowledge of the Government of Catalonia [2014 SGR 485 to Malaria Epigenetics lab, N.R-G., C.B., S.M-M., A.C.]; Secretary for Universities and Research (Catalan Government) Postdoctoral Fellowship [2011-BP-B 00060 to C.B.]. Funding for open access charge: MINECO [SAF2013-43601-R to A.C.].
Conflict of interest statement. None declared.

REFERENCES

1. Bozdech,Z., Llinas,M., Pulliam,B.L., Wong,E.D., Zhu,J. and DeRisi,J.L. (2003) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.*, **1**, E5.

2. Le Roch,K.G., Zhou,Y., Blair,P.L., Grainger,M., Moch,J.K., Haynes,J.D., De La Vega,P., Holder,A.A., Batalov,S., Carucci,D.J. *et al.* (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*, **301**, 1503–1508.
3. Painter,H.J., Campbell,T.L. and Llinas,M. (2011) The Apicomplexan AP2 family: integral factors regulating *Plasmodium* development. *Mol. Biochem. Parasitol.*, **176**, 1–7.
4. Bartfai,R., Hoeijmakers,W.A., Salcedo-Amaya,A.M., Smits,A.H., Janssen-Megens,E., Kaan,A., Treeck,M., Gilberger,T.W., Francoijs,K.J. and Stunnenberg,H.G. (2010) H2A.Z demarcates intergenic regions of the *Plasmodium falciparum* epigenome that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog.*, **6**, e1001223.
5. Gupta,A.P., Chin,W.H., Zhu,L., Mok,S., Luah,Y.H., Lim,E.H. and Bozdech,Z. (2013) Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite *Plasmodium falciparum*. *PLoS Pathog.*, **9**, e1003170.
6. Cortés,A., Crowley,V.M., Vaquero,A. and Voss,T.S. (2012) A view on the role of epigenetics in the biology of malaria parasites. *PLoS Pathog.*, **8**, e1002943.
7. Rovira-Graells,N., Gupta,A.P., Planet,E., Crowley,V.M., Mok,S., Ribas de Pouplana,L., Preiser,P.R., Bozdech,Z. and Cortés,A. (2012) Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res.*, **22**, 925–938.
8. Guizetti,J. and Scherf,A. (2013) Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell Microbiol.*, **15**, 718–726.
9. Mira-Martínez,S., Rovira-Graells,N., Crowley,V.M., Altenhofen,L.M., Llinás,M. and Cortés,A. (2013) Epigenetic switches in *clag3* genes mediate blasticidin S resistance in malaria parasites. *Cell Microbiol.*, **15**, 1913–1923.
10. Sharma,P., Wollenberg,K., Sellers,M., Zainabadi,K., Galinsky,K., Moss,E., Nguitragool,W., Neafsey,D. and Desai,S.A. (2013) An epigenetic antimalarial resistance mechanism involving parasite genes linked to nutrient uptake. *J. Biol. Chem.*, **288**, 19429–19440.
11. Kafsack,B.F., Rovira-Graells,N., Clark,T.G., Bancells,C., Crowley,V.M., Campino,S.G., Williams,A.E., Drought,L.G., Kwiatkowski,D.P., Baker,D.A. *et al.* (2014) A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature*, **507**, 248–252.
12. Duffy,M.F., Selvarajah,S.A., Josling,G.A. and Petter,M. (2014) Epigenetic regulation of the *Plasmodium falciparum* genome. *Brief. Funct. Genomics*, **13**, 203–216.
13. Lyons,D.B. and Lomvardas,S. (2014) Repressive histone methylation: a case study in deterministic versus stochastic gene regulation. *Biochim. Biophys. Acta*, **1839**, 1373–1384.
14. Chookajorn,T., Dzikowski,R., Frank,M., Li,F., Jiwani,A.Z., Hartl,D.L. and Deitsch,K.W. (2007) Epigenetic memory at malaria virulence genes. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 899–902.
15. Lopez-Rubio,J.J., Gontijo,A.M., Nunes,M.C., Issar,N., Hernandez Rivas,R. and 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.
16. Crowley,V.M., Rovira-Graells,N., de Pouplana,L.R. and 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.
17. Comeaux,C.A., Coleman,B.I., Bei,A.K., Whitehurst,N. and Duraisingh,M.T. (2011) Functional analysis of epigenetic regulation of tandem *RhopH1/clag* genes reveals a role in *Plasmodium falciparum* growth. *Mol. Microbiol.*, **80**, 378–390.
18. Jiang,L., Lopez-Barragan,M.J., Jiang,H., Mu,J., Gaur,D., Zhao,K., Felsenfeld,G. and Miller,L.H. (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.
19. Lopez-Rubio,J.J., Mancio-Silva,L. and 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.
20. Magklara,A. and Lomvardas,S. (2013) Stochastic gene expression in mammals: lessons from olfaction. *Trends Cell Biol.*, **23**, 449–456.
