Characterization of sexual commitment and the early steps of sexual development in the human malaria parasite *Plasmodium falciparum*

Oriol Llorà Batlle

TESI DOCTORAL UPF / 2019

Thesis supervisor:

Dr. Alfred Cortés Closas

Thesis tutor:

Dra. Cristina López Rodríguez

(Department of Experimental and Health Sciences, UPF)

INSTITUT DE SALUT GLOBAL BARCELONA (ISGLOBAL)

MALARIA PROGRAMME

"Life is not easy for any of us. But what of that?
We must have perseverance and above all confidence in ourselves.
We must believe that we are gifted for something, and that this thing,
at whatever cost, must be attained."

— Marie Skłodowska Curie

A la meva família,
per tot

ACKNOWLEDGMENTS

Agrair també als Falci, **Alfons**, **Pau**, **Diana**, **Laura**, **Ruth**, **Marta V.**, **Itzi**, **Xavier**, **Himanshu** i **Rebeca**, pel bon ambient (que em va fer sentir tant acollit al arribar al lab), les calçotades, els St Jordis i tot el suport en tot allò que he necessitat.

Moltes gràcies a la **Marta M.** i la **Clara** per compartir els maldecaps de la tesi, maldecaps amb el Ministerio i fer pinya per intentar millorar les condicions dels becaris. **Canaleta**, moltes gràcies per tot l'ajut rebut durant aquests anys, el teu sentit de l'humor i la *canya* que ens fots a tots pel bon funcionament del lab.

Gràcies a tot l'equip de vòlei per tots els bons moments: **Marta M.**, **Andrea**, **Javi**, **Virgi**, **Rebeca**, **Marta C.**, **Inés**, **Anastasia**, **Lucas** i **Arnau**. L'any vinent a guanyar la Casi Cracks!

I would also like to thank all the people of the Baum's lab at the Imperial College London for making my stay there a wonderful experience. Thanks, **Jake** for accepting the collaboration and your faith in the project. It's been a wonderful experience that has allowed me to meet exceptional people. I would like to express my gratitude to the Gams Team: **Kathrin**, **Alisje**, **Irene**, **Farah** and **Sabrina** for your patience with me, all your advice both inside and outside the lab and all the fantastic moments we've shared. **Mark** thank you very much for your kindness, your sense of humour and for hosting me in your place for a while. I've been missing you all a lot. I wish you all the best!

I would also like to thank **Manuel Llinás** for allowing me to visit his lab during a week and for accepting being part of the evaluation panel.

Gràcies a la meva tutora de la UPF, **Cristina López**, i als membres del comitè de seguiment de la tesis, **Alfredo Mayor**, **Jose Aramburu** i **Sara Pagans**, pels seus consells i ajuda.

Moltes gràcies als Biòlegs, **Yolanda**, **Laura**, **Lali**, **Marina**, **Clara**, **Gisela**, **Compte**, **Marc**, **Edu**, **Sergi**, **Pau** i **Quirze**, per tots els bons moments que hem compartit, viatges, sopars, festes, etc., però també per tot el vostre suport i efecte.

Moltes gràcies també als del Pòdium: **Cèlia**, **Clàudia**, **Laura**, **Nuri**, **Júlia**, **Sara**, **Marta**, **Berto**, **Pere**, **Paulí**, **Xavi**, **Guillem**, **Nil**, **Albert** i **Otto**; els amics de tota la vida, per ser-hi sempre, tota la vostra confiança i suport en els bons i mals moments.

També voldria agrair a la meva cosina **Lluïsa Cros** tot el suport, ajuda i consells al llarg d'aquests anys. És una sort poder comptar amb tu; moltes gràcies!

Finalment, moltes gràcies als meus pares, la **Gerta** i l'**Emili**, per tots els valors i l'educació que m'heu donat, eines indispensables perquè hagi pogut arribar fins aquí. Gràcies per animar-me a seguir allò que m'apassiona i saber-me motivar a fer les coses millor, per tots els vostres consells i la vostra confiança en mi. Moltes gràcies als meus germans, la **Gerta** i en **Roger**, pel seu efecte, confiança i complicitat, però sobretot per aguantar els meus *rollos* sobre la feina :) Moltes gràcies a la resta de la família, **àvies**, **tius**/**es** i **cosins/es** pel vostre suport. **Avi**, aquesta tesi també va per tu, al final no seré "ingeniero", però doctor no sona gens malament.

ABSTRACT

Transmission of malaria requires that some parasites abandon asexual replication and develop into sexual stages termed gametocytes. The discovery of the transcription factor PfAP2-G as the master regulator of sexual conversion has boosted our understanding of sexual development. The *pfap2-g* locus is controlled by heterochromatin-based silencing, with only a few parasites activating the locus and committing to sexual development at each asexual cycle. How heterochromatin forms in *pfap2-g* is currently unknown. Moreover, the initial steps of sexual conversion after *pfap2-g* activation, together with the biology of sexually committed stages, remain poorly characterized. In this thesis, we identified a new route of sexual conversion involving direct conversion within the same cycle of initial PfAP2-G expression. We also developed a conditional activation system for PfAP2-G, achieving synchronous sexual conversion of the majority of parasites, which enables the characterization of sexually committed parasites and early sexual stages. Finally, in an attempt to elucidate the mechanisms behind heterochromatin formation in *pfap2-g*, we gained insight into the role of different elements in heterochromatin nucleation.

RESUM

Per a la transmissió de la malària, alguns paràsits han de deixar de replicar-se de forma asexual i desenvolupar-se com a estadis sexuals anomenats gametòcits. El descobriment del factor de transcripció PfAP2-G com a regulador clau de la conversió sexual ha permès entendre millor el procés. El locus *pfap2-g* està silenciat per la presència d'heterocromatina i només alguns paràsits activen el locus i inicien el desenvolupament sexual a cada cicle asexual. Actualment es desconeix com es forma heterocromatina a *pfap2-g*. A més a més, els primers passos de la conversió sexual després de l'activació de *pfap2-g*, juntament amb la biologia dels estadis que acabaran desenvolupant-se sexualment, estan molt poc caracteritzats. En aquesta tesis, hem identificat una nova ruta de conversió que consisteix en la conversió sexual directa, sense un nou cicle de replicació, just després de l'activació de PfAP2-G. També hem desenvolupat un sistema d'activació condicional de PfAP2-G que indueix la conversió sexual sincrònica de la majoria dels paràsits i que permet la caracterització dels estadis sexuals primerencs. Finalment, en un intent per entendre els mecanismes darrera la formació d'heterocromatina a *pfap2-g*, hem obtingut informació rellevant sobre el paper de diferents elements en la nucleació d'heterocromatina.

PREFACE

In the last few years there have been great advances in our understanding of sexual development in malaria parasites. The identification of PfAP2-G as the master regulator of the process was, without doubt, a game-changer. Malaria sexual stages, or gametocytes, are essential for the transmission of the disease, which is one of the main concerns for malaria elimination efforts.

Back in January 2015, I joined the team led by Alfred Cortés to perform my MSc internship. PfAP2-G had been identified not long ago, providing a great opportunity to better characterize its role. My project aimed to develop a conditional activation system for *pfap2-g*, quite ambitious for a master student. After obtaining an FPU fellowship (FPU014/02456) from the Spanish government, I continued my research as a PhD student; a journey that now reaches its end. During these years, we've been able to conclude the project, with results beyond our initial expectations. The high yield of our system allowed the phenotypic characterization of early committed and sexual stages without any purification step – something impossible until now, as these stages cannot be readily separated from their asexual counterparts. I also contributed to the identification of a new route of sexual conversion, challenging the textbook view of the malaria life cycle. Moreover, we have also provided insight into the regulation of *pfap2-g*, in addition to the identification of specific regions of the *pfap2-g* coding sequence involved in its heterochromatin-based silencing. Overall, we have contributed to a better understanding of the initial steps of sexual development in *P. falciparum*.

TABLE OF CONTENTS

INTRODUCTION

1. Malaria disease

1.1 Brief history

"Nüe of the foot major yang: it lets a person have lower back pain and a heavy head. Cold rises from the back. [Patients] are cold first and afterward hot. The heat is intense harm caused by summer heat. When the heat stops, sweat leaves." (Unschuld, 2003).

Written around 2700 BC, those were the words used to describe malaria symptoms in the Chinese Canon of Medicine, Nei Jing – one of the first written records about malaria (Cox, 2010). Chinese physicians had already linked the disease with the air, or "caused by the wind" (Unschuld, 2003). However, the Greeks were probably the first (850 BC to 400 BC) to associate the disease with swamps and wetlands. This link with wetlands and miasmas persisted for more than two millennia, probably leading to the modern word malaria, from the Italian *mala aria* that means "bad air" (Cox, 2010).

It was not until the end of the 19th century, with the rise of the germ theory of disease, that researchers started to look for the causative agent of malaria. In 1880, Charles Louise Alphonse Laveran observed for the first time spherical motionless bodies inside red blood cells (RBCs) of malaria patients. This discovery pushed the field forward and by 1890 the scientific community already established that the different fever symptomatologies observed in malaria patients were due to three different species. Around that time, Italian malariologists were convinced about the role of mosquitoes in the transmission of

the disease, later confirmed by Ronald Ross in 1897 with his studies on birds (Ross, 1923). Finally in 1898 Giovanni Battista, Amico Bignami and Giuseppe Bastianelli demonstrated that human malaria was transmitted by *Anopheles spp.* mosquitoes (Battista et al., 1899; Cox, 2010).

One major breakthrough for the field was the development of the *in vitro* culture for the erythrocytic stages of the parasite in 1976 by William Trager and James B. Jensen (Trager and Jensen, 1976). The availability of unlimited amounts of parasites boosted research to another level: cell biology, genetics and biochemistry studies, together with genetic manipulation, started to emerge.

1.2 Life cycle

Malaria is a vector-borne disease caused by parasites of the genus *Plasmodium* and transmitted through the bites of infected female mosquitoes (Phillips et al., 2017). *Plasmodium spp.* are eukaryotic microorganisms of the phylum Apicomplexa: obligate intracellular parasites characterized by the presence of an apical complex involved in the invasion of the host cells (Adl et al., 2012; Morrison, 2009). To date, there are around 120 known *Plasmodium spp.* infecting mammals, birds and reptiles, but only 6 are capable of infecting humans: *Plasmodium falciparum*, *P. vivax*, *P. malariae*, *P. ovale curtisi*, *P. ovale wallikeri* and *P. knowlesi*, the latter of zoonotic transmission (Ashley et al., 2018).

Malaria parasites have a complex life cycle involving several developmental stages in two different hosts: the *Anopheles* mosquito and the vertebrate host (**Figure 1**). Infected female mosquitoes inject sporozoites into the dermis during a blood meal; these motile forms of

the parasite have the capacity to invade hepatocytes in the liver. The following 10 days parasites replicate until they form merosomes, releasing thousands of merozoites into the bloodstream. Merozoites invade erythrocytes in less than 2 minutes using specific receptor-ligand interactions, in a process that involves several steps such as initial contact, reorientation of the merozoite, microneme secretion, rhoptry secretion and formation of the parasitophorous vacuole – a plasma membrane protecting the parasite from the host cell intracellular environment (Cowman et al., 2016). *P. falciparum* parasites undergo an intraerythrocytic developmental cycle (IDC) and replicate through schizogony to release 16-32 new merozoites after ~48 h. Once inside the RBC, parasites export hundreds of proteins into the host cell membrane to remodel it, allowing the uptake of some nutrients in addition to mediating cell adhesion, important to avoid clearance by the spleen or the immune system (Spillman et al., 2015). *Plasmodium spp*. are auxotrophic for most of the amino acids and digest haemoglobin to fulfil their metabolic needs. Digestion of haemoglobin generates toxic free haem groups that parasites polymerize into hemozoin, an insoluble crystal visible in infected RBCs, to avoid haem toxicity (Phillips et al., 2017; Wunderlich et al., 2012).

The IDC is characterized by different stages: after invasion, parasites develop into ring stages (~24h), then trophozoites (~12h) and finally schizonts (~12h) that will release new merozoites (**Figure 1**). During this exponential growth parasites replicate asexually through mitotic divisions (Gerald et al., 2011). Nevertheless, some parasites commit to the sexual fate by developing into gamete precursors termed gametocytes. Male and female gametocytes are the only stages capable

of infecting the mosquito vector and therefore allow the transmission of the parasite to a new host. Inside the mosquito gut, fertilization occurs, and the zygote rapidly undergoes a meiotic division producing an ookinete. Ookinetes invade the mosquito midgut forming an oocyst that will further multiply by mitosis and release sporozoites. Sporozoites then travel to the salivary glands of the mosquito, from where they are released in the following blood meal to infect a new vertebrate host (Miller et al., 2002; Phillips et al., 2017).
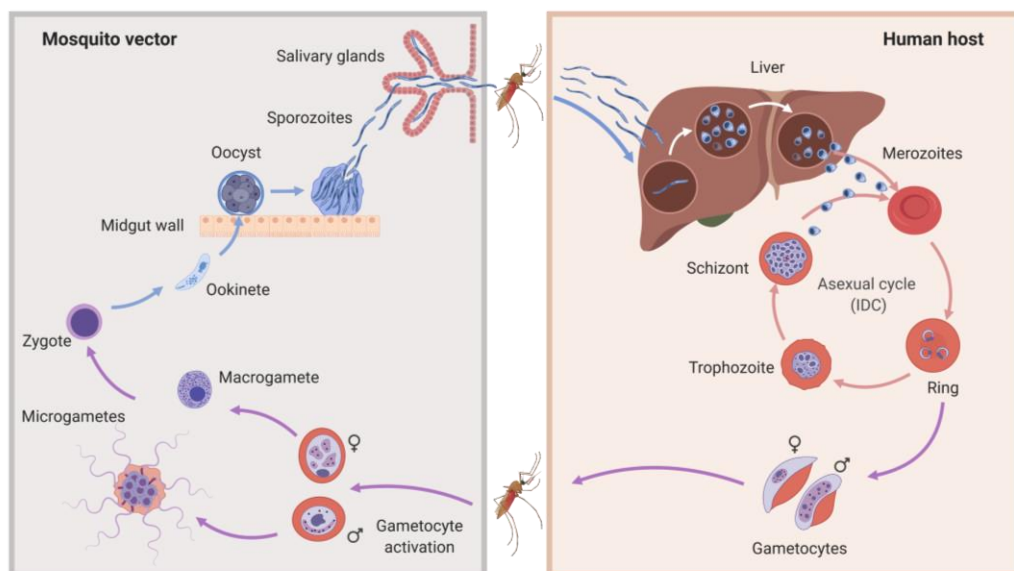
Figure 1. ***Plasmodium falciparum* life cycle.** See main text for the description of the whole life cycle involving several developmental stages within two host: the mosquito vector and the human host with the intraerythrocytic developmental cycle (IDC). Designed with Biorender.com.

1.3 Epidemiology

Malaria mainly occurs in tropical and subtropical areas, where the environmental conditions are favourable for the ~40 species of *Anophele*s mosquitoes that act as a vector (Phillips et al., 2017). The

incidence of the disease depends on environmental factors such as altitude, climate and vegetation that determine the presence of the vector. On top of that, control measures also impact its distribution, and therefore malaria incidence is also linked to poverty (Ashley et al., 2018). In countries with limited resources and weak health systems, the delivery of effective health care is impaired. Moreover, malaria disease itself has a big socioeconomic impact as it represents an obstacle for the development of the economy: loss of school and work time leads to limited education opportunities and reduced income (Phillips et al., 2017). In 2017, 219 million malaria cases occurred around the world, with the African Region accounting for 92% of the cases, followed by Southeast Asia with 5% and Eastern Mediterranean with 2% (WHO, 2018). Each year around half a million people die of malaria worldwide, with especial incidence in Africa that accounted for 93% of all malaria deaths in 2017 (WHO, 2018).

P. falciparum is responsible for most of the burden of disease worldwide, whereas *P. vivax* malaria cases mainly occur in Central and South America and Southeast Asia (Snow et al., 2017; WHO, 2018). This distribution can be attributed to the high prevalence of Duffy antigen negativity in Africa that provides partial protection against *P. vivax* invasion (Phillips et al., 2017; Twohig et al., 2019). *P. ovale* subspecies are found in Africa and Asia, being more prevalent in West Africa, whereas *P. malariae* can be found worldwide and the zoonotic *P. knowlesi* is only present in Southeast Asia (Ashley et al., 2018; WHO, 2018). Moreover, the majority of species may have different animal reservoirs (**Figure 2**).
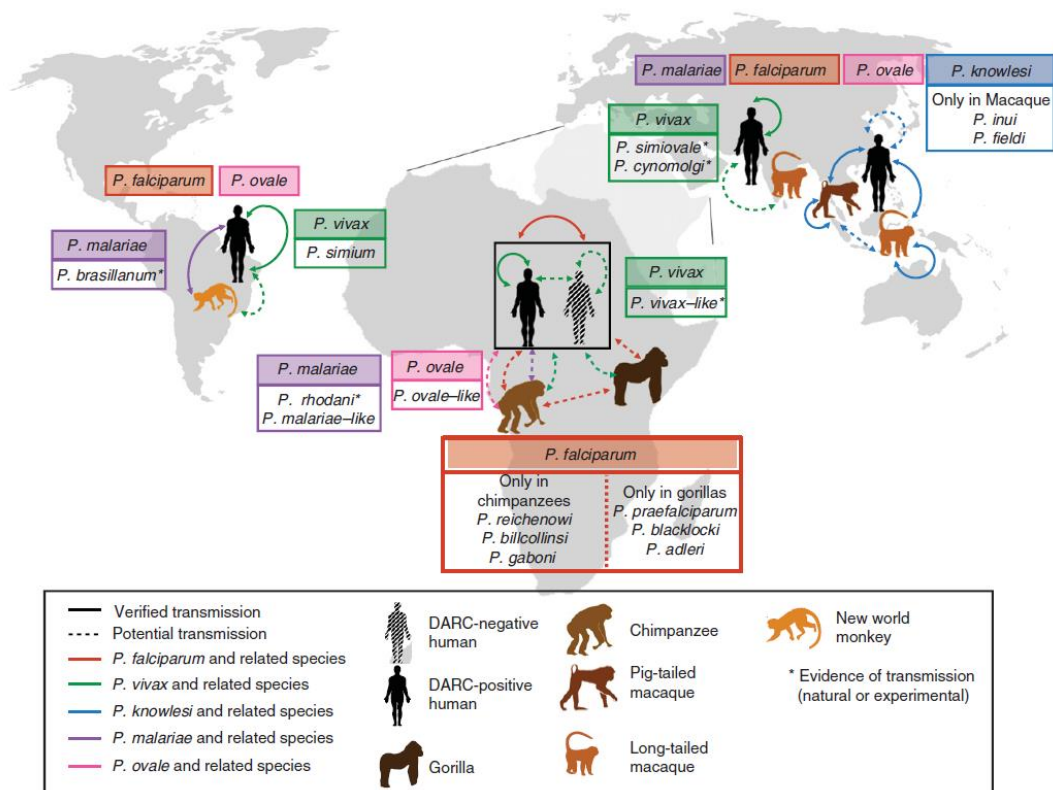
Figure 2. Global distribution of _Plasmodium spp._ in human and animal populations. _P. falciparum_ (red), _P. vivax_ (green), _P. malariae_ (purple) and the two species of _P. ovale_ (pink) are found in all endemic areas. _P. knowlesi_ (blue) is found in Southeast Asia. Arrows indicate established transmission of the different parasite species between hosts. Dotted arrows show potential transmission of parasites based on population studies. Species for which cases of human infection have been observed are marked with an asterisk. DARC: Duffy antigen receptor for chemokines. Reproduced from Lim et al., 2017.

Children under 5 years of age and pregnant women are the more vulnerable populations. It is estimated that ~1,200 African children die of malaria every day. In areas of continuous malaria transmission, the population progressively develops partial immunity after repeated exposure to the parasite. Adults in these areas achieve the control of the infection and typically only show mild symptoms; in some cases,

infections are chronic without any clinical manifestation, defined as asymptomatic malaria. On the other hand, children who have not been in contact with the parasite before are particularly vulnerable. Regarding pregnant women, the susceptibility arises from the capacity of the parasites to attach to the placental vasculature, increasing the risk of miscarriage and also linked to low weight at birth (Ataíde et al., 2014; Phillips et al., 2017).

1.4 Clinical malaria

The symptomatology of the disease is mainly associated with the exponential growth of the parasites in the blood and in the case of *P. falciparum*, its capacity to adhere to the vasculature (Cowman et al., 2016). The bursting of the parasite at the schizont stage and the release of merozoites liberates hemozoin and parasite DNA: these components activate pattern recognition receptors such as toll like receptor 9 leading to the production of pro-inflammatory cytokines (e.g. tumour necrosis factor) that trigger fever (Liehl and Mota, 2012; Parroche et al., 2007). Malaria is mainly characterized by periodic fevers that coincide with the release of new parasites in the blood, but also by other non-specific symptoms such as chills, rigor, headache, nausea and muscle pain (Phillips et al., 2017).

According to the presentation of the disease, malaria can be classified as uncomplicated or severe (Ashley et al., 2018). The clinical presentation is highly influenced by host immunity, which depends on age and transmission intensity (**Figure 3**). In low transmission areas, where most of the individuals are naïve and do not have a naturally acquired protection, the chances to develop severe malaria are higher.

Infection produces a febrile condition accompanied by the non-specific symptoms mentioned previously, and if not treated, can develop into severe malaria. Nevertheless, after a period of clinical symptoms, parasites are usually controlled at a low level, symptoms decline and eventually the infection resolves (Cowman et al., 2016).
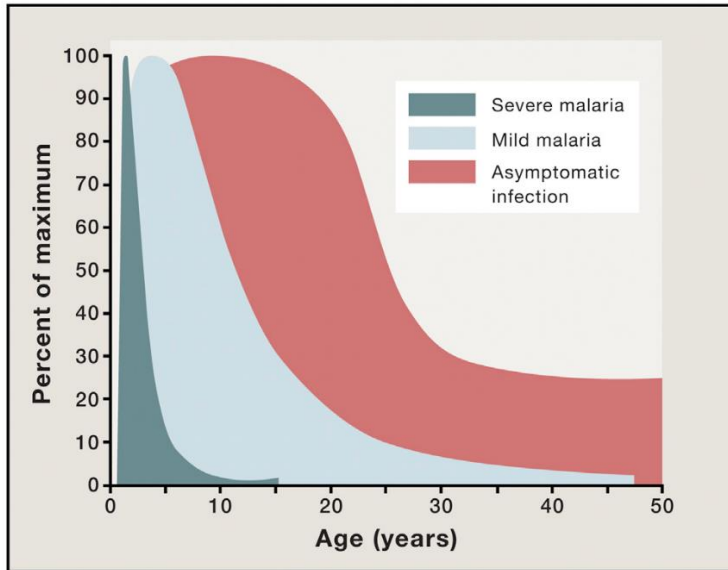
Figure 3. Age-dependence of malaria symptomatology in a population living in an endemic area of *P. falciparum* transmission. The risk of severe disease is highest in children under the age of 5 and declines rapidly. The susceptibility to symptomatic malaria continues up to a later age, such that adults experience few and mild febrile episodes; eventually, after repeated exposures, adults control the disease, but are still susceptible to asymptomatic infections. Reproduced from Cowman et al., 2016.

In some cases, infections cannot be controlled, and the disease progresses to severe malaria. This occurs in 1% of infected African children and in 10% of the cases results in death (Wassmer et al., 2015). Severe malaria is characterized by either severe anaemia, respiratory distress, metabolic acidosis or cerebral malaria (Marsh et al., 1995). A major factor for the pathogenic basis of severe malaria is

the sequestration to the tissue microvasculature, mediated by the binding of infected erythrocytes to the endothelium though surface antigens expressed by the parasite, such as the *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) (Miller et al., 2002) (**Figure 4**).

1.5 Diagnosis and treatment

The WHO criteria for the diagnosis of malaria is the presence of parasites in the blood and fever. The gold standards to detect parasites are light microscopy analysis of stained thick blood smears and rapid diagnostic tests (RDTs). Light microscopy has a limit of detection of ~5 parasites per microlitre, while for RDTs the limit is 50-100 parasites per microlitre. Both approaches are sensitive enough to screen for patients with uncomplicated or complicated malaria (typically 1,000-50,000 parasites per microlitre); however, their sensitivity is not sufficient to predict clinical relapses after interventions or to detect asymptomatic patients with submicroscopic infections (although some asymptomatic patients can have high parasitaemia). In these cases, molecular methods (e.g. PCR or loop-mediated isothermal amplification [LAMP]) have proven to be really valuable, as they can detect as few as 22 parasites per millilitre. However, in some settings these technologies are not available (Phillips et al., 2017).

Regarding treatment, different drugs are used depending on parasite species, the clinical manifestation or the epidemiological status of the area (WHO, 2015a). For uncomplicated malaria caused by *P. falciparum*, current frontline treatments are artemisinin-combination

therapies (ACTs). ACTs are based on the combination of an artemisinin and a quinine derivative.
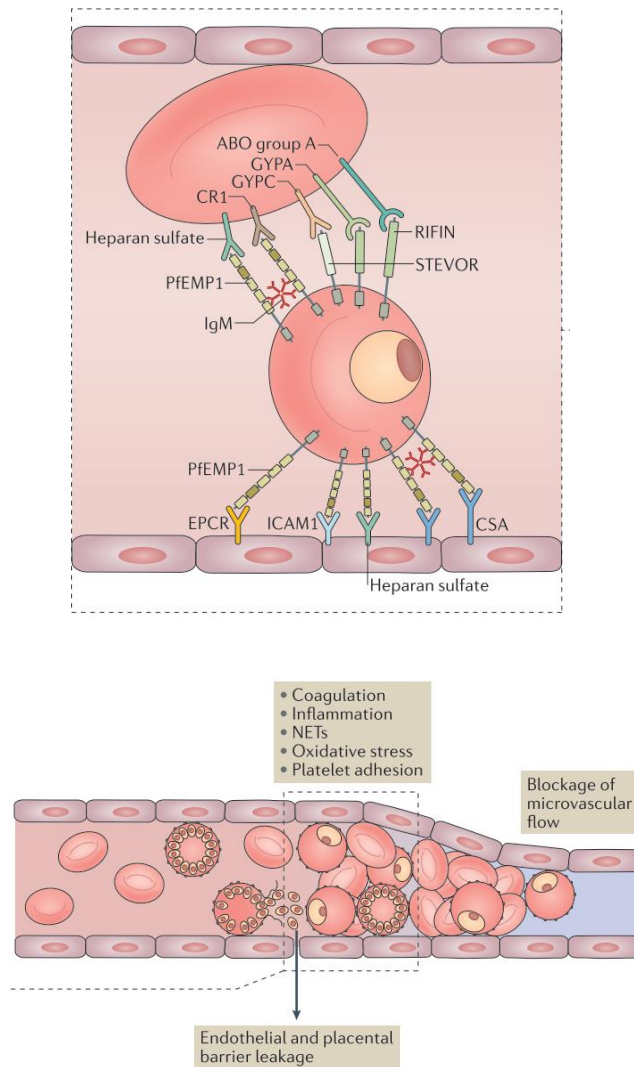
Figure 4. Sequestration of *P. falciparum* infected erythrocytes in the microvasculature through *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). Parasite-infected erythrocytes adhere to the microvasculature through the interaction between PfEMP1 and human receptors. Accumulation of infected erythrocytes causes the activation of the endothelium and the obstruction of blood flow. Infected erythrocytes can also bind uninfected erythrocytes through repetitive interspersed families of polypeptides (RIFIN), subtelomeric variable open reading frame (STEVOR) and PfEMP1 proteins, through the interaction with different host cell receptors. Reproduced from Wahlgren, Goel and Akhouri, 2017.

The most used artemisinin derivatives are dihydroartemisinin, artesunate or artemether; and common quinine derivatives are amodiaquine, piperaquine, pyronaridine, mefloquine or lumefantrine. These combinations are extremely effective as artemisinin derivatives can rapidly decrease parasitaemia, whereas the partner drug removes residual parasites and provides some prophylactic effect thanks to the long half-life of quinine derivatives (Phillips et al., 2017). Nevertheless, resistance to some of these derivatives, such as piperaquine, and partial resistance to artemisinin have been reported in some regions of Southeast Asia, urging the field to develop new effective drugs or design newer strategies to provide a cure (Tilley et al., 2016). An alternative to ACTs is the combination of atovaquone and proguanil, usually used for travellers and in cases of ACT failure; however, its implementation is not recommended in endemic countries because atovaquone resistance can emerge rapidly. Another drug that is used in combination with ACTs is primaquine, the only approved drug able to completely kill mature gametocytes. For severe malaria, the objective is to clear the parasites from circulation as fast as possible. To achieve this, mainly parenteral artemisinin derivates are used, such as artesunate or artemether. Once the patient has recovered from severe malaria symptoms and is able to eat and drink, the oral administration of ACTs is recommended (Ashley et al., 2018).

1.6 Prevention

In addition to providing a cure for the infection, antimalarial drugs can also be used as chemoprophylaxis to prevent malaria infections. Pregnant women, young children and travellers are the main targets of this strategy (Ashley et al., 2018). In the case of pregnant women, the

approach followed consists on spaced administration of sulfadoxine-pyrimethamine along pregnancy in what is called intermittent preventive therapy in pregnancy (IPTp) (Radeva-Petrova et al., 2014). However, the appearance of parasite resistance to this drug combination represents a considerable threat. The same drugs in combination with amodiaquine are used to prevent malaria infections in children. For travellers, mainly atovaquone-proguanil and doxycycline are used as prophylaxis, but also mefloquine (Ashley et al., 2018). WHO recommends using different drugs for treatment and chemoprophylaxis to reduce the emergence of drug resistance (WHO, 2015a).

Another major intervention to prevent malaria infections are vaccines. The development of naturally acquired protective responses to malaria parasites suggests that the design of an effective vaccine is possible. However, this protective immunity is acquired after repeated exposures and does not provide sterile immunity. Moreover, immunity is lost within 3 to 5 years in the absence of exposure, probably as a result of the failure to developed long-lived plasma B cells and the decline of circulating antibodies (Doolan et al., 2009). Developing a vaccine for antigens that in nature do not seem to be sufficient to trigger sterile immunity is challenging. Vaccine development efforts have focused mainly on surface antigens present in merozoites, sporozoites and infected RBCs (**Figure 5**). This is a challenging task because most of the surface proteins have variant expression or are highly polymorphic, which limits the list of possible candidates. Thus, many experimental vaccines end up conferring strain-specific protection only (Draper et al., 2018; Hoffman et al., 2015). To date,

the only vaccine that has proven to consistently provide some protection in endemic regions is the RTS,S/A01 vaccine. Based on the recognition of the circumsporozite surface protein (CSP) and combined with the hepatitis B surface antigen, it provided ~30% protection against clinical malaria during the first year, but failed to provide long-term protection (RTS.S Clinical Trials Partnership, 2015).
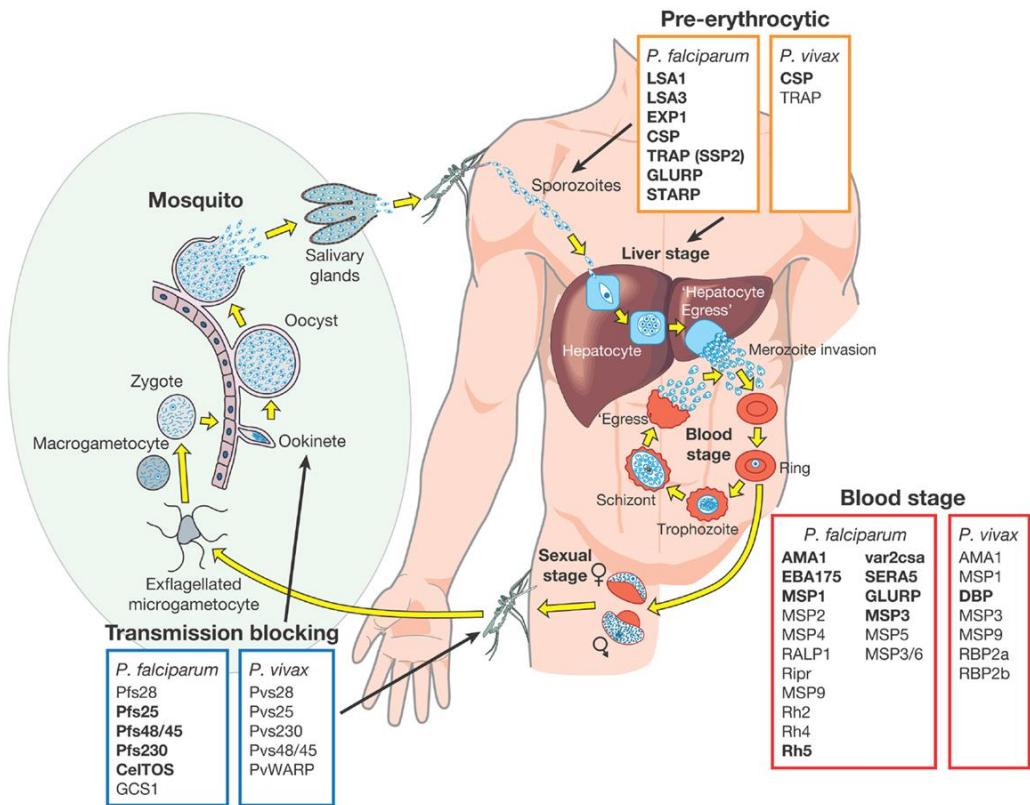
Figure 5. Candidate antigens for a malaria vaccine. Candidate antigens involved in different developmental steps of the life cycle are shown both for *P. falciparum* and *P. vivax*. Antigens indicated in bold are currently being evaluated in clinical trials. Reproduced from Barry and Arnott, 2014.

The impact of the RTS,S vaccine in reducing mortality has not been demonstrated, but it has been recommended by WHO for pilot

introduction in selected areas of 3 African countries. Despite these limitations, the study of the immune responses upon immunization can provide key information for the design of newer and better vaccines (Dobaño et al., 2019). The malaria vaccine pipeline is increasing with several candidates under evaluation: irradiated sporozoites, Pfs25 gametocyte antigen to block transmission, merozoite proteins such as Rh5, or the PfEMP1 VAR2CSA that mediates binding to the placenta (Ashley et al., 2018; Phillips et al., 2017) (**Figure 5**).

Vector control strategies are still the most effective measure to prevent malaria infections (Shaw and Catteruccia, 2019). It is estimated that they account for two thirds of the malaria cases avoided in Africa (Bhatt et al., 2015). Vector control strategies are mainly based on insecticide treated bed nets (ITNs) and indoor residual spraying (IRS) that aim to reduce vectorial capacity (Lobo et al., 2018). Vectorial capacity refers to the rate at which a vector can transmit a pathogen from a previous infection. Parasite population densities, the human biting rate, the competence for the development of the pathogen and the vector lifespan, define the vectorial capacity (Shaw and Catteruccia, 2019). ITNs combine the effect of the physical barrier with the pyrethroid-based insecticides; nevertheless they are only effective for vectors biting preferentially indoors and in the evening or at night time. IRS mainly targets surfaces where mosquitoes rest before or after the blood meal, therefore it is only effective against indoor biting (Lobo et al., 2018). Moreover, *Anopheles* mosquitoes have demonstrated a great plasticity in their behaviour, for example preferentially biting outdoors during day-light or low-light

situations, threatening transmission prevention (Sherrard-Smith et al., 2019; Thomsen et al., 2016). Another concern is mosquito resistance to insecticides: 60 of the 78 countries that monitor it have reported resistance to one or more insecticides (WHO, 2015b).

Other strategies currently under development are endectocides or genetic control, which have started to change the paradigm of vector control. Endectocides are systemic drugs administered to the host (i.e. humans) with an effect against the arthropod vector and the development of the parasite inside the vector (Foy et al., 2011). Drugs such as ivermectin have already been proven to be effective against both *Anopheles* mosquitoes and malaria parasites (Foy et al., 2019). Genetic control strategies are based on the dissemination of inheritable factors that decrease the vector population or their vectorial capacity (Lobo et al., 2018). It can be achieved by modifying the germline of the vector or through paratransgenesis of the vector microbiota (Alphey et al., 2002; Wilke and Marrelli, 2015). Essentially, the aim is to reduce population numbers through the spreading of sex-lethal alleles; or to reduce vectorial capacity by making the vector resistant to the infection. The advances with CRISPR/Cas9 technology have provided the grounds for such modifications based on gene drive approaches, allowing the spreading of the allele in the mosquito population (Alphey, 2016) (**Figure 6**). Nevertheless, polymorphisms in the targeted regions represent an obstacle for the application in the field. Moreover, these approaches harbour ethical and ecological concerns as the consequences of altering vector populations are difficult to predict (Shaw and Catteruccia, 2019).
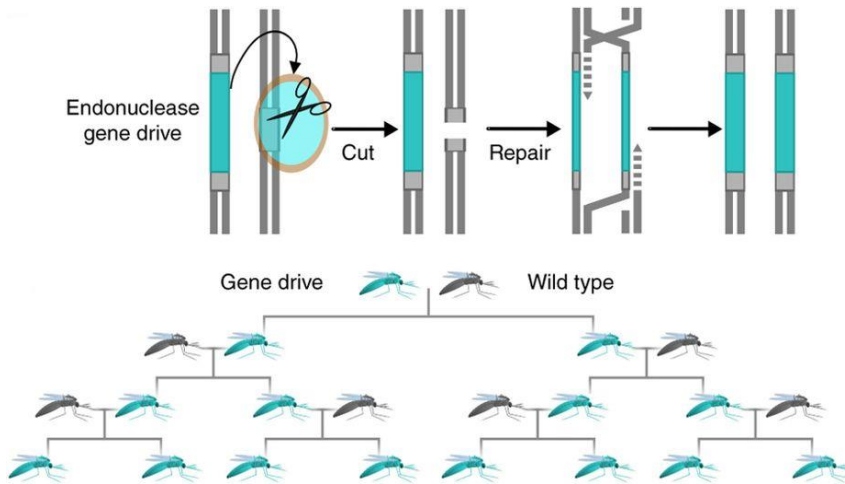
Figure 6. Gene drive strategy based on endonuclease homing. The gene drive is inherited to most of the offspring when individuals with an endonuclease gene drive (blue) mate with a wild type individual (grey). This allows the spreading of the gene drive until all the population contains it. The preferential inheritance occurs because the endonuclease cuts the homologous wild type chromosome, which is repaired by homologous recombination using the chromosome containing the gene drive as a repair template. Reproduced from DiCarlo et al., 2015.

2. Regulation of eukaryotic gene expression

2.1 General aspects of eukaryotic transcription

Gene expression starts with the transcription of the genetic information stored in the DNA to an RNA transcript that will later be translated into a functional protein (Crick, 1970). Regulation of gene expression is essential for the integrity of the cells, by determining which genes will be transcribed at a particular time depending on their needs. Therefore, a fine regulation of gene expression is essential for development, the adaptation to environmental changes and damage, and the integration of external signals (Pope and Medzhitov, 2018).

Gene expression can be regulated at three different levels: transcriptional, posttranscriptional and posttranslational. Chromatin structure plays a central role in transcriptional regulation by modulating the accessibility of DNA to the transcription machinery. This machinery consists of general transcription factors (GTFs), gene specific transcription factors (GSTFs), coactivators and corepressors that specifically recognize regulatory sequences such as promoters, enhancers, silencers or insulators (Andersson et al., 2015).

Eukaryotic genes are usually regulated by two cis-acting DNA regulatory elements, the core promoter and its proximal regulatory sequences (collectively termed as the promoter) and distally located regulatory elements, such as enhancers or silencers. The core promoter is composed of several DNA elements that serve as a binding platform for the transcription machinery.

The binding of the GTFs and the recruitment of the RNA polymerase II (Pol II) to the core promoter forms the pre-initation complex (PIC) that will initiate transcription at low levels. The proximal regulatory sequences of the promoter can be bound by GSTFs that recruit coactivators (often chromatin modifiers) to alter the chromatin structure modulating the activity of the PIC by increasing transcription or facilitating its assembly near the transcription start site (TSS) (Haberle and Stark, 2018) (**Figure 7**).
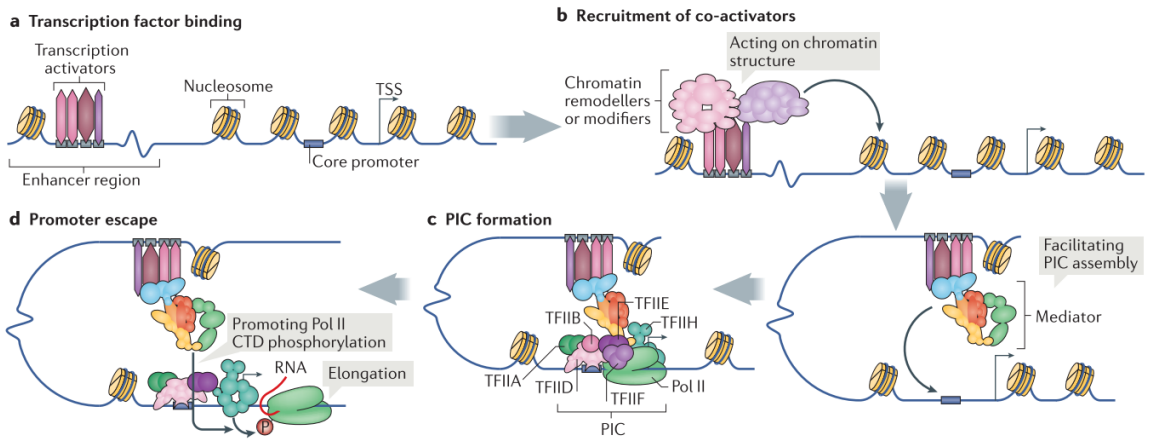
Figure 7. Model for the main steps of transcription initiation by Pol II. a, Transcription activation starts with the binding of transcription factors (referred as GSTF in the main text) on enhancer regions (or upstream activating sequences). The TSS is indicated by an arrow. **b**, Activators recruit coactivator complexes such as chromatin modifiers or remodellers that alter chromatin structure, which make it more accessible. Many coactivators cooperate and contribute to PIC formation. Mediator of Pol II transcription (Mediator) is one of the key coactivator complexes. **c**, The PIC is assembled at the core promoter and is composed of Pol II (12 subunits), GTF: transcription initiation factor IIA (TFIIA; 2 subunits), TFIIB, TFIID (TATA-box-binding protein [TBP] and 14 TBP-associated factors), TFIIE (2 subunits), TFIIF (2–3 subunits) and TFIIH (10 subunits). **d**, CDK7 phosphorylates (P) the carboxy-terminal domain (CTD) of Pol II at Ser5 triggering Pol II escape from the core promoter beginning the elongation step. Reproduced from Soutourina, 2018.

Distal regulatory regions such as enhancers contain specific DNA motifs that can recruit GSTFs to increase the formation and loading of the PIC at the core promoter and therefore increase transcription. They typically act over long distances and can be found both in intergenic regions and in genes. The three-dimensional (3D) landscape of the genome plays an essential role in the interaction between enhancers and promoters, with loops allowing direct contact between the regulatory regions (**Figure 7**). Silencers downregulate gene expression through the binding of repressors, proteins that interfere

with the assembly of the PIC, or can passively prevent the binding of GSTFs to the cis-regulatory elements. Enhancer-blocking insulators interfere with the interaction between the enhancer and the promoter when present between the two regions. Other cis-regulatory elements are barrier insulators that demarcate chromatin domains in the genome preventing the spread of heterochromatin (Haberle and Stark, 2018).

2.2 Chromatin structure and organization

In eukaryotic cells, DNA is packed in the nucleus as chromatin, a nucleoprotein complex of DNA and histones. The fundamental unit of chromatin is the nucleosome, which consists in 146 bp of DNA wrapped around a histone octamer composed of two copies of H2A, H2B, H3 and H4. The linker histone H1 binds outside of the nucleosome core forming a higher-order chromatin structure. Histone proteins have a globular C-terminal domain involved in histone-histone interactions and an N-terminal domain that stands out from the histone octamer (histone tail). These tails are highly basic (positively charged), which allows their interaction with DNA. Nucleosomes prevent transcription *per se*, mainly by physical obstruction of the transcription factors (TFs). Moreover, chromatin accessibility is modulated by posttranslational modifications (PTMs) of histones, histone variants, nucleosome remodelling and DNA methylation. Histone PTMs impact chromatin structure and accessibility directly affecting gene expression, DNA repair and replication. In general, genome regions with more accessible chromatin are transcriptionally active, whereas regions with less

accessibility are silenced (Even-Faitelson et al., 2016; Lawrence et al., 2016).

Histone PTMs include methylation, acetylation, phosphorylation, ubiquitylation, SUMOylation, ADP-ribosylation and deamination, among others (**Figure 8**). Lysine methylation and acetylation are the PTMs with higher impact on gene expression. They can influence transcriptional regulation via two characterized mechanisms: alteration of DNA-histone interactions, affecting the accessibility of TFs to DNA; and recruitment of effector proteins containing binding domains that specifically interact with modified histones (Lawrence et al., 2016). PTMs can be considered as information that is stored in histones. Specific combinations of these modifications are crucial to determine the transcriptional competence of promoters. In this context, increasing evidence supports the histone code hypothesis, which predicts that PTMs of histones, alone or in combination, direct specific and distinct DNA-templated programs by determining DNA accessibility or allowing the specific binding of transcriptional regulators (Strahl and Allis, 2000).
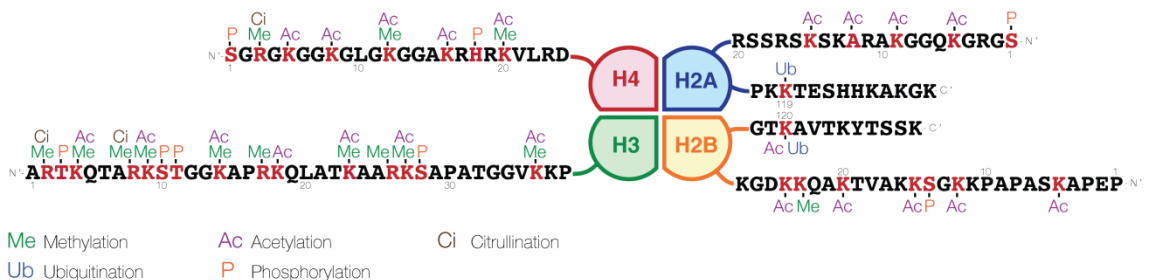
Figure 8. Histone PTMs. Schematic of the modifications that can be found in human histone tails. By Mariuswalter- Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=41382917.

- *Histone acetylation*

Histone acetylation is generally associated with gene activation, as the addition of an acetyl group to the lysine side chain neutralizes its positive charge, weakening the interaction between histones and DNA. Therefore, chromatin accessibility is enhanced, allowing the recruitment of TFs to the regulatory regions. Moreover, lysine acetylation can be recognized by the bromodomain of transcriptional coactivators. Deposition and removal of acetylation is mediated through the action of histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs use acetyl-CoA as a cofactor to transfer an acetyl group to the amino group of lysine side chains. There are two main groups of HATs: type A and type B. Type A HATs are diverse and can be classified in three groups according to their structural conformation: GNAT, MYST and p300/CBP. They usually form multiprotein complexes that determine their recruitment, activity and substrate. Type B HATs are mainly cytoplasmatic and acetylate free histones that have not been deposited into chromatin. HDACs remove the acetyl group, restoring the positive charge of the lysine and consequently stabilizing the interactions of histones with DNA; thus, HDACs typically act as transcriptional repressors. They can be classified into four main classes based on their homology with *S. cerevisiae* histone deacetylases (class I, II, III and IV). In general HDACs have low substrate specificity, as they can deacetylate multiple sites within histones and also act as part of multiprotein complexes (Bannister and Kouzarides, 2011).

- *Histone methylation*

Unlike acetylation, methylation does not alter the charge of the histone protein and can be associated with gene activation or repression depending on the specific methylated residue. Methylation mainly occurs on lysines and arginines. Specific residues can present different methylation levels: lysines can be mono-, di- or tri-methylated; whereas arginines can be mono- or di-methylated. Methylated residues can be recognized by Chromo-, Tudor, PhD, PWWP and MBT domains. Again, these proteins can further recruit other effectors or be found in complexes that alter the chromatin structure leading to activation or repression of gene expression. Histone methylation is catalysed by the action of histone lysine methyltransferases (HKMTs) that harbour the SET domain involved in the transfer of a methyl group from S-adenosylmethionine to the lysine amino group. Removal of the methylation mark is performed by histone demethylases (HDMs) that can be classified in two families: lysine specific demethylase 1 (LSD1) and Jumonji-C domain containing histone demethylases (JMJD). HDMs are highly specific and sensitive to the degree of methylation, with some enzymes only capable of removing mono- and di-methyl groups (Bannister and Kouzarides, 2011).

