ROSANA BEATRIZ DUQUE ARAUJO

Estudo sobre o mecanismo de transcrição dos genes *rif* (*repetitive interspersed family*) no *Plasmodium falciparum* e o papel da demetilase putativa LSD1, como possível modificador de histonas em *Plasmodium* spp.

Tese apresentada ao programa de Pós-Graduação em Biología da Relação Patógeno-Hospedeiro do Instituto de Ciências Biomédicas da Universidade de São Paulo, para a obtenção do Título de Doutor em Ciências.

Área de concentração: Biologia da Relação Patógeno-Hospedeiro.

Orientador: Prof. Gerhard Wunderlich.

Versão Original.

São Paulo

2019

ROSANA BEATRIZ DUQUE ARAUJO

Studies on the transcription mechanism of *rif* (*repetitive interspersed family*) genes of *Plasmodium falciparum* and the role of the putative Lysine Specific Demethylase 1 (LSD1) as possible histone modifier in *Plasmodium* spp.

Ph.D. Thesis presented to the Post-graduation program Biology of Host Pathogen Interaction at the Institute of Biomedical Sciences of the University of São Paulo, in order to obtain the degree of Doctor in Sciences.

Area: Biology of Host-Pathogen Interactions.

Supervisor: Prof. Gerhard Wunderlich.

Original version.

São Paulo

2019

UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE CIÊNCIAS BIOMÉDICAS

_____

Candidato: Rosana Beatriz Duque Araujo

Título da Tese: Estudo sobre o mecanismo de transcrição dos genes *rif* (*repetitive interspersed family*) no *Plasmodium falciparum* e o papel da demetilase putativa LSD1, como possível modificador de histonas em *Plasmodium* spp.

Orientador: Prof. Gerhard Wunderlich.

A Comissão Julgadora dos trabalhos de Defesa da Tese de Doutorado, em sessão

pública realizada a ..../................./................., considerou

( ) Aprovado(a)   ( ) Reprovado(a)

Examinador(a):Assinatura: ....

Nome: ....

Instituição: ....

Examinador(a):   Assinatura: ....

Nome: ....

Instituição: ....

Examinador(a):   Assinatura: ....

Nome: ....

Instituição: ....

Examinador(a):   Assinatura: ....

Nome: ....

Instituição: ....

Presidente: Assinatura: ....

Nome: ....

Instituição: ....

ICBUSP

CERTIFICADO DE ISENÇÃO

Certificamos que o Protocolo CEP-ICB nº **755/2015** referente ao projeto intitulado: *"Estudos sobre o modo de transcrição de genes rif de Plasmodium falciparum"* sob a responsabilidade de Rosana Beatriz Duque Araújo e orientação do(a) Prof.(a) Dr.(a) **Gerhard Wunderlich,** do Departamento de Parasitologia, foi analisado pela **CEUA** - Comissão de Ética no Uso de Animais e pela **CEPSH** - Comissão de Ética em Pesquisa com Seres Humanos, tendo sido deliberado que o referido projeto não utilizará animais que estejam sob a égide da Lei nº 11.794, de 8 de outubro de 2008, nem envolverá procedimentos regulados pela Resolução CONEP nº 466 de 2012.

São Paulo, 25 de novembro de 2015

Prof. Dr. **Anderson de Sá Nunes**
Coordenador CEUA ICB/USP

Prof. Dr. **Paolo M. A. Zanotto**
Coordenador CEPSH ICB/USP

A mi familia, Richard y Catalina.

El mejor estímulo que la vida me pudo dar.

ACKNOWLEDGEMENTS

- To Gerd, thank you for the opportunity and to open the doors of your lab. Thanks to show me that it can be possible to be a good teacher, researcher, dad and athlete. Thanks to be so collaborative, patient and comprehensive. *Gracias por apoyarme en los momentos difíciles y darme ánimos para continuar. Danke schöne.*

- Thanks to past and present UDD members: **<u>Flavia</u>**, Kamila, Jasmin, Thales, **<u>Sory</u>**, Wesley, Eliana, Isadora, João, Felipe, Marleen, Arne, Sum, **<u>Tati</u>**.