21. Deitsch,K.W., Lukehart,S.A. and Stringer,J.R. (2009) Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nat. Rev. Microbiol.*, **7**, 493–503.
22. Kyes,S., Horrocks,P. and Newbold,C. (2001) Antigenic variation at the infected red cell surface in malaria. *Annu. Rev. Microbiol.*, **55**, 673–707.
23. Gupta,A., Thiruvengadam,G. and Desai,S.A. (2015) The conserved *clag* multigene family of malaria parasites: essential roles in host-pathogen interaction. *Drug Resist. Updat.*, **18**, 47–54.
24. Cortés,A., Carret,C., Kaneko,O., Yim Lim,B.Y., Ivens,A. and Holder,A.A. (2007) Epigenetic silencing of *Plasmodium falciparum* genes linked to erythrocyte invasion. *PLoS Pathog.*, **3**, e107.
25. Nguitragool,W., Bokhari,A.A., Pillai,A.D., Rayavara,K., Sharma,P., Turpin,B., Aravind,L. and Desai,S.A. (2011) Malaria parasite *clag3* genes determine channel-mediated nutrient uptake by infected red blood cells. *Cell*, **145**, 665–677.
26. Pillai,A.D., Nguitragool,W., Lyko,B., Dolinta,K., Butler,M.M., Nguyen,S.T., Peet,N.P., Bowlin,T.L. and Desai,S.A. (2012) Solute restriction reveals an essential role for *clag3*-associated channels in malaria parasite nutrient acquisition. *Mol. Pharmacol.*, **82**, 1104–1114.
27. Desai,S.A. (2014) Why do malaria parasites increase host erythrocyte permeability? *Trends Parasitol.*, **30**, 151–159.
28. Cortés,A. (2008) Switching *Plasmodium falciparum* genes on and off for erythrocyte invasion. *Trends Parasitol.*, **24**, 517–524.
29. Chookajorn,T., Ponsuwanna,P. and Cui,L. (2008) Mutually exclusive *var* gene expression in the malaria parasite: multiple layers of regulation. *Trends Parasitol.*, **24**, 455–461.
30. Deitsch,K.W., Calderwood,M.S. and Wellems,T.E. (2001) Malaria. cooperative silencing elements in *var* genes. *Nature*, **412**, 875–876.
31. Dzikowski,R., Frank,M. and Deitsch,K. (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog.*, **2**, e22.
32. Voss,T.S., Healer,J., Marty,A.J., Duffy,M.F., Thompson,J.K., Beeson,J.G., Reeder,J.C., Crabb,B.S. and Cowman,A.F. (2006) A *var* gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature*, **439**, 1004–1008.
33. Avraham,I., Schreier,J. and Dzikowski,R. (2012) Insulator-like pairing elements regulate silencing and mutually exclusive expression in the malaria parasite *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E3678–E3686.
34. Brancucci,N.M., Witmer,K., Schmid,C.D., Flueck,C. and Voss,T.S. (2012) Identification of a cis-acting DNA-protein interaction implicated in singular *var* gene choice in *Plasmodium falciparum*. *Cell Microbiol.*, **14**, 1836–1848.
35. Cortés,A., Benet,A., Cooke,B.M., Barnwell,J.W. and Reeder,J.C. (2004) Ability of *Plasmodium falciparum* to invade Southeast Asian ovalocytes varies between parasite lines. *Blood*, **104**, 2961–2966.
36. Cortés,A. (2005) A chimeric *Plasmodium falciparum Pfnbp2b/Pfnbp2a* gene originated during asexual growth. *Int. J. Parasitol.*, **35**, 125–130.
37. Crabb,B.S., Rug,M., Gilberger,T.W., Thompson,J.K., Triglia,T., Maier,A.G. and Cowman,A.F. (2004) Transfection of the human malaria parasite *Plasmodium falciparum*. *Methods Mol. Biol.*, **270**, 263–276.
38. O'Donnell,R.A., Freitas-Junior,L.H., Preiser,P.R., Williamson,D.H., Duraisingh,M., McElwain,T.F., Scherf,A., Cowman,A.F. and Crabb,B.S. (2002) A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of *Plasmodium falciparum* chromosomes. *EMBO J.*, **21**, 1231–1239.
39. Dzikowski,R., Li,F., Amulic,B., Eisberg,A., Frank,M., Patel,S., Wellems,T.E. and Deitsch,K.W. (2007) Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Rep.*, **8**, 959–965.
40. Frank,M. and Deitsch,K. (2006) Activation, silencing and mutually exclusive expression within the *var* gene family of *Plasmodium falciparum*. *Int. J. Parasitol.*, **36**, 975–985.
41. Petter,M., Selvarajah,S.A., Lee,C.C., Chin,W.H., Gupta,A.P., Bozdech,Z., Brown,G.V. and Duffy,M.F. (2013) H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy *var* gene promoters in the malaria parasite *Plasmodium falciparum*. *Mol. Microbiol.*, **87**, 1167–1182.