- *Chromatin organization*

Broadly, chromatin can be classified according to the composition of proteins, the histone PTMs and DNA methylation into 5 types: two types of active chromatin (euchromatin), facultative heterochromatin, constitutive heterochromatin and null chromatin (**Figure 9**).

Generally, euchromatin is more accessible whereas heterochromatin is inaccessible and transcriptionally inactive.
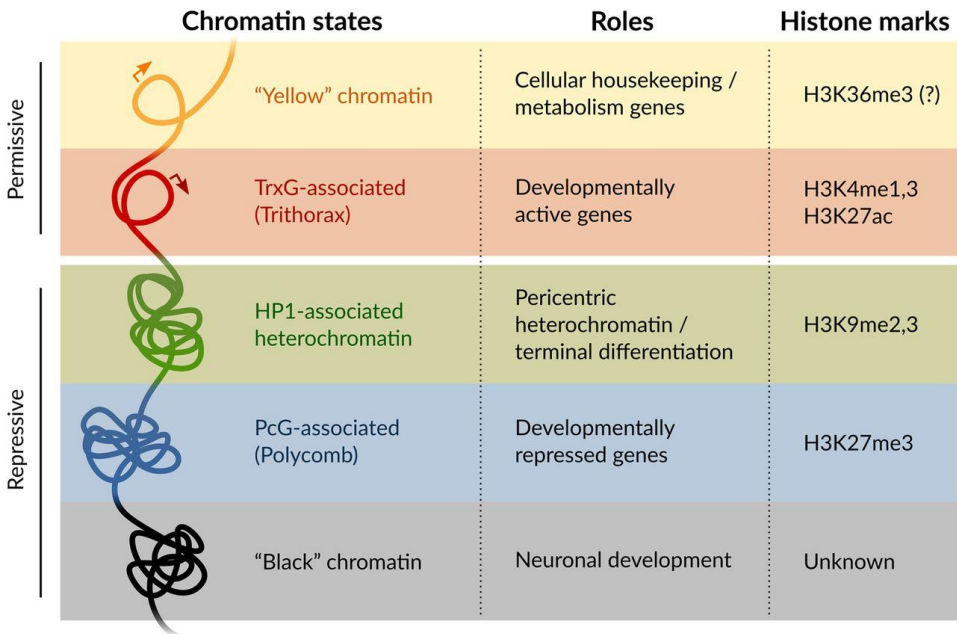
Figure 9. Classification of chromatin states. Five broad chromatin states seem to be largely conserved. There are two permissive states and three repressive states. Their known roles are listed together with the common epigenetic marks associated. Black chromatin is referred in the main text as null chromatin. Reproduced from Delandre and Marshall, 2019.

Active chromatin is the most heterogeneous, composed of many chromatin factors and histone modifications, such as different levels of H3K4, H3K36 and H3K79 methylation; but also acetylation of multiple residues in H3 and H4. Facultative heterochromatin in higher eukaryotes is mainly controlled by the Polycomb group (PcG) of proteins that deposit repressive histone marks by ubiquitylation of H2AK119 and different methylation levels at H3K27. Most of the genes regulated by PcG proteins are involved in cell fate determination and development, reflecting the facultative nature of this chromatin

state: at some point during development, or in some cell types, these genomic regions are in an active state. Constitutive heterochromatin mainly has a structural function, preserving genome integrity in telomeres and centromeres, but also silences repetitive DNA elements such as transposable elements. It is characterized by the presence of heterochromatin protein 1 (HP1), H3K9me3 and H4K20me3. By definition, constitutive heterochromatin is a highly condensed structure; nevertheless, it allows transcription of some non-coding RNAs (ncRNAs) relevant for its maintenance. Null chromatin is defined as a highly repressed state of the chromatin independent of PcG and HP1. It lacks enrichment for any specific histone modifications and the proteins associated with it are mainly H1 and lamin. Lamin is essential to tether silenced regions to the nuclear periphery compartmentalizing the nucleus (Ciabrelli and Cavalli, 2015).

A higher order of organization is the 3D folding of the chromosomes within the nucleus, which is also linked to the transcriptional output of the cell. The positioning of genes reflects their transcriptional status, with highly transcribed genes located in the interior of the nucleus, whereas repressed regions are closer to the periphery. Chromosome conformation capture techniques have started to reveal the features of genome organization and its functional relevance. At large scale, chromosomes segregate in two main compartments corresponding to active regions and inactive regions. However, at lower scales, chromosomes fold into domains that favour intradomain interactions over interactions with other domains. These domains have been defined as topologically associated domains (TADs). TADs define

regulatory landscapes by determining enhancer-promoter interactions (Szabo et al., 2019).

2.3 Mechanisms of heterochromatin formation and inheritance

- *Heterochromatin formation*

The establishment of heterochromatin is controlled through the action of different protein complexes and small RNA molecules. Heterochromatin formation requires the initial nucleation of the silencing signal, mainly by the deposition of H3K9me3 and its recognition by HP1, which later leads to its expansion and is involved in mechanisms to ensure its maintenance. Nucleation mechanisms do not explain how heterochromatin is maintained and spreads: the initial nucleation is driven by the underlying DNA sequence, whereas spreading and maintenance are independent and mainly sustained through positive feedback loops of the heterochromatin proteins (Greenstein et al., 2018).

Studies in the fission yeast *Schizosaccharomyces pombe* have provided insight into how heterochromatin establishes and is maintained: the formation of pericentromeric heterochromatin depends on the action of small interfering RNA (siRNA) involved in RNA interference (RNAi) pathways. Pervasive transcription from these highly repetitive loci results in the production of double-stranded RNAs (dsRNAs), which are processed by the DICER complex generating the siRNAs. These siRNA specifically recognize the nascent transcripts of the locus through base pairing and recruit the RNA-induced transcriptional silencing complex (RITS). RITS mainly increases the production of siRNAs driving a positive feedback loop, but also recruits the H3K9

41

HMT (Clr4 in yeast, Suv39 in mammals) for the deposition of H3K9me3 (Saksouk et al., 2015). Chp2 and Swi6 (the *S. pombe* orthologues of HP1 in higher eukaryotes) protein dimers selectively bind to H3K9me3 through the chromodomain, recruiting the SHREC complex that removes histone acetylation by the action of HDACs, further allowing the methylation of H3K9me3. This positive feedback loop is essential for the expansion of heterochromatin domains (**Figure 10**).
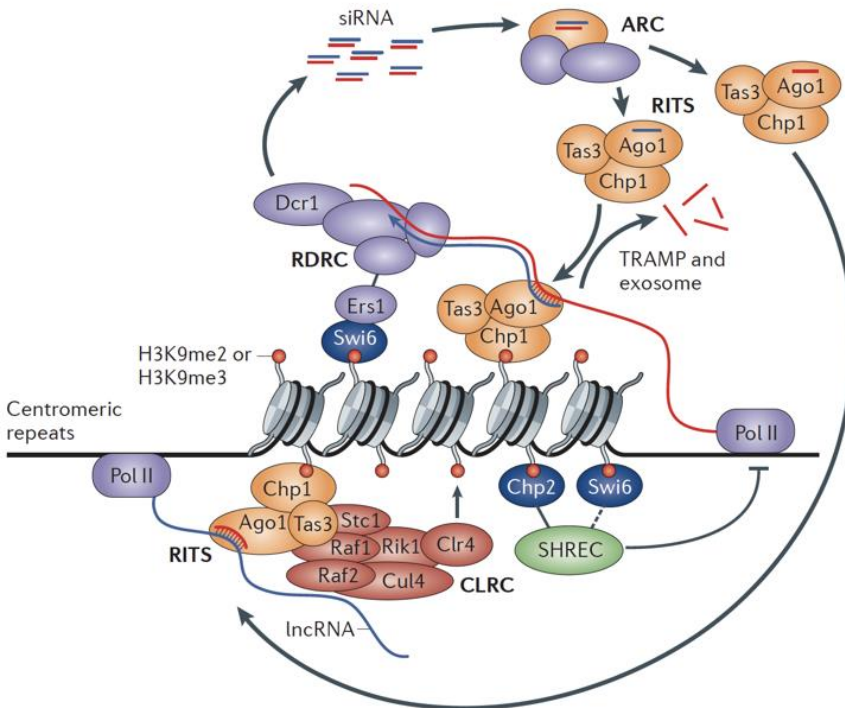
Figure 10. RNA-based heterochromatin formation in *Schizosaccharomyces pombe.* Centromeric long non-coding RNAs (lncRNAs) bound to the RITS (Ago1, Chp1 and Tas3) complex become templates for dsRNA synthesis by the RNA-dependent RNA polymerase complex (RDRC) and the processing by Dicer 1 (Dcr1) that generates siRNAs. siRNAs are processed by the ARC (Ago siRNA chaperone) complex leading to further recruitment of the RITS complex. The Chp1 subunit of the RITS complex anchors onto nucleosomes recruiting the Clr4–Rik1–Cul4 (CLRC) complex to promote H3K9 methylation. Swi6 binds to methylated H3K9

and recruits RDRC through Ers1 to increase siRNA formation. Swi6 and Chp2 restrict Pol II access recruiting the Snf2–histone deacetylase repressor complex (SHREC). The TRAMP non-canonical poly(A) polymerase and the exosome also contribute to silencing. Reproduced from Holoch and Moazed, 2015.

Once heterochromatin is nucleated at a particular locus, their components allow its expansion through a DNA sequence-independent process (Allshire and Madhani, 2018). Some of the nucleation events are not productive, with only a few that result in spreading: stochastic interactions between the heterochromatin domain and adjacent chromatin may trigger the deposition of H3K9me3 in a discontinuous manner. Population-based analysis showed that nucleation occurs after a few cell divisions (1-10 generations), whereas the spreading and establishment of an heterochromatin domain to its full size requires longer times (>20 generations) (Greenstein et al., 2018; Obersriebnig et al., 2016). To limit the spreading of heterochromatin to avoid possible erroneous gene silencing, cells have different mechanisms that block the process. Such mechanisms consist on nucleosome depleted regions with high histone turnover where TFs can bind and recruit anti-silencing factors. These regions are characterized by the presence of H2A.Z and PTMs consisting on activation marks that counteract the silencing complexes of heterochromatin. Moreover, there are intrinsic spreading inhibitors such as EpeI (a JMJD HDM), whose activity is inhibited in the interior of the heterochromatin domain, whereas it is active at the periphery where it can demethylate H3K9me3 (Allshire and Madhani, 2018). Additional control of spreading involves the action of ncRNAs such as Borderline, which is specifically transcribed at the heterochromatin boundary and interacts with Swi6, blocking its activity (Johnson and Straight, 2017) (**Figure 11**).
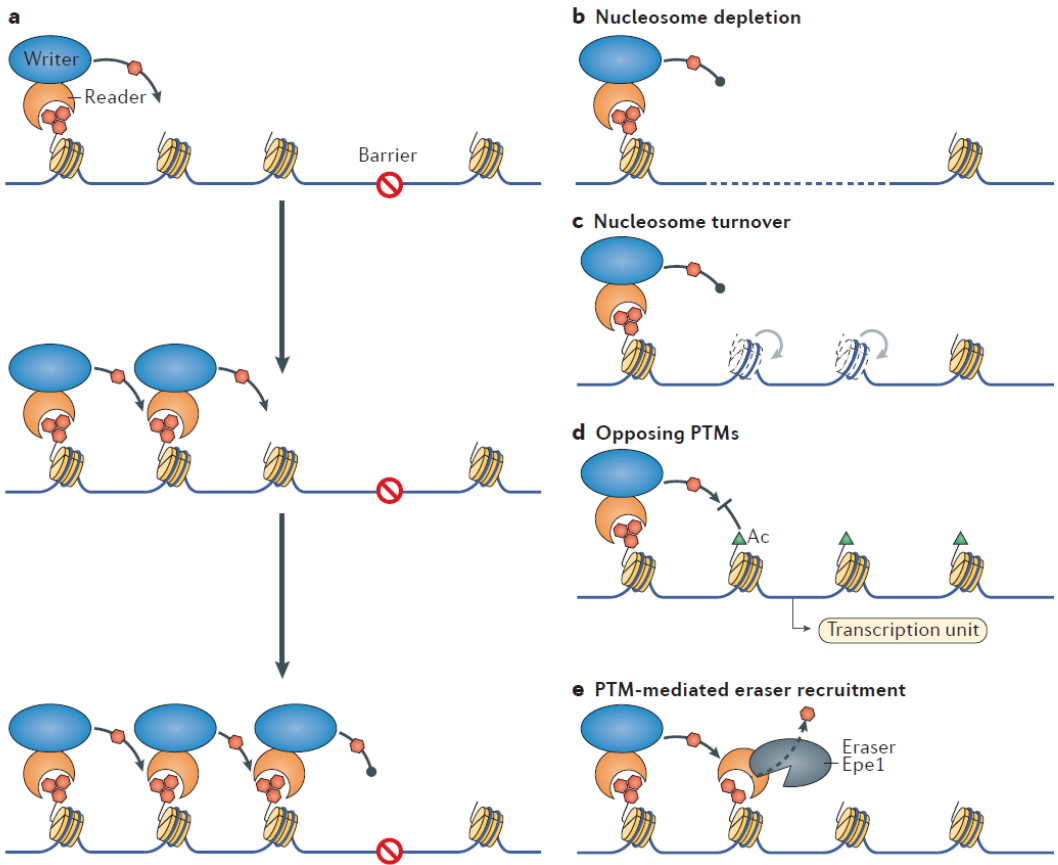
Figure 11. Heterochromatin spreading. a, Model of heterochromatin expansion based on reader and writer interactions enhancing the formation of repressive histone PTMs (red hexagons) in neighbour nucleosomes. The barrier represents different mechanisms to restrict the spreading (panels **b–e**). **b**, Sequences bound by factors that avoid nucleosome assembly create gaps that prevent heterochromatin spreading. **c**, Nucleosome turnover blocks heterochromatin expansion. **d**, Adjacent active transcription units characterized by active PTMs (green triangles) prevent the addition of repressive PTMs and heterochromatin spreading. **e**, Erasers such as the *S. pombe* demethylase enhancer of position effect 1 (Epe1) are recruited by readers of repressive PTMs at the edge of heterochromatin and prevent heterochromatin expansion. Reproduced from Allshire and Madhani, 2018.

The coupling between epigenetic readers and modifiers is essential for the spreading and maintenance of heterochromatin. The HMT Clr4 (or Suv39 in *Drosophila* and mammals) contains an N-terminal

chromodomain and a C-terminal SET domain that allows the same protein to act as a reader and a writer. The methylation of H3K9 triggers the recruitment of more HMT, but also of HP1 that contains a chromodomain and a chromoshadow domain. The chromoshadow domain of HP1 allows its dimerization, which provides a binding platform for other effector proteins. For instance, in *S. pombe*, the recruitment of the SHREC complex removes acetylation and allows H3K9 methylation (Allshire and Madhani, 2018).

Similar mechanisms regulate heterochromatin formation in worms or plants, in which DNA methylation is also required for the correct establishment of heterochromatin (Allshire and Madhani, 2018). In mammals, heterochromatin may form through similar mechanisms since the proteins involved have orthologs in metazoans. However, there is little evidence that RNAi plays a role in the establishment of heterochromatin in mammals because dsRNA originating from pericentromeric regions have not been detected (Saksouk et al., 2015). Nonetheless, the HMT Suv39 can be recruited to pericentromeric regions through RNA-dependent mechanisms (Johnson et al., 2017; Shirai et al., 2017). Of note, injection of dsRNA composed of satellite sequences in mouse embryos was able to target HP1 to pericentromeric regions, evidencing the role of RNA in heterochromatin formation, although the involvement of the RNAi pathway was not explored (Santenard et al., 2010). Moreover, mouse heterochromatin is sensitive to RNase A treatment (Maison et al., 2002).

Some RNAs seem to regulate chromatin modifications through pathways independent of RNAi. lncRNA have been shown to recruit

chromatin modifying complexes that can trigger heterochromatin formation. For instance, lncRNAs such as XIST are involved in dosage compensation by inactivating the X chromosome in female mammals through the recruitment of PRC2 to methylate H3K27 (Holoch and Moazed, 2015). Other heterochromatin formation mechanisms independent of RNAi involve protein complexes that specifically recognize DNA motifs and recruit the HMTs triggering heterochromatin nucleation. For instance, the silencing of endogenous retroelements in the genome involves the action of KRAB-Zinc finger proteins that recruit the HMT SETDB1 to methylate H3K9. Similar mechanisms account for the silencing of regions required for the maintenance of cell identity (Allshire and Madhani, 2018). Zinc finger domain TFs also associate with pericentromeric regions driving transcription, which is essential for the recruitment of Suv39 (Bulut-Karslioglu et al., 2012).

- *Heterochromatin inheritance*

Once the heterochromatin-based silencing at a particular locus has been established, it is stably inherited (**Figure 12**). The inheritance of changes in the expression status of a gene that do not involve changes in its sequence and persists in the absence of the initiating signal is defined as epigenetics. Therefore, heterochromatin, and particularly facultative heterochromatin, can carry epigenetic information. The term "epigenetics" is commonly used to refer to DNA methylation, PTMs, histone variants and ncRNAs; however, not all chemical modifications of DNA or histones can mediate a heritable change in the expression of a gene (Allshire and Madhani, 2018).
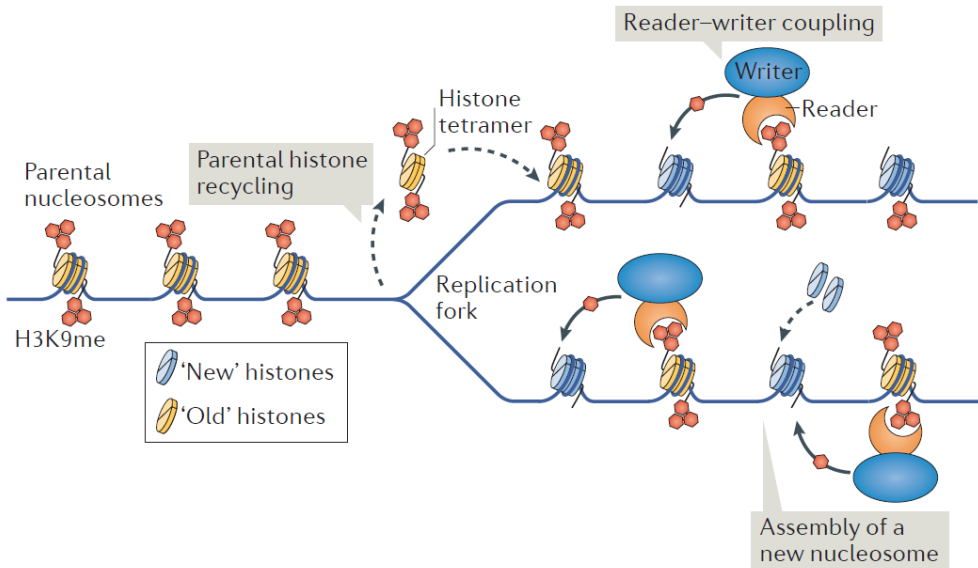
Figure 12. Reader–writer coupling maintains repressive chromatin modifications through DNA replication. During replication, parental 'old' nucleosomes are randomly reused into the two newly synthesized DNA molecules. Reader–writer coupling allows the propagation of the PTMs to the newly assembled nucleosomes. Adapted from Allshire and Madhani, 2018.

Epigenetic inheritance requires the faithful transmission of the epigenetic state of a particular locus to the offspring during DNA replication. Transmission of the epigenetic state involves specific elements, such as TFs or RNAs, which are essential to probe the information of the epigenome and re-establish it after DNA replication. Positive feedback loops, involving histone PTMs readers and writers, are also essential for the successful propagation of the chromatin state. During DNA synthesis, parental histones are sequentially removed and later assembled on the two daughter strands. Old histones are deposited on both strands of the newly replicated DNA and serve as template for the replication of the epigenetic state (Margueron and Reinberg, 2010; Moazed, 2011) (**Figure 12**). In the

specific case of constitutive heterochromatin maintenance, the positive feedback loop between the presence of H3K9me3 and its recognition by HP1, leads to further recruitment of Suv39 HMT, which allows epigenetic memory.

3. Transcriptional regulation in *Plasmodium falciparum*

3.1 General aspects

P. falciparum malaria parasites have the most AT-rich known eukaryotic genome, with an AT content of 80% that rises to 90-95% in intergenic regions. Its haploid genome is composed of 14 chromosomes containing ~5,500 protein-coding genes (Gardner et al., 2002). Given its complex life cycle and its ability to rapidly adapt to distinct environmental niches, parasites require a complex program of coordinated gene expression. Gene expression during the IDC occurs in a periodic transcriptional cascade with most of the genes expressed "just-in-time", meaning that most of them are just expressed when needed and just once every IDC (Bozdech et al., 2003) (**Figure 13**). while studies using bulk parasite populations suggested a continuous cascade of expression, the recent analysis of transcriptional profiles in single cells revealed a transcriptional cascade directed by sharp transitions of gene expression across the IDC (Reid et al., 2018). For a long time, how parasites regulate gene expression has been of major interest for the field.
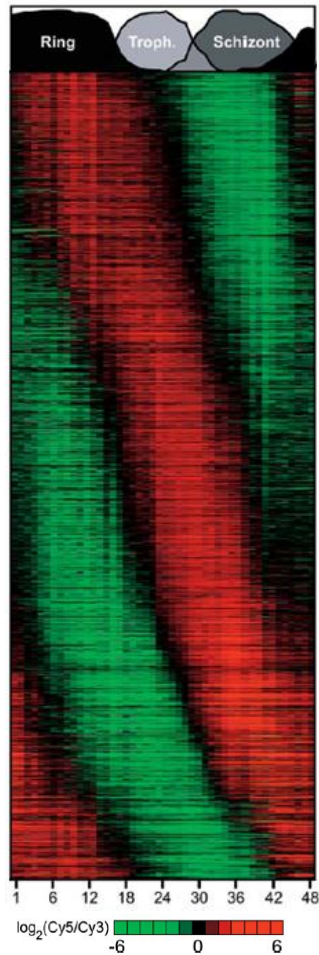
Figure 13. The *P. falciparum* IDC transcriptome. Heatmap of the IDC transcriptome with the transcriptional profiles of 2,712 genes. A cascade of continuous expression lacking clear boundaries is shown. Adapted from Bozdech et al., 2003.

Initial characterization of the parasite genome revealed that it encodes for most of the machinery involved in the assembly of the PIC at the TSSs of promoter regions. Most of the Pol II subunits and GTFs are conserved, but most GSTFs found in other eukaryotes are underrepresented, with a general low number of GSTFs (Callebaut et al., 2005; Coulson, 2004). The low number of GSTFs was

49

hypothesized to reflect the importance of epigenetics and chromatin structure for the control of gene expression in *P. falciparum,* but also of posttranscriptional events (Abel and Le Roch, 2019). How the PIC assembles in the promoter region and drives transcription from a particular TSS remains elusive (Toenhake and Bártfai, 2019). TSSs in *P. falciparum* are distributed within multiple small windows of the promoter region, with some being more prevalent than others (Adjalley et al., 2016; Kensche et al., 2016).

Regions around the TSSs present some characteristic features: a TA dinucleotide at -1 bp, GC-rich elements at -210 bp to +150 bp and homopolymeric tracks of dA- or dT- at around -50 bp (Adjalley et al., 2016; Toenhake and Bártfai, 2019). These features contribute to the establishment of a nucleosome depleted region upstream of the TSSs, allowing the assembly of the PIC (Abel and Le Roch, 2019; Bunnik et al., 2014). As a consequence of their elevated AT-content, intergenic regions are specifically enriched in the histone variant H2A.Z, which establishes weak interactions with the DNA, and affects the stability of nucleosomes in these regions. This lower stability results in a high turnover of nucleosomes in intergenic regions, increasing their accessibility and priming them for transcription (Hoeijmakers et al., 2013; Petter et al., 2013). (**Figure 14**).

Therefore, *P. falciparum* possess a highly permissive genome during the IDC and additional control could be explained by posttranscriptional regulation; indeed malaria parasites have a large repertoire of RNA binding proteins that may account for this additional level of regulation (Reddy et al., 2015; Le Roch, 2004; Vembar et al., 2016).
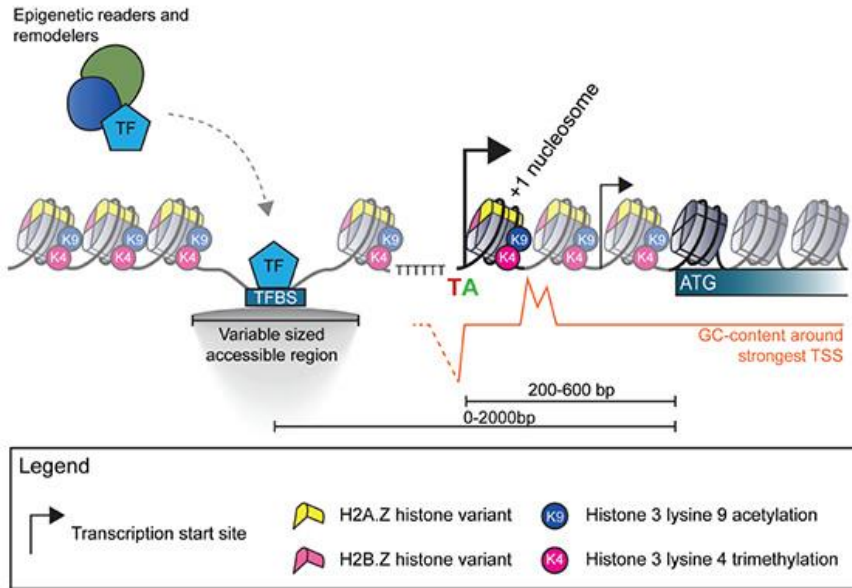
Figure 14. The *P. falciparum* transcriptional unit. Intergenic regions in *P. falciparum* are mainly occupied by H2A.Z/H2B.Z nucleosome variants (H2A.Z in yellow; H2B.Z in pink). Transcription starts within multiple TSS windows in promoter regions (black arrows). At position -1 relative to the TSS a TA dinucleotide can be detected and polymeric regions of A- or T- occur upstream of the TSS. The orange line depicts the typical GC-content around the TSS. A well-positioned nucleosome (more prominent coloring) is located at position +1. Accessible regions are detected around the TSS, containing TF binding sites (TFBSs). Blue circles with K9: H3K9ac; Pink circles with K4: H3K4me3. Adapted from Toenhake and Bártfai, 2019.

3.2 Gene specific transcription factors

Initial characterization of *P. falciparum* GSTFs was based on genome comparative approaches with the DNA binding domains found in other eukaryotic TFs. Only a small representation of conserved DNA binding domains was identified, which was surprising given the complexity of gene expression during parasite development (Aravind et al., 2003; Coulson, 2004). Further analysis revealed a lineage-specific family of proteins that contain a domain similar to Apetala2/ERF

integrase (AP2) DNA binding domain of plants. These apicomplexan-specific TFs were termed ApiAP2 and were probably acquired through horizontal gene transfer from the rhodophyte algae that was an endosymbiont of the apicomplexan ancestor (Balaji et al., 2005).

P. falciparum encodes for 27 members of the ApiAP2 family and other TFs containing the C2H2-type zinc finger domain (12 TFs), helix-turn-helix domains (8 TF), the β-scaffold domain (1 TF), the K homology domain-containing Prx regulatory element (1 TF) and an homeodomain-like domain (1 TF) (Bischoff and Vaquero, 2010; Komaki-Yasuda et al., 2013; Toenhake and Bártfai, 2019). Currently, the ApiAP2 family is by far the best characterized.

P. falciparum ApiAP2 TFs contain 1-4 AP2 domains accompanied by other functional regions, which present little homology between the different members of the family. The size of ApiAP2 TFs is variable, ranging from 200 amino acids (aa) to more than 4,000 aa (Balaji et al., 2005). For most ApiAP2 TFs, the cognate DNA motif has been identified though protein binding arrays. ApiAP2 DNA binding motifs are enriched in the upstream regulatory regions of many genes, pointing to a putative regulation by specific ApiAP2 TFs. Moreover, genes with comparable expression profiles contain similar ApiAP2 binding motifs in their regulatory regions, suggesting that they are coregulated by the same ApiAP2 TFs. The sequential activation of ApiAP2 TFs may explain the gene expression pattern along the IDC of *P. falciparum*: different ApiAP2 may take the control of the specific transcriptional program needed at each stage and the sequential activation of downstream ApiAP2 TFs could drive the transitions in gene expression. Actually, genes encoding ApiAP2 TFs contain AP2

binding motifs in their regulatory regions, revealing an intricate regulation and possibly a hierarchical organization (Campbell et al., 2010).

Most of the ApiAP2 TFs have been characterized by knock-out (KO) screens and gene-centred studies in rodent malaria parasites, revealing the specific involvement of some TFs in different developmental stages (Modrzynska et al., 2017; Zhang et al., 2017) (**Figure 15**). Most of the disruptions resulted in the blockage of progression at a specific stage, whereas some of them were refractory to disruption, suggesting an essential role in the IDC. For instance, AP2-O is essential in ookinetes and its disruption blocks oocyst formation, together with AP2-O2/3/4/5 (Modrzynska et al., 2017; Yuda et al., 2009). AP2-Sp is expressed in late oocysts and regulates known sporozoite genes, evidenced by the loss of sporogony after its disruption (Yuda et al., 2010). Another ApiAP2, termed AP2-L, has been shown to control liver-stage development (Iwanaga et al., 2012).

Apart from regulating stage transitions, some ApiAP2s seem to have structural functions in *P. falciparum*, such as AP2-Tel, which recognizes telomeric repeats and is involved in telomere-end biology (Sierra-Miranda et al., 2017); or SIP2, another ApiAP2 that binds to SPE2 motifs present in subtelomeric genes and maintains their heterochromatin-based silencing (Flueck et al., 2010).

In *P. falciparum*, the best characterized ApiAP2 TFs are PfAP2-I and PfAP2-G (Kafsack et al., 2014; Santos et al., 2017). PfAP2-I contains 3 AP2 domains recognizing different motifs, one of which specifically

binds to a motif enriched in many genes expressed in schizonts and has a function in erythrocyte invasion (Santos et al., 2017).
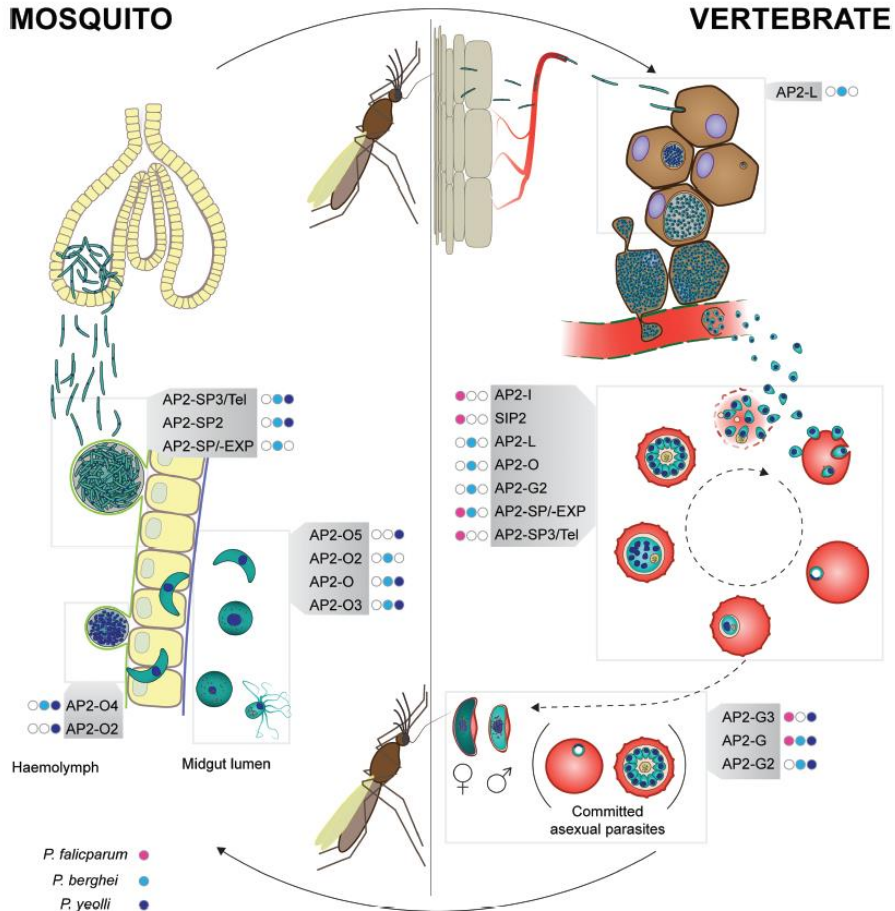
Figure 15. Role of *Plasmodium spp.* ApiAp2 TF across the life cycle. Essential ApiAP2 TFs for stage progression in *P. falciparum* (pink dot), *P. berghei* (light blue dot) and *P. yoelli* (dark blue dot) are indicated. Reproduced from Toenhake and Bártfai, 2019.

PfAP2-I associates with bromodomain-containing proteins such as PfBPD1 and PfBDP2 that recruit chromatin remodelers, assisting in transcriptional regulation (Josling et al., 2015). PfAP2-G is responsible for triggering the transcriptional program of gametocyte development

(Kafsack et al., 2014; Sinha et al., 2014) and will be further discussed in another section.

Other ApiAP2s have also been involved in sexual stage development in rodent malaria parasites, including AP2-G2 and AP2-G3, the latter also termed AP2-FG. AP2-G2 acts downstream of AP2-G repressing genes expressed in asexual stages and mosquito or liver stages, as evidenced by their up-regulation in an AP2-G2 KO line (Yuda et al., 2015). In *P. yoelii* AP2-G3/AP2-FG was suggested to act upstream of AP2-G as its KO leads to loss of gametocyte formation and a decrease in AP2-G expression (Zhang et al., 2017); however, in *P. berghei*, AP2-G3/AP2-FG presents female-specific expression and is essential for female gametocyte development, downstream of AP2-G (Yuda et al., 2019).

Overall, transcriptional regulation by ApiAp2 TFs is complex; some of them act as transcriptional activators whereas others act as repressors. Furthermore, despite being conserved, most of the TFs have different expression patterns across the different *Plasmodium* species, which may reflect a diversification of some functions. For instance, the *cis*-elements bound by AP2 domains appear to be conserved across species but the target gene sets differ extensively (Campbell et al., 2010; Modrzynska et al., 2017). This is also evidenced by important functional differences observed between some orthologue ApiAP2s in *P. falciparum* and rodent malaria parasites, such as PfAP2-EXP/PbAP2-Sp or PfAP2-Tel/AP2-SP3 orthologues. The discrepancies could reflect the experimental approach followed to characterize their function in the different species, the redundancy of the system as a mechanism to safeguard the development of the

parasite, or real differences between *Plasmodium spp.* (Jeninga et al., 2019; Modrzynska et al., 2017; Toenhake and Bártfai, 2019).

Another layer of complexity arises from the combinatorial action of ApiAP2 TFs: complex expression patterns result from hierarchical networks, where TFs work in combination and sometimes synergistically, which depends on the context provided by cis-elements (Lesage et al., 2018; Modrzynska et al., 2017; Van Noort and Huynen, 2006). One example is the overlap between PfAP2-I and PfAP2-G that share many target genes due to similarities in their specific DNA-binding motif. Actually, they have been suggested to coregulate invasion genes in cells undergoing sexual development (Josling et al., 2019). Finally, PTMs of the TFs can also impact their transcriptional activity: lysine acetylation of the ApiAp2 TFs has been shown to affect their DNA-binding capacity, for instance in PfAP2-I, reducing its affinity for DNA (Cobbold et al., 2016).

3.3 Chromatin structure and organization

Epigenetics plays an important role in the regulation of gene expression in *P. falciparum*. Epigenetic mechanisms in malaria parasites are mainly mediated through the structure of the chromatin and its composition. The nucleosome structure is conserved in *P. falciparum*, containing histones H2A, H2B, H3, H4 and the histone variants H2A.Z, H2B.Z, H3.3 and cenH3 (Miao et al., 2006). In *P. falciparum*, both constitutive and facultative heterochromatin are characterized by H3K9me3 and HP1, mainly covering telomeres, subtelomeric regions and some internal chromosome islands (Flueck et al., 2009; Lopez-Rubio et al., 2009). The parasite was suggested to lack methylation of

H3K27 (Miao et al., 2006; Trelle et al., 2009), but recently this PTM has been detected in early gametocyte stages, although the enzyme involved in its deposition remains elusive (Coetzee et al., 2017). Bioinformatics analysis revealed that the parasite possesses almost the full repertoire of histone writer domains (Coulson, 2004; Iyer et al., 2008) (**Figure 16**). Some histone PTMs, and specially histone acetylation and methylation, correlate with transcript levels and their incorporation into specific loci likely depends on the underlying DNA sequence (Cortés and Deitsch, 2017; Gupta et al., 2013).

- *Histone acetylation*

The best characterized HAT is GCN5, which is involved in the acetylation of H3K9 and H3K14 (Cui et al., 2007; Fan et al., 2004). H3K9ac, together with H3K4me3, define euchromatic regions and show dynamic changes during the IDC: H3K9ac levels correlate with temporal patterns of gene expression, whereas H3K4me3 increases throughout the IDC (Bártfai et al., 2010; Gupta et al., 2013; Salcedo-Amaya et al., 2009). Other HATs include an ortholog of HAT1 in higher eukaryotes and a member of the MYST family mediating the acetylation of H4K5, K8, K12 and K16, which are involved in transcriptional regulation, DNA damage responses and DNA replication (Fan et al., 2004; Miao et al., 2010a).

Histone deacetylation is mainly associated with transcriptional repression and is performed by HDACs. In *P. falciparum* there are 5 HDACs: HDAC1, 2 (also termed HDA1) and 3 (also termed HDA2) SIR2A and SIR2B (Cortés and Deitsch, 2017). SIR2A and SIR2B are linked to telomere maintenance and the regulation of the expression

of subtelomeric genes such as the *var* gene family (Duraisingh et al., 2005; Freitas-Junior et al., 2005; Merrick et al., 2015; Tonkin et al., 2009). HDA2 has been linked with the regulation of *var* gene silencing, in addition to sexual conversion, whereas HDA1 might be involved in repressing the asexual program right after sexual conversion (Coleman et al., 2014; Poran et al., 2017; Rono et al., 2018).

Figure 16. Histone PTMs in *P. falciparum*. Schematic of a nucleosome with the four histones (H3, H4, H2A, and H2B) present in *P. falciparum*. Histone tail modifications such as methylation (Me), acetylation (Ac) and ubiquitylation (Ub) are shown with the enzymes catalyzing their addition. Modified from Cui and Miao, 2010.

- *Histone methylation*

Ten HMTs containing the SET domain have been identified in *P. falciparum*. Most of their target residues have been identified by *in vitro* histone methylation assays or by homology with other HMTs (Cui et al., 2008). SET1, SET4 and SET6 methylate H3K4, a modification associated with transcriptional activation; SET10 methylates the same residue and specifically regulates *var* gene expression (Volz et al., 2010, 2012); SET2 methylates H3K36 at *var* gene loci through a Pol II-dependent activity (Jiang et al., 2013; Ukaegbu et al., 2014); SET3, also named PfKMT1, is involved in the methylation of H3K9 (Cui et al., 2008; Lopez-Rubio et al., 2009); SET7 was suggested to be involved in methylating H3K4 and H3K9 (Chen et al., 2016) and SET8 in methylating H4K20, which is also present in heterochromatin (Kishore et al., 2013); SET5 and SET9 methylate unknown residues (Cui et al., 2008). Additionally, the parasite possesses 5 HMTs methylating arginines, but only PRMT4 has been characterized and found to methylate H3R17, a modification usually associated with transcriptionally active sites (Horrocks et al., 2009; Miao et al., 2006). The removal of histone methylation is carried out by members of the two families of HDMs: JMJC and LSD1 (Horrocks et al., 2009). There are two members of the JMJC family: JmJC1, predicted to demethylate H3K9 and H3K36; and JmJC2, whose targets are unknown (Cui et al., 2008; Jiang et al., 2013). LSD1 usually demethylates H3K4 in other organisms but in *P. falciparum* it is not clear whether this role is conserved (Cui et al., 2008; Iyer et al., 2008).

- *Other modifications and unusual histone modifications*

Other modifications such as phosphorylation, ubiquitylation or SUMOylation have been reported in some histone residues, but its impact on transcriptional regulation is still not completely understood (Horrocks et al., 2009). In addition to well-established histone modifications, mass spectrometry analysis identified other residues modified by acetylation in H2A.Z and H2B.Z (Coetzee et al., 2017; Cui and Miao, 2010). Moreover, the parasite possesses unusual epigenetic marks at H3K14, H4K5 and H4K16, in addition to methylation of the core domain, which have not been reported in any other organism (Coetzee et al., 2017; Miao et al., 2006).

- *Nuclear architecture*

Genome organization in the nucleus of the parasite was initially described by microscopy-based techniques. The parasite nucleus is organized into two compartments: heterochromatin-repressed regions, represented by subtelomeric genes and some internal loci localizing at the nuclear periphery; and euchromatic active regions occupying the interior of the nucleus (Freitas-Junior et al., 2000; Lopez-Rubio et al., 2009; Ralph et al., 2005). More recently, chromatin conformation capture techniques, such as Hi-C, validated this view, but surprisingly *P. falciparum* lacks the typical TADs described in other organisms, revealing a much simpler nuclear organization. Like in other organisms, centromeres and telomeres cluster on opposite regions of the nucleus. Nevertheless, during the IDC and the development to transmission stages, the nuclear organization is remodelled and the architecture correlates with the level of transcriptional activity at each

stage: the nucleus expands, reaching its maximum size at the trophozoite stage, reflecting a higher accessibility of the chromatin and a higher transcriptional activity. During schizogony, chromosome territories are well defined and chromatin recompacts. Across development there are changes associated with the silencing of chromosomal regions of individual loci, which are reflected by the intrachromosomal interactions that follow the nuclear repositioning of the regions involved (Ay et al., 2014; Batugedara et al., 2017; Bunnik et al., 2019).

3.4 Clonally variant gene expression

Studies regarding the epigenetic regulation of gene expression in *P. falciparum* have focused on clonally variant gene (CVG) families. CVGs can be found both in an active or silenced state in genetically identical parasites at the same stage of development and both states are clonally transmitted (Cortés et al., 2012). The epigenetic state of the locus is stably transmitted along generations, thus CVGs lay within bistable chromatin regions (Crowley et al., 2011) (**Figure 17**).

Most of the studies on CVGs have focused on *var* genes, which are very relevant for the pathogenesis of the parasite. Clonally variant expression (CVE) is an intrinsic property of many gene families involved in processes such as antigenic variation, erythrocyte invasion, solute transport and sexual development (Cortés and Deitsch, 2017). CVE results in isogenic parasite populations that become transcriptionally heterogeneous during normal growth, such that individual parasites have different combinations of active and silenced genes (Rovira-Graells et al., 2012). Changes from one state to another

of a CVG occur at low frequencies through stochastic switches in their epigenetic makeup.

Clonally variant genes (CVGs)
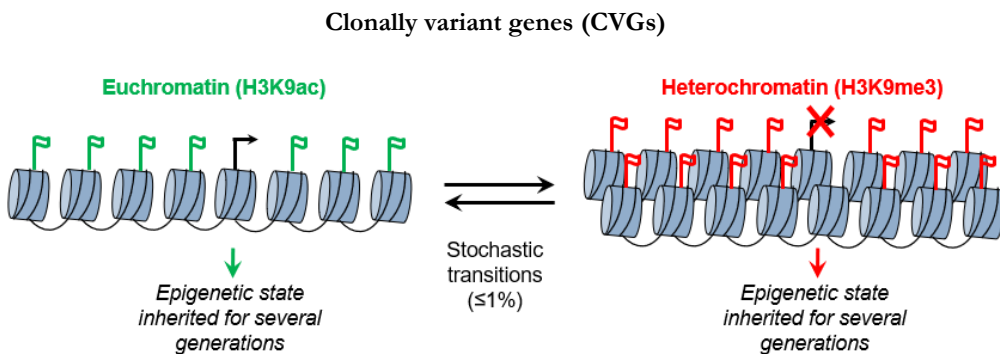
Figure 17. Epigenetic regulation of CVGs. In *P. falciparum* facultative heterochromatin controls the expression of CVGs, which are linked to the virulence and pathogenesis of the parasite. CVGs are non-repetitive loci that can adopt either an active or a repressed conformation through stochastic transitions in the epigenetic state. Once established, both states are stably inherited for several generations (bistable chromatin domains). Active chromatin (euchromatin) is associated with H3K9ac, whereas silent chromatin (heterochromatin) is associated with H3K9me3.

The mechanisms behind the switching of CVGs have not been elucidated, but may be explained by stochastic processes or by differences arising in the epigenetic regulators in response to environmental cues (Horrocks et al., 2004; Recker et al., 2011; Voss et al., 2014). Most CVGs are located in subtelomeric regions under the control of facultative heterochromatin; the silenced state is characterized by H3K9me3 and HP1, whereas the active state is characterized by H3K9ac (Crowley et al., 2011; Flueck et al., 2009; Lopez-Rubio et al., 2009; Rovira-Graells et al., 2012).

- *The bet-hedging adaptive strategy*

The diversity associated with CVE confers plasticity to parasite populations, providing the grounds for natural selection of the

parasites with highest fitness in an always-changing environment (Rovira-Graells et al., 2012). This adaptive strategy is termed bet-hedging, where populations generate phenotypic diversity through stochastic mechanisms before the environmental challenge occurs. Such strategy is commonly used in bacteria, where populations are several orders of magnitude higher than in multicellular organisms. In these populations, not all the cells have optimal fitness but they provide enough diversity to overcome unexpected changes in the environment, even if many cells are not able to survive (Carey et al., 2018; Casadesús and Low, 2013; Levy et al., 2012) (**Figure 18**).
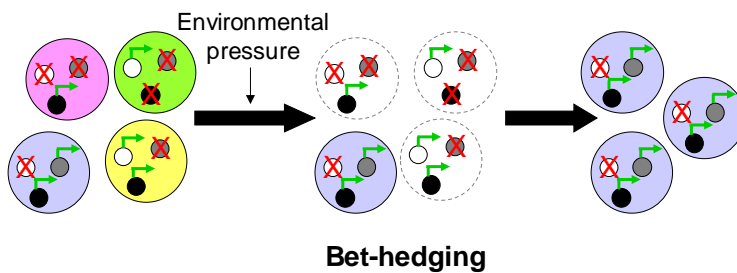
Bet-hedging

Figure 18. Bet-hedging strategy for the adaptation to environmental changes. Individual genes (small circles) can be either repressed (crossed) or active (green arrow) generating a transcriptionally heterogeneous population with different cell phenotypes (coloured big circles). Upon an expected change in the environment, parasites with combinations of expressed and silenced genes that confer increased fitness are selected and survive under the new conditions, whereas other parasites die (broken line). Adapted from Rovira-Graells et al., 2012.

In malaria parasites, there are only a few well characterized genes where transcriptional diversity drives adaptation to an environmental condition (Deitsch and Dzikowski, 2017; Mira-Martínez et al., 2013, 2017, 2019). However, it can be speculated that most CVGs might reflect a bet-hedging adaptive strategy against currently unidentified environmental challenges.

- *The var gene family and antigenic variation*

CVE plays an essential role in the antigenic variation of malaria parasites – the basis of its immune evasion strategy (Deitsch and Dzikowski, 2017). Parasites are able to maintain chronic infections even in the presence of protective antibody responses. In such conditions, the parasite population fluctuates, reflecting a cyclic outgrowth of the parasites that are not yet recognized by the immune system. These parasites will eventually be recognized and eliminated by the immune system; however, parasites with alternative surface antigens will arise and be able to survive in the host (Pasternak and Dzikowski, 2009). In general, antigenic variation in parasites is provided by CVE of surface antigens (**Figure 19**).