- Thanks to Tatiane, my lab partner, my friend and guide during this 4 years! We made it!!! *Muito obrigrada.*

- To Carsten, thanks for been always present. *Danke.*

- To Wolf, thanks for your support, every time that I need it.

- To all the German's visitors. Especially to Charlotte, Frank and Wiebke. Thanks very much for your help. *Danke.*

- To the Professor Daniel Bargieri's lab, Xio, Marisé, Juliana, Mirian and especially to Irina for you friendship and support. *Merci.*

- To the Professor Cláudio Marino's lab, to Erika, Andre and Flavia.

- To the Professor Mauro Cortez's lab, particularly to Ismael Sauter.

- To all the team of the ICB II, that make possible our day by day journey.

- To the *chicas*, Juli, Angela, Sory, Tati, Irina, Xio, Mari, Laura. Thanks for make this 4 years journey more easy.

- Thanks to my family. *Cuando hay amor no existe la distancia. Los amo.*

- To my husband Richard Girard. Thanks for you love and support. There are no words that describe how deeply grateful I am with you. *Merci beaucoup, te amo.*

- Finally, thanks to everyone who thinks they deserve to be on this page.

The key to success is getting used to doing the things you are afraid to do.
Vincent van Gogh.

RESUMO

RBD. Araujo - Estudo sobre o mecanismo de transcrição dos genes *rif* (*repetitive interspersed family*) no *Plasmodium falciparum* e o papel da demetilase putativa LSD1, como possível modificador de histonas em *Plasmodium* spp. 2019. 118 p. Tese (Doutorado em Parasitologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2019.

Todas as espécies de *Plasmodium* possuem famílias de genes variantes que podem mediar o escape do sistema imune do hospedeiro vertebrado. No *Plasmodium falciparum*, a família dos genes *rif* codifica para antígenos variantes que são parcialmente expostos na superfície das hemácias infectadas e podem funcionar como fatores de virulência por meio da interação com os receptores das células hospedeiras. O controle da expressão dos genes *rif*, no entanto, não é claro e vários estudos relataram resultados contraditórios. Aqui, abordamos o modo de transcrição de *rif* usando duas abordagens, a primeira foi um vetor plasmídico com dois marcadores de resistência a drogas. Neste vetor, uma região a *upstream rif 5′* controlava a expressão de um dos marcadores de resistência, enquanto as linhas de parasitas transfectantes são selecionadas com o segundo marcador, expresso constitutivamente. Ao testar três diferentes regiões *upstream rif 5′*, descobrimos que uma era impossível de ser ativada, a segunda era constitutivamente expressa, enquanto uma terceira região *upstream* apresento atividade que variava amplamente com a pressão da droga aplicada na cultura. Essa construção também se integrou ao genoma. Quando a transcrição global de genes *rif* nesses transfectantes foi comparada na presença ou ausência das drogas de seleção, observamos que a ativação /silenciamento de todos os outros loci *rif* não mudou profundamente entre as cepas. Na segunda abordagem, foi usado o plasmídeo p*rif*GFPHA2A-BSDglmS. Este plasmídeo possui um marcador de resistência expresso constitutivamente, permitindo a seleção de linhas de parasitas transfectadas. O segundo marcador é então usado para *selecionar/promover* a expressão do genes *rif* alvo. O segundo marcador só é transcrito se ocorrer um *knockin* no gene alvo, o que levou à criação de um gene híbrido, contendo o gene alvo, *gfp-ha* e o marcador de resistência, ambos controlados pelo promotor genômico original do gene *rif* alvo. Quando ocorreu a recombinação, devido à região de homologia da seqüência *rif*, os parasitas modificados produziram uma proteína fundida, RIFIN-GFP-HA-2A-BSD. Essa construção também tinha a seqüência glmS, que forma uma ribozima uma vez que o RNA é sintetizado, permitindo truncar condicionalmente o transcrito de *rif-gfpha-2a-bsd*. Após a integração em um dos locos *rif*, observamos que, apesar da funcionalidade do construto 2A-BSD-glms, não conseguimos obter RNA suficiente para realizar qualquer ensaio funcional. Provavelmente, o promotor *rif* alvo produziu pouco transcrito para produzir blasticidina desaminase suficiente e manter os níveis adequados de parasitemia. Em esta parte do trabalho, concluímos que ou não há *crosstalk* entre os loci *rif* ou que, diferente aos promotores *var*, apenas a região *5′ upstream* a dos genes *rif* não é capaz de promover o mecanismo de transcrição por exclusão alélica nos genes *rif*. Também é possível que a subfamília B dos genes *rif* seja regulada de uma maneira diferente em comparação com a subfamília *rif* A.