42. Flueck,C., Bartfai,R., Volz,J., Niederwieser,I., Salcedo-Amaya,A.M., Alako,B.T., Ehlgen,F., Ralph,S.A., Cowman,A.F., Bozdech,Z. *et al.*

(2009) *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog.*, **5**, e1000569.
43. Vembar,S.S., Scherf,A. and Siegel,T.N. (2014) Noncoding RNAs as emerging regulators of *Plasmodium falciparum* virulence gene expression. *Curr. Opin. Microbiol.*, **20**, 153–161.
44. Otto,T.D., Wilinski,D., Assefa,S., Keane,T.M., Sarry,L.R., Bohme,U., Lemieux,J., Barrell,B., Pain,A., Berriman,M. *et al.* (2010) New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. *Mol. Microbiol.*, **76**, 12–24.
45. Siegel,T.N., Hon,C.C., Zhang,Q., Lopez-Rubio,J.J., Scheidig-Benatar,C., Martins,R.M., Sismeiro,O., Coppee,J.Y. and Scherf,A. (2014) Strand-specific RNA-Seq reveals widespread and developmentally regulated transcription of natural antisense transcripts in *Plasmodium falciparum*. *BMC Genomics*, **15**, 150.
46. Cabili,M.N., Trapnell,C., Goff,L., Koziol,M., Tazon-Vega,B., Regev,A. and Rinn,J.L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.*, **25**, 1915–1927.
47. Scherf,A., Hernandez-Rivas,R., Buffet,P., Bottius,E., Benatar,C., Pouvelle,B., Gysin,J. and Lanzer,M. (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*. *EMBO J.*, **17**, 5418–5426.
48. Witmer,K., Schmid,C.D., Brancucci,N.M., Luah,Y.H., Preiser,P.R., Bozdech,Z. and Voss,T.S. (2012) Analysis of subtelomeric virulence gene families in *Plasmodium falciparum* by comparative transcriptional profiling. *Mol. Microbiol.*, **84**, 243–259.
49. Joergensen,L., Bengtsson,D.C., Bengtsson,A., Ronander,E., Berger,S.S., Turner,L., Dalgaard,M.B., Cham,G.K., Victor,M.E., Lavstsen,T. *et al.* (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.
50. Merrick,C.J., Jiang,R.H., Skillman,K.M., Samarakoon,U., Moore,R.M., Dzikowski,R., Ferdig,M.T. and Duraisingh,M.T. (2015) Functional analysis of sirtuin genes in multiple *Plasmodium falciparum* Strains. *PLoS One*, **10**, e0118865.
51. Duffy,M.F., Brown,G.V., Basuki,W., Krejany,E.O., Noviyanti,R., Cowman,A.F. and Reeder,J.C. (2002) Transcription of multiple var genes by individual, trophozoite-stage Plasmodium falciparum cells expressing a chondroitin sulphate A binding phenotype. *Mol. Microbiol.*, **43**, 1285–1293.
52. Calderwood,M.S., Gannoun-Zaki,L., Wellems,T.E. and Deitsch,K.W. (2003) *Plasmodium falciparum var* genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J. Biol. Chem.*, **278**, 34125–34132.
53. Amit-Avraham,I., Pozner,G., Eshar,S., Fastman,Y., Kolevzon,N., Yavin,E. and Dzikowski,R. (2015) Antisense long noncoding RNAs regulate *var* gene activation in the malaria parasite *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E982–E991.
54. Broadbent,K.M., Park,D., Wolf,A.R., Van Tyne,D., Sims,J.S., Ribacke,U., Volkman,S., Duraisingh,M., Wirth,D., Sabeti,P.C. *et al.* (2011) A global transcriptional analysis of *Plasmodium falciparum* malaria reveals a novel family of telomere-associated lncRNAs. *Genome Biol.*, **12**, R56.
55. Sierra-Miranda,M., Delgadillo,D.M., Mancio-Silva,L., Vargas,M., Villegas-Sepulveda,N., Martinez-Calvillo,S., Scherf,A. and Hernandez-Rivas,R. (2012) Two long non-coding RNAs generated from subtelomeric regions accumulate in a novel perinuclear compartment in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.*, **185**, 36–47.
56. Grewal,S.I. and Jia,S. (2007) Heterochromatin revisited. *Nat. Rev. Genet.*, **8**, 35–46.
57. Swamy,L., Amulic,B. and Deitsch,K.W. (2011) *Plasmodium falciparum var* gene silencing is determined by *cis* DNA elements that form stable and heritable interactions. *Eukaryot. Cell*, **10**, 530–539.
58. Phillips-Cremins,J.E. and Corces,V.G. (2013) Chromatin insulators: linking genome organization to cellular function. *Mol. Cell*, **50**, 461–474.