The best characterized variant surface protein is PfEMP1, encoded by the *var* gene family consisting of 60 members (Smith et al., 1995; Su et al., 1995). PfEMP1 mediates the sequestration of infected RBCs in the microvasculature through its interaction with endothelial cell receptors, which avoids splenic clearance. Moreover, different PfEMP1 proteins possess diverse binding specificities allowing sequestration within different tissues, such as the brain or the placenta. Switches in *var* gene expression allow subpopulations of parasites to proliferate in the presence of immune responses against previous PfEMP1 proteins (Bernabeu and Smith, 2017; Miller et al., 2002; Montgomery et al., 2007; Pasternak and Dzikowski, 2009) (**Figure 19**). On top of this switching characteristic of clonally variant genes, *var* genes present mutually exclusive expression, such that only a single member of the family is transcribed in individual parasites (Scherf et al., 1998).
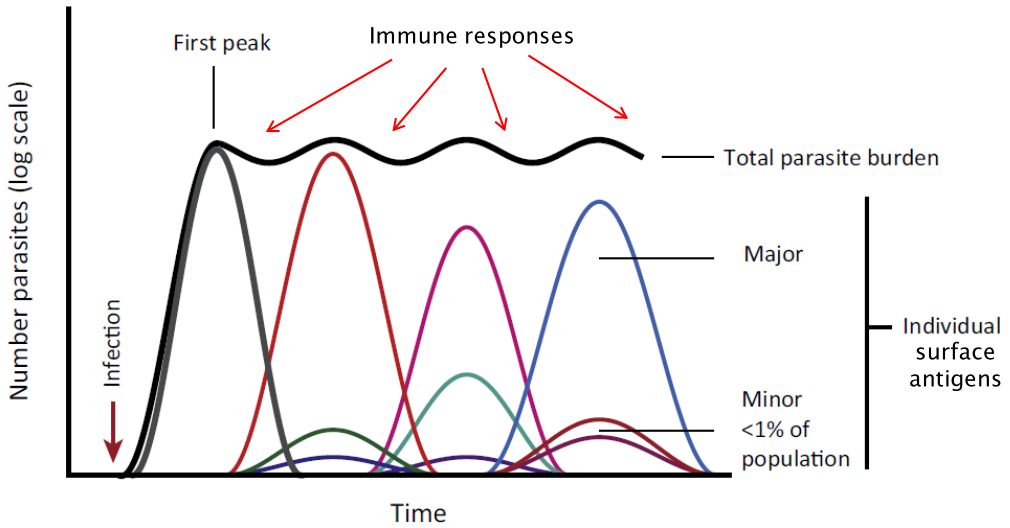
Figure 19. Impact of immune responses against variant surface antigens on parasite population dynamics. A small number of variant surface antigens ('major') can dominate parasite populations, but the total number of variants expressed is greater. In this figure the first peak is dominated by a particular variant, but in *P. falciparum* the first peak of parasitaemia is characterized by parasites expressing different variants. There are many variants that are present in low abundance ('minor'). Once the first major variant is recognized by the immune system, parasites expressing a minor variant become the major and this will occur with successive rounds of antigenic variants. Each colour represents parasites expressing different surface antigens. Modified from McCulloch and Field, 2015.

Despite the characterization of some essential players for the active and silenced state of *var* genes, the mechanism behind mutually exclusive expression remains elusive (Deitsch and Dzikowski, 2017). As all clonally variant genes, the repressed state of *var* genes is characterized by H3K9m3 and HP1 protein, and also requires HDA2 and SIR2A leading to their clustering in the nuclear periphery (Brancucci et al., 2014; Chookajorn et al., 2007; Coleman et al., 2014; Duraisingh et al., 2005; Lopez-Rubio et al., 2009; Tonkin et al., 2009). Nevertheless, *var* genes are also controlled by H3K36me3 present in the *cis*-regulatory elements and the coding sequence (CDS) of the repressed *var* genes, but absent from the TSS of the active *var* gene

65

(Jiang et al., 2013; Volz et al., 2012). The active *var* gene is also defined by the presence of the histone variants H2A.Z and H2B.Z during its transcription at the ring stage, plus H3K9ac and H3K4me3 (Bártfai et al., 2010; Hoeijmakers et al., 2013; Lopez-Rubio et al., 2007; Petter et al., 2013; Volz et al., 2012). Moreover, antisense lncRNAs are produced from the intron of the active *var* gene during the trophozoite and schizont stage, with essential functions for the maintenance of the active state (Amit-Avraham et al., 2015; Zhang et al., 2014).

- *Other clonally variant gene families involved in antigenic variation*

In addition to *var* genes, *P. falciparum* contains other multicopy gene families encoding surface antigens, such as *rif*, *stevor*, *Pfmc-2tm* and *surfins* (Kaviratne et al., 2002; Kyes et al., 1999; Sam-Yellowe, 2004; Winter et al., 2005). RIFIN are even a larger family than *var* genes, with 150–200 members, which have been associated to several functions both in asexual and sexual stages. Some RIFIN variants bind to endothelial cells and erythrocytes resulting in the formation of rosettes (clusters of uninfected erythrocytes bound to an infected erythrocyte) (Goel et al., 2015). Others can bind to immune receptors inhibiting B cells and natural killer cells (Saito et al., 2017). The STEVOR family consists of 39 members that can present stage dependent expression, such that some of them are expressed in late stages, others in merozoites, gametocytes or sporozoites, with localisation changing between stages (Lavazec et al., 2007). This might reflect their involvement in different functions at different parasite stages. In fact, they have been described to have roles in merozoite invasion, the formation of rosettes and the alteration of host cell rigidity both in asexual stages and gametocytes (Niang et al., 2009,

2014; Tibúrcio et al., 2012). The Maurer's cleft two transmembrane protein (PfMC-2TM) family consists of 13 members with a stage-specific profile, some members specifically expressed in early stages and others in late stages; however, their biological role remains elusive. Similar to PfMC-2TM, the function of SURFIN proteins is not well-understood, with just 1 of the 10 members being involved in merozoite invasion (Kaur and Hora, 2018; del Pilar Quintana et al., 2018). Additionally, some members of the *phist* and *hyp* CVG families are also exported to the erythrocyte membrane and therefore may contribute to antigenic variation; nevertheless, their functions are currently unknown (Duraisingh and Skillman, 2018; Kumar et al., 2019).

- *Roles of clonally variant genes beyond antigenic variation*

CVE also plays an important role in other cellular processes such as invasion of the host cell and solute transport through the host cell membrane (Cortés and Deitsch, 2017) (**Figure 20**). The *eba* and *Pfrh* gene families encode for adhesins released from the apical organelles of the merozoite that interact with the receptors present in the surface of erythrocytes to mediate apical attachment during invasion (Cowman et al., 2017). The multiple parasite ligand and host receptor interactions reflect alternative invasion pathways. Variant expression in these gene families may mediate phenotypic variation in the invasion process by determining alternative pathways (Cortés et al., 2007; Crowley et al., 2011; Duraisingh et al., 2003; Jiang et al., 2010; Taylor et al., 2002).
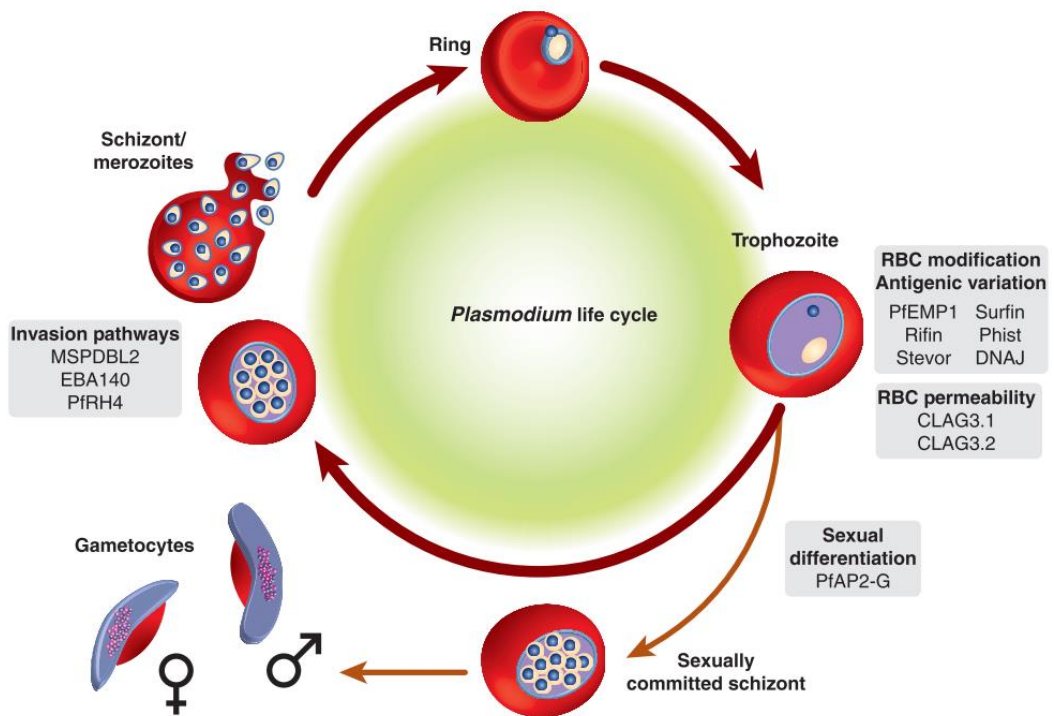
Figure 20. Roles of clonally variant gene expression in *P. falciparum* blood stages. At the trophozoite stage, parasites express clonally variant proteins involved in erythrocyte modification, antigenic variation and cell permeability. In schizonts/merozoites, clonally variant proteins can determine alternative erythrocyte invasion pathways. Sexual commitment is initiated by the activation of PfAP2-G leading to the production of gametocytes. Reproduced from Cortés and Deitsch, 2017.

P. falciparum also harbours the *clag* gene family, which has been linked to invasion and nutrient acquisition by the formation of channels in the surface of the infected erythrocyte (Cortés et al., 2007; Kaneko et al., 2005; Nguitragool et al., 2011; Pillai et al., 2012). CLAG proteins associate with RhopH2 and RhopH3 to form the RhopH complex, which is secreted from the rhoptry bulb to assist in the invasion process and later exported to the erythrocyte membrane (Ito et al., 2017).

Each RhopH complex contains only one CLAG protein out of the 5 members of the family: *clag2*, *clag3.1*, *clag3.2*, *clag8* and *clag9* (Kaneko et al., 2005). Of the 5 members, only *clag2*, *clag3.1* and *clag3.2* present CVE. In addition, *clag3.1* and *clag3.2* are expressed in a mutually exclusive manner (Cortés et al., 2007; Rovira-Graells et al., 2015). CLAG3 proteins are required for the formation of the plasmodial surface anion channel (PSAC) that increases the membrane permeability of infected erythrocytes to multiple solutes, underlying the so called new permeation pathways (NPPs), which are involved in nutrient acquisition (Nguitragool et al., 2011; Pillai et al., 2012). The expression of alternative CLAG3 results in differences in solute transport, allowing the parasite to adapt to varying concentrations of nutrients and/or toxic compounds (Mira-Martínez et al., 2019). The roles of the other members of the family remain to be determined, but they might be involved in mediating the selectivity of the NPPs to different solutes (Cortés and Deitsch, 2017; Duraisingh and Skillman, 2018; Mira-Martínez et al., 2019).

Finally, another cellular process controlled by CVE is the regulation of sexual conversion, which depends on the variantly expressed gene *pfap2-g* (Cortés and Deitsch, 2017) (**Figure 20**). Due to the relevance of the topic for this thesis, a detailed review of sexual conversion will be provided in the following sections.

4. Biology of transmission stages

4.1 Gametocyte development

The formation of male and female gametocytes is a prerequisite for the transmission of malaria (Baker, 2010). After commitment to sexual

development, *P. falciparum* parasites develop into gametocytes through five morphologically distinguishable stages (I-V) along 10-12 days, an unusually long process compared to other *Plasmodium* species (Ngotho et al., 2019) (**Figure 21**).
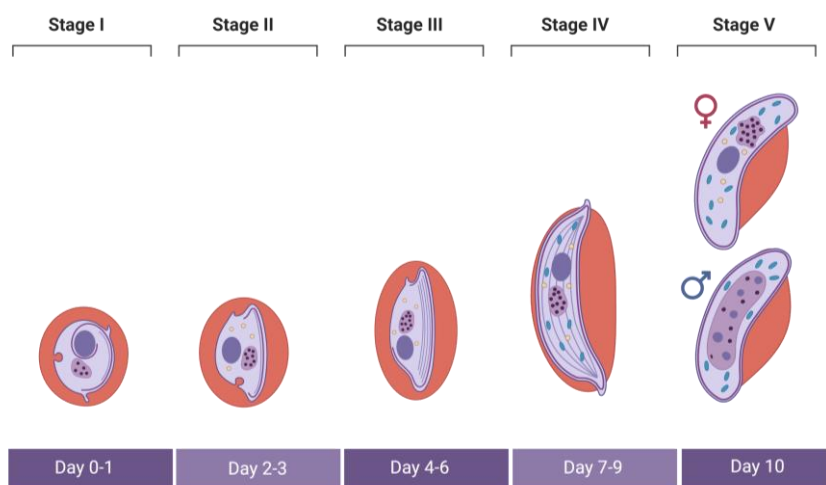
Figure 21. *P. falciparum* gametocyte development. Gametocytes develop through five distinct morphological stages: stage I is morphologically similar to trophozoites, whereas stage II can already be distinguished as the parasites start to elongate. In stage III, gametocytes elongate and the ends become rounded. Stage IV gametocytes elongate even more but have pointed ends. In stage V, gametocytes adopt their characteristic crescent shape. Stage V male and female gametocytes can be differentiated, with females being more elongated and curved than males. Designed with Biorender.com.

Gametocyte stages II-V can be easily identified in Giemsa-stained blood smears, whereas stage I gametocytes are almost identical to young trophozoites and can only be confidently identified by immunofluorescence assay (IFA) using antibodies against early gametocyte markers. Male and female gametocytes can be morphologically distinguished by trained researchers from stage IV onwards: female gametocytes have a smaller nucleus and a concentrated hemozoin pigment pattern, whereas male gametocytes

have a larger nucleus and a more diffuse pigment. Moreover, female gametocytes have a higher content of osmophilic bodies, organelles involved in the secretion of enzymes involved in the disruption of the host membrane during the egress of the gamete (Baker, 2010).

Stage V gametocytes are the only stage present in the peripheral blood, suggesting that during maturation parasites are sequestered in tissues (Meibalan and Marti, 2016). Over the last years, the bone marrow and the spleen have been confirmed as the main reservoirs of maturing gametocytes (Aguilar et al., 2014; Farfour et al., 2012; Smalley et al., 1981). Gametocytes are enriched in the bone marrow parenchyma and specifically in the erythroblastic islands (Joice et al., 2014), which may reflect a specific preference for this niche.

The mechanism of sequestration is still not understood, but there are several hypotheses regarding the homing to the bone marrow (Nilsson et al., 2015) (**Figure 22**). Sexually committed parasites or immature gametocytes have been suggested to extravasate and reach the bone marrow parenchyma, where they develop into stage V gametocytes that will later intravasate to reach the peripheral circulation. It is speculated that exported proteins may be required for sequestration and association with the erythroblast islands, while changes in the host cell deformability allow stage V gametocytes to reach the circulation again. In addition to evading recognition by the immune system or clearance by the spleen, sequestration in the bone marrow may provide a favourable environment for the development of gametocytes: richness in nutrients, abundant young erythrocytes, and an aerobic environment (Ngotho et al., 2019; Nilsson et al., 2015). These conditions might not be essential for gametocyte development

as *in vitro* cultures support the growth of viable gametocytes capable of infecting mosquitoes. Nevertheless, *in vitro* studies with erythroid progenitors showed increased gametocyte formation within young erythrocytes (Peatey et al., 2013).
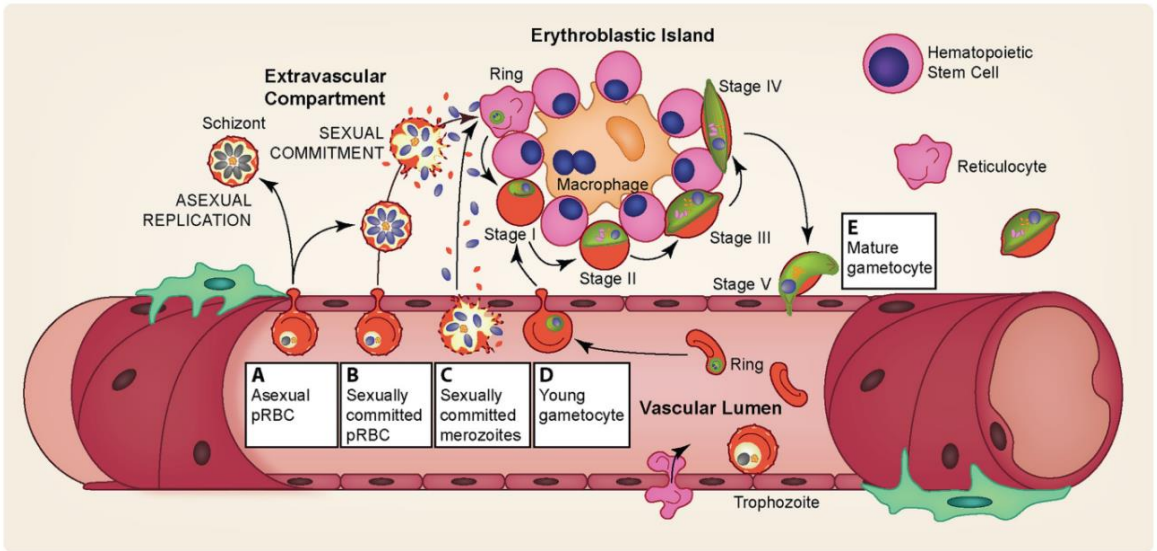
Figure 22. Models of gametocyte sequestration in the bone marrow. A, Asexual parasites adhere to the endothelium of the bone marrow and transmigrate into the extravascular space. Once inside, parasites either continue growing asexually or commit to production of gametocytes. **B**, Sexually committed parasites specifically "home" to the bone marrow through adhesion to the endothelium and transmigrate into the extravascular space. Committed merozoites invade the erythroid progenitor cells. **C**, Sexually committed schizonts can "home" to the bone marrow without transmigrating. The committed merozoites could transmigrate and invade the erythroid progenitor cells. **D**, Young gametocytes "home" to the bone marrow and transmigrate into the extravascular space. **E**, Stage V gametocytes exit the extravascular space and the return to the circulation thanks to their increased deformability. Reproduced from Nilsson et al., 2015.

The maturation of sexual stages is characterized by the formation of a multilayer membranous structure termed inner membrane complex (IMC) (Kaidoh et al., 1993). In association with the plasma membrane it forms the pellicle, a structure linked to the cytoskeleton, serving as a scaffold for parasite structures involved in gliding motility (Baum et

al., 2006). Sporozoites and ookinetes actively invade and traverse their host cells; nevertheless, whether gametocytes are capable of active motility remains elusive. Recently, movement of *P. berghei* gametocytes within the bone marrow parenchyma was reported, suggesting a possible mechanism of active motility (De Niz et al., 2018). In stage V gametocytes, part of the cytoskeleton network associated with the IMC is disassembled, leading to the rounding of the gametocyte tips and the relaxation of the membranes, which also contributes to the increased deformability of mature gametocytes. The exit of stage V gametocytes from their sequestration is enabled by this higher deformability (Aingaran et al., 2012; Dearnley et al., 2016; Parkyn Schneider et al., 2017; Tibúrcio et al., 2012).

- *Gametocytes transcriptional profile*

The early steps of sexual development are characterized by expression of early gametocyte genes, such as *Pfs16*, *Pfg27/25*, *Pfg14.744*, *Pfg14.745* and *Pfg14.748* in addition to several members of the *phist* gene family, which are required for the development of gametocytes (Alano et al., 1991; Bruce et al., 1994; Eksi et al., 2005; Kongkasuriyachai et al., 2004; Lobo et al., 1999). The transcriptomic comparison between the 3D7 line and the gametocyte-deficient line F12 identified additional early gametocyte markers, such as *pfge-3* (*mdv1*) and *pfge-4* (*etramp10.3*) (Silvestrini et al., 2005). A larger number of gametocyte markers are available for later gametocyte stages. However, until recently, the only *P. falciparum* gametocyte transcriptomic time-courses, covering mature gametocyte stages, were more than a decade old and contained a mixture of asexual parasites and gametocytes (Eksi et al., 2005; Young et al., 2005). However, by

comparing these datasets with expression data across the whole life cycle it was possible to generate transcriptional networks associated with each stage. This analysis revealed gene clusters expressed at different gametocyte stages providing a detailed profile of the transcription dynamics during gametocyte development (Pelle et al., 2015). The integration of gametocyte transcriptional and proteomic data has also provided a list of gametocyte markers in *P. falciparum* (Meerstein-Kessel et al., 2018). Recently, a more detailed time-course analysis of sexual commitment and gametocyte development has been performed (Biljon et al., 2019).

Posttranscriptional regulation also plays an essential role in gametocyte development, for instance stabilizing transcripts such that they can accumulate without the need of high transcription rates (Josling et al., 2018). RNA-binding proteins, such as PUF proteins, are involved in the regulation of transcripts during gametocytogenesis. PUF proteins generally act as translational repressors by binding to transcripts and preventing their translation (Miao et al., 2013). For instance, *PfPuf2* has been suggested to be a repressor of sexual development, as its disruption leads to increased gametocyte production. *PfPuf2* interacts with the 5' and 3' of the transcripts blocking their translation (Miao et al., 2010b). *PfPuf1* is also involved in gametocyte development as its KO impairs gametocyte development (Shrestha et al., 2016). Moreover, in female gametocytes many genes required for parasite development in the mosquito are transcribed but the messenger RNAs (mRNAs) are stored through a translational repression mechanism (Guerreiro et al., 2014; Khan et al., 2005; Mair, 2006).

- *Gametocytes proteome*

At the protein level, comparative proteomic analysis revealed the upregulation of many specific exported proteins in the early steps of gametocytogenesis. These proteins were named gametocyte exported proteins (GEXPs) and represented a tenth of all the proteins detected in gametocytes (Silvestrini et al., 2010). Most of them are from the *phist* gene family that might be involved in host cell remodelling processes, such as the restructuration of the host cytoskeleton (Kumar et al., 2019). STEVOR proteins also play a role in host cell remodelling by altering the deformability of the cell, characteristic of stage V gametocytes (Tibúrcio et al., 2012). Moreover, female gametocytes have abundant ribosomal and vesicular network proteins to reinitiate translation in the mosquito midgut. Enzymes involved in the mitochondrial Krebs cycle are also abundant (Florens et al., 2002).

- *Gametocytes metabolism*

At the metabolic level, *Plasmodium* parasites cannot generate glucose and rely on glucose uptake from the host cells. Asexual parasites mainly perform anaerobic glycolysis, with <7% of the glucose processed in the mitochondrion through the Krebs cycle (MacRae et al., 2013). However, despite lacking pyruvate dehydrogenase, parasites are thought to metabolize pyruvate into acetyl-coA through the action of α-keto acid dehydrogenase. Therefore, parasites can metabolize glucose and glutamine through a highly flexible Krebs cycle to produce energy (Ke et al., 2015; MacRae et al., 2013; Ngotho et al., 2019). In fact, gametocytes uptake more glucose than asexual parasites and mostly catabolize it in the Krebs cycle to obtain energy (MacRae

75

et al., 2013). Most of the enzymes involved in the mitochondrial Krebs cycle are upregulated in gametocytes (Khan et al., 2005; Young et al., 2005).

Lipids are essential in *Plasmodium*, being involved in processes such as protein trafficking, haemoglobin degradation or synthesis of membranes (Mi-Ichi et al., 2006). Parasites obtain fatty acids from the host cell and use them to synthesize lipids, but they also possess a fatty acid biosynthetic pathway (Grellier, 1991; Tarun et al., 2009). In gametocytes, phospholipids and glycerolipids levels decrease probably reflecting the non-proliferative nature of sexual stages and an altered substrate use. However, sphingolipids and some ceramides increase in sexual stages and have been reported to be essential for gametocyte maturation (Gulati et al., 2015; Tran et al., 2016). Enzymes involved in fatty acid elongation and phosphoethanolamine methyltransferase (PMT) have critical roles in gametocyte development (Bobenchik et al., 2013; Ikadai et al., 2013). Moreover, gametocytes use different lipid moieties from the culture media than asexual parasites during maturation (Lamour et al., 2014).

4.2 Development in the mosquito

Once ingested by the vector, gametocytes activate in the mosquito midgut and undergo sexual reproduction, allowing genetic recombination and thus providing genetic diversity to the population. After activation, male gametocytes produce eight flagellated microgametes in a process termed exflagellation, whereas females round up before the egress from the erythrocyte and develop into a single macrogamete. Fertilization results in the formation of a zygote

that after a meiotic division develops into an ookinete. Ookinetes are motile and have the capacity to invade the midgut epithelium. Within the epithelium, ookinetes develop into oocysts, where they replicate asexually for almost 2 weeks into thousands of sporozoites, later released upon oocyst rupture. These sporozoites migrate to the salivary glands from where they can be transmitted to a new host during a blood meal (Bennink et al., 2016).

Stage V gametocytes activate in the mosquito midgut by environmental stimuli: mainly a >5ºC drop of temperature and the presence of xanthurenic acid (XA), a metabolite of tryptophan catabolism. Parasites can also sense an increase of the extracellular pH from 7.2 to 8 (Billker et al., 1998; Garcia et al., 1998; Kawamoto et al., 1991). The presence of XA triggers the activation of guanylyl cyclase in gametocytes, leading to the synthesis of the second messenger cyclic GMP (cGMP) (Muhia et al., 2001). cGMP activates the cGMP-dependent protein kinase (PKG) that increases the production of inositol-(1,4,5)-trisphosphate (IP3) and the release of Ca^{2+} from the endoplasmatic reticulum (Raabe et al., 2011). The increase of intracellular Ca^{2+} activates calcium-dependent protein kinases (CDPKs), which triggers the release of the translational repression of mRNAs in female gametocytes and the assembly of the axoneme, and DNA replication and exflagellation in male gametocytes (Billker et al., 2004; Sebastian et al., 2012) (**Figure 23**).
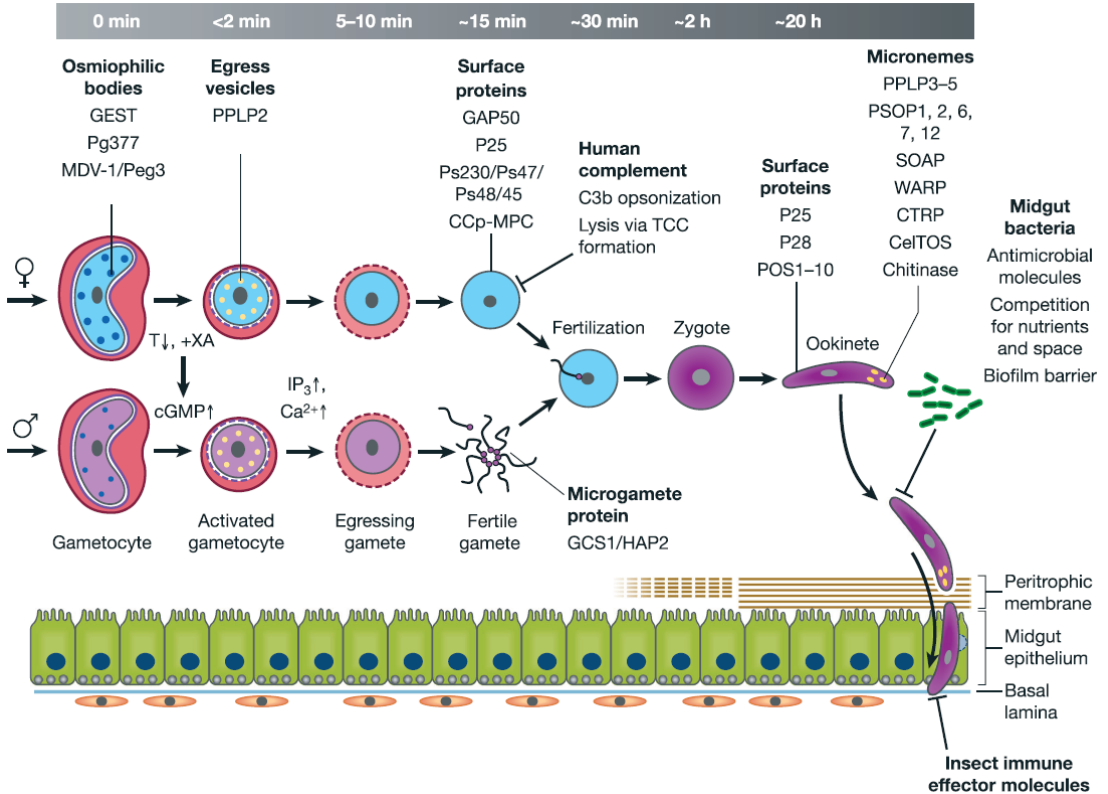
Figure 23. Signalling events triggering gametocyte activation and key proteins involved in the development within the mosquito midgut. See main text for the description of the events involved in gamete formation. T, temperature. Reproduced from Bennink et al, 2016.

During the egress of gametocytes from the erythrocyte, the parasitophorous vacuole membrane ruptures before the erythrocyte membrane. The osmiophilic bodies are responsible for the degradation of the parasitophorous vacuole, as they accumulate underneath the rupture sites and disappear simultaneously with its destruction (Sologub et al., 2011; Wirth and Pradel, 2012). Around 6 min after activation, another type of vesicles containing the plasmodial pore-forming perforins (PPLP2) are released into the cytoplasm of the erythrocyte to perforate the erythrocyte membrane, resulting in the

release of the erythrocyte cytoplasm. Around 9 min later the erythrocyte membrane is disassembled, releasing the fertile gamete (Sologub et al., 2011; Wirth and Pradel, 2012; Wirth et al., 2014). During fertilization, the nuclei of the gametes fuse and divide by meiosis, which produces a tetraploid zygote (**Figure 23**). Tetraploidy persists until oocyst replication to produce sporozoites, restoring the haploid state (Bennink et al., 2016).

The mosquito midgut represents a major bottleneck for the parasites, as they must survive for more than 20 h without a host cell providing protection from the extracellular environment. Factors of the human immune system present in the blood, the natural microflora of the midgut and the immune system of the mosquito are the main threats that parasites have to face. This environment causes a 300-fold reduction in parasite abundance (Bennink et al., 2016; Vaughan et al., 1994). Some studies also reported a fitness cost for mosquito fecundity associated to the presence of parasites (Mitchell and Catteruccia, 2017). However, a recent study suggested that *P. falciparum* has evolved a non-competitve strategy based on the use of nutrients present in the mosquito that are not needed for egg development (Werling et al., 2019). These observations contrast with previous results in rodent malaria parasite infections showing a reproductive cost (Ahmed and Hurd, 2006; Hogg and Hurd, 1997). Nevertheless, rodent malaria parasites trigger strong immune reactions that may have side-effects in egg development (Blandin et al., 2004; Dong et al., 2006). This effect may not occur in *P. falciparum* due to the surface protein Pfs47 that avoids complement activation (Molina-Cruz et al., 2013; Werling et al., 2019).

5. Regulation of sexual conversion

Our understanding of the mechanisms regulating the trade-off between asexual replication and sexual differentiation is limited. Gametocytogenesis has been proposed to be the default developmental decision in the apicomplexan ancestor, as haemoproteids (a phylum of apicomplexans) convert into gametocytes right after merozoite invasion of the avian RBCs (Sinden, 2009). Thus, mechanisms to control gametocytogenesis must have evolved at some point, conferring the ability to replicate continuously in the host, enhancing the effective population and therefore transmission. Committing to sexual conversion or continuing replicating asexually is the only developmental step where *P. falciparum* has alternative options (Josling et al., 2018). Moreover, the development of sexual stages provides an evasion mechanism in front of unfavourable conditions (Baker, 2010).

5.1 Commitment to sexual conversion

Back in 1979, different models of gametocytogenesis were proposed based on the point at which commitment may occur (Carter and Miller, 1979) (**Figure 24**). Years later, studies on sexual commitment determined that every merozoite within a schizont will become either a gametocyte or an asexual parasite (Bruce et al., 1990). Bruce and colleagues made use of the gametocyte specific antigen Pfg27 and the schizont/merozoite specific antigen MSP2 to identify the fate of the progeny of single schizonts. To achieve this, they performed plaque assays based on the infection of an immobilized monolayer of

erythrocytes with an overlaid culture, which was later processed to perform an IFA.
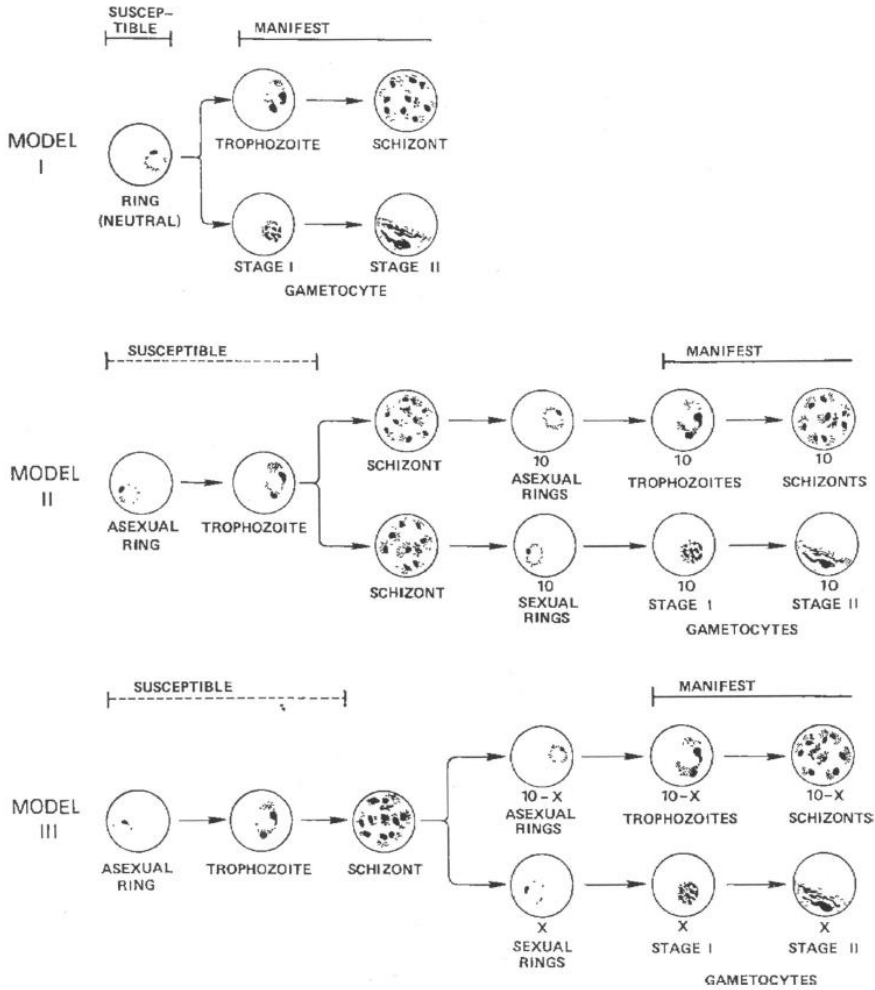
Figure 24. Models for gametocytogenesis. Three models for gametocytogenesis were proposed based on different times of commitment (susceptibility) and different patterns of sexual conversion (manifestation), according to the presence of stage I gametocytes. In model I, a merozoite or a ring can commit either into an asexual parasite or into a gametocyte. In Models II and III commitment occurs during one asexual generation, but gametocyte development occurs after reinvasion. In Model II, commitment determines whether the progeny will develop entirely into gametocytes or entirely into asexual parasites. In Model III only a proportion of the merozoites within a schizont will become gametocytes. From Carter and Miller, 1979.

By these means, the offspring of single schizonts invades the erythrocytes present in a particular area, forming clusters of parasites or the so-called plaques. Schizonts preferentially produced only pure sexual plaques or pure asexual plaques. Plaques containing both asexual and sexual parasites were attributed to double-infected erythrocytes in the overlaid culture. Therefore, the authors concluded that commitment to sexual development occurs the cycle before sexual conversion, prior to the onset of schizogony (Model II from Carter and Miller). Similar experiments tracking the sex of gametocytes within the plaques determined that all sibling parasites become either male or female gametocytes, indicating that sex determination also occurs before the schizont stage (Silvestrini et al., 2000; Smith et al., 2000).

During the IDC of *P. falciparum* only a few schizonts commit to sexual conversion, typically less than 10% of the parasites (Josling et al., 2018). The low levels of sexual conversion have historically limited the biological characterization of gametocytes and especially of the early committed stages, since they are morphologically identical to their asexual counterparts (Baker, 2010). Moreover, sexual conversion rates vary between parasite lines and even between isogenic subclones, which suggests the involvement of genetic or epigenetic mechanisms (Josling et al., 2018; Kafsack et al., 2014). Some culture-adapted lines lose their ability to develop into gametocytes, suggesting a fitness cost for maintaining sexual capacity and a selection for traits blocking sexual conversion (Claessens et al., 2017; Eksi et al., 2012; Kafsack et al., 2014; Sinha et al., 2014). One example is the partial deletion of chromosome 9 resulting in loss of gametocyte production mainly due

to the loss of the gametocyte developmental protein 1 gene (*gdv1*): complementation with an episomally expressed GDV1 restored sexual conversion (Eksi et al., 2012).

- *Identification of the master regulator of sexual conversion*

For decades, the molecular mechanism triggering sexual conversion remained elusive and was considered the "Holy Grail" of parasite biology (Baker, 2010). The genome-wide analysis of heterochromatin revealed that H3K9me3 is mainly restricted to CVGs and, interestingly, a single member of the ApiAp2 family of TFs, PFL1085w, was positive for H3K9me3, suggesting that it may present CVE (Lopez-Rubio et al., 2009). Back in 2012, Rovira-Graells and colleagues performed a genome-wide study of transcriptional variation in isogenic subclones, confirming that PFL1085w has CVE (Rovira-Graells et al., 2012). Further analysis of the transcriptional patterns of the isogenic subclones revealed a correlation between the expression of this ApiAP2 TF and early gametocyte markers (**Figure 25**).

Additional characterization of parasites lines that lost the capacity to produce gametocytes, but had an intact chromosome 9, revealed premature stop mutations in PFL1085w, and the complete deletion of the PFL1085w CDS resulted in loss of gametocyte production (Kafsack et al., 2014). Additionally, conditional degradation of the protein also demonstrated its essentiality for gametocyte production, which led to the naming of PFL1085w as PfAP2-G (Kafsack et al., 2014).
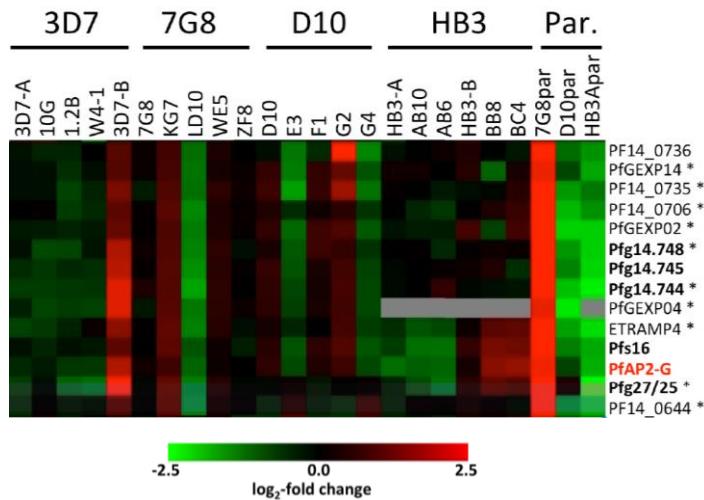
Figure 25. Variant expression of PfAP2-G correlates with the expression pattern of early gametocyte markers. Clustering analysis of the expression of gametocyte specific genes (rows) in 21 parasite clones (columns) revealing co-clustering with PfAP2-G. Early gametocyte markers (bold) and genes enriched in the early gametocyte proteome (*) are shown. Adapted from Kafsack et al., 2014.

Finally, a transcriptomic analysis revealed that PfAP2-G is responsible for triggering the expression of the sexual stage transcriptional programme (Kafsack et al., 2014). At the same time, the role of AP2-G in sexual conversion was also reported in *P. berghei* (Sinha et al., 2014). In fact, AP2-G is conserved in all *Plasmodium* species, highlighting its essential role in the life cycle of malaria parasites, ensuring the formation of gametocytes (Meibalan and Marti, 2016).

5.2 The epigenetic regulation of PfAP2-G

Epigenetic mechanisms play an essential role in the regulation of *pfap2-g* expression: during asexual growth it is epigenetically silenced by heterochromatin based on H3K9me3 and presence of HP1 protein (Flueck et al., 2009; Josling et al., 2018; Lopez-Rubio et al., 2009). Conditional degradation of HP1 at 30-40 h after erythrocyte invasion

resulted in 50% sexual conversion the following cycle, which indicates a role for HP1 in maintaining the repressed state in schizonts. Transcriptional analysis upon HP1 depletion revealed an increase of *pfap2-g* transcripts and also of early gametocyte genes (Brancucci et al., 2014). Another epigenetic regulator involved in maintaining the repressed state is HDA2: its depletion doubled the number of gametocytes produced. Thus, HDA2 might be important for deacetylating H3K9ac to allow the formation of H3K9me3 (Coleman et al., 2014) (**Figure 26**).

It is clear that H3K9me3, HP1 and HDA2 are required for the maintenance of the silenced state of *pfap2-g*; nevertheless, they are general regulators of CVGs (Cortés et al., 2012; Voss et al., 2014). This implies that additional DNA elements or DNA binding factors must be involved in the targeting of these regulators to *pfap2-g*. Therefore, additional work is needed to identify the elements that maintain the repressed state, but also the ones involved in its activation. In fact, single cell transcriptomic analysis provided an initial characterization of sexually committed schizonts, identifying several histone remodelers (ISWI, SNF2L), epigenetic enzymes (LSD1, HDA1) and ApiAP2 TFs (PF3D7_1222400, PF3D7_1139300) enriched in sexually committed cells. This provided a hypothetical model of the drivers of PfAP2-G activation (Poran et al., 2017).

Another key regulator of *pfap2-g* is GDV1, which was initially identified by studying gametocyte-deficient lines with a deletion in a subtelomeric region of chromosome 9 (Eksi et al., 2012). More recently, GDV1 was reported to be involved in the removal of HP1 from the *pfap2-g* locus and some early gametocyte gene loci (Filarsky et

al., 2018). Therefore, GDV1 is an upstream regulator of *pfap2-g* that operates by derepressing the gene through dismantling heterochromatin silencing. Moreover, GDV1 is controlled by the expression of an antisense RNA that blocks its production.
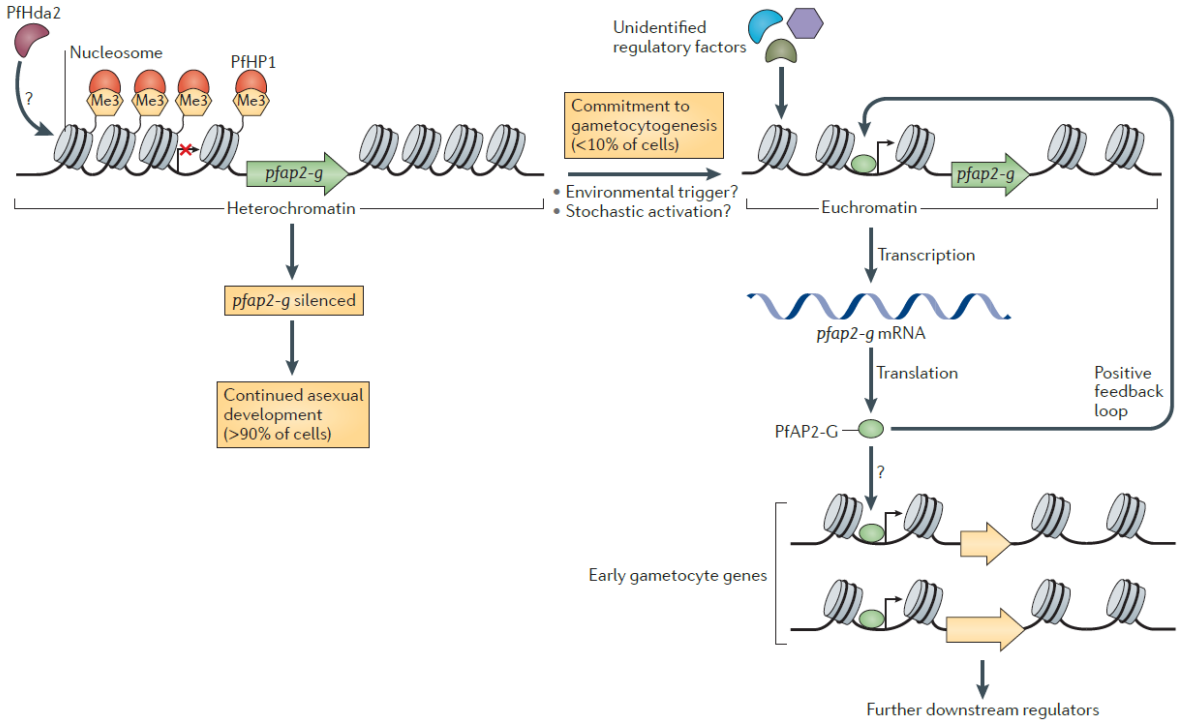
Figure 26. Regulation of *pfap2-g* in *P. falciparum*. *pfap2-g* is only transcribed in a small proportion of parasites (<10%). In most parasites, the *pfap2-g* locus is heterochromatic, characterized by H3K9me3 and PfHP1. PfHDA2 is also required for the maintenance of the silenced state. At each cycle of growth, the *pfap2-g* locus becomes euchromatic in a few parasites. *pfap2-g* expression occurs stochastically and certain environmental signals can increase the rate of activation. Other unidentified complexes are likely to be involved in the activation. PfAP2-G drives the expression of early gametocyte genes but also can bind to its own promoter in a positive feedback loop. Reproduced from Josling and Llinás, 2015.

pfap2-g was also suggested to be regulated by the mechanisms involved in mutually exclusive expression of the *var* gene family, such that *pfap2-g* would enter the *var* mutually exclusive expression program. Indeed,

recent studies using purified gametocytes detected a strong downregulation of *var* genes in stage II gametocytes compared with ring stages (Tibúrcio et al., 2013). Moreover, flanking the *pfap2-g* locus there are insulator-like pairing elements (PE), such as the ones present in *var* gene loci (Alano, 2014; Kafsack et al., 2014). PE are involved in the maintenance of *var* gene silencing by mediating long-range interactions between the *var* gene promoter and the intron promoter. These interactions are essential for *var* silencing and their arrangement is involved in the mechanism of mutually exclusive expression (Avraham et al., 2012). Nevertheless, *var* genes are regulated by SET2, involved in the deposition of H3K36me3, which is absent from *pfap2-g*, and is essential for mutually exclusive expression (Deitsch and Dzikowski, 2017; Lopez-Rubio et al., 2009; Volz et al., 2012). Evidence for mutually exclusive expression of *pfap2-g* and *var* genes is lacking and testing the hypothesis that it occurs will require further investigation (Josling and Llinás, 2015).

Upon activation, PfAP2-G binds to the upstream region of gametocyte-specific genes, driving the sexual conversion process (Josling et al., 2019; Kafsack et al., 2014). It recognizes the palindromic motif: (GxGTAC/GTACxC), identified by *in vitro* experiments (Campbell et al., 2010). PfAP2-G mainly acts as a transcriptional activator that triggers the transcriptional cascade driving sexual conversion (Kafsack et al., 2014). The AP2-G binding motif is also present within the upstream region of *pfap2-g* itself, allowing a positive feedback loop that enhances its own transcription (Josling et al., 2019; Kafsack et al., 2014; Poran et al., 2017) (**Figure 26**). During sexual differentiation, the *pfap2-g* locus relocalizes in the

nucleus abandoning the perinuclear repressive centre (Bunnik et al., 2018). Several chromatin changes at other loci also take place, such as the expansion of heterochromatin to silence genes involved in erythrocyte remodelling or erythrocyte invasion (Bunnik et al., 2018; Fraschka et al., 2018).

Despite all these observations, currently there is no evidence that AP2-G alone is sufficient to initiate sexual development as other regulators might provide additional checkpoints. Therefore, whether AP2-G is sufficient to drive sexual conversion needs to be determined by assessing whether every cell that express AP2-G develops into a gametocyte or if additional regulators are required. Similarly, it would be important to determine whether a particular threshold of AP2-G is required to trigger gametocytogenesis (Bechtsi and Waters, 2017).

5.3 Stochastic and induced sexual conversion

Despite the current understanding of the epigenetic regulation of *pfap2-g*, the mechanisms triggering the activation of the locus remain elusive. *pfap2-g* is controlled by similar mechanisms to other CVGs, suggesting that its activation can also be driven by stochastic switches and may explain the low frequency of *pfap2-g* activation. Variation in gametocyte production among parasite lines (and even clones) may reflect the intrinsic stochasticity of *pfap2-g* activation (Kafsack et al., 2014). Actually, *pfap2-g* activation could reflect a bet-hedging strategy for the population: ensuring transmission of a few parasites upon unpredictable changes of the environment that may affect survival in the host (Bechtsi and Waters, 2017). The stochastic activation of *pfap2-*

g is essential to maintain production of gametocytes, even if it is only at low levels (Josling et al., 2018).