Entre os fatores que podem afetar a transcrição de genes variantes em *P. falciparum* estão a metilação do DNA, modificações de cromatina, proteínas de ligação à cromatina e talvez RNAs não codificantes. *P. falciparum* possui duas famílias desmetilases de lisina de histona que empregam dois mecanismos moleculares diferentes de desmetilação, e no seu genoma contém pelo menos uma LSD1. A proteína *Pf*LSD1 desmetila especificamente histona 3 mono ou dimetilada na lisina 4 ou 9. Neste trabalho, exploramos a função da LSD1 putativa do *P. falciparum* na dinâmica de transcrição de genes variantes e outras famílias de genes, bem como sua interação com outras proteínas. Foi gerada a linhagem de parasitas transgênicos *PfLSD1GFPHAglmS*. A eficiência do *knockdown* ao nível transcricional foi avaliada via RT-qPCR e ao nível proteico por ensaios de Western blot. Foi realizada uma análise de

espectrometria de massa (MS) após o *pull-down* da *Pf*LSD1 marcada com HA para explorar as interações proteicas, e para explorar a resposta transcricional ao nível global do parasita quando o *Pf*LSD1 era abolida, foi feita uma análise de seqüenciamento por RNA-Seq. Observamos que o *knockdown* de *Pf*LSD1 leva a uma diminuição significativa do nível de transcrição do *Pflsd1*, no entanto, não afetou a viabilidade do parasita pelo menos durante o estágio eritrocitário. O resultado do ensaio de MS sugeriu uma interação entre a *Pf*LSD1 e a histona H3 e outras proteínas de remodelação da cromatina no estágio de trofozoíto. Na análise por RT-qPCR e RNA-Seq, não foi possível observar nenhuma alteração significativa no perfil transcricional dos genes variantes. Em vez disso, os genes que mostram alteração mais significativa foram os genes relacionados ao metabolismo do DNA e desenvolvimento sexual. Nossos resultados sugerem que a *Pf*LSD1 não participa da dinâmica transcricional da variação antigênica. No entanto, no *knockout* do roedor *P. berghei* LSD1 a linhagem de parasitas *P.berghei PbOokluc-LSD1^{KO}* mostrou uma formação atrasada ou incompleta de formas de oocineto. De acordo com os dados do transcriptoma de estudos, sugerimos que as proteínas LSD1 dos parasitas *Plasmodium* podem participar da regulação epigenética dos estágios sexuais dos parasitas, possivelmente das formas oocineto.

ABSTRACT

RBD. Araujo - Studies on the transcription mechanism of *rif* (*repetitive interspersed family*) genes of *Plasmodium falciparum* and the role of the putative Lysine Specific Demethylase 1 (LSD1) as possible histone modifier in *Plasmodium* spp. 2019. 118 p. Ph.D. Thesis (Parasitology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo 2019.