Along the years, several factors have been linked to induce sexual conversion. Initial studies on infected individuals suggested a possible association with clinical symptoms (Meibalan and Marti, 2016). In fact, under *in vitro* conditions higher parasite densities result in higher conversion rates (Bruce et al., 1990), a relationship that may not be applicable in human infections, where lower parasite densities were associated with higher gametocyte numbers (Drakeley et al., 2006). These differences can be explained by host factors such as immunity, anaemia, haemoglobin variants or drug treatments (Drakeley et al., 2006; Gouagna et al., 2010; Nacher et al., 2002; Trager and Gill, 1992). Other factors have been reported to increase gametocyte production under *in vitro* conditions, such as the presence of young erythrocytes; reticulocytes or erythroid precursors; and antimalarial drugs (Buckling et al., 1999; Peatey et al., 2013; Trager et al., 1999). The epidemiological setting also influences gametocyte production as evidenced by lower sexual conversion rates in high transmission areas than in low transmission areas, probably reflecting a selective pressure for the optimisation of transmission (Rono et al., 2018).

For a long time, the use of spent medium (or parasite-conditioned medium) was known to increase gametocyte production (Dyer and Day, 2003; Mantel et al., 2013; Williams, 1999), which was exploited in several protocols to stress the culture and induce sexual conversion (Brancucci et al., 2015; Delves et al., 2016; Duffy et al., 2016; Fivelman et al., 2007; Roncalés et al., 2012). Spent medium is enriched in extracellular vesicles secreted by infected erythrocytes, and they were

initially suggested to be the triggers of sexual conversion (Mantel et al., 2013; Regev-Rudzki et al., 2013). However, a specific component of the human serum was recently identified as being responsible for the effect of spent medium: lysophosphatidylcholine (LysoPC) (Brancucci et al., 2017).

Mass spectrometry analysis of culture media and spent medium revealed that LysoPC was present in the medium fractions that blocked sexual conversion. Depletion of LysoPC from the media before 34-38 h after invasion leads to irreversible induction of sexual conversion in a large proportion of the parasites. LysoPC is metabolized by the parasite as a source of choline and converted into phosphatidylcholine via the Kennedy pathway (Brancucci et al., 2017). Phosphatidylcholine is especially relevant during schizogony, being the most abundant phospholipid in the parasite membrane. When LysoPC is depleted from the culture media, parasites can produce choline *de novo* by phosphorylation and methylation of ethanolamine through the activity of ethanolamine kinase (EK) and PMT using S-adenosylmethionine (SAM) as a methyl donor (Bobenchik et al., 2013; Pessi et al., 2004). In fact, upon LysoPC depletion there is an upregulation of *pfap2-g*, *ek*, *pmt*, *sam synthetase* and other factors involved in the metabolic adjustment of the cell (Brancucci et al., 2017). The requirement of SAM to maintain the Kennedy pathway active in the absence of LysoPC has been suggested to directly impact the epigenetic silencing of *pfap2-g*. The pool of SAM also serves as a methyl donor for HMTs, such as the one involved in the methylation of H3K9. The use of SAM to produce phosphatidylcholine might reduce its availability in the nucleus, affecting the activity of the HMTs

and the repression of their target genes (Llinás, 2017). Nevertheless, other genes such as *sir2a*, *hda2*, *lsd1* and *gdv1* responded before *pfap2-g* upon LysoPC depletion, suggesting their potential involvement in regulating its expression (Brancucci et al., 2017). While some mechanistic details still remain unclear, it has been well established that LysoPC depletion impacts both cell metabolism and sexual conversion via changes in the expression of metabolic enzymes and *pfap2-g*.

Interestingly, the depletion of choline from the culture media increased the number of GDV1-positive cells and therefore sexual conversion, which linked the signalling of lipid metabolism with *gdv1* regulation (Filarsky et al., 2018). It is likely that the absence of LysoPC is transduced into the silencing of the antisense RNA allowing the production of GDV1 and the activation of *pfap2-g*. Nevertheless, the authors did not provide a direct link between the expression of the antisense and the depletion of choline (**Figure 27**).

For the first time, an environmental cue triggering sexual conversion has been characterized at the molecular level, revealing a signalling process that results in a metabolic response linked to sexual conversion (Brancucci et al., 2017). At the physiological level, LysoPC levels can reflect parasite biomass, with higher parasitemias leading to reduced levels of LysoPC in the plasma, possibly triggering sexual conversion (Brancucci et al., 2017; Lakshmanan et al., 2012; Orikiiriza et al., 2017). This may operate as a quorum sensing mechanism. Moreover, LysoPC levels are lower at tissues like the bone marrow, which may induce sexual conversion and partly explain gametocyte abundance in this tissue (Brancucci et al., 2017). However, parasites

cannot rely on the environment to ensure transmission; therefore, a basal rate of sexual conversion, based on stochastic activation, is needed to sustain the formation of gametocytes at each cycle.
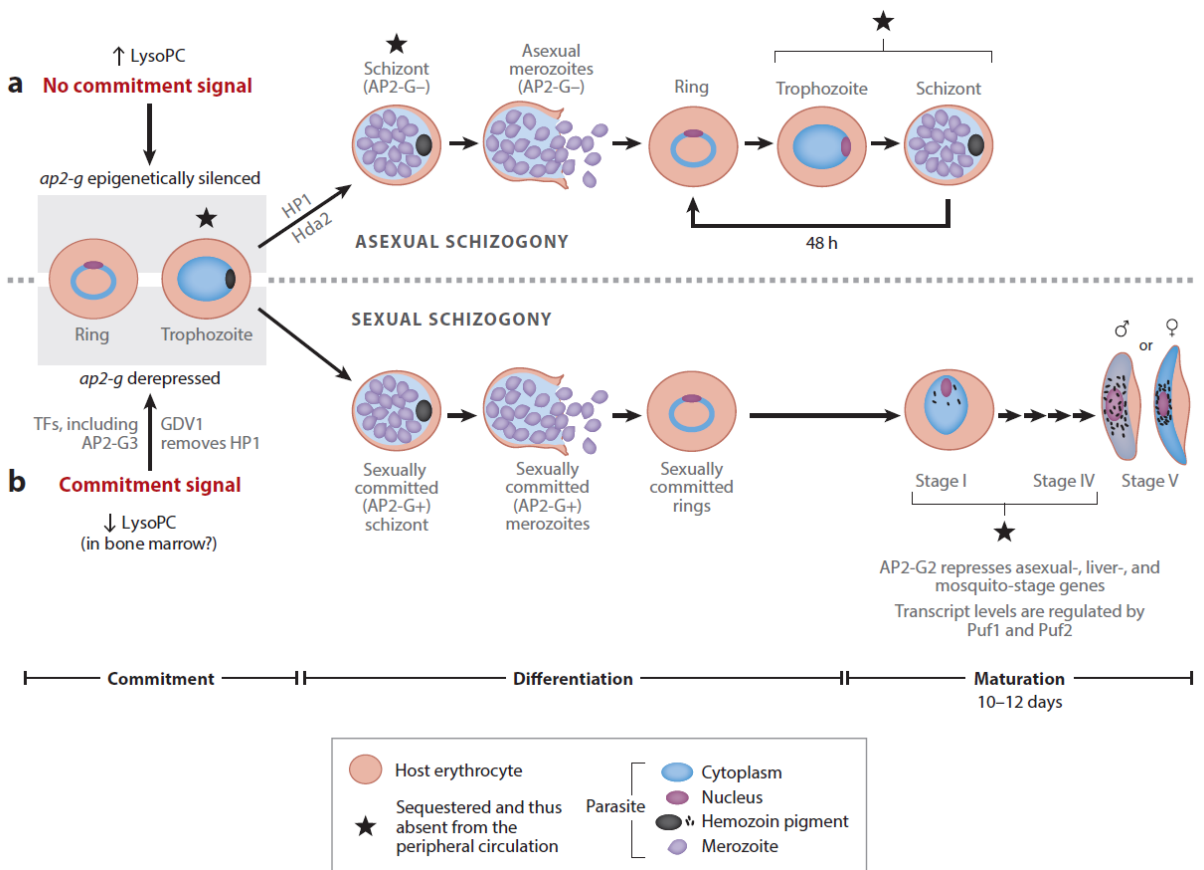
Figure 27. Sexual commitment is environmentally regulated. a, In the presence of high levels of LysoPC, the *pfap2-g* locus is epigenetically silenced by H3K9me3 and HP1 involving Hda2. **b,** When LysoPC levels are lower, parasites are more likely to commit to sexual conversion. Sexual commitment requires the removal of HP1 by GDV1, allowing *pfap2-g* transcription. Downstream of AP2-G, other epigenetic and transcriptional regulators are expressed, controlling the gametocyte transcriptional program. Reproduced from Josling et al., 2018.

Future work should determine the factors involved in the stochastic activation of *pfap2-g*, how the basal rates are defined in different

parasite lines and how environmental cues are transduced into *pfap2-g* activation. Another important question is why only a small proportion of parasites can overcome the epigenetic silencing of *pfap2-g* and convert into gametocytes even under inducing conditions. After depletion of choline Brancucci et al. detected 50% of GDV1-positive parasites, but only 30% sexual conversion, suggesting that not all the cells that express GDV1 will be able to activate the *pfap2-g* locus, which probably reflects an extra layer of regulation that determines the activation of *pfap2-g*.

The identification of the genetic switch controlling sexual conversion represents an important milestone, allowing the experimental control of sexual conversion. This would enable the characterization of early committed and early sexual stages that remain poorly characterized. The control of sexual conversion would provide the basis for a detailed transcriptional analysis of these poorly characterized stages, identifying much-needed additional markers for sexually committed schizonts and rings. Such markers would be very valuable for epidemiological studies to assess parasite investment in transmission, by predicting transmission potential 12 days later. This would allow to test the effect of an intervention such as drug administration or vector control strategies (Josling and Llinás, 2015). In fact, PfAP2-G itself could be used as a marker of sexually committed parasites in epidemiological settings. However, most epidemiological studies tracking transmission stages are based on the detection of sex specific markers expressed in stage III-V gametocytes, such as *Pfs25*, *Pf77* and *Pfg377* for females; *Pfs230p*, *Pf13* and *PfMGET* for males and *Pfg27* for both sexes (Tadesse et al., 2019).

HYPOTHESES AND OBJECTIVES

Despite all the recent advances in our understanding of the regulation of *pfap2-g* and sexual conversion, there are still many unresolved questions. A detailed characterization of sexually committed schizonts and sexual rings is lacking, mainly due to the low levels of sexual conversion and the difficulties to obtain sufficiently pure preparations of these stages. Moreover, there is still no evidence that PfAP2-G expression is sufficient to initiate sexual development. Likewise, some aspects of the mechanisms involved in the regulation of the transcriptional state of the gene remain unknown. For instance, it is still unclear how heterochromatin specifically assembles at the *pfap2-g* locus.

We hypothesize that the use of PfAP2-G as a marker of committed cells can provide insight into the early steps of sexual conversion. Moreover, a conditional activation system for PfAP2-G may enable controlled conversion of the majority of parasites in a culture, allowing a detailed characterization of committed parasites and sexual stages. We also hypothesize that heterochromatin nucleation is driven by a sequence-dependent mechanism, as RNAi-based mechanisms mediating nucleation in other organisms are absent in *P. falciparum*.

The specific objectives to test these hypotheses are the following:

- To study sexual conversion using PfAP2-G as a marker of sexually committed cells.
- To develop a conditional activation system for PfAP2-G.
- To characterize committed schizonts and sexual rings.
- To identify the specific sequences involved in heterochromatin nucleation at the *pfap2-g* locus.

RESULTS

1. Article 1: Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*.

Bancells C, Llorà-Batlle O, Poran A, Nötzel C, Rovira-Graells N, Elemento O, et al. Revisiting the initial steps of sexual development in the malaria parasite Plasmodium falciparum. Nature microbiology. 2019;4(1):144–54. DOI: 10.1038/s41564-018-0291-7

2. Article 2: Conditional expression of *pfap2-g* for controlled massive sexual conversion in *Plasmodium falciparum*.

Llorà-Batlle O, Michel-Todó L, Witmer K, Toda H, Fernández-Becerra C, Baum J, et al. Conditional expression of PfAP2-G for controlled massive sexual conversion in Plasmodium falciparum. Science advances. 2020;6(24). DOI: 10.1126/sciadv.aaz5057

3. Article 3: Specific regions of the *Plasmodium falciparum pfap2-g* gene can nucleate heterochromatin.

Llorà-Batlle, O., Michel-Todó, L. & Cortés, A. Specific regions of the *Plasmodium falciparum pfap2-g* gene can nucleate heterochromatin. (**In preparation**)

Specific regions of the *Plasmodium falciparum pfap2-g* gene can nucleate heterochromatin formation

Oriol Llorà-Batlle[1], Lucas Michel-Todó[1] and Alfred Cortés[1,2,*]

[1]ISGlobal, Hospital Clinic – Universitat de Barcelona, Barcelona 08036, Catalonia, Spain
[2]ICREA, Barcelona 08010, Catalonia, Spain

*Correspondence: alfred.cortes@isglobal.org (A.C.)

1

ABSTRACT

About one tenth of the genes of the malaria parasite *Plasmodium falciparum* show clonally variant expression, such that they can be found in either an active or a silenced state in genetically identical parasites at the same stage of development. Clonally variant genes play fundamental roles in host-parasite interactions. Both the active and the silenced states of these genes can be transmitted through multiple rounds of asexual replication by epigenetic mechanisms, with switches between the two states occurring at low frequency. The active state of clonally variant genes is characterized by euchromatin at the promoter, whereas the silenced state is associated with facultative heterochromatin based on tri-methylation of histone H3 lysine 9 (H3K9me3) and the heterochromatin protein 1 (HP1). How this repressive chromatin structure is targeted to specific loci in the genome is not known. Here we studied heterochromatin nucleation in *P. falciparum* focusing on the *pfap2-g* gene, which encodes the master regulator of sexual conversion. This gene is in a silenced, heterochromatic state in asexual parasites, with activation resulting in conversion into sexual stages necessary for transmission. We found that DNA fragments of the *pfap2-g* coding sequence ectopically integrated in a previously euchromatic region of the genome can drive *de novo* formation of heterochromatin, whereas fragments from the *pfap2-g* upstream region, including the predicted promoter, were unable to do so. This result is consistent with DNA sequence-dependent heterochromatin nucleation in malaria parasites. Even for nucleation-competent sequences, heterochromatin only formed in a fraction of the parasites, indicating that the process is inefficient and heterochromatin silencing in malaria largely depends on inheritance from previous generations, rather than on *de novo* formation.

2

INTRODUCTION

Regulation of gene expression is essential for cell development and also for cell plasticity. One of the major mechanisms to regulate gene expression is by modulating the accessibility of the transcription machinery to the DNA (Andersson et al., 2015). In the nucleus, DNA is packed as chromatin, a complex of DNA and histones. Histones can be modified by posttranslational modifications (PTMs) that determine their interactions with DNA and also with other proteins. Histone PTMs can modulate chromatin accessibility, allowing or preventing the interaction of the transcription machinery or other factors with the DNA (Lawrence et al., 2016).

Chromatin can be broadly classified into euchromatin (more accessible) or heterochromatin (less accessible). The latter can be classified into constitutive heterochromatin, which is permanently found in a heterochromatic state, and facultative heterochromatin, which is defined as transcriptionally silent chromatin that has the potential to turn into euchromatin, allowing transcription. Constitutive heterochromatin is typically found in telomeres and subtelomeric regions, where it plays a central role in the maintenance of chromosome integrity. It is also involved in silencing repetitive elements such as transposons. In contrast, facultative heterochromatin participates in the regulation gene expression, mainly in developmentally-regulated genes (Allshire and Madhani, 2018).

In higher eukaryotes, facultative heterochromatin is mainly based on the Polycomb Repressive Complex 2 (PRC2), which mediates tri-methylation of lysine 27 of histone H3 (H3K27me3). This histone mark is recognized by PRC1, which ubiquitylates H2AK119 and leads to chromatin compaction (Trojer and Reinberg, 2007). On the other hand, constitutive heterochromatin is characterized by deposition of H3K9me3, which is recognized by heterochromatin protein 1 (HP1) that represses transcription and prevents recombination of repetitive DNA elements (Nishibuchi and Déjardin, 2017). Although the H3K9me3 modification has been typically associated with constitutive heterochromatin, there is also evidence for its involvement in facultative heterochromatin. For instance, H3K9me3 can assist in the regulation

3

of gene family clusters, such as zinc finger transcription factors, olfactory receptors or neurotransmitter genes (Becker et al., 2016).

How heterochromatin based on H3K27me3 or H3K9me3 is nucleated at specific regions of the genome is still not completely understood (Becker et al., 2016; Trojer and Reinberg, 2007; Wiles and Selker, 2017). For facultative heterochromatin based on H3K27me3, in *Drosophila* the Polycomb Responsive Elements recruit PRC2 to silence specific loci, but many organisms lack these elements. In yeast and plants, some loci use sequence-specific PcG targeting mechanisms such as long non-coding RNAs or transcription factor binding motifs to trigger heterochromatin formation (Wiles and Selker, 2017). The characterization of constitutive heterochromatin formation in *Schizosaccharomyces pombe* revealed a connection between the RNA interference (RNAi) pathway and the targeting of the H3K9 histone methyltransferase (HMT) to repetitive pericentromeric loci. Transcripts arising from these highly repetitive loci are used as a template for the production of double-stranded RNAs (dsRNAs), which are processed by the DICER complex to produce small interference RNAs (siRNAs). The siRNAs bind to the nascent transcripts of the locus and recruit the RNA-induced transcriptional silencing complex (RITS), engaging the H3K9 HMT Clr4 (ortholog of Suv39 in mammals) for the deposition of H3K9me3. Dimers of Chp2 (ortholog of HP1 in mammals) bind to H3K9me3, recruiting the SHREC complex that contains a histone deacetylase and enables further methylation of H3K9me3. This positive feedback loop is essential for the maintenance and spreading of heterochromatin domains (Holoch and Moazed, 2015; Saksouk et al., 2015). In higher eukaryotes, the involvement of non-coding RNAs in H3K9me3-based heterochromatin nucleation is not as clearly established as in yeast (Saksouk et al., 2015). However, RNAi-independent heterochromatin formation can involve protein complexes that recruit the HMTs and trigger nucleation. For instance, the silencing of retroelements involves KRAB-Zinc finger DNA binding proteins that recruit the H3K9 HMT SETDB1 (Imbeault et al., 2017). Zinc finger TFs also drive transcription of pericentromeric regions, which is essential for Suv39 recruitment (Bulut-Karslioglu et al., 2012). Additionally, H3K9me3-based heterochromatin spreading and maintenance in higher eukaryotes also depends

4

on a positive feedback loop, which in this case involves binding of HP1 to H3K9me3, and recruitment of Suv39 by HP1 (Allshire and Madhani, 2018).

In the malaria parasite *Plasmodium falciparum*, facultative heterochromatin controls the expression of clonally variant genes, which play a fundamental role in host-parasite interactions by mediating processes such as immune evasion, solute transport, cell invasion or sexual development (Cortés and Deitsch, 2017). Clonally variant genes are located in subtelomeric regions and in a few internal heterochromatin islands. These genes are characterized by bistable chromatin domains, which can adopt either an active (euchromatin) or a repressed (facultative heterochromatin) state (Crowley et al., 2011). Once established, both states are stably inherited for several generations, with switches between the two states occurring infrequently. Clonally variant expression generates cellular heterogeneity within isogenic populations. This heterogeneity is the basis of an adaptive strategy termed 'bet-hedging', where the population invests in cell diversity to ensure the survival of some individuals in front of unexpected environmental changes (Rovira-Graells et al., 2012; Satory et al., 2011).

Facultative heterochromatin in *P. falciparum* is characterized by H3K9me3 and HP1. These heterochromatin marks are present in subtelomeric repeats and also in coding parts of the genome (Flueck et al., 2009; Lopez-Rubio et al., 2009), where they largely coincide with clonally variant genes (Rovira-Graells et al., 2012). The parasite lacks the PcG proteins, which was thought to explain the reported absence of H3K27 methylation (Miao et al., 2006; Trelle et al., 2009). However, recent data suggest that H3K27me3 may be present in sexual stages (Coetzee et al., 2017), although its functional relevance has not been established yet. In recent years, the characterization of heterochromatin components such as the malarial orthologs of Suv39 and HP1 (Brancucci et al., 2014; Lopez-Rubio et al., 2009; Pérez-Toledo et al., 2009) has improved our understanding of the regulation of clonally variant genes. Nevertheless, how heterochromatin formation is targeted to some particular loci, resulting in variant expression, remains unknown. *P. falciparum* lacks all the machinery involved in the RNAi pathway (Mueller et al., 2014) and the recruitment of HP1 is independent of RNA components (Flueck et al., 2009). Therefore, we

5

hypothesize that heterochromatin nucleation in particular genomic regions is driven by a sequence-dependent mechanism, rather than through the formation of siRNAs.

To characterize heterochromatin formation in *P. falciparum* clonally variant genes, we focused on *pfap2-g*. This gene encodes an ApiAp2 gene-specific transcription factor (Campbell et al., 2010; Jeninga et al., 2019) that is the master regulator of sexual conversion in *P. falciparum* (Kafsack et al., 2014). During the asexual replication of the parasite in the human blood, the *pfap2-g* locus remains silenced by the presence of heterochromatin involving H3K9me3 and HP1 (Brancucci et al., 2014; Flueck et al., 2009; Kafsack et al., 2014; Lopez-Rubio et al., 2009). We selected *pfap2-g* as a model to study heterochromatin formation in *P. falciparum* because the active state of the locus is never inherited, as parasites that activate the locus stop replicating. Therefore, the vast majority of the population has this gene silenced, which is expected to facilitate the characterization of the molecular events involved in the nucleation of heterochromatin. Moreover, most clonally variant genes are clustered in subtelomeric regions, where the influence of the heterochromatin environment of telomeres cannot be excluded. In contrast, *pfap2-g* is located at an internal chromosome region, forming an isolated heterochromatic island. This suggests that heterochromatin nucleation at the *pfap2-g* locus is independent from its proximal chromosomal environment.

Here we characterized the *pfap2-g* promoter and integrated overlapping fragments of the *pfap2-g* locus in an ectopic location to test their ability to mediate *de novo* heterochromatin formation. We also tested the effect of deleting parts of the *pfap2-g* locus on its heterochromatin domain. We found that some regions of the *pfap2-g* coding sequence (CDS) can mediate heterochromatin nucleation, but the process is inefficient, as it only occurs in a fraction of the parasites.

6

RESULTS

Characterization of the *pfap2-g* regulatory region

We first characterized the *pfap2-g* regulatory region by mapping the TSSs by 5'
Rapid Amplification of cDNA Ends (5'RACE). Previous genome-wide studies on
transcription initiation did not identify any TSSs for *pfap2-g* (Adjalley et al.,
2016), likely because of the very low proportion of parasites that activate *pfap2-
g* at each cycle of multiplication in most parasite lines. Therefore, for our gene-
targeted approach we used as starting material RNA form two parasite lines
with unusually high sexual conversion rates: E5 and the E5-PfAP2-G-DD lines.
E5 is a subclone of 3D7-B with a sexual conversion rate ~15%, whereas E5-
PfAP2-G-DD has a destabilization domain appended to PfAP2-G that allows
controlling sexual conversion and shows ~30% sexual conversion in the
presence of the stabilizing ligand Shield 1 (Kafsack et al., 2014). While
experiments using different DNA polymerases for amplification showed some
variability (Fig. 1A-B), as expected in extremely AT-rich non-coding sequences,
we consistently identified two main TSSs blocks. The first block (Block 1) is
located around -1,900 bp from the ATG and the second block (Block 2) around -
1,600 bp (Fig. 1B).

To validate the predicted TSSs identified by 5'RACE, we performed a reverse
transcriptase coupled to quantitative PCR (RT-qPCR) analysis at different times
during the asexual blood cycle with primers located upstream of Block 1,
between the two blocks, downstream of both blocks and within the CDS of
pfap2-g (Fig. 1C). In the E5 line we observed a complex pattern that may reflect
technical limitations attributable to the extreme AT-richness of this region, or
intrinsic transcriptional complexity (e.g. production of non-coding RNAs).
However, it was obvious that essentially no signal was detected upstream of
Block1, indicating that transcripts starting upstream from Block 1 are rare (Fig.
1D). As expected, the signal was stronger in cultures at the early ring stage,
when the gene is more abundantly transcribed (Bancells et al., 2019). A similar
analysis with the E5-PfAP2-G-DD line revealed that in the absence of functional
PfAP2-G (cultures without Shield 1), transcripts containing sequences upstream
of Block 1 were as abundant as transcripts containing the other regions

7

analyzed. In contrast, in cultures treated with Shield 1 to stabilize PfAP2-G, the pattern was similar to the E5 line (Fig. 1D). These results suggest that in parasites in which the PfAP2-G positive feedback loop is active (Kafsack et al., 2014; Poran et al., 2017), *pfap2-g* transcripts predominantly initiate at the TSSs in blocks 1 and 2, whereas basal PfAP2-G-independent transcription likely initiates further upstream.

In line with these results, several PfAP2-G binding motifs to which PfAP2-G has been experimentally shown to bind *in vivo* (Josling et al., 2019), are present upstream of the TSSs identified, in relatively close proximity to Block 1 (Fig. 1E). Moreover, the TSSs identified contain some of the common features of TSSs in *P. falciparum*, such as being within nucleosome depleted regions and around low complexity regions enriched in tandem repeats (Adjalley et al., 2016; Kensche et al., 2016; Toenhake and Bártfai, 2019) (Fig. 1E).

Screening for the *cis*-regulatory elements involved in heterochromatin formation in *pfap2-g*

We hypothesized that some specific regions have the capacity to direct *de novo* deposition of H3K9me3 and the establishment of heterochromatin at the *pfap2-g* locus. To test this hypothesis, we designed a screening approach based on the integration of overlapping fragments of the *pfap2-g* regulatory region and part of the CDS into a euchromatic locus, followed by chromatin immunoprecipitation (ChIP) coupled to qPCR (ChIP-qPCR) with antibodies against H3K9me3 to evaluate the capacity of each fragment to nucleate heterochromatin *de novo* (Fig. 2A-B). To specifically detect heterochromatin formation in the ectopically integrated fragments, without interference from the endogenous locus, we first deleted the entire promoter region and part of the coding sequence of *pfap2-g* in the non-gametocyte producing 3D7 subclone 1.2B. The inability of this parasite line to activate *pfap2-g* and form gametocytes is attributable to a truncation in PfGDV1, an upstream activator of *pfap2-g* necessary to remove HP1 from this locus (Filarsky et al., 2018). Using CRISPR/Cas9 technology with two single guide RNAs (sgRNAs), we readily obtained parasites in which the region targeted was replaced by a non-

functional *gfp* gene. One of the correctly edited subclones, H11, was used for all subsequent transfections (Fig. S1A-C).

We initially analyzed four large overlapping fragments (with sizes from 1,166 bp to 1,318 bp) covering 4 Kb of the *pfap2-g* upstream region and part of the CDS, plus a control fragment from the promoter region of the euchromatic gene *ama1*. All fragments were integrated using CRISPR/Cas9 technology into the PF3D7_1144400 gene (Fig.2A-B), which is an euchromatic locus expressed only in mosquito stages and therefore without a function in asexual blood stages. The integration efficiency was nearly 100%, as almost no wild-type (WT) locus was detected by diagnostic PCR analysis in any of the transgenic lines (Fig. 2B, S1D).

To assess *de novo* nucleation of heterochromatin within the PF3D7_1144400 locus we performed H3K9me3 ChIP-qPCR analysis with primer pairs targeting the regions flanking the integrated fragment (within the 5' and 3' homology regions [HRs] used to integrate the fragment, 5'HR and 3'HR) and primers located within the fragments (Internal) (Fig. 2B). Each experiment also included primers against a heterochromatic *var* gene (ID: PF3D7_1240300, previously PFL1950w), that was used as a ChIP positive control, and primers against the euchromatic gene serine-tRNA ligase (*serrs*) that was used as a negative control and to calculate the fold enrichment in H3K9me3 levels. Of note, H3K9me3 ChIP-qPCR analysis of the parental lines 1.2B and H11 validated PF3D7_1144400 as an euchromatic locus. In the 1.2B line, the presence of H3K9me3 in the endogenous regions corresponding to the fragments of the *pfap2-g* locus analyzed were also validated, with H3K9me3 levels similar or even higher than in the control heterochromatic genes. In contrast, in the H11 line these regions were found to be absent, as expected (Fig. S1E).

Of the four fragments, only F1, and to a lesser extent F2, promoted heterochromatin formation, but H3K9me3 levels were lower than in the control *var* gene. As expected, the *ama1* fragment did not promote heterochromatin formation (Fig. 2C). The F1 fragment covers the beginning of the CDS and a small part of the upstream region. To further dissect the key regions within F1 and F2 that promote heterochromatin formation, we generated five new

9

transgenic lines with new fragments integrated at the PF3D7_1144400 locus. Four of them covered different parts of F1 (F1.1, F1.2, F1.3 and Rest F1), and one overlapped with the end of F1 and covered the region immediately downstream of F1 in the CDS (F0) (Fig. 2A-B). From the new fragments, we could only detect heterochromatin formation unambiguously in F0, and a weak H3K9me3 signal at near background levels in F1.3 and F1.2 (Fig. 2C). This result supports the idea that heterochromatin nucleation largely occurs in the coding region of the gene, and indicates that long fragments are needed for efficient nucleation. Next, to assess the spreading capacity of heterochromatin in this ectopic location, we designed primers flanking the HRs (5'ext. and 3'ext.). Indeed, H3K9me3 levels further away than the HRs roughly matched the levels observed with internal primers or proximal primers within the HR (Fig. 2D), suggesting that heterochromatin is able to spread form the nucleation site to the HRs and beyond.

Unexpectedly, qPCR analysis of the ChIP input samples revealed that multiple copies of the fragments were present in all the transgenic lines, except H11_F4 (Fig. S2A-B). This may be explained by either presence of episomal copies of the donor plasmids used for transfection, or by integration of multiple copies of the plasmids. Considering that donor plasmids are linearized before transfection, both possibilities were not expected to occur. However, we and others (Till Voss, personal communication) have recently observed that integration of multiple copies, including the full plasmid rather than only the regions intended for integration, frequently occurs as a consequence of concatemer formation prior to integration. To determine if integration of multiple copies occurred in our transgenic lines carrying the fragments, we performed PCR analysis of genomic DNA with external primers using longer extension time. In all the strains we could detect bands compatible with multiple integrations of the plasmids, except in H11_F4 (Fig. S2C). Moreover, Southern Blot analysis revealed a complex pattern of bands compatible with integration of multiple copies of the plasmids in all transgenic lines except in H11_F4, which had a single band of the expected size for correct single copy integration (Fig. S2D). Thus, while we cannot fully exclude the possibility that donor plasmid

10

episomes are present, the copy number of the fragments can be explained by integration of plasmid concatemers.

If episomal copies of the donor plasmids were present in the transgenic lines, they could be detected by the internal 5'HR and 3'HR primers used in the ChIP-qPCR analysis, potentially interfering with the quantification of H3K9me3 in the integrated fragments. However, the primers flanking the HRs (5'ext. and 3'ext.) only detect the genomic region (Fig. S2B) and yielded similar results, excluding potential artifacts related with the presence of episomes. On the other hand, we cannot completely rule out the possibility that integration of multiple copies of the fragments affects their ability to nucleate heterochromatin, but the observation that only F1 and F0 are clearly positive for H3K9me3 suggests that these regions have a specific capacity to nucleate heterochromatin.

Heterochromatin nucleation occurs only in a subset of parasites

The H3K9me3 levels observed in F1 and F0 were lower than the levels in control heterochromatic genes. Maintaining them in culture for longer times (up to two months) did not increase the H3K9me3 levels (data not shown). To test the possibility that heterochromatin is only formed in a subset of parasites, we obtained subclones of the H11_F1 strain and analyzed them by ChIP-qPCR. In the subclones 5E and 11A, the PF3D7_1144400 locus was fully euchromatic; in 9B, only the 5' region was heterochromatic; whereas in 3E, both the 5' and the 3' regions were clearly heterochromatic (Fig. 3A). Therefore, we confirmed that the levels of H3K9me3 observed in the bulk population reflect population heterogeneity, with some parasites carrying the locus in an euchromatic state and others in a fully heterochromatic state. Of note, all the subclones contained two copies of the integrated region (Fig. 3B).

To map precisely the extension of the heterochromatin formed at the PF3D7_1144400 locus, we assessed the distribution of H3K9me3 and H3K9ac in the 3E subclone by ChIP-Seq. Heterochromatin spanned over the full CDS of the locus, such that the limits of the H3K9me3-enriched domain roughly coincided with the start and end of the gene (Fig. 3C). The alternative modification at H3K9, H3K9ac, was severely depleted throughout the heterochromatic region. This analysis also validated the presence of two copies

11

of F1 in the 3E subclone, as in the input sample the coverage at the F1 fragment and the HRs was double than in the rest of the gene (Fig. 3C).

A deletion-based approach to assess the role of different regions of *pfap2-g* in heterochromatin nucleation

The analysis of the genome wide distribution of H3K9me3 in the 3E sublcone also provided an opportunity to analyze the heterochromatin profile at the *pfap2-g* locus after the deletion of the promoter region and part of the CDS. Unexpectedly, heterochromatin was completely absent from the locus, indicating that the region deleted is involved not only in nucleating heterochromatin formation but also is necessary for its maintenance (Fig. 3D).

Next we tested if disruption of smaller regions is sufficient to disrupt the full *pfap2-g* heterochromatin domain. We deleted the sequence corresponding to F1 and a region of similar size immediately upstream (F2-3) in the 1.2B line (KO F2-3 and KO F1 lines) and assessed the presence of heterochromatin at the *pfap2-g* locus by ChIP-qPCR (Fig. 3E, S3A-C). Deletion of F2-3 altered the H3K9me3 levels at the *pfap2-g* upstream region, but not in the CDS, whereas deletion of F1 had no effect on H3K9me3 levels at either position (Fig. 3F). These results suggest that the deletion of a small portion of the locus is not sufficient to destabilize the whole heterochromatin domain. Since the surrounding regions at both sides are in a heterochromatic state, spreading can possibly maintain the heterochromatin domain. Furthermore, the results with the KO F2-3 line suggest that heterochromatin at the upstream region is less stable, consistent with its inability to mediate *de novo* nucleation. In any case, the impossibility to disrupt the full heterochromatin domain by deleting small fragments indicates that a deletion-based screening to identify specific sequences necessary for the maintenance of the *pfap2-g* heterochromatin domain is not feasible.

DISCUSSION

Heterochromatin-based silencing of specific malaria genes is essential for many fundamental processes in parasite biology, including the regulation of the balance between asexual growth and transmission. Here we provide the first insight into how heterochromatin is formed at some specific positions within the

12

genome. By integrating specific regions of the *pfap2-g* locus in an ectopic location, we demonstrate that specific sequences can nucleate the formation of heterochromatin, indicating that in malaria parasites heterochromatin nucleation is a sequence-dependent process.

Unlike other clonally variant genes, *pfap2-g* is always heterochromatic in asexually-growing parasites. In the gametocyte-deficient parasite line used in this study (1.2B), all the parasites in the population are asexual. We reasoned that using a gene that is typically found in a heterochromatic state may facilitate the identification of the elements involved in heterochromatin nucleation. However, our screening of the *cis*-regulatory elements involved in heterochromatin nucleation suggests that *de novo* heterochromatin formation in *P. falciparum* is an infrequent event. Even for nucleation-competent sequences, heterochromatin is only established in a subset of the parasites, as evidenced by the relatively low levels of H3K9me3 in the bulk transgenic populations and heterogeneity between subclones. This argues against a deterministic scenario in which the presence of specific sequences always results in formation of heterochromatin, and suggests that in asexual blood stages presence of heterochromatin at the *pfap2-g* locus mainly depends on epigenetic inheritance from previous generations, rather than on *de novo* formation. The same may be true for other clonally variant genes. One possible scenario is that heterochromatin only nucleates *de novo* during transmission stages (e.g. mosquito stages), requiring factors absent from asexual stages, such that there is a reset of the epigenetic memory during transmission stages and heterochromatin is established *de novo* before starting a new blood infection. A reset of the expression pattern of clonally variant genes during transmission stages has been demonstrated for *P. falciparum var* and *clag3* genes, and also for *P. chabaudi* variant genes (Bachmann et al., 2016; Mira-Martínez et al., 2017; Spence et al., 2013, 2015), although in neither case it was determined whether heterochromatin was completely dismantled. Alternatively, heterochromatin may be maintained throughout transmission stages, dispensing the need of *de novo* heterochromatin formation. The heterochromatin patterns at *pfap2-g* and other clonally variant genes during gametocyte and mosquito stages, showing variable extension of

13

heterochromatin domains but not complete dismantling (Fraschka et al., 2018; Gómez-Díaz et al., 2017), supports this latter model. In this model, heterochromatin nucleation at *pfap2-g* may be an ancient event and heterochromatin would be maintained by the well characterized positive feed-back loops throughout the full life cycle.

A practical implication of the low levels of H3K9me3 detected in the nucleation-competent integrated fragments in bulk transgenic populations is that serial deletion screening to precisely delimitate the position of sequences able to nucleate heterochromatin is not feasible. This was especially evident when attempting to narrow down the key regions using smaller fragments, as H3K9me3 promoted by such fragments occurred at near background levels. In spite of this limitation, we were able to determine that some regions within the *pfap2-g* CDS are able to nucleate heterochromatin at the PF3D7_1144400 locus. The ability of a DNA sequence integrated in an ectopic location to recapitulate the epigenetic makeup of the endogenous sequence provides evidence for the sequence containing the information necessary for the acquisition of the histone modifications: thus, our results suggest that heterochromatin nucleation in malaria parasites is a sequence-dependent mechanism. Further studies should use complementary approaches to identify the specific elements present in the CDS that are able to trigger heterochromatin formation. One possible approach would be to use electrophoretic mobility shift assays with DNA sequences that cover the fragments with nucleation capacity identified here. This may reveal both the protein complexes involved in the deposition of H3K9me3 and the underlying DNA motifs.

The majority of the transgenic lines contained multiple copies of the fragment integrated, including the lines with nucleation-competent fragments. This is likely a consequence of concatemer formation before integration. We cannot rule out the possibility that multi-copy integration influences the nucleation of heterochromatin, as in *Drosophila* multiple copies of a transgene are sufficient to trigger heterochromatin-based silencing (Dorer and Henikoff, 1994). In fact, H3K9me3-based heterochromatin is specially common in repetitive elements such as telomeres, pericentromeric regions or retrotransposons, where

14

heterochromatin is necessary to maintain genome integrity (Janssen et al., 2018). However, the transgenic lines carrying the control AMA1 fragment and the *pfap2-g* upstream region fragment F3 had an average of ~3 copies of the plasmid integrated, and were completely unable to trigger heterochromatin formation. Nevertheless, efforts are underway to obtain transgenic lines with correct single-copy integration. For this aim, one possible approach consists on including the yCFU negative selectable marker in the donor plasmid, allowing the elimination of parasites carrying plasmid integrations; another option is the use of two sgRNAs, which was recently shown to prevent unintended plasmid integrations in *Plasmodium yoelii* (Walker and Lindner, 2019).

The identification of some regions of the *pfap2-g* CDS with nucleation capacity was unexpected, as the regulatory information is usually encoded in the promoter region and, in the case of *var* and *clag3* genes, promoter regions controlling the expression of a reporter gene were sufficient to recapitulate the endogenous epigenetic state (Rovira-Graells et al., 2015; Voss et al., 2006). Furthermore, the active or silenced state of clonally variant genes typically correlates with the euchromatic or heterochromatic state at their promoters (Chookajorn et al., 2007; Crowley et al., 2011; Jiang et al., 2010; Lopez-Rubio et al., 2007). Our results suggest a potential involvement of the CDS in heterochromatin nucleation at clonally variant genes, suggesting that expansion into promoter regions rather than nucleation controls the transcriptional state of clonally variant genes. In contrast to nucleation, maintenance of heterochromatin may involve sequences in both the promoter and the CDS, or a certain heterochromatin extension. Our analysis of heterochromatin in parasite lines with different deletions in the *pfap2-g* locus revealed that deletion of a large fragment of the upstream region and the beginning of the CDS led to complete loss of heterochromatin at the locus, but deletion of smaller fragments did not disrupt the full heterochromatin domain. This results support the idea that the overall size of the heterochromatin domain is important for its maintenance.

Heterochromatin profiling of gametocytes revealed that the active state of *pfap2-g* is characterized by euchromatin at the promoter and heterochromatin at the CDS (Fraschka et al., 2018). Similar results were obtained using a

15

gametocyte-inducible line (Llorà-Batlle et al., *under review*). We propose that the CDS may act as an heterochromatin reservoir from where it can spread to silence the promoter before parasites enter again the asexual intraerythrocytic cycle. Data from transmission stages of the parasite suggest a gradual recovery of heterochromatin at the promoter region of *pfap2-g*, with almost a complete recovery at sporozoites (Gómez-Díaz et al., 2017). Heterochromatin spreading is well established in model organisms such as *Dropsohila* (Eissenberg et al., 1990), and has also been demonstrated in *P. falciparum* (Rovira-Graells et al., 2015; Voss et al., 2006). Spreading is limited by barrier insulators (Phillips-Cremins and Corces, 2013), which have been predicted to occur in *P. falciparum* (Crowley et al., 2011) but have not been characterized yet.

Altogether, the results presented here provide new insight into the heterochromatin-based regulation of *pfap2-g*. Evidence points at the CDS of *pfap2-g* as the essential element for the nucleation of heterochromatin and for its maintenance during stages at which the gene is active. Further studies should determine whether the key role of the CDS in heterochromatin nucleation and maintenance is specific for *pfap2-g* or common to all clonally variant genes, and identify the specific factors and DNA elements involved.

ACKNOWLEDGMENTS

16

AUTHOR CONTRIBUTIONS

O.L.-B. performed the experiments. L-M.-T. and O.L.-B. performed the bioinformatics analysis. O.L.-B. and A.C. conceived the project, designed the experiments, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

METHODS

Parasite cultures

The 3D7 subclones E5 (derived from the 3D7-B stock), E5-PfAP2-G-DD and 1.2B (derived from the 3D7-A stock) have been previously described and characterized (Cortés, 2005; Kafsack et al., 2014). Parasites were cultured in B+ erythrocytes at 3% hematocrit under standard conditions, with RPMI-1640-based culture medium supplemented with 0.5% Albumax II (Invitrogen). Cultures were regularly synchronized by sorbitol lysis to eliminate late asexual stages (trophozoites and schizonts). For some experiments, cultures were tightly synchronized to a 5 h age window by purification of schizont stages using Percoll (Sigma-Aldrich) gradients (63% Percoll) followed by sorbitol lysis 5 h later. For the stabilization of PfAP2-G in the E5-PfAP2-G-DD line, 0.5 µM AquaShield-1 (Shld, Cheminpharma) was added to the cultures.

5'RACE

Transcription start sites were identified using the FirstChoice 5′ RLM-RACE kit (Thermo Fisher) according to the manufacturer's instructions. In brief, a mixture of total RNAs from E5 and E5-PfAP2-G-DD (in the presence of Shld) cultures at different stages was treated with calf intestinal alkaline phosphatase to remove free 5'-phosphate groups from non-mRNA molecules. Samples were then treated with the tobacco acid pyrophosphatase to remove the 5'-cap structure of full-length mRNAs, allowing the ligation of an RNA adapter with T4 RNA ligase. We used a redesigned RNA adapter because the one provided with the kit resulted in artefacts that incorporated exogenous sequences of another chromosome into the *pfap2-g* PCR-amplified product (Table S1). cDNA synthesis was performed using random hexamers followed by nested PCR using *pfap2-g* specific primers and primers specific for the RACE adapter (Table

17

S1). Nested PCRs were performed with the *LA* Taq DNA Polymerase (Takara), with 25 cycles of amplification for each PCR reaction and an annealing temperature of 60ºC, or with the KAPA HiFi PCR Kit (KAPA Biosystems, KK2103), with 20 cycles of amplification in the first PCR, 30 cycles in the second PCR and an annealing temperature of 62ºC. PCR products were cloned into the pCR-2.1 plasmid using the TA Cloning Kit (Invitrogen) and sequenced with M13 forward and reverse primers.

RNA extraction and transcriptional analysis by RT-qPCR

RNA was extracted using the TRIzol (Invitrogen) method, DNAse-treated (QIAGEN) and reverse-transcribed using the AMV Reverse Transcription Kit (Promega) with a mixture of random primers and oligo (dT). Transcript abundance was measured by real-time quantitative PCR (qPCR) using the standard curve method in a 7900HT Fast Real-Time PCR System and the Power SYBR Green Master Mix (Applied Biosystems). For the standard curve we used genomic DNA from the same parasite line being analysed unless otherwise stated. Transcript levels of the serine-tRNA ligase (*serrs*, ID: PF3D7_0717700) were used for normalization. All primers used for qPCR analysis are described in Table S1.

Plasmids

The pL7-KO-ap2g plasmid used to delete the *pfap2-g* upstream region and part of the coding sequence was derived from the pL6-egfp-yfcu plasmid (Ghorbal et al., 2014). The original HR2 was removed by digesting with *EcoR*I and *Nco*I, generating blunt ends by treatment with T4 DNA polymerase (NEB) and religating with T4 DNA ligase (Roche). The *pfap2-g* HR1 (positions -3929 bp to -3521 bp) was amplified with primers p6 and p7 and cloned into *Not*I/*Pst*I sites. The *pfap2-g* HR2 (positions +991 bp to +1973 bp) was amplified with primers p8 and p9 and cloned into *Sac*II/*Afl*II sites. An incomplete *gfp* sequence, lacking the start codon, was PCR amplified with primers p10 and p11 and cloned between HR1 and HR2 using *Pst*I/*Spe*I sites. Last, to clone the guide, the plasmid was digested with *Btg*ZI and two annealed oligonucleotides (p12 and p13) containing the guide sequence (located at -3503 bp to -3522 bp relative to the *pfap2-g* start codon) were cloned using the In-Fusion HD Cloning Kit (Clontech). To generate plasmid pDC2-Cas9-hDHFR-ap2g-3' expressing the

18

second guide in addition to Cas9 and the *hdhfr* selectable marker, a guide covering positions +921 bp to +940 bp was prepared by annealing oligonucleotides p14 and p15 and cloned into a *Bbs*I site of plasmid pDC2-Cas9-hDHFR (Lim et al., 2016) using the In-Fusion system.

For the integration of the fragments into the PF3D7_1144400 locus, the pL1144400-Fx plasmids ("x" indicates each different fragment) were generated from the pL6-egfp-yfcu plasmid. The original HR2 was removed as previously described. The PF3D7_1144400 HR1 (positions +1144 bp to +1562 bp relative to the start codon) was amplified with primers p16 and p17 incorporating *Nco*I/*Spe*I sites in the reverse primer. The PF3D7_1144400 HR2 (positions +1646 bp to +2199 bp) was amplified with primers p18, incorporating a *Spe*I site, and p19, incorporating an *Afl*II site, and cloned together with HR1 into *Sac*II/*Afl*II sites. The *yfcu* cassette was removed using restriction sites *Not*I and *Sac*II, blunt ending and religating as described above. The *hdfhr* cassette and the sgRNA expression cassette were removed using restriction sites *Aat*II and *Afl*II because the guide and selectable marker were expressed from the pDC2-Cas9-HDHFR plasmid. Test fragments were PCR amplified and cloned into *Nco*I/*Spe*I sites: F1 (positions -181 bp to +992 bp relative to the *pfap2-g* start codon) with primers p20 and p21; F2 (positions -1200 bp to +118 bp) with primers p22 and p23; F3 (positions -2147 bp to -891 bp) with primers p24 and p25; F4 (positions -3000 bp to -1834 bp) with primers p26 and p27; F1.3 (positions -181 bp to +425 bp) with primers p20 and p30; F1.2 (positions -6 bp to +425 bp) with primers p31 and p30; F1.1 (positions +210 bp to +425 bp) with primers p32 and p30; Rest F1 (positions +341 bp to +992 bp) with primers p33 and p21; F0 (positions +748 bp to +1951 bp) with primers p34 and p35; and AMA1 (positions -966 bp to +249 bp relative to the *ama1* start codon) with primers p28 and p29. To generate the plasmid pDC2-Cas9-HDHFRyFCU-1144400 a guide covering positions +1566 bp to +1585 bp from the PF3D7_1144400 start codon was prepared by annealing oligonucleotides p36 and p37 and cloned into a *Bbs*I site of plasmid pDC2-Cas9-hDHFRyFCU (Knuepfer et al., 2017) using the In-Fusion system.

For the deletion of F1, the pUC19-KO-F1 plasmid was generated by amplifying HR1 (positions -448 bp to -159 bp relative to the *pfap2-g* start codon) with

19

primers p38 and p39 and HR2 (positions +993 bp to +1974 bp) with primers p40 and p41, and cloning the two fragments and the incomplete *gfp* sequence into a *Bam*HI site of the pUC19 vector, using the In-Fusion system. To generate the plasmid pDC2-Cas9-HDHFR-5'F1 a guide covering positions -139 bp to -158 bp was prepared by annealing oligonucleotides p42 and p43 and cloned into a *Bbs*I site of plasmid pDC2-Cas9-hDHFR. For the KO of region F2-3, the pUC19-KO-F2-3 plasmid was generated by amplifying HR1 (positions -1857 bp to -1427 bp) with primers p44 and p45 and HR2 (positions -129 bp to +234 bp) with primers p46 and p47 and cloning the two fragments and the incomplete *gfp* sequence into a *Bam*HI site of the pUC19 vector, using the In-Fusion system. To generate the plasmid pDC2-Cas9-HDHFR-ap2g-1.4 a guide covering positions -1412 bp to -1393 bp was prepared by annealing oligonucleotides p48 and p49 and cloned into a *Bbs*I site of plasmid pDC2-Cas9-hDHFR.

Guides were designed using the EuPaDGT web-based tool (Tarleton and Peng, 2015). PCR amplifications from *P. falciparum* genomic DNA were performed using *LA* Taq DNA Polymerase (Takara). To clone the plasmids, we used *Escherichia coli* DH5α or Stellar Competent Cells (Clontech) for difficult cloning. Oligonucleotides were from Integrated DNA Technologies (IDT). All primers and oligonucleotides are described in Table S1.

Generation of transgenic lines

All transfections were performed by electroporation of cultures at the ring stage according to standard procedures (Crabb et al., 2004). For the deletion of the *pfap2-g* regulatory region and part of the CDS, 1.2B parasites were transfected with 60 µg of pDC2-Cas9-HDHFR-ap2g-3' and 12 µg of pL7-KO-ap2g linearized using a *Sca*I site located in the backbone of the plasmid. Once parasites were recovered, subclones were obtained by limiting dilution of the transfected population. For the integration of the fragments within the PF3D7_1144400 locus, cultures of the H11 subclone were transfected with 60 µg of pDC2-Cas9-U6-HDHFR-1144400 and 12 µg of pL1144400-Fx (containing the different fragments) linearized using a *Sca*I site. For the KO of F1, we transfected 1.2B cultures with 60 µg of pDC2-Cas9-HDHFR-5'F1, 60 µg of pDC2-Cas9-HDHFR-ap2g-3' and 12 µg of pUC19-KO-F1 linearized using an *Xmn*I site located in the backbone of the plasmid. For the KO of F2-3, we transfected 1.2B cultures with

20

60 µg of pDC2-Cas9-HDHFR-ap2g-1.4, 60 µg of pDC2-Cas9-HDHFR-5'F1 and 12 µg of pUC19-KO-F2-3 linearized using an *Xmn*I site located in the backbone of the plasmid.

Starting the day after transfection, parasites were selected for 4 days with 10 nM of WR99210 (Jacobus Pharmaceuticals). Genomic DNA from the transfected parasites for diagnostic PCR or Southern blot analysis was prepared by standard methods. Diagnostic PCR to assess the correct integration of the transgenes was performed using the *LA* Taq DNA Polymerase (Takara). Diagnostic qPCRs were performed to assess the presence of episomal pDC2-Cas9-HDHFR plasmids: copy number was determined relative to a single copy gene (*serrs*) , using genomic DNA of the W4-2 line (containing a single copy of the *hdhfr* gene) (Cortés et al., 2007) for the standard curve. The quantification by RT-qPCR of the HRs or inserted fragments in the ChIP input samples was used to determine the relative copy number of the transgene within the transgenic lines. Copy number was estimated relative to single copy regions (*serrs* and PF3D7_1240300). The primers used for diagnostic PCR and qPCR are described in Table S1.

Southern Blot

Southern blot was performed according to standard procedures. Genomic DNA was digested with the restriction enzymes *Aat*II and *Hinc*II, plus *Afl*II to cleave episomal plasmids. Digestions were resolved in 1.25% agarose gels and transferred to Amersham Hybond N+ nylon membranes (GE Healthcare) to be detected with specific probes labelled with α-^{32}P dATP (Perkin Elmer). The primers used to amplify the probe were the same as the ones used for the amplification of the HR1 of PF3D7_1144400 (Table S1). Hybridization and washes were performed at 62°C and the membrane was exposed to an X-ray film.

ChIP-qPCR and ChIP-seq

Chromatin extraction from cultures at the late trophozoite/schizont stage was performed as previously described using the MAGnifyTM Chromatin Immunoprecipitation System (Life Technologies) after the crosslinking and washing steps (Rovira-Graells et al., 2015), with minor modifications.

21

For ChIP-qPCR, samples were sonicated using a Bioruptor® Plus sonication device (Diagenode) at high power, with 6 cycles of 30 seconds ON/30 seconds OFF followed by spinning and recovery of the aqueous phase, which was sonicated for 3 additional cycles of 30 seconds ON/30 seconds OFF to obtain sheared chromatin with a size around 400-500 bp. Immunoprecipitations were performed overnight at 4°C with 0.5 μg of chromatin and 2 μg of antibodies against H3K9me3 (Diagenode pAb-193–050), previously coupled to protein A/G magnetic beads provided in the kit. Washing, de-crosslinking and elution were performed following the MAGnifyTM Chromatin Immunoprecipitation System recommendations. Eluted ChIP DNA was dilutes 1:2 for further downstream analysis by qPCR.

For ChIP-Seq, samples were sonicated using a M220 sonicator (Covaris) at 10% duty factor, 200 cycles per burst, 140 W of peak incident power for 10 min. Immunoprecipitations were performed overnight at 4°C with 4 μg of chromatin and 8 μg of antibodies against H3K9me3 (Diagenode pAb-193–050) or H3K9ac (Diagenode pAb-004-050), previously coupled to protein A/G magnetic beads provided in the kit. Washing, de-crosslinking and elution were performed following the MAGnifyTM Chromatin Immunoprecipitation System recommendations, but avoiding high temperatures that may result in denaturation of extremely AT-rich intergenic regions: de-crosslinking, proteinase K treatment and elution were performed at 45°C (for 2 h, over-night, and 1.5 h, respectively).

Libraries for Illumina sequencing were prepared from 5 ng of immunoprecipitated DNA using a protocol adapted for genomes with an extremely high AT-richness (Kensche et al., 2016). In brief, after end-repair and addition of 3′ A-overhangs, NEBNext Multiplex Oligos for Illumina (NEB, E7500) were ligated. Purification steps were performed with Agencourt AMPure XP beads (Beckman Coulter, A63880). Libraries were amplified using the KAPA HiFi PCR Kit (KAPA Biosystems, KK2103) in KAPA HiFi Fidelity Buffer (5X) with the following conditions: 95°C for 3 min; 9 cycles at 98°C for 20 s and 62°C for 2 min 30 s; and 62 °C for 5 min. Amplified libraries were purified using 0.9X AMPure XP beads to remove adapter dimers. The library size was analyzed in

22

a 4200 TapeStation System (Agilent Technologies). We obtained 12-16 million 125 bp paired-end reads per sample using a HiSeq2500 System (Illumina).

Data analysis

For the analysis of ChIP-seq data, we first removed repetitive kmers from the reads using BBDuK (v36.99) with parameters ktrim=r, k=22, and mink=6. After this initial filtering, reads where aligned to modified 3D7 reference genomes reflecting the expected alterations introduced by CRISPR/Cas9 genome editing (i.e. KO of the *pfap2-g* regulatory region and part of the CDS, integration of F1 in PF3D7_1144400). Reads alignment was performed using Bowtie (v1.2) (Langmead et al., 2009) with parameters *--very-sensitive* and *--local,* trimming 4 bp from both the 5' and 3' ends, and restricting fragment size to 50-200 bp. After aligning, duplicate reads where removed using PicardTools MarkDuplicates (v2.9.4). For each sample, per-base coverage was calculated using BEDtools (v2.27.1) (Quinlan and Hall, 2010) and divided by the number of million reads (RPM). For each base, H3K9me3 enrichment was calculated as ChIP (RPM)/input (RPM) with a pseudocount of 0.1 to avoid division by 0. Data was visualized using IGV (v2.4.10) (Thorvaldsdottir et al., 2013).

Data and software availability

The accession number for the ChIP-seq data reported in this paper is GEO: XXX (deposition in progress).

REFERENCES

Adjalley, S.H., Chabbert, C.D., Klaus, B., Pelechano, V., and Steinmetz, L.M. (2016). Landscape and Dynamics of Transcription Initiation in the Malaria Parasite Plasmodium falciparum. Cell Rep. *14*, 2463–2475.

Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. Nat. Rev. Mol. Cell Biol. *19*, 229–244.

Andersson, R., Sandelin, A., and Danko, C.G. (2015). A unified architecture of transcriptional regulatory elements. Trends Genet. *31*, 426–433.

Bachmann, A., Petter, M., Krumkamp, R., Esen, M., Held, J., Scholz, J.A.M., Li, T., Sim, B.K.L., Hoffman, S.L., Kremsner, P.G., et al. (2016). Mosquito Passage Dramatically Changes var Gene Expression in Controlled Human Plasmodium falciparum Infections. PLOS Pathog. *12*, e1005538.

Bancells, C., Llorà-Batlle, O., Poran, A., Nötzel, C., Rovira-Graells, N., Elemento, O., Kafsack, B.F.C., and Cortés, A. (2019). Revisiting the initial steps

23

of sexual development in the malaria parasite Plasmodium falciparum. Nat. Microbiol. *4*, 144–154.

Becker, J.S., Nicetto, D., and Zaret, K.S. (2016). H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. Trends Genet. *32*, 29–41.

Brancucci, N.M.B., Bertschi, N.L., Zhu, L., Niederwieser, I., Chin, W.H., Wampfler, R., Freymond, C., Rottmann, M., Felger, I., Bozdech, Z., et al. (2014). Heterochromatin Protein 1 Secures Survival and Transmission of Malaria Parasites. Cell Host Microbe *16*, 165–176.

Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., de la Rosa-Velazquez, I.A., van de Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M., et al. (2012). A transcription factor–based mechanism for mouse heterochromatin formation. Nat. Struct. Mol. Biol. *19*, 1023–1030.

Campbell, T.L., De Silva, E.K., Olszewski, K.L., Elemento, O., and Llinás, M. (2010). Identification and Genome-Wide Prediction of DNA Binding Specificities for the ApiAP2 Family of Regulators from the Malaria Parasite. PLoS Pathog. *6*, e1001165.

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. *104*, 899–902.

Coetzee, N., Sidoli, S., van Biljon, R., Painter, H., Llinás, M., Garcia, B.A., and Birkholtz, L.-M. (2017). Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual Plasmodium falciparum parasites. Sci. Rep. *7*, 607.

Cortés, A. (2005). A chimeric Plasmodium falciparum Pfnbp2b/Pfnbp2a gene originated during asexual growth. Int. J. Parasitol. *35*, 125–130.

Cortés, A., and Deitsch, K.W. (2017). Malaria epigenetics. Cold Spring Harb. Perspect. Med. *7*, 1–23.

Cortés, A., Carret, C., Kaneko, O., Yim Lim, B.Y.S., Ivens, A., and Holder, A.A. (2007). Epigenetic Silencing of Plasmodium falciparum Genes Linked to Erythrocyte Invasion. PLoS Pathog. *3*, e107.

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 <I>Plasmodium falciparum<I>. In Parasite Genomics Protocols, (New Jersey: Humana Press), pp. 263–276.

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.

Dorer, D.R., and Henikoff, S. (1994). Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. Cell *77*, 993–1002.

Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V.,

24

and Elgin, S.C. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. Proc. Natl. Acad. Sci. *87*, 9923–9927.

Filarsky, M., Fraschka, S.A., Niederwieser, I., Brancucci, N.M.B., Carrington, E., Carrió, E., Moes, S., Jenoe, P., Bártfai, R., and Voss, T.S. (2018). GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. Science (80-.). *359*, 1259–1263.

Flueck, C., Bartfai, R., Volz, J., Niederwieser, I., Salcedo-Amaya, A.M., Alako, B.T.F., 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.

Fraschka, S.A., Filarsky, M., Hoo, R., Niederwieser, I., Yam, X.Y., Brancucci, N.M.B., Mohring, F., Mushunje, A.T., Huang, X., Christensen, P.R., et al. (2018). Comparative Heterochromatin Profiling Reveals Conserved and Unique Epigenome Signatures Linked to Adaptation and Development of Malaria Parasites. Cell Host Microbe 1–14.

Ghorbal, M., Gorman, M., Macpherson, C.R., Martins, R.M., Scherf, A., and Lopez-Rubio, J.-J. (2014). Genome editing in the human malaria parasite Plasmodium falciparum using the CRISPR-Cas9 system. Nat. Biotechnol. *32*, 819–821.

Gómez-Díaz, E., Yerbanga, R.S., Lefèvre, T., Cohuet, A., Rowley, M.J., Ouedraogo, J.B., Corces, V.G., Bozdech, Z., Roch, K.G. Le, Llinás, M., et al. (2017). Epigenetic regulation of Plasmodium falciparum clonally variant gene expression during development in Anopheles gambiae. Sci. Rep. *7*, 40655.

Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. Nat. Rev. Genet. *16*, 71–84.

Imbeault, M., Helleboid, P.-Y., and Trono, D. (2017). KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature *543*, 550–554.

Janssen, A., Colmenares, S.U., and Karpen, G.H. (2018). Heterochromatin: Guardian of the Genome. Annu. Rev. Cell Dev. Biol. *34*, 265–288.

Jeninga, M., Quinn, J., and Petter, M. (2019). ApiAP2 Transcription Factors in Apicomplexan Parasites. Pathogens *8*, 47.

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. *107*, 2224–2229.

Josling, G.A., Venezia, J., Orchard, L., Russell, T.J., Painter, H.J., and Llinas, M. (2019). Regulation of sexual differentiation is linked to invasion in malaria parasites. BioRxiv 533877.

Kafsack, B.F.C., Rovira-Graells, N., Clark, T.G., Bancells, C., Crowley, V.M., Campino, S.G., Williams, A.E., Drought, L.G., Kwiatkowski, D.P., Baker, D., et

25

al. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. Nature *507*, 248–252.

Kensche, P.R., Hoeijmakers, W.A.M., Toenhake, C.G., Bras, M., Chappell, L., Berriman, M., and Bártfai, R. (2016). The nucleosome landscape of Plasmodium falciparum reveals chromatin architecture and dynamics of regulatory sequences. Nucleic Acids Res. *44*, 2110–2124.

Knuepfer, E., Napiorkowska, M., van Ooij, C., and Holder, A.A. (2017). Generating conditional gene knockouts in Plasmodium - a toolkit to produce stable DiCre recombinase-expressing parasite lines using CRISPR/Cas9. Sci. Rep. *7*, 3881.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet. *32*, 42–56.

Lim, M.Y.-X., LaMonte, G., Lee, M.C.S., Reimer, C., Tan, B.H., Corey, V., Tjahjadi, B.F., Chua, A., Nachon, M., Wintjens, R., et al. (2016). UDP-galactose and acetyl-CoA transporters as Plasmodium multidrug resistance genes. Nat. Microbiol. *1*, 16166.

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.

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.

Miao, J., Fan, Q., Cui, L., Li, J., Li, J., and Cui, L. (2006). The malaria parasite Plasmodium falciparum histones: Organization, expression, and acetylation. Gene *369*, 53–65.

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, G.P., et al. (2017). Expression of the Plasmodium falciparum Clonally Variant clag3 Genes in Human Infections. J. Infect. Dis. *215*, 938–945.

Mueller, A.-K., Hammerschmidt-Kamper, C., and Kaiser, A. (2014). RNAi in Plasmodium. Curr. Pharm. Des. *20*, 278–283.

Nishibuchi, G., and Déjardin, J. (2017). The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. Chromosom. Res. *25*, 77–87.

Pérez-Toledo, K., Rojas-Meza, A.P., Mancio-Silva, L., Hernández-Cuevas, N.A., Delgadillo, D.M., Vargas, M., Martínez-Calvillo, S., Scherf, A., and Hernandez-Rivas, R. (2009). Plasmodium falciparum heterochromatin protein 1

26

binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. Nucleic Acids Res. *37*, 2596–2606.

Phillips-Cremins, J.E., and Corces, V.G. (2013). Chromatin Insulators: Linking Genome Organization to Cellular Function. Mol. Cell *50*, 461–474.

Poran, A., Nötzel, C., Aly, O., Mencia-Trinchant, N., Harris, C.T., Guzman, M.L., Hassane, D.C., Elemento, O., and Kafsack, B.F.C. (2017). Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. Nature *551*, 95–99.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rovira-Graells, N., Gupta, A.P.A., 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.

Rovira-Graells, N., Crowley, V.M., Bancells, C., Mira-Martínez, S., Ribas de Pouplana, L., and Cortés, A. (2015). Deciphering the principles that govern mutually exclusive expression of Plasmodium falciparum clag3 genes. Nucleic Acids Res. *43*, 8243–8257.

Saksouk, N., Simboeck, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin *8*, 3.

Satory, D., Gordon, A.J.E., Halliday, J.A., and Herman, C. (2011). Epigenetic switches: Can infidelity govern fate in microbes? Curr. Opin. Microbiol. *14*, 212–217.

Spence, P.J., Jarra, W., Lévy, P., Reid, A.J., Chappell, L., Brugat, T., Sanders, M., Berriman, M., and Langhorne, J. (2013). Vector transmission regulates immune control of Plasmodium virulence. Nature *498*, 228–231.

Spence, P.J., Brugat, T., and Langhorne, J. (2015). Mosquitoes Reset Malaria Parasites. PLOS Pathog. *11*, e1004987.

Tarleton, R., and Peng, D. (2015). EuPaGDT: a web tool tailored to design CRISPR guide RNAs for eukaryotic pathogens. Microb. Genomics *1*, e000033.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Toenhake, C.G., and Bártfai, R. (2019). What functional genomics has taught us about transcriptional regulation in malaria parasites. Brief. Funct. Genomics *00*, 1–12.

Trelle, M.B., Salcedo-Amaya, A.M., Cohen, A.M., Stunnenberg, H.G., and Jensen, O.N. (2009). Global Histone Analysis by Mass Spectrometry Reveals a High Content of Acetylated Lysine Residues in the Malaria Parasite Plasmodium falciparum. J. Proteome Res. *8*, 3439–3450.

Trojer, P., and Reinberg, D. (2007). Facultative Heterochromatin: Is There a Distinctive Molecular Signature? Mol. Cell *28*, 1–13.

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.

Walker, M.P., and Lindner, S.E. (2019). Ribozyme-mediated, multiplex CRISPR gene editing and CRISPR interference (CRISPRi) in rodent-infectious Plasmodium yoelii. J. Biol. Chem. *294*, 9555–9566.

Wiles, E.T., and Selker, E.U. (2017). H3K27 methylation: a promiscuous repressive chromatin mark. Curr. Opin. Genet. Dev. *43*, 31–37.

28
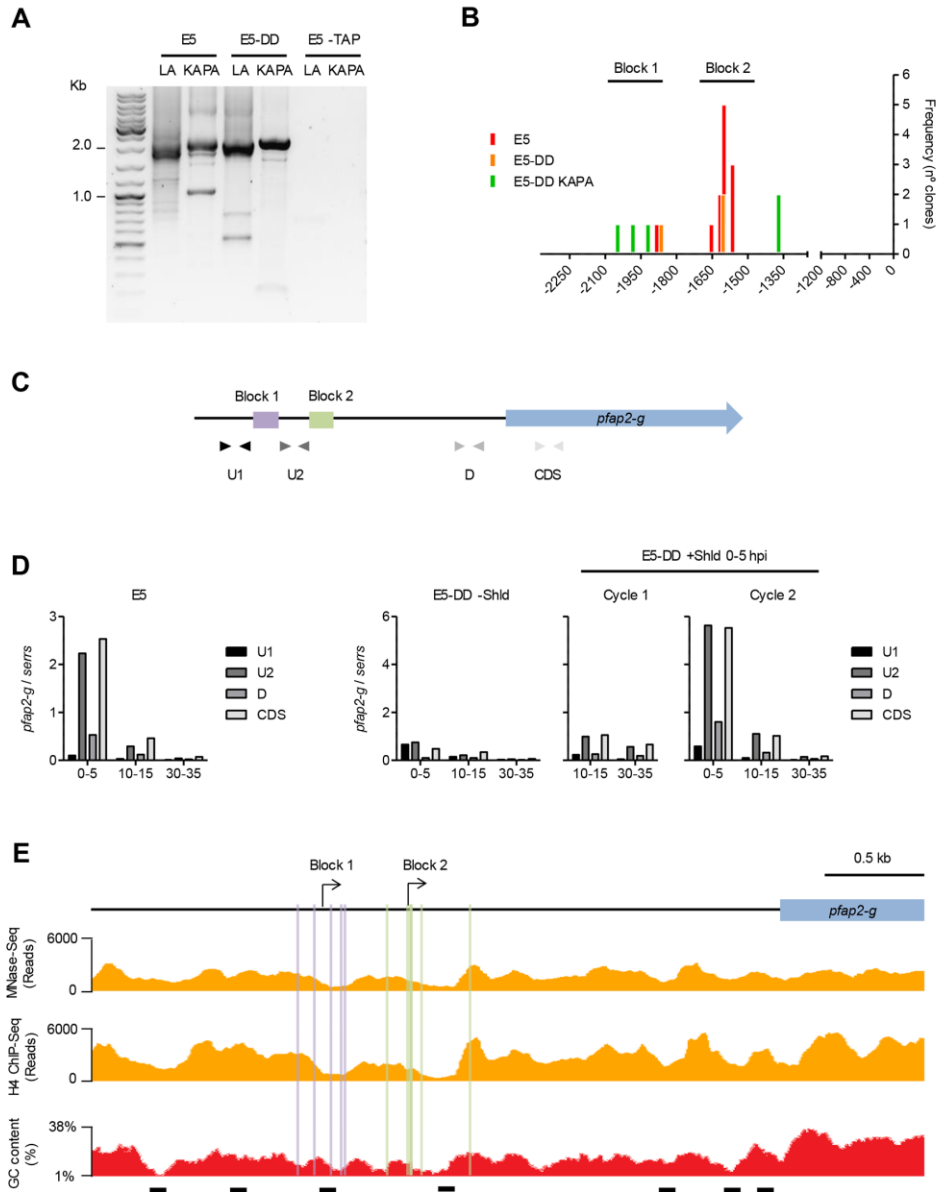
Figure 1. Identification of the *pfap2-g* TSSs.

(A) Electrophoresis analysis of the nested PCR products obtained from the 5'RACE analysis of *pfap2-g* in the E5 and E5-PfAP2-G-DD (E5-DD) lines, and a

29

negative control in which RNA was not treated with the Tobacco Acid Pyrophosphatase (TAP). The E5-PfAP2-G-DD line was cultured in the presence of Shield 1. PCR was performed with *LA* or KAPA DNA polymerases, as indicated.

(B) Position relative to the start codon and frequency (number of clones) of the TSSs identified with the different strains and polymerases. The position of the two main TSS blocks, located around -1,900 (Block 1) and -1,600 bp (Block 2), is indicated.

(C) Schematic of the primers (arrowheads) used for RT-qPCR validation of the TSSs identified, showing their position relative to the TSS blocks.

(D) Validation of the TSSs by RT-qPCR with samples from tightly synchronized E5 and E5-PfAP2-G-DD cultures, the latter maintained either in the absence (-Shld) or presence (+Shld) of Shield 1 and analyzed at the cycle of stabilization (Cycle 1) or in the next cycle (Cycle 2). In cultures with Shield, the compound was added at 0-5 h post-invasion of Cycle 1, resulting in an increase in *pfap2-g* transcripts mainly at the next cycle. Transcript levels are normalized against serine-tRNA ligase (*serrs*).

(E) Schematic (at scale) of the position of individual TSSs identified by 5'RACE (vertical lines) relative to other features. Yellow tracks are profiles of nucleosome occupancy at the *pfap2-g* locus derived from MNase-Seq (upper track) or ChIP-Seq against H4 (bottom track) analysis (Kensche et al., 2016). The red track shows the GC content across the locus. Black and green boxes indicate the position of tandem repeats and the predicted PfAP2-G DNA binding motifs (from PlasmoDB), respectively. PfAP2-G binding at these positions was recently validated by ChIP-seq analysis (Josling et al., 2019).

30
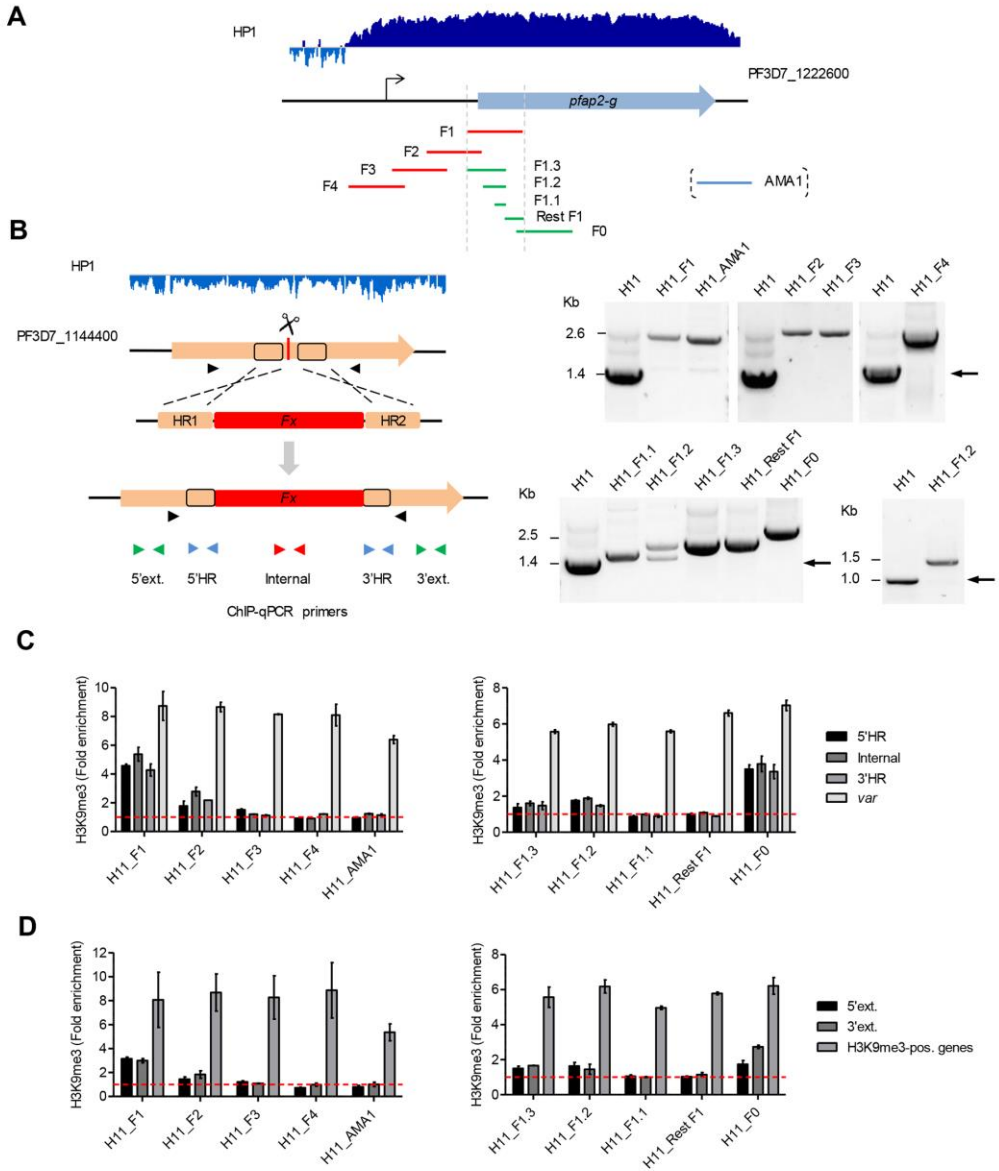
Figure 2. Identification of sequences that can nucleate heterochromatin formation in the *pfap2-g* locus.

(A) Schematic of the position of the fragments tested for heterochromatin nucleation capacity. The fragments initially designed to cover the

heterochromatic domain in the upstream region and part of the CDS are shown in red, whereas the fragments used in a second round of screening are shown in green. The HP1 profile marking the position of the heterochromatin domain (Fraschka et al., 2018) is shown on top. The AMA1 fragment was designed from the promoter region of the *ama1* gene (not shown).

(B) Integration of the fragments into the euchromatic gene PF3D7_1144400 using CRISPR/Cas9 technology. The scissors indicate the position targeted by the guide RNA. Fx refers to the fragment to be integrated; HR refers to homology regions. Black arrowheads show the position of the primers used for diagnostic PCR analysis of genomic DNAs. Coloured arrowheads show the position of the primers used for ChIP-qPCR. The panels at the right show the diagnostic PCR analyses to validate the integration of the constructs. The position of the band expected for wild type parasites is indicated by an arrow. This band was absent from all the transgenic lines with fragments integrated at the PF3D7_1144400 locus. All lines showed a clearly predominant band of the expected size for correct integration of the corresponding fragment, except for the line with the F1.2 fragment that showed a double band (likely a PCR artefact). Analysis with a different primer pair (bottom right gel), also external to the HRs, demonstrated correct integration of this fragment.

(C) H3K9me3 levels in the transgenic lines carrying the integrated fragments, assessed by ChIP-qPCR with the primers located in the HRs flanking the fragment (5'HR and 3'HR) and within the fragment (Internal) (see panel B). Primers against the heterochromatic *var* gene PF3D7_1240300 (previously PFL1950w) were used as a positive control. H3K9me3 fold-enrichment levels for each primer pair are the percentage of the input recovered in the H3K9me3 ChIP (% input) relative to the % input in the euchromatic gene serine-tRNA ligase (*serrs*). The dashed line indicates same level of enrichment as in *serrs*. Data are presented as the average and s.e.m. of two biological replicates.

(D) H3K9me3 levels in the same lines as in panel C but using primers located outside the HRs 5'ext. and 3'ext. (see panel B). Primers against the *var* gene PF3D7_1240300 and *clag3.2* (H3K9me3-pos. genes) were used as positive control of the ChIP (average of the two genes is shown). Data are presented as the average and s.e.m. of two biological replicates.

32
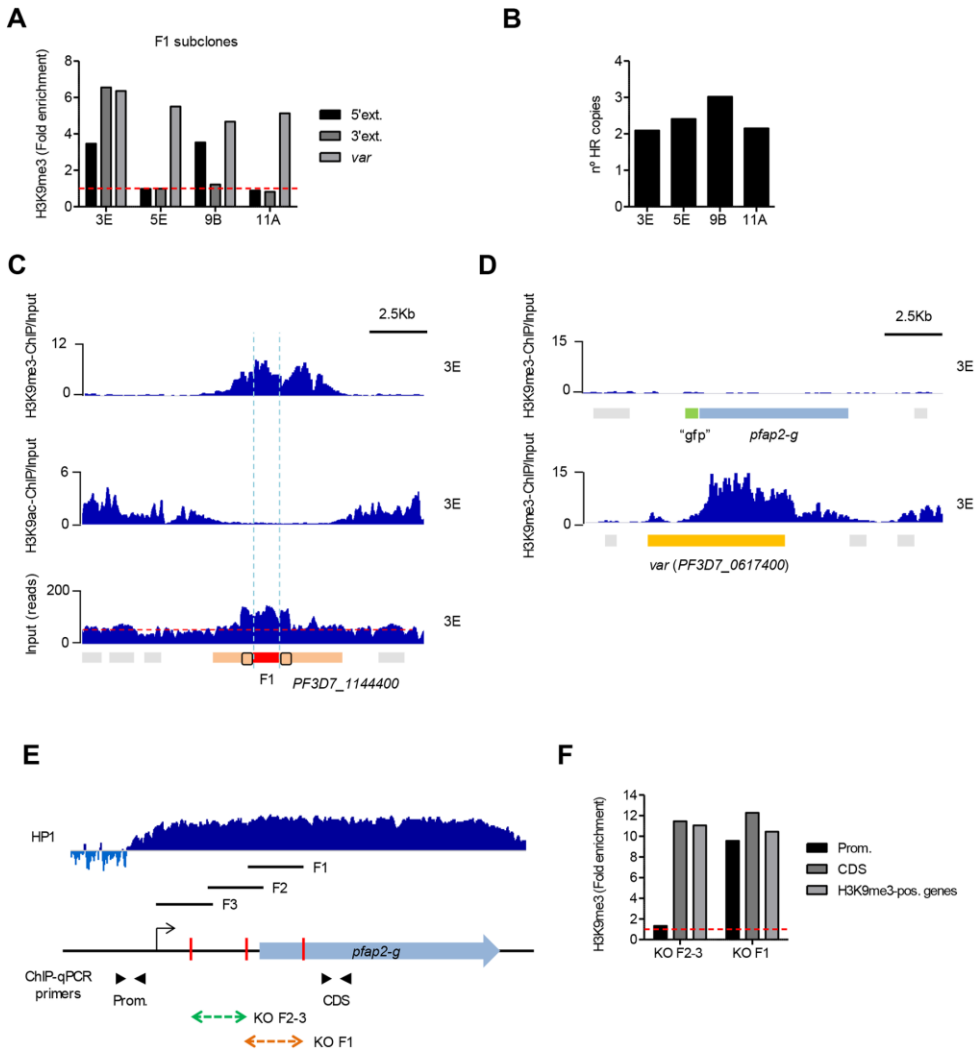
Figure 3. Heterochromatin formation in H11_F1 subclones and lines with deletions in the *pfap2-g* locus.

(A) ChIP-qPCR analysis of H11_F1 subclones with primers 5'ext. and 3'ext. ChIP-qPCR analysis was performed as in Fig. 2C. Data is from a single biological replicate.

(B) qPCR analysis of the number of copies of PF3D7_1144400 HRs in H11_F1 subclones (reflecting the number of copies of F1) using ChIP input samples.

33

1.2B (wild type) genomic DNA was used for the standard curve. Copy number was estimated relative to single copy regions (*serrs* and PF3D7_1240300). Values are the average of values for HR1 and HR2 in a single biological replicate.

(C) ChIP-Seq profiles of H3K9me3 and H3K9ac at the PF3D7_1144400 locus in the 3E sublcone. Values are relative to the input. The bottom track shows the number of reads in the input sample to illustrate the increased coverage at the HRs and F1 fragment. The horizontal dashed line represents the average reads coverage for single-copy sequences. The vertical dashed lines mark the position of F1.

(D) ChIP-Seq profiles of H3K9me3 normalized by input at the *pfap2-g* locus and at a representative heterochromatic *var* gene (PF3D7_0617400) in the subclone 3E. The *var* gene is shown as an H3K9me3 positive control.

(E) Schematic of the deletions in parasite lines KO-F2-3 and KO-F1. The position of the fragments deleted in each line relative to fragments F1 to F3 is shown. The positions targeted by the single guide RNAs (two for each transgenic line) are shown as vertical red bars. Black arrows show the position of the primers used for the ChIP-qPCR analysis.

(F) H3K9me3 levels in the transgenic lines KO F2-3 (subclone 9D) and KO F1 (subclone 3G) assessed by ChIP-qPCR with primers targeting the promoter region (Prom.) and the CDS of *pfap2-g*. ChIP-qPCR analysis was performed as in Fig. 2D. Data is from a single biological replicate.

34

SUPPLEMENTARY INFORMATION

Specific regions of the *Plasmodium falciparum pfap2-g* gene can nucleate heterochromatin formation

Oriol Llorà-Batlle, Lucas Michel-Todó and Alfred Cortés

35
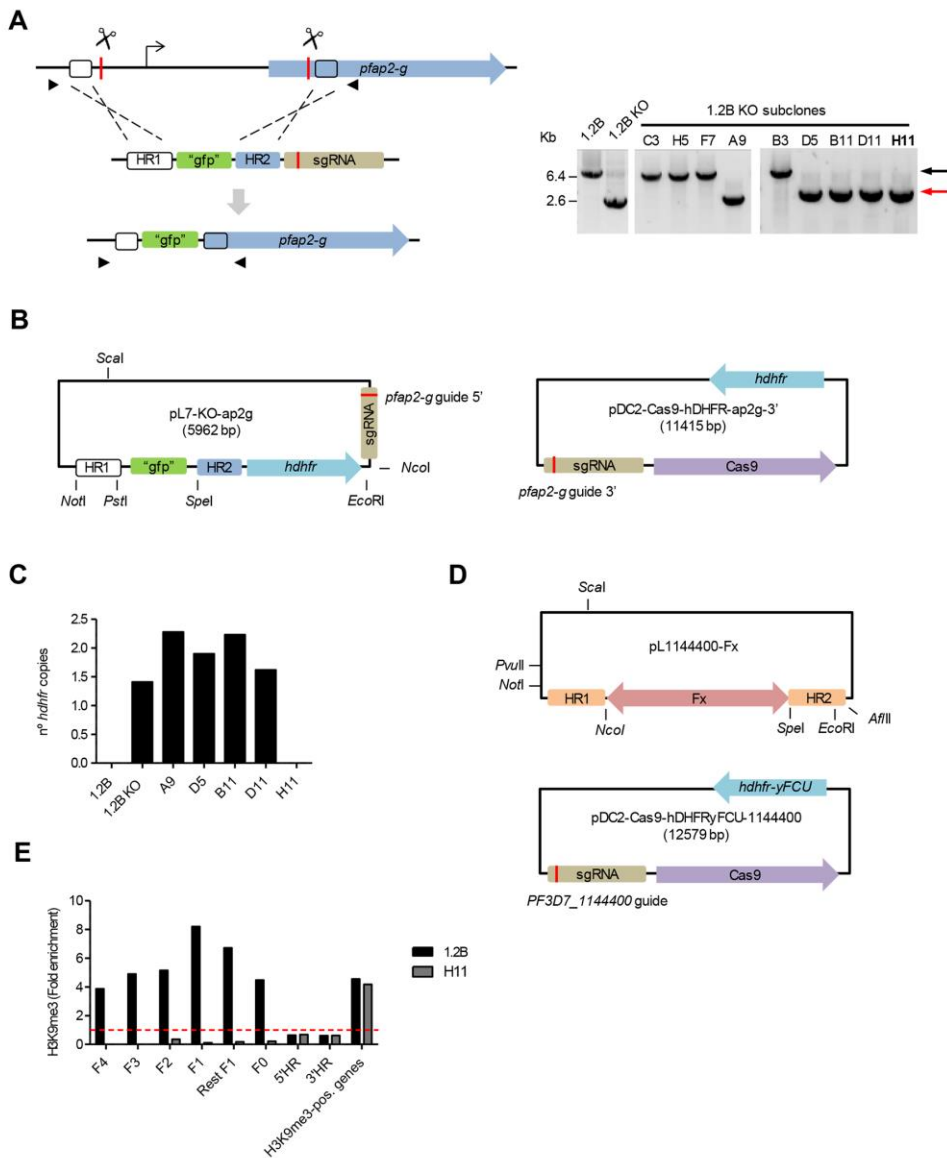
Figure S1. Generation and preliminary characterization of the *pfap2-g* KO line and lines carrying *pfap2-g* integrated ectopically.

(A) Schematic (not to scale) of the strategy used to delete the *pfap2-g* upstream region and part of the CDS using the CRISPR/Cas9 system with two single guide RNAs. The scissors indicate the positions targeted by the guide RNAs.

36

The donor plasmid contained homology regions (HRs) flanking an incomplete *gfp* sequence lacking the start codon ("gfp"). Black arrowheads show the position of the primers used for diagnostic PCR. The gels at the right are the diagnostic PCR analysis of the parental 1.2B line (1.2B WT), the 1.2B *pfap2-g* KO line (1.2B KO) and its subclones. The position of the band expected for WT and correctly edited parasites is indicated by black and red arrows, respectively.

(B) Schematic (not to scale) of the plasmids used to generate the *pfap2-g* KO line. Restriction sites used for the generation of the plasmids are shown.

(C) qPCR analysis of the number of copies of the *hdhfr* selectable marker in the parental 1.2B line, the 1.2B *pfap2-g* KO line before subcloning (1.2B KO) and its subclones. Quantification was performed against a standard curve prepared with genomic DNA of the W4-1 line previously shown by Southern blot to have a single copy of the *hdfhr* gene integrated (Cortés et al., 2007). Only the H11 clone was free of the marker.

(D) Schematic (not to scale) of the plasmids used for the integration of the fragments into the PF3D7_1144400 locus. Restriction sites used for the generation of the plasmids are shown.

(E) H3K9me3 ChIP-qPCR analysis of the parental 1.2B line and the H11 subclone to determine heterochromatin levels at the endogenous sites. The same primers used in Fig.2C were used for this analysis, together with primers against the heterochromatic *var* gene PF3D7_1240300 (previously PFL1950w) and *clag3.2* (H3K9me3-pos genes) that were used as positive controls (the average between the two genes is shown). ChIP-qPCR analysis was performed as in Fig. 2C. Data is from a single biological replicate.

37
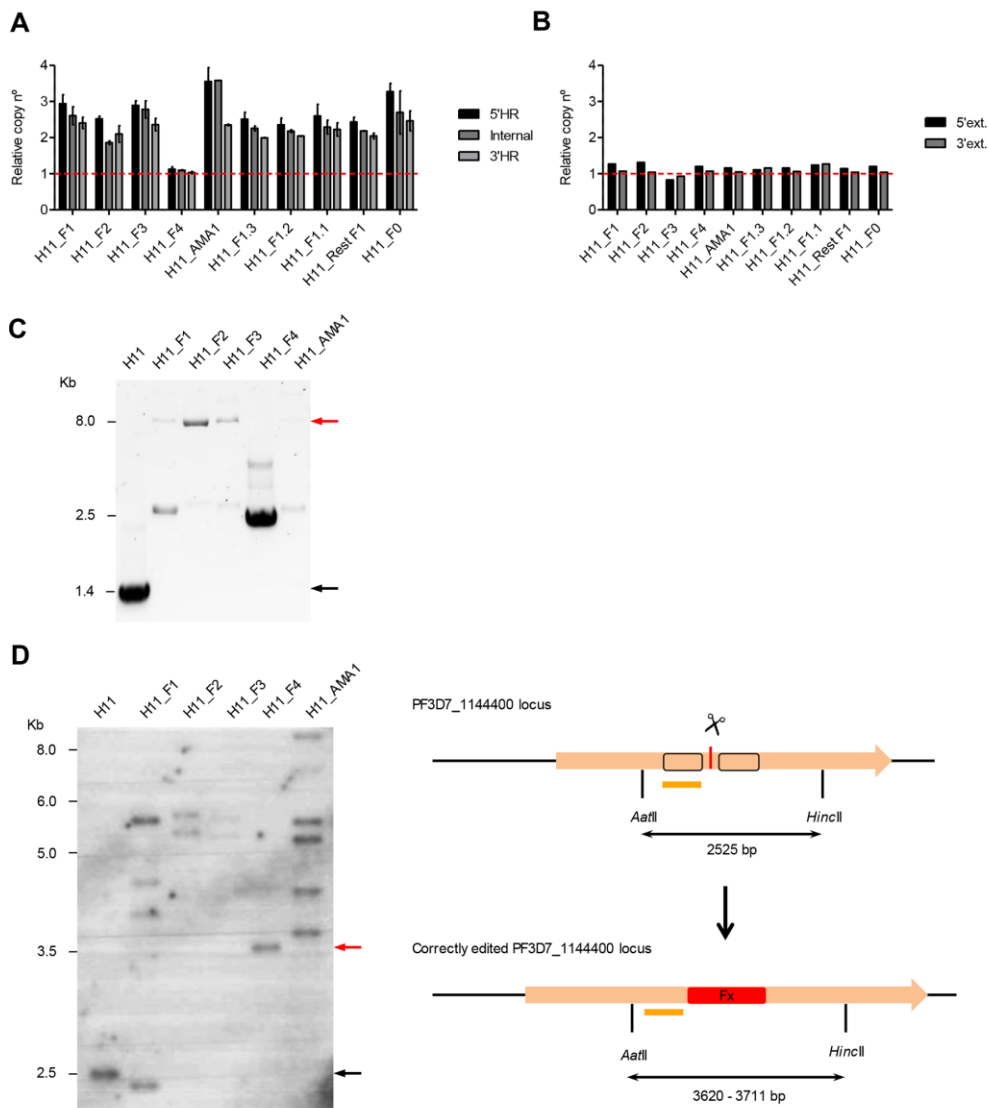
Figure S2. Copy number of the constructs integrated.

(A) qPCR analysis of the number of copies of the integrated fragments (Internal) and PF3D7_1144400 HRs (5'HR and 3'HR) in the transgenic lines, using the ChIP input samples. 1.2B (wild type) genomic DNA was used for the standard curve. Copy number was estimated relative to single copy regions (*serrs* and PF3D7_1240300). In all cases, the estimated copy number was similar for HRs

38

and the fragment itself, suggesting that plasmid concatemers were integrated. Data are presented as the average and s.e.m. of two biological replicates.

(B) Analogous analysis with primer pairs in the regions flanking the HRs (5'ext. and 3'ext.). Data is from a single biological replicate.

(C) Diagnostic PCR with the same primers as in Fig. 2B but using longer extension time to enable the detection of integrated concatemers. The red arrow marks the size expected for integration of two copies of the plasmid, whereas the black arrow indicates the size of the expected band for the wild type locus.

(D) Southern blot analysis of the parental H11 line and the lines with integrated fragments F1-F4 and AMA1. The schematic shows the expected sizes of genomic DNA digested with *Aat*II, *Hinc*II and *Afl*II (the latter only cleaving the plasmid) and hybridized with a PF3D7_1144400-specific probe (orange line) in wild type parasites and parasites with the fragments (Fx) correctly integrated (single copy). The scissors indicate the position targeted by the guide RNA, whereas black boxes indicate the position of the HRs. The red arrow indicates the position of the band expected for the correctly edited locus, and the black arrow the position of the band expected for the wild type locus. While the H11-F4 line showed the expected band for correct single integration, other lines showed a complex pattern consistent with integration of plasmid concatemers with copies of the plasmid in different orientations.

39
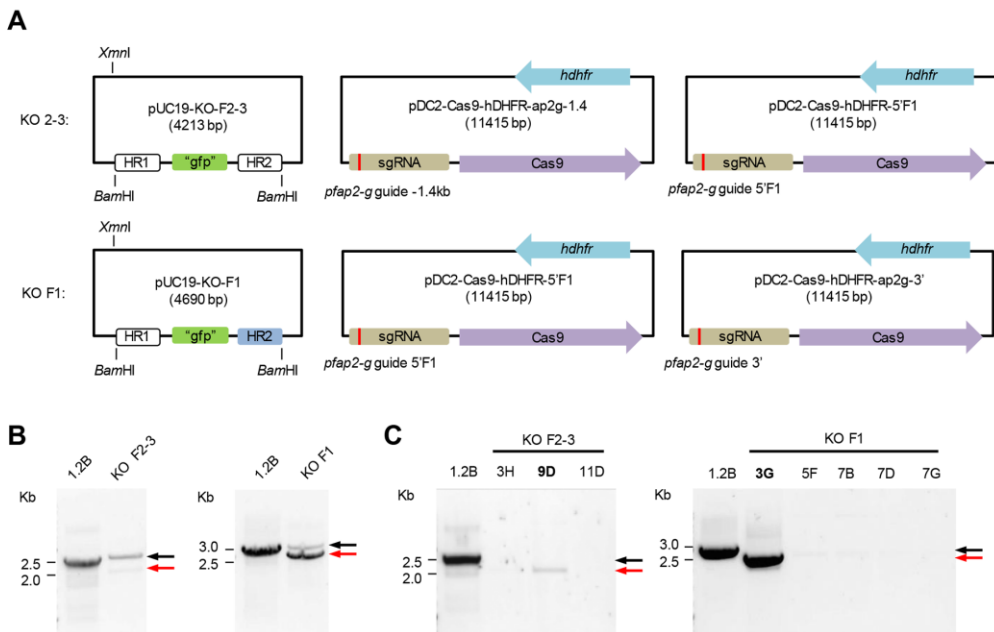
Figure S3. Generation of parasites with deletions in the *pfap2-g* locus.