All *Plasmodium* species possess variant gene families which may mediate immune escape in the vertebrate host. In *Plasmodium falciparum*, the *rif* gene family encodes variant antigens which are partly exposed on the infected red blood cell surface and may function as virulence factors through interaction with host cell receptors. The expression control of *rif* genes, however, is unclear and several studies reported contradictory results. Herein, we addressed the mode of *rif* transcription using two approaches, the first one was a plasmid vector with two drug resistance markers. In this vector a *rif* 5' upstream region controlled the expression of one drug resistance marker while transfectant parasite lines are selected with the second, constitutively expressed marker. When testing three different *rif 5'* upstream regions, we found that one was impossible to be activated, the second was constitutively expressed while a third region showed activity that largely varied with the drug pressure applied in the culture. This construct had also integrated in the genome. When the global transcription of *rif* genes in these transfectants was compared in the presence or absence of drugs, we observed that the activation/silencing of all other *rif* loci did not change profoundly between strains. In the second approach, the *prifGFPHA2A-bsdglmS* plasmid was employed. This plasmid has one constitutively expressed resistance marker permitting the selection of transfected parasite lines. The second marker is then used to select/promote the expression of targeted *rif* genes. The second marker is only transcribed if a knockin in the target gene occurred, which led to the creation of a hybrid gene, containing the target gene, *gfp-ha* and the resistance, both of which controlled by the original genomic *rif* target gene promoter. When a single recombination event occurred, due to the *rif* homology region, the modified parasites produced a fused protein, RIFIN-GFP-HA-2A-BSD. This construct also had the glmS sequence, which forms a ribozyme once the RNA is synthesized, permitting to conditionally deplete the *rif-gfpha-2a-bsd* transcript. After the integration in one of the *rif* loci, we observed that despite the functionality of the 2A-BSD-glms construct, we were unable to obtain enough RNA to perform any functional assay. Probably, the targeted *rif* promoter produced too few transcript to produce enough blasticidin deaminase and to maintain the adequate levels of parasitemia. For this part of the work, we conclude that either there is no crosstalk between *rif* loci or that - unlike *var* promoters - solely the 5' upstream region of *rif* genes is not able to engage in a yet elusive system of allelic exclusion of *rif* gene transcription. Also is possible that the B subfamily of *rif* genes is differentially regulated compared to the *rif* A subfamily.

Between the factors that could affect variant gene transcription in *P. falciparum* are DNA methylation, dynamic chromatin modifications, chromatin binding proteins and perhaps non-coding RNAs. *P. falciparum* possess two families of histone lysine demethylases that employ two different molecular mechanisms of demethylation, and its genome contains at least one LSD1. The *Pf*LSD1 protein specifically demethylates mono- or dimethylated histone H3 lysine 4 and H3 lysine 9. In this work, we explore the function of the putative *P. falciparum* LSD1 on variant gene transcription dynamics and other gene families, as well its interaction with other proteins. We generated *Pf*LSD1GFPHAglmS-transgenic parasites lines. The efficiency of the transcriptional knockdown was evaluated via RT-qPCR and western blot assays. A mass spectrometry analysis after pull-down of HA-tagged *Pf*LSD1 was

performed to explore its protein interactions and RNA-Seq analysis to explore the global transcriptional response of the parasite when *Pf*LSD1 was knocked down. We observed that *Pf*LSD1 knockdown leads to a significant decrease of its transcript level, however, it did not affect the parasite viability at least during the erythrocyte stage. An interaction between HA-tagged LSD1 and the histone H3 and different chromatin remodeling proteins was suggested by MS assay at the trophozoite stage. In RT-qPCR and RNA-Seq analysis we couldn't observe any significant alteration in the transcriptional profile of variant genes. Instead, the genes that show most significant alteration were genes related with DNA metabolism and sexual development. Our results suggest that the *Pf*LSD1 does not participate in antigenic variation dynamics. Nevertheless, the knockout of rodent *P. berghei* LSD1 in a *P. berghei PbOokluc-LSD1^{KO}* parasite line, led to a delayed or incomplete formation of ookinete forms. In agreement with transcriptome data from other studies, we suggest that *Plasmodium* LSD1 proteins may participate in the epigenetic regulation of sexual stages of *Plasmodium* parasites, possibly downstream of ookinete forms.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACT Artemisinin combination therapy