(A) Schematic (not to scale) of the plasmids used for the deletion of the different regions of the locus. The strategy was similar to that shown in Fig. S1, but cotransfecting three plasmids: two plasmids expressing the two different single guide RNAs (sgRNA), and a donor plasmid. Restriction sites used for the generation of the plasmids are shown. The donor plasmid consisted on an incomplete *gfp* sequence lacking the start codon ("gfp") flanked by the HRs (different for the deletion of each fragment).

(B) Diagnostic PCR to validate the edition of the *pfap2-g* locus in the transfected lines (KO F2-3 and KO F1). The position of the bands expected for WT and correctly edited parasites are indicated by black and red arrows, respectively.

(C) Equivalent diagnostic PCR analysis of KO F2-3 and KO F1 subclones.

40

Table S1. Oligonucleotide sequences. List of the oligonucleotides used in the study classified according to their use: cloning procedures, diagnostic PCR or qPCR. In lowercase: non-annealing part of the oligonucleotide; in bold: restriction sites included in the oligonucleotide; underlined: guide sequences.

Oligonucelotides used for 5'RACE		
Number	Name	Sequence (5'-3')
p1	5'RACE_RNA_adapter	GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGC
p2	5'RACE_Outer	GATGGCGATGAATGAACACTG
p3	Pfap2-g_+345_R	GATACATTCTCGTTACTCTGCA
p4	5'RACE_Inner	GAATGAACACTGCGTTTGCTG
p5	Pfap2-g_+272_R	CATGCTCTCTTCCCATTCGAA
Primers for cloning steps		
Number	Name	Sequence (5'-3')
p6	Pfap2g_-3929_HR1_KO_F_NotI	tggtgt**gcggccgc**CGTACATATATACAAATAGAGAT
p7	Pfap2g_-3521_HR1_KO_R_PstI	tggtgt**ctgcag**ACAGGTATTGTACGCCTTTTA
p8	Pfap2g_+991_HR2_KO_F_SpeI	tgttgg**actagt**CCTTGAAAAGAATATAGAAGAAC
p9	Pfap2g_+1973_HR2_KO_R_AflII	tggtgt**cttaag**CAATTTATCAGCATCGTCATCA
p10	GFP_+3_F_PstI	tggtgt**ctgcag**AGTAAAGGAGAAGAACTTTTCA
p11	GFP_+714_R_SpeI	tggtgt**actagt**TTATTTGTATAGTTCATCCATGC
p12	Pfap2g_-3503_Guide_F	TAAGTATATAATATT<u>TTCTTTTTAAGGTTGCGTAC</u>GTTTTAGAGCTAGAA
p13	Pfap2g_-3522_Guide_R	TTCTAGCTCTAAAAC<u>GTACGCAACCTTAAAAAGAA</u>AATATTATATACTTA
p14	Pfap2-g_+940_Guide_F	TAAGTATATAATATT<u>TCATAAGAAATATGATTCAT</u>GTTTTAGAGCTAGAA
p15	Pfap2-g_+921_Guide_R	TTCTAGCTCTAAAAC<u>ATGAATCATATTTCTTAT</u>GAAATATTATATACTTA
p16	1144400_+1144_HR1_F_SacII	tgg**ccgcgg**GGAAAGACAATCAACAAACCTT
p17	1144400_+1562_HR1_R_SpeINcoI	tggtgt**actagtctgccatgg**TCACTTGAAATCGTTTTGTTATTA
p18	1144400_+1646_HR2_F_SpeI_OL	tggttg**actagt**ATTGTATAAATGAAGAGACATACA
p19	1144400_+2199_HR2_R_AflII_OL	tggtgt**cttaag**TGATTATTATGTGTAACTTCAGTA
p20	ap2g_-181_F1_F_NcoI	tggtgt**ccatgg**GTAGGTACATTCAAATATCTCC
p21	ap2g_+992_F1_R_SpeI	tggtgt**actagt**TTCTTCTATATTCTTTTCAAGGAT
p22	ap2g_-1200_F2_F_NcoI	tggtgt**ccatgg**GATACTTTATGATTAGTTGATATG
p23	ap2g_+118_F2_R_SpeI	tggagaa**ctagt**TTCCTGGGATGTAATCAAAAGT
p24	ap2g_-2147_F3_F_NcoI	tggtgt**ccatgg**ATAATGCTCTGAGAATATATACTT
p25	ap2g_-891_F3_R_SpeI	tgttgg**actagt**AATTTATATATCTCGACACCTTTA
p26	ap2g_-3000_F4_F_NcoI	tggtgt**ccatgg**TATTGACGTGCATACTATATGTA
p27	ap2g_-1834_F4_R_SpeI	tggagaa**ctagt**AATGTATTTCATGATATTAGTGTAA
p28	ama1_-966_F_NcoI	tggtgt**ccatgg**GTTATGTGTAGAATATTTACAAAG
p29	ama1_+249_R_SpeI	tgttgg**actagt**ACCTTCGTGGTCTATTGGATA
p30	Pfap2-g_+425_nucleo_R_SpeI	tggtgt**actagt**TCTATGATTTAACAACATATCCAA
p31	Pfap2-g_-6_2nucleo_F_NcoI	tggtgt**ccatgg**AAGAAGATGACCGCTAAGATAT
p32	Pfap2-g_+210_1nucleo_F_NcoI	tggtgt**ccatgg**ATTGAAGAAAGTTGTTTTGACATT
p33	Pfap2-g_+341_RestF1_F_NcoI	tggtgt**ccatgg**TTGCAGAGTAACGAGAATGTAT
p34	Pfap2-g_+748_F0_F_NcoI	tggtgt**ccatgg**GAAATGTAAAAGAGTATTCTCATA
p35	Pfap2-g_+1951_F0_R_SpeI	tggtgt**actagt**CAATTTATCAGCATCGTCATCA
p36	1144400_+1566_Guide_F	TAAGTATATAATATT<u>ATTATGTGTTCCAAGAACGA</u>GTTTTAGAGCTAGAA
p37	1144400_+1585_Guide_R	TTCTAGCTCTAAAAC<u>TCGTTCTTGGAACACATAA</u>TAATATTATATACTTA
p38	Pfap2-g_-448_HR1_KOF1_F	cggtacccgggatccATATGTCCTATAGGTGTCAAAC
p39	Pfap2-g_-159_HR1_KOF1_R	tcttctcctttactctgcagGGAGATATTTGAATGTACCTAC
p40	Pfap2-g_+993_HR2_KOF1_F	atacaaataaactagtCCTTGAAAAGAATATAGAAGAAC
p41	Pfap2-g_+1974_HR2_KOF1_R	gactctagaggatccCAATTTATCAGCATCGTCATCA
p42	Pfap2-g_-159_Guide_F	TTCTAGCTCTAAAAC<u>TTATATTGGCACTAATT</u>TAGAATATTATATACTTA
p43	Pfap2-g_-139_Guide_R	TAAGTATATAATATT<u>CTAAATTAGTGCCAATATAA</u>GTTTTAGAGCTAGAA
p44	Pfap2-g_-1857_HR1_KOF2-3_F	cggtacccgggatccTTACACTAATATCATGAAATACATT
p45	Pfap2-g_-1427_HR1_KOF2-3_R	tcttctcctttactctgcagACAAAATCTATAATCTTATATATGAA
p46	Pfap2-g_-129_HR2_KOF2-3_F	atacaaataaactagtAATTTGAAGTACCAACATATACAT
p47	Pfap2-g_+234_HR2_KOF2-3_R	gactctagaggatccAATGTCAAAACAACTTTCTTCAAT
p48	Pfap2-g_-1412_Guide_F	TAAGTATATAATATT<u>TTTAATAATACGTATGCTT</u>GGTTTTAGAGCTAGAA
p49	Pfap2-g_-1393_Guide_R	TTCTAGCTCTAAAAC<u>CAAGCATACGTATTATTAAAAAT</u>ATTATATACTTA

Primers for diagnostic PCR		
Number	**Name**	**Sequence (5'-3')**
p6	Pfap2g_-3929_HR1_KO_F_NotI	tggtgtgcggccgcCGTACATATATACAAATAGAGAT
p50	Pfap2-g _+2474_R	TTACTCCATTAGGTGCATTCAT
p51	1144400_+865_F	GATATTATGAGTACTGATAGTGA
p52	1144400_+2319_R	TGACTCGTTCAGGTTATTCTTA
p53	1144400_+1799_R	CAGATCTATATTTCATTTGTTCTT
p54	Pfap2-g_-2147_UpTSS1_F	TAATGCTCTGAGAATATATACTTA
p55	Pfap2-g_CDS_+345_R	GATACATTCTCGTTACTCTGC
p56	Pfap2-g_5'UTR_-460_F	GCTTCTTTAATGTTGTATGTATG

Primers for qPCR		
Number	**Name**	**Sequence (5'-3')**
p54	Pfap2-g_-2147_UpTSS1_F	TAATGCTCTGAGAATATATACTTA
p58	Pfap2-g_-1979_UpTSS1_R	CAATTAGGATAAATAAATATTCCAT
p59	Pfap2-g_-1820_UpTSS2_F	ATGGCTTTATTTATTCTTAATTGT
p60	Pfap2-g_-1644_UpTSS2_R	TATCTCTCACGGTTCATATCTT
p61	Pfap2-g_-1409_DwnTSS_F	TAATACGTATGCTTGTGGATATT
p62	Pfap2-g_-1226_DwnTSS_R	ATATGGAAACTAATAATAAATTGTTA
p63	PF3D7_1222600_F1 (CDS)	AACAACGTTCATTCATTCAATAAATAAGG
p64	PF3D7_1222600_R1 (CDS)	ATGTTAATGTTCCCAAACAACCG
p65	Seryl_qPCR_F (serrs)	AAGTAGCAGGTCATCGTGGTT
p66	Seryl_qPCR_R (serrs)	TTCGGCACATTCTTCCATAA
p67	hdhfr_F	AGTAGAAGGTAAACAGAATCT
p68	hdhfr_R	GGCATCATCTAGACTTCTGG

Primers for ChIP-qPCR		
Number	**Name**	**Sequence (5'-3')**
p69	5'HR_1144400_+1424_F	TAATGCAAAACGATAGTAATCTTA
p70	5'HR_1144400_+1526_R	TCTTCATTAGGATCATATCCATT
p71	3'HR_1144400_+1676_F	TGAAACTTGAAAACTGTGATGAA
p72	3'HR_1144400_+1799_R	CAGATCTATATTTCATTTGTTCTT
p73	PFL1950wF	CTATGTTGTATTATTCGATATTTTC
p74	PFL1950wR	AGAATAGGAAAATACAAATTATAGC
p75	Pfap2-g_CDS_+259F_OL	GAAGAGAGCATGCAATGAAGT
p76	Pfap2-g_CDS_+382R_OL	TTGTCCATGCAACTATTCGATA
p61	Pfap2-g_-1409_DwnTSS_F	TAATACGTATGCTTGTGGATATT
p62	Pfap2-g_-1226_DwnTSS_R	ATATGGAAACTAATAATAAATTGTTA
p77	Ap2-g_-2967_F_OLB	ATTATTACCTTCGGTACCTTAAT
p78	Ap2-g_-2861_R_OLB	AATGCACTTTTTGAGTACAGTTA
p79	PFL1085w_-2553_A2.5F	CAATATAATACCATACTTCAAAAC
p80	PFL1085w_-2460_A2.5R	AGTATGGTTTTGTACTCCTTTTA
p81	AMA1_+23_F_OLB	ATTATTGAGCGCCTTTGAGTTT
p82	AMA1_+145_R_OLB	GTGTAATGGATATTCGTATTCTT
p83	Pfap2-g_+653_F	TTTCCCATACTATATGCTGAAAT
p84	Pfap2-g_+774_R	CATATGAGAATACTCTTTTACATT
p85	Pfap2-g_+754_F	GTAAAAGAGTATTCTCATATGTAA
p86	Pfap2-g_+876_R	AACATAAAAGTACATGACATCATT
p87	clag3.2_5_F	TAGGCGAAAATAAAAACGAAAATG
p88	clag3.1_clag3.2_5_R	CATGGATTTTAATTGTTCAATATTG
p89	5'ext_1144400_+734_F	ATGAAGTAGCTACTACAATACAT
p90	5'ext_1144400_+870_R	AATATCCATAAGAACATCCTTGT
p91	3'ext_1144400_+2177_F	CTGAAGTTACACATAATAATCAAA
p52	3'ext_1144400_+2319_R	TGACTCGTTCAGGTTATTCTTA
p54	Pfap2-g_-2147_UpTSS1_F	TAATGCTCTGAGAATATATACTTA
p58	Pfap2-g_-1979_UpTSS1_R	CAATTAGGATAAATAAATATTCCAT
p92	PF3D7_1222600_F1	AACAACGTTCATTCATTCAATAAATAAGG
p93	PF3D7_1222600_R1	ATGTTAATGTTCCCAAACAACCG

42

DISCUSSION

In this thesis, we have identified a new route of sexual conversion that has revealed unexpected features of sexual conversion, challenging the textbook view of the parasite life cycle. We have also developed an inducible system for PfAP2-G, achieving the highest sexual conversion rate reported so far. The system allowed the characterization of the poorly studied sexually committed parasites and early sexual stages. Moreover, our results have provided the first insight into the mechanisms of heterochromatin formation in *pfap2-g*, revealing an unexpected role of the coding sequence. In the following lines we discuss about the regulation sexual commitment, sexual conversion and heterochromatin formation in the light of our results, but also about their implications and future perspectives.

1. On the regulatory basis of SCC and NCC

By using PfAP2-G as a marker of commitment, we identified a new route of sexual conversion in *P. falciparum*, challenging the current view of the life cycle. This new route involves early expression of PfAP2-G in ring stages, which allows the parasite to convert into stage I gametocytes without undergoing an additional round of replication. It is still unclear whether early activation of PfAP2-G involves the dismantling of the heterochromatin domain at the *pfap2-g* locus in ring stages or in the previous schizont stage. The occurrence of same cycle conversion (SCC) in natural infections, together with its functional implications, is currently unknown. Furthermore, the involvement of SCC and next cycle conversion (NCC) in stochastic or environmentally induced sexual conversion remains unexplored.

Nevertheless, we hypothesize that SCC may only reflect basal stochastic sexual conversion, whereas NCC may reflect both environmentally induced and stochastic sexual conversion.

One of the main observations supporting this hypothesis is the temporal dynamics of GDV1 expression and its regulation. GDV1 is mainly expressed in late IDC stages, with maximal expression at the schizont stage (Eksi et al., 2012; Filarsky et al., 2018). Moreover, depletion of LysoPC from the media increases GDV1 expression, probably by downregulating the expression of the *gdv1* antisense RNA (Filarsky et al., 2018). The sensing of LysoPC levels appears to occur at the trophozoite stage, probably involving transport via the PSAC channel, which is not present at earlier stages (Brancucci et al., 2017; Ito et al., 2017). Therefore, the environmentally induced sexual conversion by LysoPC can only be sensed in trophozoites leading to increased expression of GDV1 and thereafter PfAP2-G in schizonts. This time of PfAP2-G activation is incompatible with SCC. This was also evidenced in experiments tracking single schizonts and its offspring after LysoPC depletion (Brancucci et al., 2018). These committed parasites will reinvade and form sexual rings that will convert to stage I gametocytes the cycle after commitment (NCC) (**Figure 28**). Of note, throughout the thesis, we have defined sexual commitment as the point at which expression of the earliest available commitment marker occurs, which is presence of PfAP2-G protein.
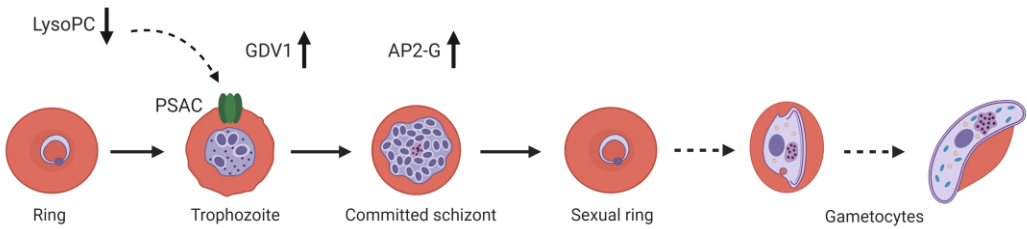
Figure 28. Inducible sexual conversion may only operate through the NCC. The drop in LysoPC levels is sensed by the parasite through its transport involving the PSAC channel, which is first expressed at the surface of trophozoite-infected RBCs. Low levels of LysoPC induce an increase in GDV1 protein levels that results in removal of HP1 from the *pfap2-g* locus, triggering its expression. In the committed schizont all the merozoites express PfAP2-G, developing into sexual rings upon reinvasion and later into gametocytes. Designed with Biorender.com.

In the absence of inducing signals, sexual conversion still occurs, likely reflecting basal stochastic activation of the *pfap2-g* locus. The stochastic activation of the locus may account for SCC: our experiments involving plaque assays revealed that mixed plaques reflect SCC events, such that the "decision" to commit to sexual conversion can occur independently in sibling parasites. This "decision" is likely to be stochastic, as it is an independent event arising in parasites growing under the same conditions. At which point this stochastic activation occurs is unexplored, however, in the following lines, we first discuss the possible mechanism of switching and then propose different models that may explain SCC based on them.

- *Mechanisms of epigenetic switching*

Stochastic events are by definition unpredictable, but at the population level they can result in defined outcomes such as a stable frequency of switching (Tawfik, 2010). In fact, these stable frequencies of switching

are essential for the maintenance of cell-to-cell variability that are the basis of bet-hedging survival strategies (Bury-Moné and Sclavi, 2017; Herman et al., 2014). Specifically, the on and off switching rate of CVGs appears to be a stable intrinsic property of these gene families, as reported for *var* genes (Fastman et al., 2012; Frank et al., 2007; Horrocks et al., 2004; Recker et al., 2011) and also indirectly for *pfap2-g* (Kafsack et al., 2014).

This intrinsic property suggests that the switching rate is inherited across generations and is probably encoded by the epigenetic state of the locus. In the case of *pfap2-g*, only the silenced state is inherited, as sexual stages abandon the replicative cycle. Therefore, the probability of activation must be somehow transmitted by the heterochromatin domain. It is possible that heterochromatin domains have different degrees of "strength", understood as their resistance to activation, which may be reflected in their extension. While the heterochromatin state would in all cases result in silencing of the gene, variable extension of the domain may determine the probability of activation: heterochromatin domains with shorter extensions may be more prone to switching. We propose that the transmission of the epigenetic memory is susceptible to "errors", and these "errors" might be more frequent in regions with "weaker" heterochromatin domains (**Figure 29**).

Actually, the accuracy and fidelity of biological processes are limited, something inevitable due to the characteristics of biological components and processes. However, this "messiness" provides the grounds for evolution, providing variability or noise for natural selection to operate (Tawfik, 2010). One good example of "errors" in

biology are DNA polymerases, which make mistakes in 1 nucleotide every 100,000, with proofreading mechanisms repairing 99% of them (Pray, 2008). Therefore, the transmission of the epigenetic memory is also likely to be susceptible to mistakes.

In fact, during DNA replication, nucleosomes are removed and later recycled to be incorporated into the new DNA (Margueron and Reinberg, 2010). DNA replication requires double amounts of nucleosomes; therefore, parental histones represent, at most, half of the histones in the new chromatin. The current model postulates that parental histones serve as a template for the modification of new histones, through the coupling of epigenetic readers and writers. Furthermore, the distribution of parental histones is stochastic, meaning that some regions of the new DNA may exclusively contain new or old histones, limiting the accuracy of epigenetic inheritance (O'Kane and Hyland, 2019). Actually, inheritance of gene silencing has been proposed to depend on a threshold quantity of repressive histone modifications over broad genomic regions. Such regions act as "buffering" areas tolerating errors that may occur during DNA replication (Xu et al., 2011).

Therefore, "weaker" heterochromatin domains, or less extensive heterochromatic regions, may provide histone methylation levels below the threshold required for a successful epigenetic inheritance, which may lead to epigenetic switches. In fact, errors in the epigenetic memory themselves could provide a mechanism to progressively weaken heterochromatin domains along subsequent DNA replication rounds (**Figure 29**).
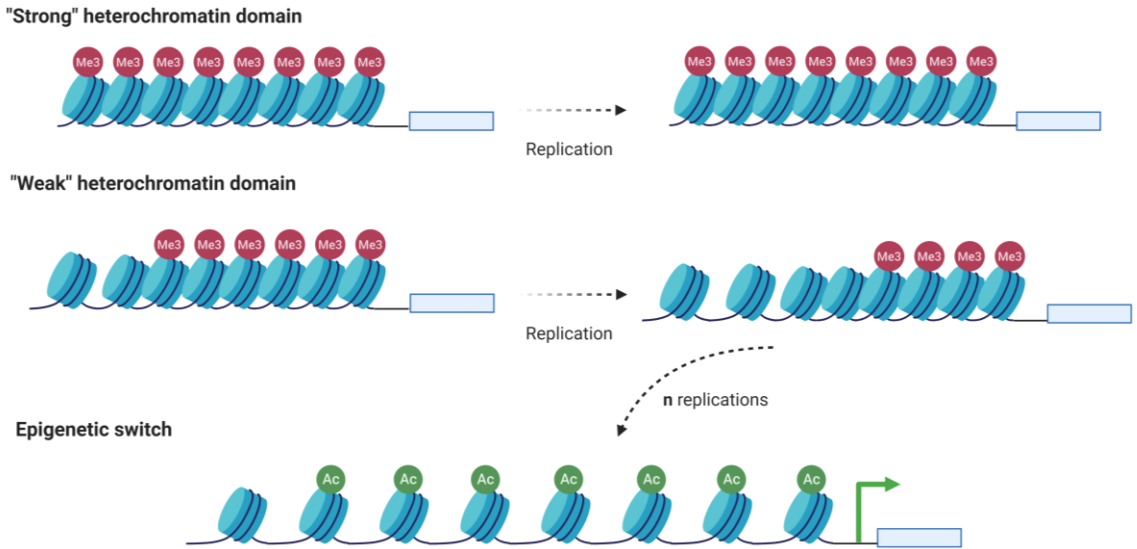
Figure 29. The extension of an heterochromatin domain may determine its resistance to an epigenetic switch. Hypothetical model in which "stronger" heterochromatin domains occupy a more extensive region, which provides a more stable epigenetic inheritance. Nevertheless, mistakes in the transmission of the epigenetic memory can weaken a domain by reducing its extension. Less extensive domains are more prone to "errors" in the transmission of the epigenetic memory, favouring epigenetic switches. The red and green circles over the histones represent H3K9me3 and H3K9ac, respectively. The green arrow represents active transcription. Designed with Biorender.com.

Alternatively, these "weaker" heterochromatin domains could be more susceptible to an epigenetic switch independently of the replication process, probably involving the interaction of epigenetic regulators or transcription factors.

- *Stochastic switching of pfap2-g and SCC*

We propose that the switch in the epigenetic state of *pfap2-g* involved in SCC can occur at different moments during the IDC, such as during schizogony or at the ring stage.

A switch during schizogony could be explained by the failure of the epigenetic memory during DNA replication. DNA replication in *P. falciparum* starts at the interface between late trophozoites and schizonts, through asynchronous schizogony that involves multiple nuclear divisions, forming a multinucleate syncytium, prior to budding and cytokinesis (Francia and Striepen, 2014). Therefore, errors in the transmission of the epigenetic state of *pfap2-g* at any point during schizogony, can be transmitted in subsequent nuclear divisions, resulting in a schizont with heterogeneous merozoites (**Figure 30**). In merozoites in which a switch in *pfap2-g* has occurred during replication, *pfap2-g* expression could start after invasion, probably controlled by the action of a ring-specific ApiAP2 TF together with the PfAP2-G feedback loop (**Figure 31.1**). Following a definition of commitment based on the presence of PfAP2-G protein, this epigenetic heterogeneity in sibling merozoites would still be defined as SCC.
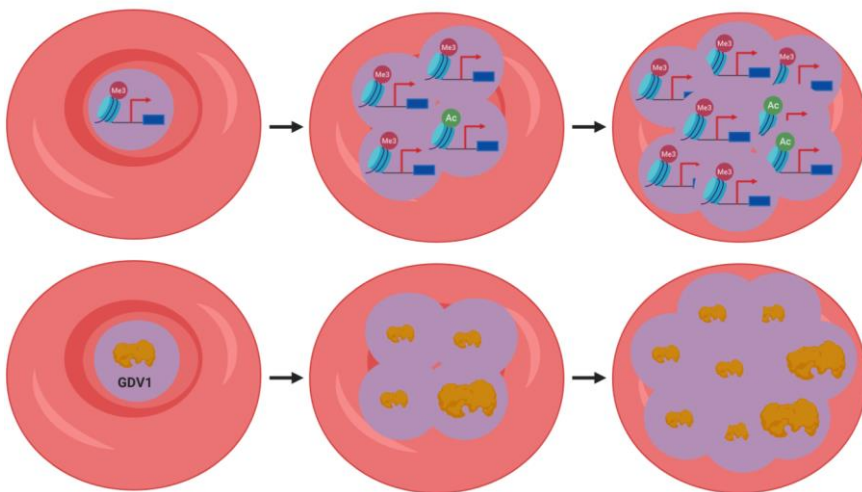
Figure 30. Models of *pfap2-g* switching at the schizont stage that may account for SCC. During schizogony, a mistake in the transmission of the epigenetic memory can occur during DNA replication, resulting in a switch of the

241

epigenetic state in one of the daughter cells. This switch into an active state is stably transmitted through subsequent replication rounds. Alternatively, uneven GDV1 distribution can promote the switch into an active state of the *pfap2-g* promoter in some of the merozoites. Purple circles represent merozoites. The red and green circles over the histones represent H3K9me3 and H3K9ac, respectively. The red arrow represents an inactive *pfap2-g* TSS. GDV1 is represented as an orange molecule and the size represents its protein levels. Designed with Biorender.com.

An alternative model for the generation of epigenetic heterogeneity could be the uneven distribution of GDV1 protein into sibling merozoites. If during schizogony some merozoites inherit GDV1 protein levels that surpass a determined threshold, the *pfap2-g* locus could switch into an active state (**Figure 30** and **Figure 31.2**). The main argument against both models is that schizonts are able to express PfAP2-G protein; therefore, merozoites with an active locus should express the TF, resulting in schizonts with merozoites positive and negative for PfAP2-G, something that our IFAs did not detect. It is unclear how merozoites could not express PfAP2-G with a permissive *pfap2-g* locus. However, initial expression of PfAP2-G may require the involvement of a ring-specific TF, whereas expression in schizonts would require the presence of the positive feedback loop.

SCC could also be explained by a switch of *pfap2-g* into an active state in ring stages (**Figure 31.3**). However, the exact mechanisms that might be involved are unknown, as they would be likely independent of DNA replication and GDV1, which is expressed from trophozoites onwards. The transition might involve a stochastic event or specific epigenetic regulators that target the locus in ring stages.

DISCUSSION

- *NCC can also arise from basal pfap2-g activation*

In the first article included in this thesis, we proposed that parasites expressing *pfap2-g* early at the ring stage, and that reach a determined threshold of PfAP2-G before the trophozoite stage, will undergo direct conversion into stage I gametocytes (SCC). If the committed parasites do not reach this threshold, they continue growing asexually as committed trophozoites and schizonts until they convert at the following cycle (NCC). Therefore, conversion through NCC can also proceed after commitment in ring stages, providing a basal rate of sexual conversion (**Figure 31.4**).

In the absence of the positive feedback loop, *pfap2-g* expression essentially occurs in ring stages, as shown by its temporal expression pattern in the F12 line. Together with the observation that PfAP2-G-positive rings can arise from PfAP2-G-negative schizonts, this suggests that the onset of *pfap2-g* expression likely occurs at the ring stage. The existence of a threshold level of PfAP2-G protein to drive conversion has not been demonstrated. Thus, given that *pfap2-g* expression is maximal in rings, it remains possible that cells expressing PfAP2-G in rings can develop into gametocytes, without the need to reach a determined threshold. In such scenario, NCC could not be explained by commitment at the ring stage, as all parasites would develop into gametocytes through SCC, and the occurrence of NCC would only depend on the activity of GDV1 protein at trophozoites (**Figure 31.5**).
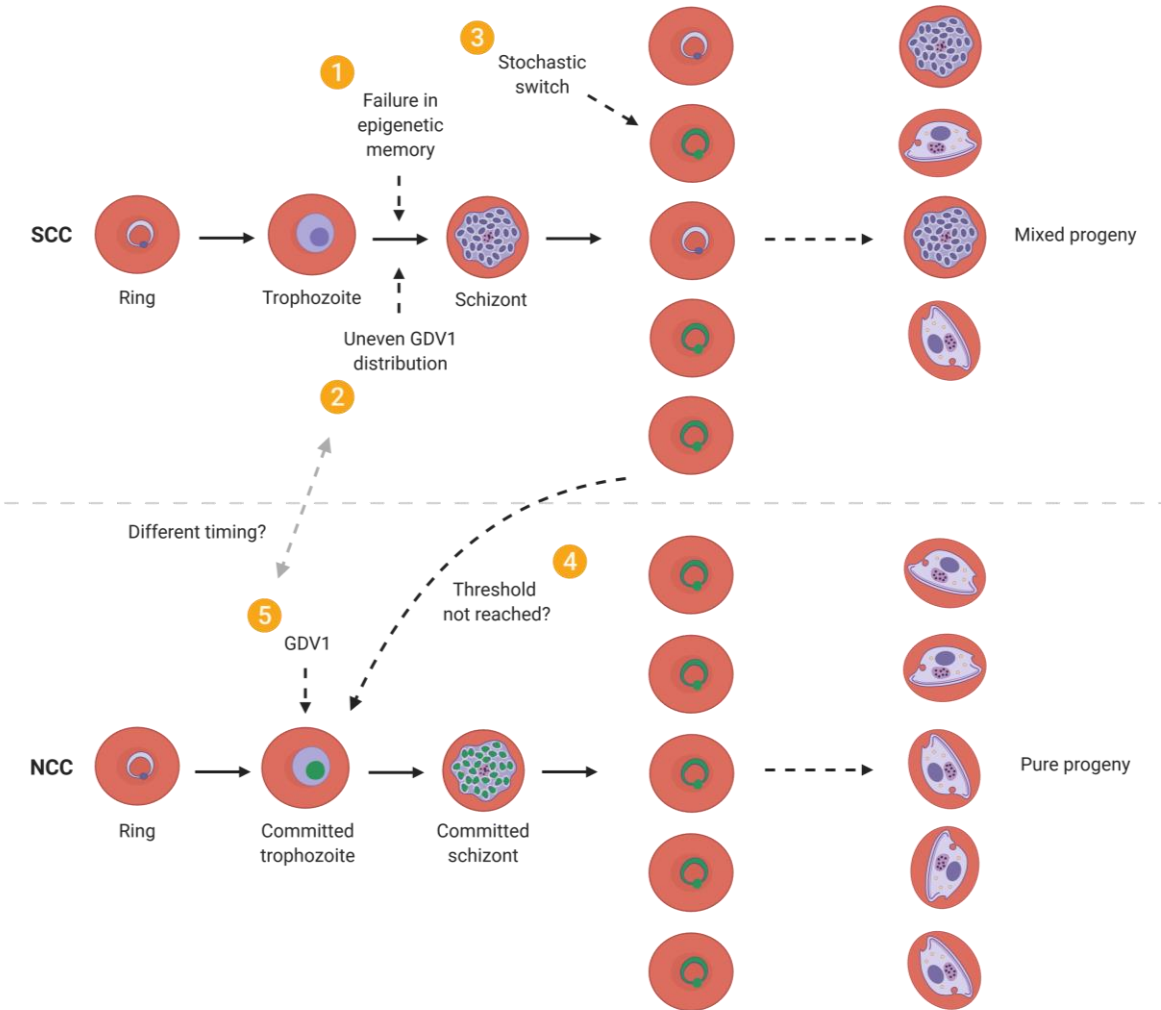
Figure 31. Hypothetical model of the switching in the epigenetic stage of *pfap2-g* leading to SCC or NCC. "Errors" in the transmission of the epigenetic memory (**1**) or uneven distribution of GDV1 protein (**2**) may result in epigenetically heterogeneous merozoites. Upon reinvasion, parasites expressing PfAP2-G above a determined threshold will develop into gametocytes (SCC). *De novo* activation of PfAP2-G in rings can also occur as a consequence of direct dismantling of heterochromatin at *pfap2-g* in rings (**3**). If PfAP2-G levels do not reach the threshold, parasites continue as committed forms and convert the following cycle (NCC) (**4**). NCC can also reflect a switch of the epigenetic state of *pfap2-g* in trophozoites involving GDV1 and leading to the formation of committed schizonts. After reinvasion, parasites develop into gametocytes (**5**). The green parasites represent committed stages expressing PfAP2-G. Designed with Biorender.com

According to the strand-specific RNA-Seq data, the *gdv1* antisense RNA is highly expressed across the whole IDC (Bártfai et al., 2010; Toenhake et al., 2018), suggesting a tight control of GDV1 protein levels. However, even under non-inducing conditions, such as in the presence of LysoPC, GDV1 is expressed in some cells (Filarsky et al., 2018). Therefore, GDV1 could activate the *pfap2-g* locus in trophozoites, explaining the presence of committed trophozoites, and trigger conversion at the following cycle, a route that is enhanced under inducing conditions.

Overall, we propose a model in which NCC reflects both basal and inducible sexual conversion with commitment occurring in rings or trophozoites. On the other hand, SCC only reflects basal sexual conversion with the switch in the epigenetic state of *pfap2-g* occurring either in schizonts or in ring stages, with commitment in ring stages (**Figure 31**).

- *The role of GDV1 in stochastic and induced sexual conversion*

Surprisingly, not all cells expressing GDV1 activate *pfap2-g* expression under non-inducing conditions, as evidenced by a sexual conversion rate below 5% in a population with 18% of parasites positive for GDV1 (Filarsky et al., 2018). This suggests that a threshold in GDV1 protein levels may be needed to dismantle the heterochromatin domain at *pfap2-g*. Even under inducing conditions, not all GDV1-expressing parasites convert into gametocytes (30% conversion vs 50% GDV1 positive) (Filarsky et al., 2018). This could reflect a delicate balance between opposing forces operating at the *pfap2-g* locus: on the one hand, GDV1 destabilizing the heterochromatin

environment; on the other hand, the HP1, HDA2 and H3K9me3 HMT feedback loops reinforcing the heterochromatin state (Josling et al., 2018).

As opposed to a deterministic model in which GDV1 expression results in *pfap2-g* activation, it is possible that stochasticity may play a role, favouring the balance towards one side or the other. The action of GDV1 may also provide a source of variability in the "strength" of the *pfap2-g* heterochromatin domain. In those cells where presence of GDV1 is not sufficient to trigger the activation of *pfap2-g*, the heterochromatin domain could become "weaker", with higher chances of switching in subsequent cycles of growth. Therefore, GDV1 may also influence and determine the rate of stochastic activation at the following generations by altering the *pfap2-g* heterochromatin domain.

The role of GDV1 in induced conversion is well established, as choline depletion leads to a rapid increase in the percentage of GDV1-expressing cells, concomitant with the increase in sexual conversion (Filarsky et al., 2018). In contrast, the role of GDV1 in basal sexual conversion is less clear, but some data suggests that it also plays a role. For instance, in parasite lines lacking functional GDV1 protein, such as the 3D7 deficient line (Eksi et al., 2012), gametocytes are not produced, suggesting that *pfap2-g* is not expressed. Actually, in our 1.2B line, which contains a premature stop codon in GDV1, *pfap2-g* is not expressed (Rovira-Graells et al., 2012). This highlights the involvement of GDV1 in the stochastic activation of *pfap2-g*, in addition to its role in induced sexual conversion. The absence of GDV1 for prolonged times probably favours the balance towards a

"stronger" silencing of *pfap2-g* through a more stable heterochromatin domain.

In fact, the ChIP-Seq analysis of 1.2Bind+Rapa_prol parasites revealed a broader heterochromatin domain than in E5ind+Rapa_prol. Ongoing comparative ChIP-Seq studies on parasite clones also support this idea. Additionally, 1.2Bind parasites yield lower sexual conversion after rapamycin treatment and a higher reduction of sexual conversion levels after the removal of WR99210 selection for 5 weeks, compared to E5ind. Altogether, these results reflect a "stronger" heterochromatin domain in 1.2B. Nevertheless, in these parasites, GDV1 has been absent for long periods of time. The conditional depletion of GDV1 would be very valuable to assess the immediate effect of GDV1 depletion on the heterochromatin at the *pfap2-g* locus and on the stochastic sexual conversion.

Interestingly, silencing of *pfap2-g* expression is mediated by heterochromatin in all *Plasmodium spp.* (Fraschka et al., 2018), but GDV1 is absent in rodent malaria parasites (Eksi et al., 2012). Therefore, the stochastic activation of *pfap2-g* could also be independent of GDV1, at least in rodent malaria parasites. Rodent malaria parasites were suggested to be insensitive to environmentally induced sexual conversion, as they lack some of the proteins involved in the response to LysoPC depletion, such as an alpha/beta hydrolase (PF3D7_1001600), the tryptophan rich antigen P-Art (PF3D7_1002200), an ApiAP2 (PF3D7_1222400) or GDV1 (Brancucci et al., 2017).

Overall, the identification of two alternative routes of sexual conversion has provided the basis to better interpret cell fate determination in *P. falciparum*. The added complexity that these two routes provide will push the field for a better characterization of the molecular mechanisms behind sexual commitment. Here we have proposed different models of epigenetic switching and how they can fit into the two routes of sexual conversion. Nevertheless, there are several questions that need to be addressed experimentally. The exact point(s) at which the epigenetic switch occurs leading to *pfap2-g* activation, and later to sexual conversion, is unknown. The development of single-cell ChIP-Seq procedures could provide insight into the switching model involving a failure in the epigenetic memory during schizogony. If that was the case, it is unclear how a permissive *pfap2-g* is not coupled to presence of PfAP2-G protein. Another key point would be the validation of the PfAP2-G threshold needed for conversion through SCC.

2. On the *pfap2-g* transcriptional unit

The identification of the TSSs of *pfap2-g* has provided insight into the potential use of alternative TSSs in the presence or in the absence of the PfAP2-G positive feedback loop. In the absence of functional PfAP2-G, the relative abundance of transcripts including the region immediately downstream of the TSSs identified is similar to the abundance of transcripts including regions upstream of these TSSs, as determined by RT-qPCR. This suggests that other TSSs may be occurring upstream and that they might be relevant for the initial transcription of *pfap2-g*. To test this hypothesis, other approaches such as 5'RACE of the F12 line or the E5-PfAP2-G-DD line in the absence

of Shld, would be needed to confirm the predominance of upstream TSSs in the absence of the positive feedback loop.

The location of the TSSs, together with the transcriptional pattern across the IDC, can provide some clues regarding the potential candidate TFs involved in *pfap2-g* expression. Several predicted ApiAP2 DNA binding motifs are present around the TSSs identified, which may account for a regulatory function (Campbell et al., 2010). Among the ApiAP2s with a cognate motif near the *pfap2-g* TSSs, several have been predicted to be involved in gametocyte production or to be expressed in committed cells (**Figure 32**) (Josling et al., 2018): PfAP2-Exp is upregulated in PfAP2-G positive cells (Poran et al., 2017), with expression mainly occurring in late schizonts (Martins et al., 2017); SIP2 is upregulated in committed cells and stabilized in gametocytes (Painter et al., 2017; Poran et al., 2017), with expression in late stages (Flueck et al., 2010); PfAP2-O is upregulated and stabilized in gametocytes (Lu et al., 2017; Painter et al., 2017), and highly expressed in schizonts; PF3D7_1139300 is upregulated in committed cells (Poran et al., 2017) and is mainly expressed in ring stages; and PF3D7_0613800 is upregulated in committed parasites upon LysoPC depletion (Brancucci et al., 2017; Poran et al., 2017) and is mainly expressed in late stages.

Other ApiAP2, for which the predicted binding motifs are absent in the regulatory region of *pfap2-g*, have been linked to gametocyte production and are thus also candidates to be involved in the expression of *pfap2-g*. These include PF3D7_1222400 and PF3D7_1429200, which are upregulated in committed parasites upon

LysoPC depletion (Brancucci et al., 2017; Lu et al., 2017; Poran et al., 2017); and PfAP2-FG (Ikadai et al., 2013).
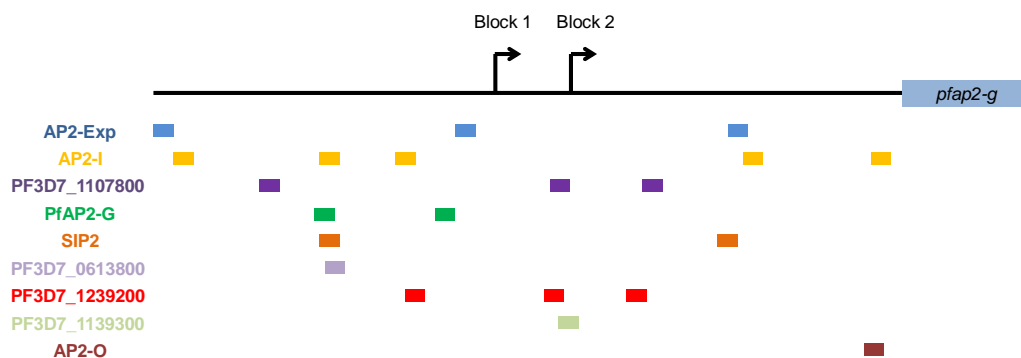
Figure 32. Position of the predicted ApiAP2 DNA binding motifs of TFs linked to sexual conversion along the promoter region of *pfap2-g*. The positions of the predicted DNA binding motifs of the ApiAP2 TFs are depicted as small boxes. The arrows represent the positions of the blocks of TSSs identified. Data from PlasmoDB.

According to our model, initial basal transcription of *pfap2-g* occurs at the ring stage, probably involving a ring-specific TF. Among the candidates presented above, only PF3D7_1139300 is expressed in rings. Our inducible system provides a great opportunity to dissect the dichotomy between cause-effect: until now, discerning whether higher expression of a particular TF in early committed parasites was reflecting a potential role upstream of *pfap2-g*, or if PfAP2-G was activating the expression of the TF, was almost impossible. With our inducible lines, we can identify all the events occurring downstream of PfAP2-G, discerning whether the TFs increased in committed parasites are targets of PfAP2-G or not.

In our inducible system, PF3D7_1139300 transcript levels increased in committed schizonts after the activation of *pfap2-g*, suggesting that it is a target of PfAP2-G, which is also reflected by ChIP-Seq analysis

(Josling et al., 2019). Therefore, the ApiAP2 TF involved in *pfap2-g* expression at the ring stage remains unknown. However, it is plausible that a constitutive TF expressed at rings, whose expression does not change during commitment or sexual development, triggers *pfap2-g* expression at this stage when the locus is accessible (i.e. when the promoter is in an euchromatic state).

From all the other candidates, expressed mainly in late stages, only SIP2 and PfAP2-O did not respond to the activation of *pfap2-g*, suggesting that they may operate upstream of it. Nevertheless, SIP2 can be excluded as no binding at the *pfap2-g* locus was observed by ChIP-on-ChIP approaches (Flueck et al., 2010), leaving PfAP2-O as the only potential candidate for regulating *pfap2-g* in late stages. However, its predicted DNA binding motifs are located close to the CDS, suggesting that it is unlikely to be involved in triggering *pfap2-g* expression.

It is important to mention that TFs upregulated upon *pfap2-g* activation cannot be completely excluded, as *pfap2-g* may also activate its upstream activators to promote a positive feedback loop to increase its transcription. A detailed characterization of the whole ApiAP2 family, together with the validation of their predicted DNA binding motifs *in vivo*, will represent an important milestone, disentangling the intricate regulatory network of ApiAP2 TFs.

Other proteins have also been proposed to regulate *pfap2-g* expression due to their upregulation in committed parasites, such as the epigenetic regulators ISWI, SNF2L, LSD2 and HDA1 (Poran et al., 2017). In agreement with this study, our inducible system suggests that

LSD2 and HDA1 are targets of PfAP2-G, whereas ISWI and SNF2L likely operate upstream of it. Other epigenetic regulators involved in triggering *pfap2-g* expression, such as HP1 or GDV1, are indeed not upregulated in our dataset, highlighting the suitability of our data to discern between events occurring before, or after, the point of commitment.

3. On the posttranscriptional regulation of *pfap2-g*

As previously discussed, the inducible activation system for PfAP2-G has provided the field with a great opportunity to study and characterize committed schizonts and sexual rings. Moreover, it has also provided some insight into the regulation of *pfap2-g*.

The recently available ChIP-Seq data for PfAP2-G identified many genes bound by the TF at stage I gametocytes, which were not bound in previous stages (Josling et al., 2019). Therefore, PfAP2-G may have a role beyond sexual commitment. Indeed, our IFA analysis for PfAP2-G revealed that PfAP2-G signal is lost form the nucleus in late stage I gametocytes. The absence of PfAP2-G beyond stage I gametocytes suggests that it is mainly needed during the first steps of sexual conversion in addition to triggering commitment.

For this reason, the use of a strong and constitutive promoter to drive expression of *pfap2-g* in our inducible system was a risky decision: the overexpression of PfAP2-G could have detrimental effects for the cell. Fortunately, the activity of the *calmodulin* (*cam*) promoter integrated at the *pfap2-g* locus was far from the activity of the endogenous *cam* promoter and resembled the activity of the active endogenous *pfap2-g* promoter. The heterochromatin environment within the locus likely

influences the activity of the *cam* promoter, maintaining *pfap2-g* transcripts within physiological levels (**Figure 33**). It is possible that in some parasites higher expression levels are achieved, leading to an excess of PfAP2-G which may have detrimental effects. This could occur in the few parasites that get stuck in their development after the treatment with rapamycin.
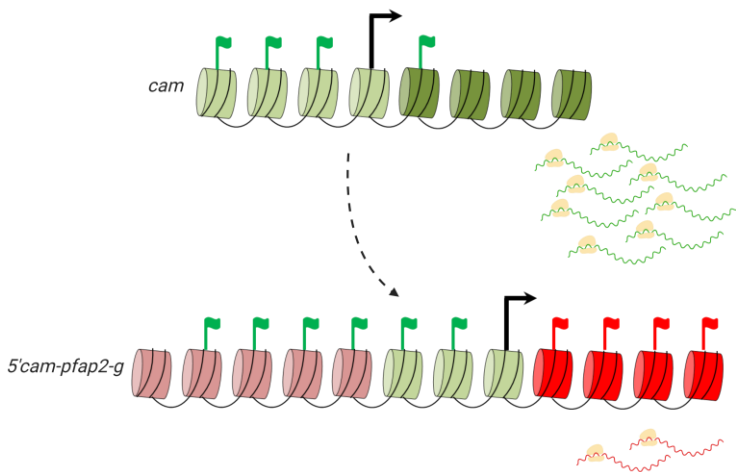
Figure 33. The activity of the *cam* promoter at the *pfap2-g* locus may be influenced by the heterochromatin environment. The constitutive and strong *cam* promoter provides high transcript levels (wavy lines with a ribosome attached) of the *cam* gene. The *cam* promoter integrated at the *pfap2-g* locus is possibly affected by the heterochromatin environment, resulting in lower levels of transcription. Pale green and pale red histones represent the promoter regions of *cam* and *pfap2-g*, respectively. Green and red histones represent the CDS of each gene. The green flags and the red flags represent activating histone marks and repressing histone marks, respectively. Designed with Biorender.com.

Moreover, the use of a constitutive promoter may represent an obstacle, as expressing PfAP2-G beyond the required stages may have deleterious consequences for gametocyte development. Nevertheless, and to our surprise, *pfap2-g* transcripts levels decreased dramatically beyond stage I gametocytes, whereas *cam* transcripts were still detected

in these stages. The heterochromatin profile of stage II-V gametocytes evidences the absence of HP1 at the promoter region of *pfap2-g* (Fraschka et al., 2018), making unlikely a possible silencing of the *cam* promoter by heterochromatin.

As the heterochromatin environment cannot fully explain the silencing of *pfap2-g* beyond stage I gametocytes, we propose that posttranscriptional regulatory mechanisms might be involved in controlling *pfap2-g* transcript levels in gametocyte stages. Our inducible lines have a modified 5'UTR, but the 3'UTR is intact, which suggests that in wild type parasites the 3'UTR may be important for the posttranscriptional regulation of the gene. In fact, *pfap2-g* transcript levels are stabilized after its transcription, as reported by a study describing real-time transcription and steady state transcripts along the IDC (Painter et al., 2018), suggesting the involvement of posttranscriptional mechanisms in its regulation. RNA-binding proteins of the Puf family have been described to be involved in gametocyte development (Miao et al., 2013; Shrestha et al., 2016). Indeed, in PfPUF2-KO parasites *pfap2-g* transcript levels seem to increase (Josling and Llinás, 2015; Miao et al., 2010b).