ALI Acute lung injury

AMA1 Apical membrane antigen 1

ApiAp2 Apicomplexan Apetala Family 2

ARDS Acute respiratory distress syndrome

ARF Acute renal failure

AT Annealing Temperature

bp Base pair

cDNA Complementary deoxyribose nucleic acid

ChIP Chromatin Inmunoprecipitation

CM Cerebral malaria

Co-IP Co-Immunoprecipitation

CSA Chondroitin Sulphate A

CSP Circumsporozoite protein

DAPI 4,6-diamidino-2-phenylindole

DARC Duffy antigen receptor of chemokines

DBL Duffy binding-like

DBP Duffy binding protein

DHFR Dihydrofolate Reductase

DNA Deoxyribose nucleic acid

EBA Erythrocyte binding antigen

EBL Erythrocyte binding-like

FBS Fetal bovine serum

g Relative centrifugal force

gDNA Genomic deoxyribose nucleic acid

GFP Green fluorescent protein

H3 Histone 3

HA Hemagglutinin

HAT Histone acetyltransferase

HDAC Histone deacetylase

hDHFR Human Dihydrofolate Reductase

HMT Histone methyltransferase

Ig Immunoglobulin

IP Immunoprecipitation

iRBC infected red blood cell

IRS Indoor residual spraying

ITN Insecticide treated nets

JmjC2 Jumonji C-domain containing protein

KD Knock down

kDa Kilo Dalton

KO Knockout

LSD 2 Lysine-specific histone demethylase 2

LSD1 Lysine-specific histone demethylase 1

mAb Monoclonal antibody

Mb Megabase

MS Mass Spectrometry

O/N Overnight

OT Optimal Temperature

P. Plasmodium

Pb P. berghei

PBS Phosphate-buffered saline

PCR Polymerase chain reaction

Pf P. falciparum

PFA Paraformaldehyde

PfEMP1 *P. falciparum* erythrocyte membrane protein 1

PTEX *Plasmodium* Translocon of Exported Proteins

Pv P. vivax

PV Parasitophorous vacuole

RBC Red blood cell

rif Repetitive Interspersed Family

RPM Rotation Per Minute

RT Room temperature

SDS-PAGE Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SDS-PAGE Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Seq Sequencing

SUB Subtilisin

SWI/SNF SWItch/Sucrose Non-Fermentable

TF Transcriptional Factor

Var variant gene of *P. falciparum*

VSA variant surface antigens

WHO World Health Organization

WT Wild type

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Malaria: historical aspects.

The origin of the name Malaria comes from Medieval Italian: mala-ária "*bad-air*". Then, this disease was also denominated ague or marsh fever, due to its correlation with places where it appeared (swamps and marshland) (1).

The first report of the causative agent of malaria was given by the French army doctor Charles Louis Alphonse Laveran at the military hospital of Constantine in Algeria in 1880. Laveran observed a protist inside the red blood cells of an infected person and hypothesized that malaria was caused by this organism. Later, a Cuban doctor, Carlos Finlay, working with yellow fever patients at the Havana, provided results which hinted to mosquitoes as the transmissors of this disease (2).The complete life cycle of the malaria parasite in vertebrates and mosquitoes was then demonstrated by Sir Ronald Ross in 1897, at the Presidency General Hospital in Calcutta (3).These findings were confirmed by a medical board headed by Walter Reed in 1900.