Other posttranscriptional regulators that might be involved in regulating *pfap2-g* are ribonucleases. Actually, the knockdown of PfRNase II, a non-canonical exoribonuclease, leads to increased expression of *pfap2-g* and gametocyte markers, suggesting an upregulation of *pfap2-g* and sexual conversion (Zhang et al., 2014). Moreover, PfRNase II expression peaks in stage I gametocytes, correlating with the decrease in *pfap2-g* transcript levels (Biljon et al., 2019). Therefore, PfRNase II is a good candidate to regulate the

expression of *pfap2-g*, preventing the presence of PfAP2-G in the following gametocyte stages. This predicted control mechanism would block the positive feedback loop of PfAP2-G, resulting in a further decrease in *pfap2-g* transcripts levels. Nevertheless, in stage IV-V gametocytes, PfRNase II levels decrease again, coinciding with a slight increase of PfAP2-G (Biljon et al., 2019); this suggests that once activated, the *pfap2-g* promoter keeps some level of activity along gametocyte development.

4. On the functional implications of SCC and NCC

As previously discussed, the existence of the two routes could provide the parasite with different alternatives for sexual conversion, safeguarding transmission and allowing adjustments to overcome environmental stresses. SCC allows the parasite to convert immediately after invasion, ensuring that gametocytes would be produced even if the population encounters unfavourable conditions that severely reduce its numbers. On the other hand, NCC allows increasing gametocyte numbers and therefore the probability of transmission. Overall, the occurrence of both routes is likely to provide an evolutionary advantage, as evidenced by its conservation in *P. berghei* (Kent et al., 2018) and probably across *Plasmodium spp.*. This suggests a strong selection for the maintenance of both routes; however, the exact biological relevance of each route in natural infections and how often each of them is used remains unknown. Due to its epidemiological implications, future studies should try to uncover the biological relevance of the two routes.

- *The sequestration of SCC and NCC gametocytes*

With the current data, we can only speculate on functional differences between the gametocytes produced through SCC and NCC. Conversion through NCC involves the formation of committed schizonts, releasing committed merozoites that will develop into sexual stages. Our characterization of committed schizonts revealed that, contrary to the expectations, they are almost identical to non-committed schizonts, with just a few transcriptional differences. Committed schizonts and merozoites were expected to have some phenotypic peculiarities, which may be involved in the sequestration of sexual stages in the spleen or the bone marrow. For instance, a preferential invasion of erythrocyte precursors was proposed as a mechanism for the sequestration in the bone marrow (Josling et al., 2019; Nilsson et al., 2015; De Niz et al., 2018) (**Figure 34**).

Our results suggest that sequestration in the bone marrow cannot be explained by an increased tropism of committed merozoites for erythrocyte precursors like reticulocytes. Although we cannot exclude the possibility that they have tropism for other RBC precursors, it is likely that sequestration is directed by another mechanism.

Another piece of data argues against sequestration being mainly determined at the committed schizont or committed merozoite stage. SCC gametocytes skip the committed schizont stage, therefore the potential factors present in committed schizonts and merozoites, which could explain tissue sequestration, are not applicable to SCC gametocytes. Immature gametocytes are not present in the peripheral circulation, suggesting that gametocytes produced by SCC also

sequester in these tissues. One explanation is the involvement of sexual rings or early gametocytes in the sequestration process, already proposed previously (Nilsson et al., 2015; De Niz et al., 2018), rather than committed schizonts or merozoites stages (**Figure 35**).
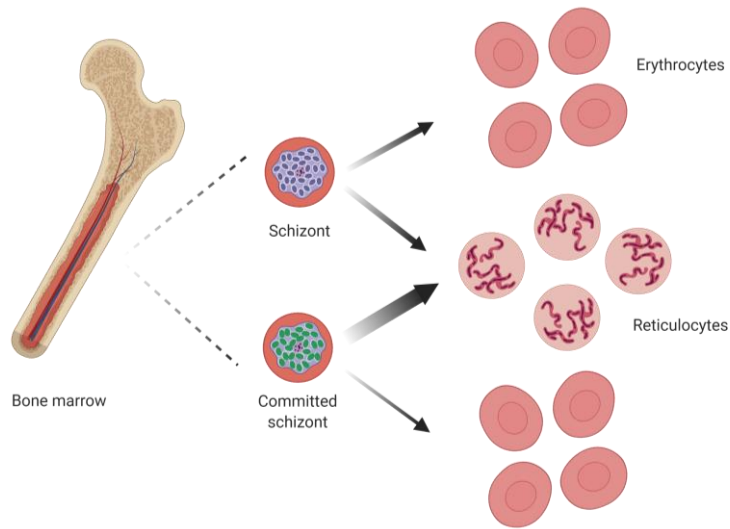
Figure 34. Hypothetical model for the sequestration of sexual stages in the bone marrow. Sequestration in the bone marrow could be mediated by a higher preference for reticulocyte invasion of sexually committed merozoites. Sexually committed merozoites are represented in green. The higher preference for reticulocytes is shown by a thicker arrow. Designed with Biorender.com.

Alternatively, if committed merozoites can influence the tissue of gametocyte sequestration, gametocytes produced by SCC and NCC may sequester to different tissues (**Figure 35**). We have demonstrated that committed schizonts downregulate some of the components of the PSAC, which is translated into a decreased permeability of gametocytes to some solutes. Given that SCC avoids the committed schizont stage, gametocytes arising from SCC events probably express the PSAC in their surface. If presence of the PSAC is confirmed in SCC, it could be exploited as a marker to identify gametocytes

produced by SCC in the different tissues. The screening of the relative abundance of SCC and NCC gametocytes in these tissues would shed light into a possible preference of committed merozoites for one of them.
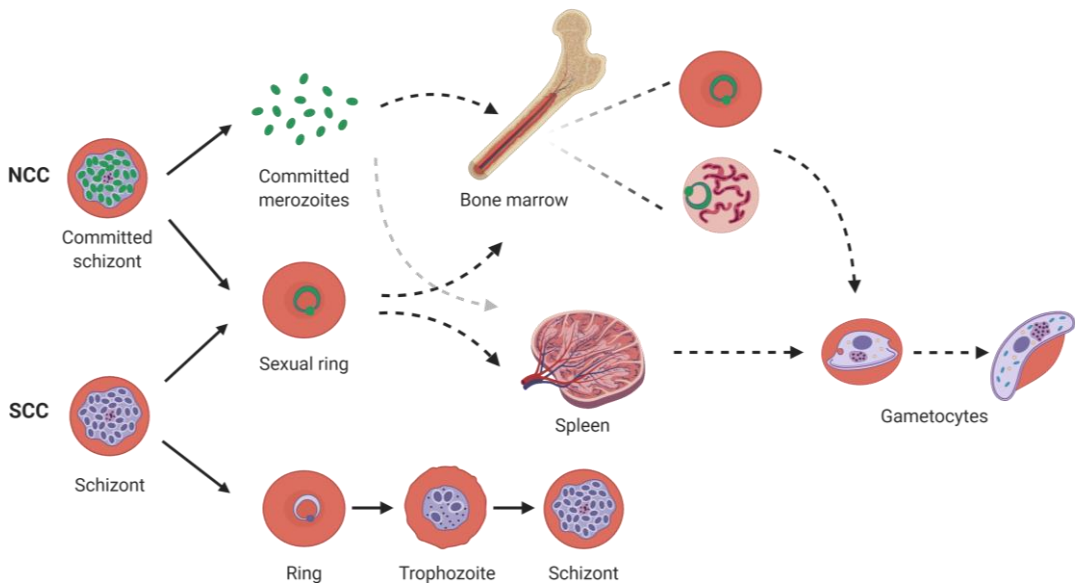
Figure 35. Gametocytes arising from SCC and NCC might sequester through different mechanisms. Committed merozoites have been hypothesized to express some ligands allowing the enrichment of sexual stages in tissues where they sequester, especially in the bone marrow. SCC skips the committed merozoite stage; therefore, SCC gametocytes may lack this specific tropism. Given that no immature gametocytes are present in the circulation, SCC gametocytes are likely to be able to sequester. Therefore, sexual rings might also direct sequestration, potentially with a different relative preference for tissues than committed merozoites. Parasites in green are sexually committed or sexual rings. Designed with Biorender.com.

- *The role of PSAC in gametocytes*

The silencing of the PSAC could reflect a shift in the metabolic needs of sexual stages or a defence mechanism against toxic compounds. However, SCC gametocytes do not have a choice, as the PSAC components are inherited from the previous cycle. The trafficking

though the PSAC is a passive mechanism, leading to the accumulation of nutrients that gametocytes may not need. This raises the possibility that SCC gametocytes can benefit from other nutrient sources or that PSAC may be somehow blocked to avoid the entrance of such nutrients, which may pose a fitness cost for SCC gametocytes.

- *The emergence of SCC*

Examining the mechanisms of sexual conversion in other apicomplexan parasites can provide some clues on the relevance of the SCC and NCC routes. PfAP2-G is conserved in *Plasmodium spp.*, with orthologues present in other apicomplexan parasites such as *Babesia spp.*, *Theileria spp.* and *Eimeria spp.* (Kafsack et al., 2014), which suggests a possible conservation of its role.

In *Babesia spp.* and *Theileria spp.* a few parasites convert into sexual forms during multiplication in the blood (Jalovecka et al., 2018). In *Eimeria spp.* no gametocytes are produced until a determined number of asexual multiplications are reached (a number that varies between different species), after which all parasites convert (Walker et al., 2013). Other apicomplexans, such as *Haemoproteus spp.*, convert directly into gametocytes after one replication through schizogony in the blood (Atkinson and van Riper III, 1991) (**Figure 36**).

Replication within the host cell for prolonged times, rather than directly developing into sexual stages, increases the number of gametocytes and enhances the probability of transmitting. Therefore, mechanisms allowing a tight control of this developmental transition were likely to evolve. The silencing of AP2-G, or a related TF regulating sexual commitment, could be the basis of this mechanism.

259

In parasites such as *Eimeria spp.*, silencing could be programmed to last a determined number of cycles or be overridden through a quorum sensing mechanism, whereas in *Babesia spp.*, *Theileria spp.* and *Plasmodium spp.* the silencing is overridden in a few parasites at each cycle. It would be interesting to assess the potential involvement of heterochromatin in the silencing of the AP2-G orthologue in these parasites. In *Plasmodium spp.*, the spontaneous switches in the epigenetic state of the locus provide a basal rate of activation and constant production of gametocytes.
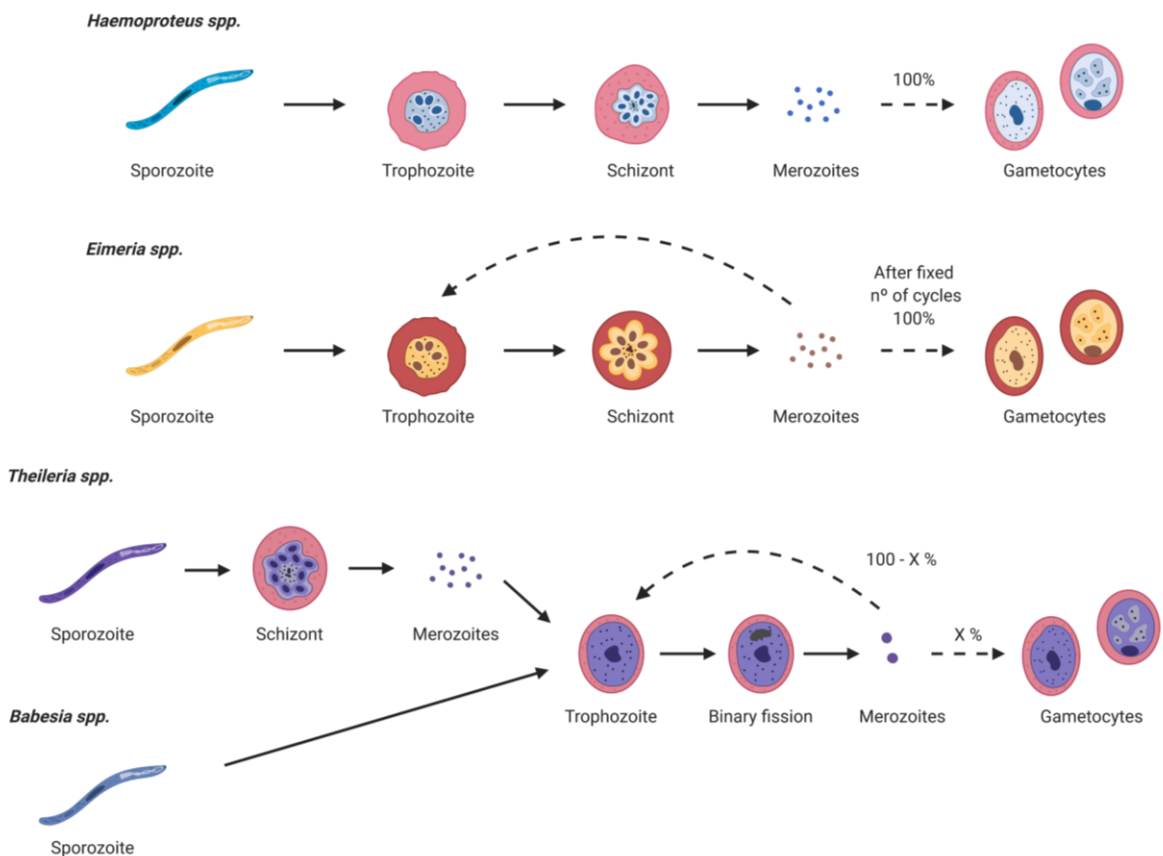
Figure 36. Models of sexual conversion in other apicomplexan parasites. Trophozoites in these parasites refer to the stage preceding nuclear division, which in *Plasmodium spp.* are equivalent to rings and trophozoites. In *Haemoproteus spp.* all

parasites develop into gametocytes after a single round of the IDC. In *Eimeria spp.* parasites infect the gut and grow asexually for a determined number of cycles before developing all into gametocytes. In *Theileria spp.* and *Babesia spp.* parasites can sustain asexual growth through binary fission and at each cycle a small percentage of the population develop into gametocytes. Designed with Biorender.com.

Altogether, the patterns of sexual conversion in other apicomplexan parasites suggest that by default, following commitment, sexual conversion occurs after a round of schizogony (NCC). The emergence of SCC in *Plasmodium spp.* could reflect a by-product of the epigenetic regulation of AP2-G: if the silencing is lost during schizogony or early in ring stages, parasites can convert directly into gametocytes. Nevertheless, as SCC may provide some potential advantages for the parasite, it has been maintained across the genus. Therefore, determining how sexual commitment is regulated in other apicomplexan parasites containing an orthologue for AP2-G, could provide some insight into the origin of the alternative routes of sexual conversion in *Plasmodium spp.*.

5. On the formation of heterochromatin in *P. falciparum*

The specific involvement of facultative heterochromatin in the control of CVGs suggests that these loci are able to initiate heterochromatin formation. We hypothesized about a possible involvement of the underlying DNA sequence in the process and validated that some regions of the *pfap2-g* CDS are able to form heterochromatin when inserted into an euchromatic locus. Nevertheless, the process is highly inefficient, which may reflect that heterochromatin formation at these regions is a very rare event. *De novo* heterochromatin nucleation in *P. falciparum* may be infrequent, perhaps just occurring sporadically

during evolution, but once formed it can be stably maintained across generations. According to the model discussed below, a switch into an active conformation just involves the removal of heterochromatin form the promoter region, which can be later recovered by expansion of heterochromatin from the surrounding regions. We cannot exclude the possibility that the formation of H3K9me3 occurs in a step-wise process, however methylation of H3K9 in *P. falciparum* is catalysed by the action of PfSET3 (or PfKMT1), which is involved in mono- to tri-methylation (Cui et al., 2008; Lopez-Rubio et al., 2009). Therefore, methylation of H3K9 should be a straightforward process that once initiates proceeds all the way to the final state.

Our results point to the CDS of *pfap2-g* as an essential element for the nucleation of heterochromatin at this locus. Moreover, heterochromatin in *pfap2-g* is never dismantled from the CDS (Fraschka et al., 2018; Gómez-Díaz et al., 2017). Therefore, the epigenetic state of the CDS, or nearby regions, in CVGs may act as a form of epigenetic bookmark that defines the locus as a CVG, whereas the epigenetic state of the promoter region typically correlates with their transcriptional activity (Crowley et al., 2011; Lopez-Rubio et al., 2009). The heterochromatin in the CDS, or neighbour regions may act as a reservoir, from where heterochromatin can spread into, or retract from, the promoter region, which leads to a switch in the epigenetic state of the gene (**Figure 37**). We propose that 3 distinct elements define heterochromatic regions in *P. falciparum*: the nucleation site, the reservoir zones and zones where heterochromatin can spread. The regions involved in the nucleation do not necessarily

have to be located within the region acting as reservoir, although in the case of *pfap2-g* the CDS appears to perform both functions.
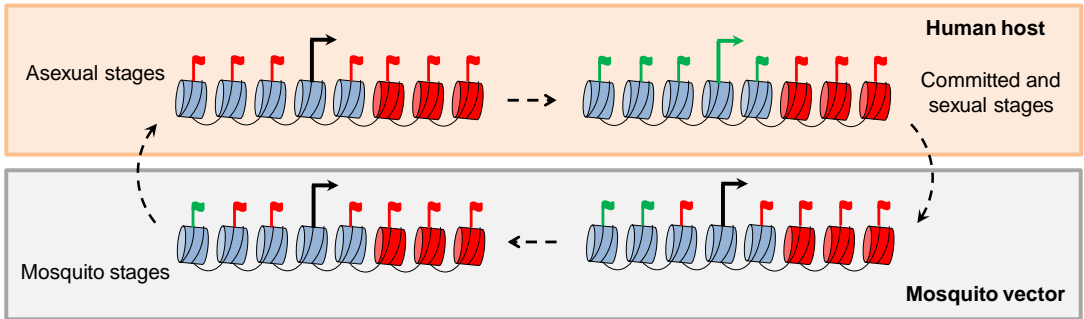
Figure 37. The CDS of *pfap2-g* may act as a reservoir of heterochromatin, allowing the recovery of the silenced state during the infection of the mosquito. Activation of *pfap2-g* just requires the dismantling of heterochromatin at the promoter region. The active state of the locus is characterized by maintaining heterochromatin in the CDS. During the infection of the mosquito, parasites progressively recover the heterochromatin at the promoter region (Gómez-Díaz et al., 2017), allowing the asexual growth of the parasite in the blood. Red histones represent the CDS of the gene. Red flags and green flags represent repressive and activating histone marks, respectively. The arrow shows the position of the TSS.

Genome wide analysis of H3K9me3 occupancy, together with the expression profiles of parasite clones, could provide valuable insight into the distribution of heterochromatin in active and silenced CVGs, validating the hypothesis presented here.

The H3K9me3 ChIP-Seq analysis of the inducible lines suggests that the presence of heterochromatin in the CDS does not interfere with the expression of *pfap2-g*. Therefore, for a switch in the epigenetic state of the locus, dismantling of heterochromatin at the promoter region is sufficient. The presence of intragenic H3K9me3 in active genes has also been identified in other *P. falciparum* CVGs (Lopez-Rubio et al., 2007, 2009) and in other organisms:

- In humans, H3K9me3 is enriched in the 3' exon of Zinc Finger genes and overlaps with H3K36me3, a hallmark of transcription elongation. Given that these genes can be transcribed, the authors suggest that H3K9me3 could provide a more compact chromatin structure that prevents homologous recombination between homologous Zinc Finger genes (Blahnik et al., 2011). The presence of H3K9me3 in the CDS has also been observed in expressed genes of human cancer cell lines (Wiencke et al., 2008).

- In *Drosophila*, the chromosome 4 is highly enriched in transposable elements silenced by heterochromatin, but also in transcribed genes. These transcribed genes are enriched in H3K9me3 in their CDS, which correlates with lack of Pol II pausing (Riddle et al., 2012). Actually, Pol II elongation and H3K9me3/HP1 have been previously described to coexist in actively transcribed genes in humans, but H3K9me3/HP1 are usually lost after transcription (Vakoc et al., 2005).

- In *Arabidopsis*, intragenic transposable elements are silenced by the presence of heterochromatin. However, full length transcripts from these elements can still be produced through the action of a specific demethylase (Saze et al., 2013).

  Overall, these results suggest that presence of heterochromatin in CDSs does not necessarily imply a transcriptionally repressed state. Alternatively, heterochromatin in *P. falciparum* could be less compact than in other organisms, facilitating the progression of the transcription machinery and allowing expression of genes with heterochromatic CDS. For instance, in humans, phosphorylation of the N-terminus of HP1 promotes the formation of phase-separated

droplets. This phenomenon leads to the sequestration of the compact chromatin in droplets, helping to compartmentalize heterochromatin components in cells (Larson and Narlikar, 2018; Larson et al., 2017). Phosphorylation of HP1 in *P. falciparum* has been identified in phosphoproteomic datasets (Collins et al., 2014; Lasonder et al., 2012; Pease et al., 2013; Treeck et al., 2011), but its functional relevance is unknown.

The ectopic formation of heterochromatin at the PF3D7_1144400 locus has provided evidence of heterochromatin spreading in *P. falciparum*, as previously reported (Rovira-Graells et al., 2015; Voss et al., 2006). Barrier insulators have not been identified, but they are predicted to be located in intergenic regions to maintain the epigenetic context without affecting neighbour regions (Ali et al., 2016; Avraham et al., 2012). Our approach may provide some insight into the potential elements involved in limiting the spreading of heterochromatin, but we only observed spreading along the CDS. One possible explanation is based on the presence of an essential gene (encoding for SUMO-activating enzyme subunit 1) located downstream of PF3D7_1144400: in parasites where heterochromatin spreads into the regulatory region of this gene, the silencing would be deleterious. Nevertheless, it does not explain the absence of spreading in the upstream region, as the nearest gene is not essential. These results may suggest that expansion through intergenic regions may be more inefficient than through the CDS, probably due to their AT richness. Further studies should characterize the spreading capacity of heterochromatin in *P. falciparum*, assisting in the identification of barrier insulator elements.

CONCLUSIONS

1. In *P. falciparum*, sexual conversion can occur by two different routes: a direct route consisting on conversion within the same cycle as initial PfAP2-G expression (SCC), or the previously described route, where PfAP2-G-expressing parasites complete a replicative cycle as committed forms before converting into gametocytes after reinvasion (NCC).

2. The SCC route reflects *de novo* PfAP2-G expression in ring stages.

3. By using PfAP2-G as a marker of commitment, we identified sexually committed parasites preceding the previously known committed schizont stage, such as committed trophozoites.

4. We developed a conditional activation system for *pfap2-g* that provides 90% synchronous sexual conversion. To our knowledge, this is the highest sexual conversion rate ever observed. The high yield of the system allowed the characterization of the early events that follow commitment.

5. The inducible system allows the most accurate description to date of the initial transcriptomic changes that occur during sexual development.

6. The high level of purity of our preparations enabled the identification of genes downregulated at the initial phases of sexual development, which was not possible with previous

approaches. The RhopH complex is downregulated in committed schizonts, correlating with the altered solute transport in sexual stages.

7. Sexually committed merozoites do not have a preference for reticulocyte invasion, which suggests that sequestration in the bone marrow is directed by other mechanisms.

8. Expression of PfAP2-G rescues the production of sexual parasites in a gametocyte non-producer line, formally demonstrating that PfAP2-G is sufficient to trigger sexual conversion.

9. The ChIP-seq analysis of the inducible lines provided new insight into the regulation of sexual conversion. The expression of *pfap2-g* in sexually committed and early sexual stages just requires the dismantling of heterochromatin at the promoter region.

10. Some regions of the *pfap2-g* CDS can recapitulate heterochromatin formation into an ectopic location. Nevertheless, it just occurs in a few parasites, which suggests that *de novo* heterochromatin formation in *P. falciparum* may be a rare event.

BIBLIOGRAPHY

Abel, S., and Le Roch, K.G. (2019). The role of epigenetics and chromatin structure in transcriptional regulation in malaria parasites. Brief. Funct. Genomics *00*, 1–12.

Adjalley, S.H., Chabbert, C.D., Klaus, B., Pelechano, V., and Steinmetz, L.M. (2016). Landscape and Dynamics of Transcription Initiation in the Malaria Parasite Plasmodium falciparum. Cell Rep. *14*, 2463–2475.

Adl, S.M., Simpson, A.G.B., Lane, C.E., Lukeš, J., Bass, D., Bowser, S.S., Brown, M.W., Burki, F., Dunthorn, M., Hampl, V., et al. (2012). The revised classification of eukaryotes. J. Eukaryot. Microbiol. *59*, 429–493.

Aguilar, R., Magallon-Tejada, A., Achtman, A.H., Moraleda, C., Joice, R., Cistero, P., Li Wai Suen, C.S.N., Nhabomba, A., Macete, E., Mueller, I., et al. (2014). Molecular evidence for the localization of Plasmodium falciparum immature gametocytes in bone marrow. Blood *123*, 959–966.

Ahmed, A.M., and Hurd, H. (2006). Immune stimulation and malaria infection impose reproductive costs in Anopheles gambiae via follicular apoptosis. Microbes Infect. *8*, 308–315.

Aingaran, M., Zhang, R., Law, S.K., Peng, Z., Undisz, A., Meyer, E., Diez-Silva, M., Burke, T.A., Spielmann, T., Lim, C.T., et al. (2012). Host cell deformability is linked to transmission in the human malaria parasite Plasmodium falciparum. Cell. Microbiol. *14*, 983–993.

Alano, P. (2014). The sound of sexual commitment breaks the silencing of malaria parasites. Trends Parasitol. *30*, 509–510.

Alano, P., Premawansa, S., Bruce, M.C., and Carter, R. (1991). A stage specific gene expressed at the onset of gametocytogenesis in Plasmodium falciparum. Mol. Biochem. Parasitol. *46*, 81–88.

Ali, T., Renkawitz, R., and Bartkuhn, M. (2016). Insulators and domains of gene expression. Curr. Opin. Genet. Dev. *37*, 17–26.

Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. Nat. Rev. Mol. Cell Biol. *19*, 229–244.

Alphey, L. (2016). Can CRISPR-Cas9 gene drives curb malaria? Nat. Biotechnol. *34*, 149–150.

Alphey, L., Beard, C. Ben, Billingsley, P., Coetzee, M., Crisanti, A., Curtis, C., Eggleston, P., Godfray, C., Hemingway, J., Jacobs-Lorena, M., et al. (2002). Malaria Control with Genetically Manipulated Insect Vectors. Science. *298*, 119–121.

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. *112*, E982–E991.

Andersson, R., Sandelin, A., and Danko, C.G. (2015). A unified architecture of transcriptional regulatory elements. Trends Genet. *31*, 426–433.

Aravind, L., Iyer, L.M., Wellems, T.E., and Miller, L.H. (2003).

Plasmodium Biology. Cell *115*, 771–785.

Ashley, E.A., Pyae Phyo, A., and Woodrow, C.J. (2018). Malaria. Lancet *391*, 1608–1621.

Ataíde, R., Mayor, A., and Rogerson, S.J. (2014). Malaria, primigravidae, and antibodies: knowledge gained and future perspectives. Trends Parasitol. *30*, 85–94.

Atkinson, C.T., and van Riper III, C. (1991). Pathogenicity and epizootiology of avian haematozoa: Plasmodium, Haemoproteus, and Leucocytozoon. In Bird-Parasite Interactions: Ecology, Evolution, and Behavior, (London: Oxford University Press), pp. 19–48.

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*, E3678-86.

Ay, F., Bunnik, E.M., Varoquaux, N., Bol, S.M., Prudhomme, J., Vert, J.-P., Noble, W.S., and Le Roch, K.G. (2014). Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. Genome Res. *24*, 974–988.

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

Balaji, S., Madan Babu, M., Iyer, L.M., and Aravind, L. (2005). Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-

integrase DNA binding domains. Nucleic Acids Res. *33*, 3994–4006.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Barry, A.E., and Arnott, A. (2014). Strategies for Designing and Monitoring Malaria Vaccines Targeting Diverse Antigens. Front. Immunol. *5*, 359.

Bártfai, R., Hoeijmakers, W.A.M., Salcedo-Amaya, A.M., Smits, A.H., Janssen-Megens, E., Kaan, A., Treeck, M., Gilberger, T.-W., Françoijs, 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.

Battista, G., Bignami, A., and Bastianelli, G. (1899). Ulteriore ricerche sul ciclo dei parassiti malarici umani sul corpo del zanzarone. Atti R. Accad Lincei *8*, 21–28.

Batugedara, G., Lu, X.M., Bunnik, E.M., and Le Roch, K.G. (2017). The Role of Chromatin Structure in Gene Regulation of the Human Malaria Parasite. Trends Parasitol. *33*, 364–377.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.-W., Green, J.L., Holder, A.A., and Cowman, A.F. (2006). A Conserved Molecular Motor Drives Cell Invasion and Gliding Motility across Malaria Life Cycle Stages and Other Apicomplexan Parasites. J. Biol. Chem. *281*, 5197–5208.

Bechtsi, D.P., and Waters, A.P. (2017). Genomics and epigenetics of

sexual commitment in Plasmodium. Int. J. Parasitol. *47*, 425–434.

Bennink, S., Kiesow, M.J., and Pradel, G. (2016). The development of malaria parasites in the mosquito midgut. Cell. Microbiol. *18*, 905–918.

Bernabeu, M., and Smith, J.D. (2017). EPCR and Malaria Severity: The Center of a Perfect Storm. Trends Parasitol. *33*, 295–308.

Bhatt, S., Weiss, D.J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K.E., Moyes, C.L., Henry, A., Eckhoff, P.A., et al. (2015). The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. Nature *526*, 207–211.

Biljon, R. van, Wyk, R. van, Painter, H., Orchard, L., Reader, J., Niemand, J., Llinas, M., and Birkholtz, L.-M. (2019). Hierarchical transcriptional control regulates Plasmodium falciparum sexual differentiation. BioRxiv 633222.

Billker, O., Lindo, V., Panico, M., Etienne, A.E., Paxton, T., Dell, A., Rogers, M., Sinden, R.E., and Morris, H.R. (1998). Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. Nature *392*, 289–292.

Billker, O., Dechamps, S., Tewari, R., Wenig, G., Franke-Fayard, B., and Brinkmann, V. (2004). Calcium and a Calcium-Dependent Protein Kinase Regulate Gamete Formation and Mosquito Transmission in a Malaria Parasite. Cell *117*, 503–514.

Bischoff, E., and Vaquero, C. (2010). In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of Plasmodium

falciparum. BMC Genomics *11*, 34.

Blahnik, K.R., Dou, L., Echipare, L., Iyengar, S., O'Geen, H., Sanchez, E., Zhao, Y., Marra, M.A., Hirst, M., Costello, J.F., et al. (2011). Characterization of the Contradictory Chromatin Signatures at the 3' Exons of Zinc Finger Genes. PLoS One *6*, e17121.

Blandin, S., Shiao, S.-H., Moita, L.F., Janse, C.J., Waters, A.P., Kafatos, F.C., and Levashina, E.A. (2004). Complement-Like Protein TEP1 Is a Determinant of Vectorial Capacity in the Malaria Vector Anopheles gambiae. Cell *116*, 661–670.

Bobenchik, A.M., Witola, W.H., Augagneur, Y., Nic Lochlainn, L., Garg, A., Pachikara, N., Choi, J.-Y., Zhao, Y.O., Usmani-Brown, S., Lee, A., et al. (2013). Plasmodium falciparum phosphoethanolamine methyltransferase is essential for malaria transmission. Proc. Natl. Acad. Sci. *110*, 18262–18267.

Bozdech, Z., Llinás, 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*, 85–100.

Brancucci, N.M.B., Bertschi, N.L., Zhu, L., Niederwieser, I., Chin, W.H., Wampfler, R., Freymond, C., Rottmann, M., Felger, I., Bozdech, Z., et al. (2014). Heterochromatin Protein 1 Secures Survival and Transmission of Malaria Parasites. Cell Host Microbe *16*, 165–176.

Brancucci, N.M.B., Goldowitz, I., Buchholz, K., Werling, K., and Marti, M. (2015). An assay to probe Plasmodium falciparum growth, transmission stage formation and early gametocyte development. Nat.

Protoc. *10*, 1131–1142.

Brancucci, N.M.B., Gerdt, J.P., Wang, C., De Niz, M., Philip, N., Adapa, S.R., Zhang, M., Hitz, E., Niederwieser, I., Boltryk, S.D., et al. (2017). Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite Plasmodium falciparum. Cell *171*, 1532-1544.e15.

Brancucci, N.M.B., De Niz, M., Straub, T.J., Ravel, D., Sollelis, L., Birren, B.W., Voss, T.S., Neafsey, D.E., and Marti, M. (2018). Probing Plasmodium falciparum sexual commitment at the single-cell level. Wellcome Open Res. *3*, 70.

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

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

Buckling, A., Ranford-Cartwright, L.C., Miles, A., and Read, A.F. (1999). Chloroquine increases Plasmodium falciparum gametocytogenesis in vitro. Parasitology *118 Pt 4*, 339–346.

Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., de la Rosa-Velazquez, I.A., van de Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M., et al. (2012). A transcription factor–based mechanism for mouse heterochromatin formation. Nat. Struct. Mol. Biol. *19*, 1023–1030.

Bunnik, E.M., Polishko, A., Prudhomme, J., Ponts, N., Gill, S.S., Lonardi, S., and Le Roch, K.G. (2014). DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite Plasmodium falciparum. BMC Genomics *15*, 347.

Bunnik, E.M., Cook, K.B., Varoquaux, N., Batugedara, G., Prudhomme, J., Cort, A., Shi, L., Andolina, C., Ross, L.S., Brady, D., et al. (2018). Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. Nat. Commun. *9*, 1910.

Bunnik, E.M., Venkat, A., Shao, J., McGovern, K.E., Batugedara, G., Worth, D., Prudhomme, J., Lapp, S.A., Andolina, C., Ross, L.S., et al. (2019). Comparative 3D genome organization in apicomplexan parasites. Proc. Natl. Acad. Sci. *116*, 3183–3192.

Bury-Moné, S., and Sclavi, B. (2017). Stochasticity of gene expression as a motor of epigenetics in bacteria: from individual to collective behaviors. Res. Microbiol. *168*, 503–514.

Callebaut, I., Prat, K., Meurice, E., Mornon, J.-P., and Tomavo, S. (2005). Prediction of the general transcription factors associated with RNA polymerase II in Plasmodium falciparum: conserved features and differences relative to other eukaryotes. BMC Genomics *6*, 100.

Campbell, T.L., De Silva, E.K., Olszewski, K.L., Elemento, O., and Llinás, M. (2010). Identification and Genome-Wide Prediction of DNA Binding Specificities for the ApiAP2 Family of Regulators from the Malaria Parasite. PLoS Pathog. *6*, e1001165.

Carey, J.N., Mettert, E.L., Roggiani, M., Myers, K.S., Kiley, P.J., and

Goulian, M. (2018). Regulated Stochasticity in a Bacterial Signaling Network Permits Tolerance to a Rapid Environmental Change. Cell *173*, 196-207.e14.

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

Casadesús, J., and Low, D. a. (2013). Programmed heterogeneity: Epigenetic mechanisms in bacteria. J. Biol. Chem. *288*, 13929–13935.

Chen, P.B., Ding, S., Zanghì, G., Soulard, V., DiMaggio, P.A., Fuchter, M.J., Mecheri, S., Mazier, D., Scherf, A., and Malmquist, N.A. (2016). Plasmodium falciparum PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites. Sci. Rep. *6*, 21802.

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. *104*, 899–902.

Ciabrelli, F., and Cavalli, G. (2015). Chromatin-driven behavior of topologically associating domains. J. Mol. Biol. *427*, 608–625.

Claessens, A., Affara, M., Assefa, S.A., Kwiatkowski, D.P., and Conway, D.J. (2017). Culture adaptation of malaria parasites selects for convergent loss-of-function mutants. Sci. Rep. *7*, 41303.

Cobbold, S.A., Santos, J.M., Ochoa, A., Perlman, D.H., and Llinás, M. (2016). Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. Sci. Rep. *6*, 19722.

Coetzee, N., Sidoli, S., van Biljon, R., Painter, H., Llinás, M., Garcia, B.A., and Birkholtz, L.-M. (2017). Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual Plasmodium falciparum parasites. Sci. Rep. *7*, 607.

Coleman, B.I., Skillman, K.M., Jiang, R.H.Y., Childs, L.M., Altenhofen, L.M., Ganter, M., Leung, Y., Goldowitz, I., Kafsack, B.F.C., Marti, M., et al. (2014). A Plasmodium falciparum histone deacetylase regulates antigenic variation and gametocyte conversion. Cell Host Microbe *16*, 177–186.

Collins, M.O., Wright, J.C., Jones, M., Rayner, J.C., and Choudhary, J.S. (2014). Confident and sensitive phosphoproteomics using combinations of collision induced dissociation and electron transfer dissociation. J. Proteomics *103*, 1–14.

Cortés, A., and Deitsch, K.W. (2017). Malaria epigenetics. Cold Spring Harb. Perspect. Med. *7*, 1–23.

Cortés, A., Carret, C., Kaneko, O., Yim Lim, B.Y.S., Ivens, A., and Holder, A.A. (2007). Epigenetic Silencing of Plasmodium falciparum Genes Linked to Erythrocyte Invasion. PLoS Pathog. *3*, e107.

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.

Coulson, R.M.R. (2004). Comparative Genomics of Transcriptional Control in the Human Malaria Parasite Plasmodium falciparum. Genome Res. *14*, 1548–1554.

Cowman, A.F., Healer, J., Marapana, D., and Marsh, K. (2016). Malaria: Biology and Disease. Cell *167*, 610–624.

Cowman, A.F., Tonkin, C.J., Tham, W.-H., and Duraisingh, M.T. (2017). The Molecular Basis of Erythrocyte Invasion by Malaria Parasites. Cell Host Microbe *22*, 232–245.

Cox, F.E. (2010). History of the discovery of the malaria parasites and their vectors. Parasit. Vectors *3*, 5.

Crick, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561–563.

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.

Cui, L., and Miao, J. (2010). Chromatin-Mediated epigenetic regulation in the malaria parasite Plasmodium falciparum. Eukaryot. Cell *9*, 1138–1149.

Cui, L., Miao, J., Furuya, T., Li, X., Su, X.Z., and Cui, L. (2007). PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in Plasmodium falciparum. Eukaryot. Cell *6*, 1219–1227.

Cui, L., Fan, Q., Cui, L., and Miao, J. (2008). Histone lysine methyltransferases and demethylases in Plasmodium falciparum. Int. J. Parasitol. *38*, 1083–1097.

Dearnley, M., Chu, T., Zhang, Y., Looker, O., Huang, C., Klonis, N.,

Yeoman, J., Kenny, S., Arora, M., Osborne, J.M., et al. (2016). Reversible host cell remodeling underpins deformability changes in malaria parasite sexual blood stages. Proc. Natl. Acad. Sci. *113*, 4800–4805.

Deitsch, K.W., and Dzikowski, R. (2017). Variant Gene Expression and Antigenic Variation by Malaria Parasites. Annu. Rev. Microbiol. *71*, 625–641.

Delandre, C., and Marshall, O.J. (2019). United colours of chromatin? Developmental genome organisation in flies. Biochem. Soc. Trans. *47*, 691–700.

Delves, M.J., Straschil, U., Ruecker, A., Miguel-Blanco, C., Marques, S., Baum, J., and Sinden, R.E. (2016). Routine in vitro culture of P. falciparum gametocytes to evaluate novel transmission-blocking interventions. Nat. Protoc. *11*, 1668–1680.

DiCarlo, J.E., Chavez, A., Dietz, S.L., Esvelt, K.M., and Church, G.M. (2015). Safeguarding CRISPR-Cas9 gene drives in yeast. Nat. Biotechnol. *33*, 1250–1255.

Dobaño, C., Santano, R., Vidal, M., Jiménez, A., Jairoce, C., Ubillos, I., Dosoo, D., Aguilar, R., Williams, N.A., Díez-Padrisa, N., et al. (2019). Differential Patterns of IgG Subclass Responses to Plasmodium falciparum Antigens in Relation to Malaria Protection and RTS,S Vaccination. Front. Immunol. *10*, 439.

Dong, Y., Aguilar, R., Xi, Z., Warr, E., Mongin, E., and Dimopoulos, G. (2006). Anopheles gambiae Immune Responses to Human and Rodent Plasmodium Parasite Species. PLoS Pathog. *2*, e52.

Doolan, D.L., Dobano, C., and Baird, J.K. (2009). Acquired Immunity to Malaria. Clin. Microbiol. Rev. *22*, 13–36.

Drakeley, C., Sutherland, C., Bousema, J.T., Sauerwein, R.W., and Targett, G.A.T. (2006). The epidemiology of Plasmodium falciparum gametocytes: weapons of mass dispersion. Trends Parasitol. *22*, 424–430.

Draper, S.J., Sack, B.K., King, C.R., Nielsen, C.M., Rayner, J.C., Higgins, M.K., Long, C.A., and Seder, R.A. (2018). Malaria Vaccines: Recent Advances and New Horizons. Cell Host Microbe *24*, 43–56.

Duffy, S., Loganathan, S., Holleran, J.P., and Avery, V.M. (2016). Large-scale production of Plasmodium falciparum gametocytes for malaria drug discovery. Nat. Protoc. *11*, 976–992.

Duraisingh, M.T., and Skillman, K.M. (2018). Epigenetic Variation and Regulation in Malaria Parasites. Annu. Rev. Microbiol. *72*, annurev-micro-090817-062722.

Duraisingh, M.T., Triglia, T., Ralph, S.A., Rayner, J.C., Barnwell, J.W., McFadden, G.I., and Cowman, A.F. (2003). Phenotypic variation of Plasmodium falciparum merozoite proteins directs receptor targeting for invasion of human erythrocytes. EMBO J. *22*, 1047–1057.

Duraisingh, M.T., Voss, T.S., Marty, A.J., Duffy, M.F., Good, R.T., Thompson, J.K., Freitas-Junior, L.H., Scherf, A., Crabb, B.S., and Cowman, A.F. (2005). Heterochromatin Silencing and Locus Repositioning Linked to Regulation of Virulence Genes in Plasmodium falciparum. Cell *121*, 13–24.

Dyer, M., and Day, K.P. (2003). Regulation of the rate of asexual growth and commitment to sexual development by diffusible factors from in vitro cultures of Plasmodium falciparum. Am. J. Trop. Med. Hyg. *68*, 403–409.

Eksi, S., Haile, Y., Furuya, T., Ma, L., Su, X., and Williamson, K.C. (2005). Identification of a subtelomeric gene family expressed during the asexual-sexual stage transition in Plasmodium falciparum. Mol. Biochem. Parasitol. *143*, 90–99.

Eksi, S., Morahan, B.J., Haile, Y., Furuya, T., Jiang, H., Ali, O., Xu, H., Kiattibutr, K., Suri, A., Czesny, B., et al. (2012). Plasmodium falciparum Gametocyte Development 1 (Pfgdv1) and Gametocytogenesis Early Gene Identification and Commitment to Sexual Development. PLoS Pathog. *8*, e1002964.

Even-Faitelson, L., Hassan-Zadeh, V., Baghestani, Z., and Bazett-Jones, D.P. (2016). Coming to terms with chromatin structure. Chromosoma *125*, 95–110.

Fan, Q., An, L., and Cui, L. (2004). Plasmodium falciparum Histone Acetyltransferase, a Yeast GCN5 Homologue Involved in Chromatin Remodeling. Eukaryot. Cell *3*, 264–276.

Farfour, E., Charlotte, F., Settegrana, C., Miyara, M., and Buffet, P. (2012). The extravascular compartment of the bone marrow: A niche for Plasmodium falciparum gametocyte maturation? Malar. J. *11*, 2–5.

Fastman, Y., Noble, R., Recker, M., and Dzikowski, R. (2012). Erasing the Epigenetic Memory and Beginning to Switch—The Onset of Antigenic Switching of var Genes in Plasmodium falciparum. PLoS

One *7*, e34168.

Filarsky, M., Fraschka, S.A., Niederwieser, I., Brancucci, N.M.B., Carrington, E., Carrió, E., Moes, S., Jenoe, P., Bártfai, R., and Voss, T.S. (2018). GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. Science. *359*, 1259–1263.

Fivelman, Q.L., McRobert, L., Sharp, S., Taylor, C.J., Saeed, M., Swales, C.A., Sutherland, C.J., and Baker, D.A. (2007). Improved synchronous production of Plasmodium falciparum gametocytes in vitro. Mol. Biochem. Parasitol. *154*, 119–123.

Florens, L., Muster, N., Tabb, D.L., Yates, J.R., Washburn, M.P., Wolters, D., Raine, J.D., Sinden, R.E., Anthony, R.M., Haynes, J.D., et al. (2002). A proteomic view of the Plasmodium falciparum life cycle. Nature *419*, 520–526.

Flueck, C., Bartfai, R., Volz, J., Niederwieser, I., Salcedo-Amaya, A.M., Alako, B.T.F., 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.

Flueck, C., Bartfai, R., Niederwieser, I., Witmer, K., Alako, B.T.F., Moes, S., Bozdech, Z., Jenoe, P., Stunnenberg, H.G., and Voss, T.S. (2010). A Major Role for the Plasmodium falciparum ApiAP2 Protein PfSIP2 in Chromosome End Biology. PLoS Pathog. *6*, e1000784.

Foy, B.D., Kobylinski, K.C., Silva, I.M. da, Rasgon, J.L., and Sylla, M. (2011). Endectocides for malaria control. Trends Parasitol. *27*, 423–428.

Foy, B.D., Alout, H., Seaman, J.A., Rao, S., Magalhaes, T., Wade, M., Parikh, S., Soma, D.D., Sagna, A.B., Fournet, F., et al. (2019). Efficacy and risk of harms of repeat ivermectin mass drug administrations for control of malaria (RIMDAMAL): a cluster-randomised trial. Lancet *393*, 1517–1526.

Francia, M.E., and Striepen, B. (2014). Cell division in apicomplexan parasites. Nat. Rev. Microbiol. *12*, 125–136.

Frank, M., Dzikowski, R., Amulic, B., and Deitsch, K. (2007). Variable switching rates of malaria virulence genes are associated with chromosomal position. Mol. Microbiol. *64*, 1486–1498.

Fraschka, S.A., Filarsky, M., Hoo, R., Niederwieser, I., Yam, X.Y., Brancucci, N.M.B., Mohring, F., Mushunje, A.T., Huang, X., Christensen, P.R., et al. (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.

Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature *407*, 1018–1022.

Freitas-Junior, L.H., Hernandez-Rivas, R., Ralph, S. a., Montiel-Condado, D., Ruvalcaba-Salazar, O.K., Rojas-Meza, A.P., Mâncio-Silva, L., Leal-Silvestre, R.J., Gontijo, A.M., Shorte, S., et al. (2005). Telomeric Heterochromatin Propagation and Histone Acetylation Control Mutually Exclusive Expression of Antigenic Variation Genes

in Malaria Parasites. Cell *121*, 25–36.

Garcia, G.E., Wirtz, R.A., Barr, J.R., Woolfitt, A., and Rosenberg, R. (1998). Xanthurenic Acid Induces Gametogenesis in Plasmodium , the Malaria Parasite. J. Biol. Chem. *273*, 12003–12005.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., et al. (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. Nature *419*, 498–511.

Gerald, N., Mahajan, B., and Kumar, S. (2011). Mitosis in the human malaria parasite plasmodium falciparum. Eukaryot. Cell *10*, 474–482.

Goel, S., Palmkvist, M., Moll, K., Joannin, N., Lara, P., R Akhouri, R., Moradi, N., Öjemalm, K., Westman, M., Angeletti, D., et al. (2015). RIFINs are adhesins implicated in severe Plasmodium falciparum malaria. Nat. Med. *21*, 314–317.

Gómez-Díaz, E., Yerbanga, R.S., Lefèvre, T., Cohuet, A., Rowley, M.J., Ouedraogo, J.B., Corces, V.G., Bozdech, Z., Roch, K.G. Le, Llinás, M., et al. (2017). Epigenetic regulation of Plasmodium falciparum clonally variant gene expression during development in Anopheles gambiae. Sci. Rep. *7*, 40655.

Gouagna, L.C., Bancone, G., Yao, F., Yameogo, B., Dabiré, K.R., Costantini, C., Simporé, J., Ouedraogo, J.B., and Modiano, D. (2010). Genetic variation in human HBB is associated with Plasmodium falciparum transmission. Nat. Genet. *42*, 328–331.

Greenstein, R.A., Jones, S.K., Spivey, E.C., Rybarski, J.R., Finkelstein,

I.J., and Al-Sady, B. (2018). Noncoding RNA-nucleated heterochromatin spreading is intrinsically labile and requires accessory elements for epigenetic stability. Elife *7*, 1–33.

Grellier, P. (1991). Lipid traffic between high density lipoproteins and Plasmodium falciparum-infected red blood cells. J. Cell Biol. *112*, 267–277.

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

Gulati, S., Ekland, E.H., Ruggles, K. V., Chan, R.B., Jayabalasingham, B., Zhou, B., Mantel, P.-Y., Lee, M.C.S., Spottiswoode, N., Coburn-Flynn, O., et al. (2015). Profiling the Essential Nature of Lipid Metabolism in Asexual Blood and Gametocyte Stages of Plasmodium falciparum. Cell Host Microbe *18*, 371–381.

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.

Haberle, V., and Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. Nat. Rev. Mol. Cell Biol. *19*, 621–637.

Herman, J.J., Spencer, H.G., Donohue, K., and Sultan, S.E. (2014). How stable 'should' epigenetic modifications be? Insights from

adaptive plasticity and bet hedging. Evolution. *68*, 632–643.

Hoeijmakers, W.A.M., Salcedo-Amaya, A.M., Smits, A.H., Françoijs, K.-J., Treeck, M., Gilberger, T.-W., Stunnenberg, H.G., and Bártfai, R. (2013). H2A.Z/H2B.Z double-variant nucleosomes inhabit the AT-rich promoter regions of the P lasmodium falciparum genome. Mol. Microbiol. *87*, 1061–1073.

Hoffman, S.L., Vekemans, J., Richie, T.L., and Duffy, P.E. (2015). The March Toward Malaria Vaccines. Am. J. Prev. Med. *49*, S319–S333.

Hogg, J.C., and Hurd, H. (1997). The effects of natural Plasmodium falciparum infection on the fecundity and mortality of Anopheles gambiae s. l. in north east Tanzania. Parasitology *114 Pt 4*, 325–331.

Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. Nat. Rev. Genet. *16*, 71–84.

Horrocks, P., Pinches, R., Christodoulou, Z., Kyes, S.A., Newbold, C.I., Horrocks, P., and Pinches, R. (2004). Variable var transition rates underlie antigenic variation in malaria. Proc. Natl. Acad. Sci. U. S. A. *101*, 11129–11134.

Horrocks, P., Wong, E., Russell, K., and Emes, R.D. (2009). Control of gene expression in Plasmodium falciparum - Ten years on. Mol. Biochem. Parasitol. *164*, 9–25.

Ikadai, H., Shaw Saliba, K., Kanzok, S.M., McLean, K.J., Tanaka, T.Q., Cao, J., Williamson, K.C., and Jacobs-Lorena, M. (2013). Transposon mutagenesis identifies genes essential for Plasmodium falciparum gametocytogenesis. Proc. Natl. Acad. Sci. *110*, E1676–

E1684.

Ito, D., Schureck, M.A., and Desai, S.A. (2017). An essential dual-function complex mediates erythrocyte invasion and channel-mediated nutrient uptake in malaria parasites. Elife *6*, 1–24.

Iwanaga, S., Kaneko, I., Kato, T., and Yuda, M. (2012). Identification of an AP2-family Protein That Is Critical for Malaria Liver Stage Development. PLoS One *7*, e47557.

Iyer, L.M., Anantharaman, V., Wolf, M.Y., and Aravind, L. (2008). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. Int. J. Parasitol. *38*, 1–31.

Jalovecka, M., Hajdusek, O., Sojka, D., Kopacek, P., and Malandrin, L. (2018). The Complexity of Piroplasms Life Cycles. Front. Cell. Infect. Microbiol. *8*, 248.

Jeninga, M., Quinn, J., and Petter, M. (2019). ApiAP2 Transcription Factors in Apicomplexan Parasites. Pathogens *8*, 47.

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. *107*, 2224–2229.

Jiang, L., Mu, J., Zhang, Q., Ni, T., Srinivasan, P., Rayavara, K., Yang, W., Turner, L., Lavstsen, T., Theander, T.G., et al. (2013). PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. Nature *499*, 223–227.

Johnson, W.L., and Straight, A.F. (2017). RNA-mediated regulation of

heterochromatin. Curr. Opin. Cell Biol. *46*, 102–109.

Johnson, W.L., Yewdell, W.T., Bell, J.C., McNulty, S.M., Duda, Z., O'Neill, R.J., Sullivan, B.A., and Straight, A.F. (2017). RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. Elife *6*, 1–32.

Joice, R., Nilsson, S.K., Montgomery, J., Dankwa, S., Egan, E., Morahan, B., Seydel, K.B., Bertuccini, L., Alano, P., Williamson, K.C., et al. (2014). Plasmodium falciparum transmission stages accumulate in the human bone marrow. Sci. Transl. Med. *6*, 244re5.

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

Josling, G.A., Petter, M., Oehring, S.C., Gupta, A.P., Dietz, O., Wilson, D.W., Schubert, T., Längst, G., Gilson, P.R., Crabb, B.S., et al. (2015). A Plasmodium Falciparum Bromodomain Protein Regulates Invasion Gene Expression. Cell Host Microbe *17*, 741–751.

Josling, G.A., Williamson, K.C., and Llinás, M. (2018). Regulation of Sexual Commitment and Gametocytogenesis in Malaria Parasites. Annu. Rev. Microbiol. *72*, 501–519.

Josling, G.A., Venezia, J., Orchard, L., Russell, T.J., Painter, H.J., and Llinas, M. (2019). Regulation of sexual differentiation is linked to invasion in malaria parasites. BioRxiv 533877.

Kafsack, B.F.C., Rovira-Graells, N., Clark, T.G., Bancells, C., Crowley, V.M., Campino, S.G., Williams, A.E., Drought, L.G., Kwiatkowski,

D.P., Baker, D., et al. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. Nature *507*, 248–252.

Kaidoh, T., Nath, J., Okoye, V., and Aikawa, M. (1993). Novel structure in the pellicular complex of Plasmodium falciparum gametocytes. J. Eukaryot. Microbiol. *40*, 269–271.

Kaneko, O., Lim, B.Y.S.Y., Iriko, H., Ling, I.T., Otsuki, H., Grainger, M., Tsuboi, T., Adams, J.H., Mattei, D., Holder, A.A., et al. (2005). Apical expression of three RhopH1/Clag proteins as components of the Plasmodium falciparum RhopH complex. Mol. Biochem. Parasitol. *143*, 20–28.

Kaur, J., and Hora, R. (2018). '2TM proteins': an antigenically diverse superfamily with variable functions and export pathways. PeerJ *6*, e4757.

Kaviratne, M., Khan, S.M., Jarra, W., and Preiser, P.R. (2002). Small Variant STEVOR Antigen Is Uniquely Located within Maurer's Clefts in Plasmodium falciparum -Infected Red Blood Cells. Eukaryot. Cell *1*, 926–935.

Kawamoto, F., Alejo-Blanco, R., Fleck, S.L., and Sinden, R.E. (1991). Plasmodium berghei: ionic regulation and the induction of gametogenesis. Exp. Parasitol. *72*, 33–42.

Ke, H., Lewis, I.A., Morrisey, J.M., McLean, K.J., Ganesan, S.M., Painter, H.J., Mather, M.W., Jacobs-Lorena, M., Llinás, M., and Vaidya, A.B. (2015). Genetic Investigation of Tricarboxylic Acid Metabolism during the Plasmodium falciparum Life Cycle. Cell Rep.

11, 164–174.

Kensche, P.R., Hoeijmakers, W.A.M., Toenhake, C.G., Bras, M., Chappell, L., Berriman, M., and Bártfai, R. (2016). The nucleosome landscape of Plasmodium falciparum reveals chromatin architecture and dynamics of regulatory sequences. Nucleic Acids Res. *44*, 2110–2124.

Kent, R.S., Modrzynska, K.K., Cameron, R., Philip, N., Billker, O., and Waters, A.P. (2018). Inducible developmental reprogramming redefines commitment to sexual development in the malaria parasite Plasmodium berghei. Nat. Microbiol. *3*, 1206–1213.

Khan, S.M., Franke-Fayard, B., Mair, G.R., Lasonder, E., Janse, C.J., Mann, M., and Waters, A.P. (2005). Proteome Analysis of Separated Male and Female Gametocytes Reveals Novel Sex-Specific Plasmodium Biology. Cell *121*, 675–687.

Kishore, S.P., Stiller, J.W., and Deitsch, K.W. (2013). Horizontal gene transfer of epigenetic machinery and evolution of parasitism in the malaria parasite Plasmodium falciparum and other apicomplexans. BMC Evol. Biol. *13*, 37.

Komaki-Yasuda, K., Okuwaki, M., Nagata, K., Kawazu, S., and Kano, S. (2013). Identification of a Novel and Unique Transcription Factor in the Intraerythrocytic Stage of Plasmodium falciparum. PLoS One *8*, e74701.

Kongkasuriyachai, D., Fujioka, H., and Kumar, N. (2004). Functional analysis of Plasmodium falciparum parasitophorous vacuole membrane protein (Pfs16) during gametocytogenesis and

gametogenesis by targeted gene disruption. Mol. Biochem. Parasitol. *133*, 275–285.

Kumar, V., Behl, A., Sharma, R., Sharma, A., and Hora, R. (2019). Plasmodium helical interspersed subtelomeric family-an enigmatic piece of the Plasmodium biology puzzle. Parasitol. Res. PMID 31418110.

Kyes, S.A., Rowe, J.A., Kriek, N., and Newbold, C.I. (1999). Rifins: A second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. Proc. Natl. Acad. Sci. *96*, 9333–9338.

Lakshmanan, V., Rhee, K.Y., Wang, W., Yu, Y., Khafizov, K., Fiser, A., Wu, P., Ndir, O., Mboup, S., Ndiaye, D., et al. (2012). Metabolomic Analysis of Patient Plasma Yields Evidence of Plant-Like α-Linolenic Acid Metabolism in Plasmodium falciparum. J. Infect. Dis. *206*, 238–248.

Lamour, S.D., Straschil, U., Saric, J., and Delves, M.J. (2014). Changes in metabolic phenotypes of Plasmodium falciparum in vitro cultures during gametocyte development. Malar. J. *13*, 468.

Larson, A.G., and Narlikar, G.J. (2018). The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation. Biochemistry *57*, 2540–2548.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature *547*, 236–240.

Lasonder, E., Treeck, M., Alam, M., and Tobin, A.B. (2012). Insights into the Plasmodium falciparum schizont phospho-proteome. Microbes Infect. *14*, 811–819.

Lavazec, C., Sanyal, S., and Templeton, T.J. (2007). Expression switching in the stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. Mol. Microbiol. *64*, 1621–1634.

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet. *32*, 42–56.

Lesage, K.M., Huot, L., Mouveaux, T., Courjol, F., Saliou, J., and Gissot, M. (2018). Cooperative binding of ApiAP2 transcription factors is crucial for the expression of virulence genes in Toxoplasma gondii. Nucleic Acids Res. *46*, 6057–6068.

Levy, S.F., Ziv, N., and Siegal, M.L. (2012). Bet Hedging in Yeast by Heterogeneous, Age-Correlated Expression of a Stress Protectant. PLoS Biol. *10*, e1001325.

Liehl, P., and Mota, M.M. (2012). Innate recognition of malarial parasites by mammalian hosts. Int. J. Parasitol. *42*, 557–566.

Lim, C., Dankwa, S., Paul, A.S., and Duraisingh, M.T. (2017). Host Cell Tropism and Adaptation of Blood-Stage Malaria Parasites: Challenges for Malaria Elimination. Cold Spring Harb. Perspect. Med. *7*, a025494.

Llinás, M. (2017). Less Lipid, More Commitment. Cell *171*, 1474–1476.

Lobo, C.A., Fujioka, H., Aikawa, M., and Kumar, N. (1999). Disruption of the Pfg27 locus by homologous recombination leads to loss of the sexual phenotype in P. falciparum. Mol. Cell *3*, 793–798.

Lobo, N.F., Achee, N.L., Greico, J., and Collins, F.H. (2018). Modern Vector Control. Cold Spring Harb. Perspect. Med. *8*, a025643.

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.

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.

Lu, X.M., Batugedara, G., Lee, M., Prudhomme, J., Bunnik, E.M., and Le Roch, K.G. (2017). Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite Plasmodium falciparum. Nucleic Acids Res. *45*, 7825–7840.

MacRae, J.I., Dixon, M.W., Dearnley, M.K., Chua, H.H., Chambers, J.M., Kenny, S., Bottova, I., Tilley, L., and McConville, M.J. (2013). Mitochondrial metabolism of sexual and asexual blood stages of the malaria parasite Plasmodium falciparum. BMC Biol. *11*, 67.

Mair, G.R. (2006). Regulation of Sexual Development of Plasmodium by Translational Repression. Science. *313*, 667–669.

Maison, C., Bailly, D., Peters, A.H.F.M., Quivy, J.-P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. Nat. Genet. *30*, 329–334.

Mantel, P.Y., Hoang, A.N., Goldowitz, I., Potashnikova, D., Hamza, B., Vorobjev, I., Ghiran, I., Toner, M., Irimia, D., Ivanov, A.R., et al. (2013). Malaria-infected erythrocyte-derived microvesicles mediate cellular communication within the parasite population and with the host immune system. Cell Host Microbe *13*, 521–534.

Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. Nat. Rev. Genet. *11*, 285–296.

Marsh, K., Forster, D., Waruiru, C., Mwangi, I., Winstanley, M., Marsh, V., Newton, C., Winstanley, P., Warn, P., Peshu, N., et al. (1995). Indicators of Life-Threatening Malaria in African Children. N. Engl. J. Med. *332*, 1399–1404.

Martins, R.M., Macpherson, C.R., Claes, A., Scheidig-Benatar, C., Sakamoto, H., Yam, X.Y., Preiser, P., Goel, S., Wahlgren, M., Sismeiro, O., et al. (2017). An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite Plasmodium falciparum. Sci. Rep. *7*, 14042.

McCulloch, R., and Field, M.C. (2015). Quantitative sequencing confirms VSG diversity as central to immune evasion by Trypanosoma brucei. Trends Parasitol. *31*, 346–349.

Meerstein-Kessel, L., van der Lee, R., Stone, W., Lanke, K., Baker,

D.A., Alano, P., Silvestrini, F., Janse, C.J., Khan, S.M., van de Vegte-Bolmer, M., et al. (2018). Probabilistic data integration identifies reliable gametocyte-specific proteins and transcripts in malaria parasites. Sci. Rep. *8*, 410.

Meibalan, E., and Marti, M. (2016). Biology of Malaria Transmission. Cold Spring Harb. Perspect. Med. a025452.

Merrick, C.J., Jiang, R.H.Y., 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.

Mi-Ichi, F., Kita, K., and Mitamura, T. (2006). Intraerythrocytic Plasmodium falciparum utilize a broad range of serum-derived fatty acids with limited modification for their growth. Parasitology *133*, 399–410.

Miao, J., Fan, Q., Cui, L., Li, J., Li, J., and Cui, L. (2006). The malaria parasite Plasmodium falciparum histones: Organization, expression, and acetylation. Gene *369*, 53–65.

Miao, J., Fan, Q., Cui, L., Li, X., Wang, H., Ning, G., Reese, J.C., and Cui, L. (2010a). The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite Plasmodium falciparum. Mol. Microbiol. *78*, 883–902.

Miao, J., Li, J., Fan, Q., Li, X., Li, X., and Cui, L. (2010b). The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite Plasmodium falciparum. J. Cell Sci. *123*, 1039–1049.

Miao, J., Fan, Q., Parker, D., Li, X., Li, J., and Cui, L. (2013). Puf Mediates Translation Repression of Transmission-Blocking Vaccine Candidates in Malaria Parasites. PLoS Pathog. *9*, e1003268.

Miller, L.H., Baruch, D.I., Marsh, K., and Doumbo, O.K. (2002). The pathogenic basis of malaria. Nature *415*, 673–679.

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.

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, G.P., et al. (2017). Expression of the Plasmodium falciparum Clonally Variant clag3 Genes in Human Infections. J. Infect. Dis. *215*, 938–945.

Mira-Martínez, S., Pickford, A.K., Rovira-Graells, N., Guetens, P., Tintó-Font, E., Cortés, A., and Rosanas-Urgell, A. (2019). Identification of Antimalarial Compounds That Require CLAG3 for Their Uptake by Plasmodium falciparum -Infected Erythrocytes. Antimicrob. Agents Chemother. *63*, 1–14.

Mitchell, S.N., and Catteruccia, F. (2017). Anopheline Reproductive Biology: Impacts on Vectorial Capacity and Potential Avenues for Malaria Control. Cold Spring Harb. Perspect. Med. *7*, a025593.

Moazed, D. (2011). Mechanisms for the inheritance of chromatin states. Cell *146*, 510–518.

Modrzynska, K., Pfander, C., Chappell, L., Yu, L., Suarez, C., Dundas, K., Gomes, A.R., Goulding, D., Rayner, J.C., Choudhary, J., et al. (2017). A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle. Cell Host Microbe *21*, 11–22.

Molina-Cruz, A., Garver, L.S., Alabaster, A., Bangiolo, L., Haile, A., Winikor, J., Ortega, C., van Schaijk, B.C.L., Sauerwein, R.W., Taylor-Salmon, E., et al. (2013). The Human Malaria Parasite Pfs47 Gene Mediates Evasion of the Mosquito Immune System. Science. *340*, 984–987.

Montgomery, J., Mphande, F.A., Berriman, M., Pain, A., Rogerson, S.J., Taylor, T.E., Molyneux, M.E., and Craig, A. (2007). Differential var gene expression in the organs of patients dying of falciparum malaria. Mol. Microbiol. *65*, 959–967.

Morrison, D.A. (2009). Evolution of the Apicomplexa: where are we now? Trends Parasitol. *25*, 375–382.

Muhia, D.K., Swales, C.A., Deng, W., Kelly, J.M., and Baker, D.A. (2001). The gametocyte-activating factor xanthurenic acid stimulates an increase in membrane-associated guanylyl cyclase activity in the human malaria parasite Plasmodium falciparum. Mol. Microbiol. *42*, 553–560.

Nacher, M., Singhasivanon, P., Silachamroon, U., Treeprasertsuk, S., Tosukhowong, T., Vannaphan, S., Gay, F., Mazier, D., and Looareesuwan, S. (2002). Decreased hemoglobin concentrations, hyperparasitemia, and severe malaria are associated with increased

Plasmodium falciparum gametocyte carriage. J. Parasitol. *88*, 97–101.

Ngotho, P., Soares, A.B., Hentzschel, F., Achcar, F., Bertuccini, L., and Marti, M. (2019). Revisiting gametocyte biology in malaria parasites. FEMS Microbiol. Rev. *43*, 401–414.

Nguitragool, W., Bokhari, A.A.B., 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.

Niang, M., Yan Yam, X., and Preiser, P.R. (2009). The Plasmodium falciparum STEVOR Multigene Family Mediates Antigenic Variation of the Infected Erythrocyte. PLoS Pathog. *5*, e1000307.

Niang, M., Bei, A.K., Madnani, K.G., Pelly, S., Dankwa, S., Kanjee, U., Gunalan, K., Amaladoss, A., Yeo, K.P., Bob, N.S., et al. (2014). STEVOR Is a Plasmodium falciparum Erythrocyte Binding Protein that Mediates Merozoite Invasion and Rosetting. Cell Host Microbe *16*, 81–93.

Nilsson, S.K., Childs, L.M., Buckee, C., and Marti, M. (2015). Targeting Human Transmission Biology for Malaria Elimination. PLOS Pathog. *11*, e1004871.

De Niz, M., Meibalan, E., Mejia, P., Ma, S., Brancucci, N.M.B., Agop-Nersesian, C., Mandt, R., Ngotho, P., Hughes, K.R., Waters, A.P., et al. (2018). Plasmodium gametocytes display homing and vascular transmigration in the host bone marrow. Sci. Adv. *4*, eaat3775.

Van Noort, V., and Huynen, M.A. (2006). Combinatorial gene

regulation in Plasmodium falciparum. Trends Genet. *22*, 73–78.

O'Kane, C.J., and Hyland, E.M. (2019). Yeast epigenetics: the inheritance of histone modification states. Biosci. Rep. *39*, BSR20182006.

Obersriebnig, M.J., Pallesen, E.M.H., Sneppen, K., Trusina, A., and Thon, G. (2016). Nucleation and spreading of a heterochromatic domain in fission yeast. Nat. Commun. *7*, 11518.

Orikiiriza, J., Surowiec, I., Lindquist, E., Bonde, M., Magambo, J., Muhinda, C., Bergström, S., Trygg, J., and Normark, J. (2017). Lipid response patterns in acute phase paediatric Plasmodium falciparum malaria. Metabolomics *13*, 41.

Painter, H.J., Carrasquilla, M., and Llinás, M. (2017). Capturing in vivo RNA transcriptional dynamics from the malaria parasite Plasmodium falciparum. Genome Res. *27*, 1074–1086.

Painter, H.J., Chung, N.C., Sebastian, A., Albert, I., Storey, J.D., and Llinás, M. (2018). Genome-wide real-time in vivo transcriptional dynamics during Plasmodium falciparum blood-stage development. Nat. Commun. *9*, 2656.

Parkyn Schneider, M., Liu, B., Glock, P., Suttie, A., McHugh, E., Andrew, D., Batinovic, S., Williamson, N., Hanssen, E., McMillan, P., et al. (2017). Disrupting assembly of the inner membrane complex blocks Plasmodium falciparum sexual stage development. PLOS Pathog. *13*, e1006659.

Parroche, P., Lauw, F.N., Goutagny, N., Latz, E., Monks, B.G.,

Visintin, A., Halmen, K.A., Lamphier, M., Olivier, M., Bartholomeu, D.C., et al. (2007). Malaria hemozoin is immunologically inert but radically enhances innate responses by presenting malaria DNA to Toll-like receptor 9. Proc. Natl. Acad. Sci. *104*, 1919–1924.

Pasternak, N.D., and Dzikowski, R. (2009). PfEMP1: An antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite Plasmodium falciparum. Int. J. Biochem. Cell Biol. *41*, 1463–1466.

Pease, B.N., Huttlin, E.L., Jedrychowski, M.P., Talevich, E., Harmon, J., Dillman, T., Kannan, N., Doerig, C., Chakrabarti, R., Gygi, S.P., et al. (2013). Global Analysis of Protein Expression and Phosphorylation of Three Stages of Plasmodium falciparum Intraerythrocytic Development. J. Proteome Res. *12*, 4028–4045.

Peatey, C.L., Watson, J.A., Trenholme, K.R., Brown, C.L., Nielson, L., Guenther, M., Timmins, N., Watson, G.S., and Gardiner, D.L. (2013). Enhanced Gametocyte Formation in Erythrocyte Progenitor Cells: A Site-Specific Adaptation by Plasmodium falciparum. J. Infect. Dis. *208*, 1170–1174.

Pelle, K.G., Oh, K., Buchholz, K., Narasimhan, V., Joice, R., Milner, D.A., Brancucci, N., Ma, S., Voss, T.S., Ketman, K., et al. (2015). Transcriptional profiling defines dynamics of parasite tissue sequestration during malaria infection. Genome Med. *7*, 19.

Pessi, G., Kociubinski, G., and Mamoun, C. Ben (2004). A pathway for phosphatidylcholine biosynthesis in Plasmodium falciparum involving phosphoethanolamine methylation. Proc. Natl. Acad. Sci. U.

S. A. *101*, 6206–6211.

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.

Phillips, M.A., Burrows, J.N., Manyando, C., van Huijsduijnen, R.H., Van Voorhis, W.C., and Wells, T.N.C. (2017). Malaria. Nat. Rev. Dis. Prim. *3*, 17050.

del Pilar Quintana, M., Ch'ng, J.H., Zandian, A., Imam, M., Hultenby, K., Theisen, M., Nilsson, P., Qundos, U., Moll, K., Chan, S., et al. (2018). SURGE complex of Plasmodium falciparum in the rhoptry-neck (SURFIN4.2-RON4-GLURP) contributes to merozoite invasion. PLoS One *13*, 1–25.

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.

Pope, S.D., and Medzhitov, R. (2018). Emerging Principles of Gene Expression Programs and Their Regulation. Mol. Cell *71*, 389–397.

Poran, A., Nötzel, C., Aly, O., Mencia-Trinchant, N., Harris, C.T., Guzman, M.L., Hassane, D.C., Elemento, O., and Kafsack, B.F.C. (2017). Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. Nature *551*, 95–99.

Pray, L.A. (2008). DNA Replication and Causes of Mutation. Nat. Educ. *1*, 214.

Raabe, A.C., Wengelnik, K., Billker, O., and Vial, H.J. (2011). Multiple roles for Plasmodium berghei phosphoinositide-specific phospholipase C in regulating gametocyte activation and differentiation. Cell. Microbiol. *13*, 955–966.

Radeva-Petrova, D., Kayentao, K., ter Kuile, F.O., Sinclair, D., and Garner, P. (2014). Drugs for preventing malaria in pregnant women in endemic areas: any drug regimen versus placebo or no treatment. Cochrane Database Syst. Rev. *10*, 1465–1858.

Ralph, S.A., Scheidig-Benatar, C., and Scherf, A. (2005). Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. Proc. Natl. Acad. Sci. *102*, 5414–5419.

Recker, M., Buckee, C.O., Serazin, A., Kyes, S., Pinches, R., Christodoulou, Z., Springer, A.L., Gupta, S., and Newbold, C.I. (2011). Antigenic Variation in Plasmodium falciparum Malaria Involves a Highly Structured Switching Pattern. PLoS Pathog. *7*, e1001306.

Reddy, B.N., Shrestha, S., Hart, K.J., Liang, X., Kemirembe, K., Cui, L., and Lindner, S.E. (2015). A bioinformatic survey of RNA-binding proteins in Plasmodium. BMC Genomics *16*, 890.

Regev-Rudzki, N., Wilson, D.W., Carvalho, T.G., Sisquella, X., Coleman, B.M., Rug, M., Bursac, D., Angrisano, F., Gee, M., Hill, A.F., et al. (2013). Cell-cell communication between malaria-infected

red blood cells via exosome-like vesicles. Cell *153*, 1120–1133.

Reid, A.J., Talman, A.M., Bennett, H.M., Gomes, A.R., Sanders, M.J., Illingworth, C.J.R., Billker, O., Berriman, M., and Lawniczak, M.K. (2018). Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites. Elife *7*, 1–29.

Riddle, N.C., Jung, Y.L., Gu, T., Alekseyenko, A.A., Asker, D., Gui, H., Kharchenko, P. V., Minoda, A., Plachetka, A., Schwartz, Y.B., et al. (2012). Enrichment of HP1a on Drosophila Chromosome 4 Genes Creates an Alternate Chromatin Structure Critical for Regulation in this Heterochromatic Domain. PLoS Genet. *8*, e1002954.

Le Roch, K.G. (2004). Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. Genome Res. *14*, 2308–2318.

Roncalés, M., Vidal-Mas, J., Leroy, D., and Herreros, E. (2012). Comparison and Optimization of Different Methods for the In Vitro Production of Plasmodium falciparum Gametocytes. J. Parasitol. Res. *2012*, 1–7.

Rono, M.K., Nyonda, M.A., Simam, J.J., Ngoi, J.M., Mok, S., Kortok, M.M., Abdullah, A.S., Elfaki, M.M., Waitumbi, J.N., El-Hassan, I.M., et al. (2018). Adaptation of Plasmodium falciparum to its transmission environment. Nat. Ecol. Evol. *2*, 377–387.

Ross, R. (1923). Memoirs. With a Full Account of the Great Malaria Problem and Its Solution. Am. J. Public Health (N. Y). *13*, 959–960.

Rovira-Graells, N., Gupta, A.P.A., 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.

Rovira-Graells, N., Crowley, V.M., Bancells, C., Mira-Martínez, S., Ribas de Pouplana, L., and Cortés, A. (2015). Deciphering the principles that govern mutually exclusive expression of Plasmodium falciparum clag3 genes. Nucleic Acids Res. *43*, 8243–8257.

RTS.S Clinical Trials Partnership (2015). Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. Lancet *386*, 31–45.

Saito, F., Hirayasu, K., Satoh, T., Wang, C.W., Lusingu, J., Arimori, T., Shida, K., Palacpac, N.M.Q., Itagaki, S., Iwanaga, S., et al. (2017). Immune evasion of Plasmodium falciparum by RIFIN via inhibitory receptors. Nature *552*, 101–105.

Saksouk, N., Simboeck, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin *8*, 3.

Salcedo-Amaya, A.M., van Driel, M.A., Alako, B.T., Trelle, M.B., van den Elzen, A.M.G., Cohen, A.M., Janssen-Megens, E.M., van de Vegte-Bolmer, M., Selzer, R.R., Iniguez, A.L., et al. (2009). Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proc. Natl. Acad. Sci. *106*, 9655–9660.

Sam-Yellowe, T.Y. (2004). A Plasmodium Gene Family Encoding Maurer's Cleft Membrane Proteins: Structural Properties and

305

Expression Profiling. Genome Res. *14*, 1052–1059.

Santenard, A., Ziegler-Birling, C., Koch, M., Tora, L., Bannister, A.J., and Torres-Padilla, M.-E. (2010). Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. Nat. Cell Biol. *12*, 853–862.

Santos, J.M., Josling, G., Ross, P., Joshi, P., Orchard, L., Campbell, T., Schieler, A., Cristea, I.M., and Llinás, M. (2017). Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor. Cell Host Microbe *21*, 731-741.e10.

Saze, H., Kitayama, J., Takashima, K., Miura, S., Harukawa, Y., Ito, T., and Kakutani, T. (2013). Mechanism for full-length RNA processing of Arabidopsis genes containing intragenic heterochromatin. Nat. Commun. *4*, 2301.

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.

Sebastian, S., Brochet, M., Collins, M.O., Schwach, F., Jones, M.L., Goulding, D., Rayner, J.C., Choudhary, J.S., and Billker, O. (2012). A Plasmodium Calcium-Dependent Protein Kinase Controls Zygote Development and Transmission by Translationally Activating Repressed mRNAs. Cell Host Microbe *12*, 9–19.

Shaw, W.R., and Catteruccia, F. (2019). Vector biology meets disease control: using basic research to fight vector-borne diseases. Nat.

Microbiol. *4*, 20–34.

Sherrard-Smith, E., Skarp, J.E., Beale, A.D., Fornadel, C., Norris, L.C., Moore, S.J., Mihreteab, S., Charlwood, J.D., Bhatt, S., Winskill, P., et al. (2019). Mosquito feeding behavior and how it influences residual malaria transmission across Africa. Proc. Natl. Acad. Sci. 201820646.

Shirai, A., Kawaguchi, T., Shimojo, H., Muramatsu, D., Ishida-Yonetani, M., Nishimura, Y., Kimura, H., Nakayama, J., and Shinkai, Y. (2017). Impact of nucleic acid and methylated H3K9 binding activities of Suv39h1 on its heterochromatin assembly. Elife *6*, e25317.

Shrestha, S., Li, X., Ning, G., Miao, J., and Cui, L. (2016). The RNA-binding protein Puf1 functions in the maintenance of gametocytes in Plasmodium falciparum. J. Cell Sci. *129*, 3144–3152.

Sierra-Miranda, M., Vembar, S.-S., Delgadillo, D.M., Ávila-López, P.A., Herrera-Solorio, A.-M., Lozano Amado, D., Vargas, M., and Hernandez-Rivas, R. (2017). Pf AP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds to Plasmodium falciparum telomeres. Cell. Microbiol. *19*, e12742.

Silvestrini, F., Alano, P., and Williams, J.L. (2000). Commitment to the production of male and female gametocytes in the human malaria parasite Plasmodium falciparum. Parasitology *121 Pt 5*, 465–471.

Silvestrini, F., Bozdech, Z., Lanfrancotti, A., Giulio, E. Di, Bultrini, E., Picci, L., DeRisi, J.L., Pizzi, E., and Alano, P. (2005). Genome-wide identification of genes upregulated at the onset of gametocytogenesis in Plasmodium falciparum. Mol. Biochem. Parasitol. *143*, 100–110.

Silvestrini, F., Lasonder, E., Olivieri, A., Camarda, G., van Schaijk, B., Sanchez, M., Younis Younis, S., Sauerwein, R., and Alano, P. (2010). Protein export marks the early phase of gametocytogenesis of the human malaria parasite Plasmodium falciparum. Mol. Cell. Proteomics *9*, 1437–1448.

Sinden, R.E. (2009). Malaria, sexual development and transmission: retrospect and prospect. Parasitology *136*, 1427–1434.

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

Smalley, M.E., Abdalla, S., and Brown, J. (1981). The distribution of Plasmodium falciparum in the peripheral blood and bone marrow of Gambian children. Trans. R. Soc. Trop. Med. Hyg. *75*, 103–105.

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., Pinches, R., Newbold, C.I., and Miller, L.H. (1995). Switches in expression of plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell *82*, 101–110.

Smith, T.G., Lourenço, P., Carter, R., Walliker, D., and Ranford-Cartwright, L.C. (2000). Commitment to sexual differentiation in the human malaria parasite, Plasmodium falciparum. Parasitology *121 Pt 2*, 127–133.

Snow, R.W., Sartorius, B., Kyalo, D., Maina, J., Amratia, P., Mundia, C.W., Bejon, P., and Noor, A.M. (2017). The prevalence of

Plasmodium falciparum in sub-Saharan Africa since 1900. Nature *550*, 515–518.

Sologub, L., Kuehn, A., Kern, S., Przyborski, J., Schillig, R., and Pradel, G. (2011). Malaria proteases mediate inside-out egress of gametocytes from red blood cells following parasite transmission to the mosquito. Cell. Microbiol. *13*, 897–912.

Soutourina, J. (2018). Transcription regulation by the Mediator complex. Nat. Rev. Mol. Cell Biol. *19*, 262–274.

Spillman, N.J., Beck, J.R., and Goldberg, D.E. (2015). Protein Export into Malaria Parasite–Infected Erythrocytes: Mechanisms and Functional Consequences. Annu. Rev. Biochem. *84*, 813–841.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Su, X., Heatwole, V.M., Wertheimer, S.P., Guinet, F., Herrfeldt, J.A., Peterson, D.S., Ravetch, J.A., and Wellems, T.E. (1995). The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of plasmodium falciparum-infected erythrocytes. Cell *82*, 89–100.

Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. Sci. Adv. *5*, eaaw1668.

Tadesse, F.G., Meerstein-kessel, L., Gonçalves, B.P., Drakeley, C., Ranford-cartwright, L., and Bousema, T. (2019). Gametocyte Sex Ratio: The Key to Understanding Plasmodium falciparum Transmission? Trends Parasitol. *35*, 226–238.

Tarun, A.S., Vaughan, A.M., and Kappe, S.H.I. (2009). Redefining the role of de novo fatty acid synthesis in Plasmodium parasites. Trends Parasitol. *25*, 545–550.

Tawfik, D.S. (2010). Messy biology and the origins of evolutionary innovations. Nat. Chem. Biol. *6*, 692–696.

Taylor, H.M., Grainger, M., and Holder, A.A. (2002). Variation in the Expression of a Plasmodium falciparum Protein Family Implicated in Erythrocyte Invasion. Infect. Immun. *70*, 5779–5789.

Thomsen, E.K., Koimbu, G., Pulford, J., Jamea-Maiasa, S., Ura, Y., Keven, J.B., Siba, P.M., Mueller, I., Hetzel, M.W., and Reimer, L.J. (2016). Mosquito behaviour change after distribution of bednets results in decreased protection against malaria exposure. J. Infect. Dis. *215*, jiw615.

Tibúrcio, M., Niang, M., Deplaine, G., Perrot, S., Bischoff, E., Ndour, P.A., Silvestrini, F., Khattab, A., Milon, G., David, P.H., et al. (2012). A switch in infected erythrocyte deformability at the maturation and blood circulation of Plasmodium falciparum transmission stages. Blood *119*, 172–181.

Tibúrcio, M., Silvestrini, F., Bertuccini, L., Sander, A.F., Turner, L., Lavstsen, T., and Alano, P. (2013). Early gametocytes of the malaria parasite Plasmodium falciparum specifically remodel the adhesive properties of infected erythrocyte surface. Cell. Microbiol. *15*, 647–659.

Tilley, L., Straimer, J., Gnädig, N.F., Ralph, S.A., and Fidock, D.A. (2016). Artemisinin Action and Resistance in Plasmodium falciparum.

Trends Parasitol. *32*, 682–696.

Toenhake, C.G., and Bártfai, R. (2019). What functional genomics has taught us about transcriptional regulation in malaria parasites. Brief. Funct. Genomics *00*, 1–12.

Toenhake, C.G., Fraschka, S.A.-K., Vijayabaskar, M.S., Westhead, D.R., van Heeringen, S.J., and Bártfai, R. (2018). Chromatin Accessibility-Based Characterization of the Gene Regulatory Network Underlying Plasmodium falciparum Blood-Stage Development. Cell Host Microbe *23*, 557-569.e9.

Tonkin, C.J., Carret, C.K., Duraisingh, M.T., Voss, T.S., Ralph, S.A., Hommel, M., Duffy, M.F., Silva, L.M. da, Scherf, A., Ivens, A., et al. (2009). Sir2 Paralogues Cooperate to Regulate Virulence Genes and Antigenic Variation in Plasmodium falciparum. PLoS Biol. *7*, e1000084.

Trager, W., and Gill, G.S. (1992). Enhanced gametocyte formation in young erythrocytes by Plasmodium falciparum in vitro. J. Protozool. *39*, 429–432.

Trager, W., and Jensen, J. (1976). Human malaria parasites in continuous culture. Science. *193*, 673–675.

Trager, W., Gill, G.S., Lawrence, C., and Nagel, R.L. (1999). Plasmodium falciparum: Enhanced Gametocyte Formationin Vitroin Reticulocyte-Rich Blood. Exp. Parasitol. *91*, 115–118.

Tran, P.N., Brown, S.H.J., Rug, M., Ridgway, M.C., Mitchell, T.W., and Maier, A.G. (2016). Changes in lipid composition during sexual

development of the malaria parasite Plasmodium falciparum. Malar. J. *15*, 73.

Treeck, M., Sanders, J.L., Elias, J.E., and Boothroyd, J.C. (2011). The Phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii Reveal Unusual Adaptations Within and Beyond the Parasites' Boundaries. Cell Host Microbe *10*, 410–419.

Trelle, M.B., Salcedo-Amaya, A.M., Cohen, A.M., Stunnenberg, H.G., and Jensen, O.N. (2009). Global Histone Analysis by Mass Spectrometry Reveals a High Content of Acetylated Lysine Residues in the Malaria Parasite Plasmodium falciparum. J. Proteome Res. *8*, 3439–3450.

Twohig, K.A., Pfeffer, D.A., Baird, J.K., Price, R.N., Zimmerman, P.A., Hay, S.I., Gething, P.W., Battle, K.E., and Howes, R.E. (2019). Growing evidence of Plasmodium vivax across malaria-endemic Africa. PLoS Negl. Trop. Dis. *13*, e0007140.

Ukaegbu, U.E., Kishore, S.P., Kwiatkowski, D.L., Pandarinath, C., Dahan-Pasternak, N., Dzikowski, R., and Deitsch, K.W. (2014). Recruitment of PfSET2 by RNA polymerase II to variant antigen encoding loci contributes to antigenic variation in P. falciparum. PLoS Pathog. *10*, e1003854.

Unschuld, P. (2003). Huang Di Nei Jing Su Wen: Nature, Knowledge, Imagery in an Ancient Chinese Medical Text (Berkeley and Los Angeles, California: University of California Press).

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 Lysine 9 Methylation and HP1γ Are Associated with

Transcription Elongation through Mammalian Chromatin. Mol. Cell *19*, 381–391.

Vaughan, J.A., Noden, B.H., and Beier, J.C. (1994). Sporogonic development of cultured Plasmodium falciparum in six species of laboratory-reared Anopheles mosquitoes. Am. J. Trop. Med. Hyg. *51*, 233–243.

Vembar, S.S., Droll, D., and Scherf, A. (2016). Translational regulation in blood stages of the malaria parasite *Plasmodium spp.*: systems-wide studies pave the way. Wiley Interdiscip. Rev. RNA *7*, 772–792.

Volz, J., Carvalho, T.G., Ralph, S.A., Gilson, P., Thompson, J., Tonkin, C.J., Langer, C., Crabb, B.S., and Cowman, A.F. (2010). Potential epigenetic regulatory proteins localise to distinct nuclear subcompartments in Plasmodium falciparum. Int. J. Parasitol. *40*, 109–121.

Volz, J.C., Bártfai, R., Petter, M., Langer, C., Josling, G.A., Tsuboi, T., Schwach, F., Baum, J., Rayner, J.C., Stunnenberg, H.G., et al. (2012). PfSET10, a Plasmodium falciparum Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division. Cell Host Microbe *11*, 7–18.

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.

Voss, T.S., Bozdech, Z., and Bártfai, R. (2014). Epigenetic memory takes center stage in the survival strategy of malaria parasites. Curr.

Opin. Microbiol. *20*, 88–95.

Wahlgren, M., Goel, S., and Akhouri, R.R. (2017). Variant surface antigens of Plasmodium falciparum and their roles in severe malaria. Nat. Rev. Microbiol. *15*, 479–491.

Walker, R.A., Ferguson, D.J.P., Miller, C.M.D., and Smith, N.C. (2013). Sex and Eimeria: A molecular perspective. Parasitology *140*, 1701–1717.

Wassmer, S.C., Taylor, T.E., Rathod, P.K., Mishra, S.K., Mohanty, S., Arevalo-Herrera, M., Duraisingh, M.T., and Smith, J.D. (2015). Investigating the Pathogenesis of Severe Malaria: A Multidisciplinary and Cross-Geographical Approach. Am. J. Trop. Med. Hyg. *93*, 42–56.

Werling, K., Shaw, W.R., Itoe, M.A., Westervelt, K.A., Marcenac, P., Paton, D.G., Peng, D., Singh, N., Smidler, A.L., South, A., et al. (2019). Steroid Hormone Function Controls Non-competitive Plasmodium Development in Anopheles. Cell *177*, 315-325.e14.

WHO (2015a). Guidelines for the treatment of malaria, third edition (World Health Organization).

WHO (2015b). World Malaria Report 2015 (World Health Organization).

WHO (2018). World Malaria Report 2018 (World Health Organization).

Wiencke, J.K., Zheng, S., Morrison, Z., and Yeh, R.-F. (2008). Differentially expressed genes are marked by histone 3 lysine 9

trimethylation in human cancer cells. Oncogene *27*, 2412–2421.

Wilke, A.B.B., and Marrelli, M.T. (2015). Paratransgenesis: a promising new strategy for mosquito vector control. Parasit. Vectors *8*, 342.

Williams, J.L. (1999). Stimulation of Plasmodium falciparum gametocytogenesis by conditioned medium from parasite cultures. Am. J. Trop. Med. Hyg. *60*, 7–13.

Winter, G., Kawai, S., Haeggström, M., Kaneko, O., von Euler, A., Kawazu, S., Palm, D., Fernandez, V., and Wahlgren, M. (2005). SURFIN is a polymorphic antigen expressed on Plasmodium falciparum merozoites and infected erythrocytes. J. Exp. Med. *201*, 1853–1863.

Wirth, C.C., and Pradel, G. (2012). Molecular mechanisms of host cell egress by malaria parasites. Int. J. Med. Microbiol. *302*, 172–178.

Wirth, C.C., Glushakova, S., Scheuermayer, M., Repnik, U., Garg, S., Schaack, D., Kachman, M.M., Weißbach, T., Zimmerberg, J., Dandekar, T., et al. (2014). Perforin-like protein PPLP2 permeabilizes the red blood cell membrane during egress of P lasmodium falciparum gametocytes. Cell. Microbiol. *16*, 709–733.

Wunderlich, J., Rohrbach, P., and Dalton, J.P. (2012). The malaria digestive vacuole. Front. Biosci. (Schol. Ed). *4*, 1424–1448.

Xu, M., Wang, W., Chen, S., and Zhu, B. (2011). A model for mitotic inheritance of histone lysine methylation. EMBO Rep. *13*, 60–67.

Young, J.A., Fivelman, Q.L., Blair, P.L., de la Vega, P., Le Roch, K.G., Zhou, Y., Carucci, D.J., Baker, D.A., and Winzeler, E.A. (2005). The

Plasmodium falciparum sexual development transcriptome: A microarray analysis using ontology-based pattern identification. Mol. Biochem. Parasitol. *143*, 67–79.

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

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

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

Yuda, M., Kaneko, I., Iwanaga, S., Murata, Y., and Kato, T. (2019). Female- specific gene regulation in malaria parasites by an AP2- family transcription factor. Mol. Microbiol. mmi.14334.

Zhang, C., Li, Z., Cui, H., Jiang, Y., Yang, Z., Wang, X., Gao, H., Liu, C., Zhang, S., Su, X., et al. (2017). Systematic CRISPR-Cas9-Mediated Modifications of Plasmodium yoelii ApiAP2 Genes Reveal Functional Insights into Parasite Development. MBio *8*, e01986-17.

Zhang, Q., Siegel, T.N., Martins, R.M., Wang, F., Cao, J., Gao, Q., Cheng, X., Jiang, L., Hon, C.-C., Scheidig-Benatar, C., et al. (2014). Exonuclease-mediated degradation of nascent RNA silences genes linked to severe malaria. Nature *513*, 431–435.

ABBREVIATIONS

5'RACE	5' Rapid Amplification of cDNA Ends
ACT	Artemisinin-combination therapy
BC	Before Christ
CDPK	Calcium-dependent protein kinase
CDS	Coding sequence
cGMP	Cyclic GMP
CSP	Circumsporozite surface protein
CVE	Clonally variant expression
CVG	Clonally variant gene
DNA	Deoxyribonucleic acid
dsRNA	Double-stranded RNA
EK	Ethanolamine kinase
GDV1	Gametocyte developmental protein 1
GEXP	Gametocyte exported proteins
GSTF	Gene specific transcription factor
GTF	General transcription factor
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HDM	Histone demethylase
HKMT	Histone lysine methyltransferase
HP1	Heterochromatin protein 1
IDC	Intraerythrocytic developmental cycle
IFA	Immunofluorescence assay
IMC	Inner membrane complex
IP3	Inositol-(1,4,5)-trisphosphate
IPTp	Intermittent preventive therapy in pregnancy

IRS	indoor residual spraying
ITN	Insecticide treated bed net
JMJD	Jumonji-C domain containing histone demethylase
KO	Knock-out
LAMP	Loop-mediated isothermal amplification
lncRNA	Long non-coding RNA
LSD1	Lysine specific demethylase 1
LysoPC	Lysophosphatidylcholine
mRNA	Messenger RNA
NCC	Next cycle conversion
ncRNA	Non-coding RNA
NPP	New permeation pathway
PcG	Polycomb group
PCR	Polymerase chain reaction
PE	Pairing elements
PfEMP1	*P. falciparum* erythrocyte membrane protein 1
PfMC-2TM	*P. falciparum* Maurer's cleft two transmembrane protein
PIC	Pre-initation complex
PKG	cGMP-dependent protein kinase
PMT	Phosphoethanolamine methyltransferase
Pol II	RNA polymerase II
PPLP2	Plasmodial pore-forming perforins
PSAC	Plasmodial surface anion channel
PTM	Posttranslational modifications
RBC	Red blood cell
RDT	Rapid diagnostic test
RIFIN	Repetitive interspersed families of polypeptides
RITS	RNA-induced transcriptional silencing complex

RNA   Ribonucleic acid

RNAi   RNA interference

SAM   S-adenosylmethionine

SCC   Same cycle conversion

siRNA   Small interfering RNA

STEVOR   Subtelomeric variable open reading frame

TAD   Topologically associated domain

TF   Transcription factor

TSS   Transcription start site

WHO   World Health Organization

XA   Xanthurenic acid

ANNEXES

During my thesis I have also been involved in the writing of a scientific review about the mechanisms underlying transcriptional variation in malaria parasites, which has been published in *Briefings in Functional Genomics*. Additionally, I have assisted in the development of reporter lines for sexual conversion. These lines are based on the expression of a tandem tomato fluorescent marker under the control of the *gexp02* promoter, one of the first genes responding to *pfap2-g* activation. The manuscript has been peer-reviewed in *Scientific Reports* and is currently under consideration.

Llorà-Batlle, O.*, Tintó-Font, E.* & Cortés, A. Transcriptional variation in malaria parasites: why and how. *Brief. Funct. Genomics* (2019). **00,** 1–13

* These authors contributed equally

Portugaliza, H. P., Llorà-Batlle, O., Rosanas-Urgell, A. & Cortés, A. Reporter lines based on the *gexp02* promoter enable early quantification of sexual conversion rates in the malaria parasite *Plasmodium falciparum.* (**Under review in *Scientific Reports***)