Malaria has been a player in many different historical and socio-economical events, such as in several military campaigns (4). For example, malaria was the main health hazard for US troops in the Southern Pacific during the Second World War. About 500,000 men were infected and about 60,000 American soldiers died of malaria during campaigns in Africa and the South Pacific (5), (6). Over the late 19th and early 20th centuries, it was a major factor in the slow economic development of the South American nations (7), establishing a vicious circle still valid today, where poverty increases the risk of malaria, and malaria increases the poverty. The economic impact of malaria includes costs of health care, loss of working days lost due to sickness, days lost in education, in specific cases decreased productivity due to brain damage from cerebral malaria, and loss of investment and tourism.

The fight against malaria englobes several points of action which are: i) chemotherapy against the parasite in infected human hosts (8) ii) the use of insecticide-treated bed-nets in endemic areas, iii) vector elimination using insecticides (9), iv) awareness campaigns for populations at risk (10) and v) the development of vaccines (11).

Probably the first deliberate antimalarial treatment known in the occidental part of the world consisted of an extract from the bark of cinchona tree from the Peruvian Andes. It was introduced to Europe by the Jesuits around 1640 (12). Interestingly, the active agent quinine was isolated by the French chemists Pierre Joseph Pelletier and Joseph Bienaimé Caventou only in 1820 (13). Quinine was the major malarial medication until the 1920s. In the 1940s, chloroquine replaced quinine as the treatment of both uncomplicated and severe malaria (14). Also, the quinine treatment was used until resistance appeared, first in Southeast Asia and South America in the 1950s and then globally in the 1980s (14) (15). Resistance is now wide spread against all classes of antimalarial drugs including the first–line treatment drug artemisinin. Treatment of resistant strains became increasingly dependent on this class of drugs. Treatment failure upon treatment with artemisinin has been detected in Cambodia, Myanmar, Thailand, and Vietnam (16), and there has been emerging loss of artemisinin function in Laos (17). Some *Plasmodium* strains found on the Cambodia-Thailand border are refractory to combination therapies and may, therefore, become untreatable (18).

Vector control programs, together with the monitoring and treatment of infected humans, successfully eliminated or greatly reduced the malaria occurrence in Europe and parts of the United States, where in the early 20th century using methods as draining of wetland breeding grounds for agriculture, changes in

water management practices, and advances in sanitation, including greater use of glass windows and screens in dwellings and the appropriate use of the pesticide DDT, eliminated the disease from the remaining pockets in the South of United States in the 1950s as part of the National Malaria Eradication Program.

In others countries where malaria was also endemic, DDT was initially used exclusively to combat malaria, but its use quickly spread to agriculture. In time, pest control, rather than disease control, came to dominate DDT use, and this large-scale agricultural use led to the evolution of resistant mosquitoes in many regions.

Malaria vaccines have been a subject of hearty research. The first promising studies were done in 1967 by immunizing mice with live, radiation-attenuated sporozoites, which provided significant protection to the mice upon subsequent infection with normal, viable sporozoites (19). Later in the 1970s, a considerable effort was taken to develop a similar vaccination strategies for humans (20). The first vaccine, called RTS,S, was approved by European regulators in 2015 (21). The RTS,S/AS01 (commercial name *Mosquirix*) is the world's first licensed malaria vaccine and also the first vaccine licensed for use against a human parasitic disease of any kind. A pilot project for vaccination has been launched on April 23, of this year in Malawi. Ghana and Kenya are going to join the program in the coming month (22).

As malaria has been and remains a global health problem, the fight against the disease receives significant financial investments from different governmental or non-governmental organizations such as: African Leaders Malaria Alliance (ALMA), Bill and Melinda Gates Foundation, Comic Relief, Catholic Relief Services (CRS), GAVI, Goodbye Malaria Project, Lutheran World Relief (LWR), Malaria No More, *Médecins Sans Frontières* (MSF), Rotary International, United Against Malaria (UAM), among others.

Although in 2017 the total investment in malaria was about 3.1 billion US$, the disease still is a challenge for public health, predominantly in underdeveloped countries. Indeed, Malaria is considered endemic in around 104 countries causing 219 million cases and 435 000 deaths in that year, mostly children under five years or pregnant women (23).
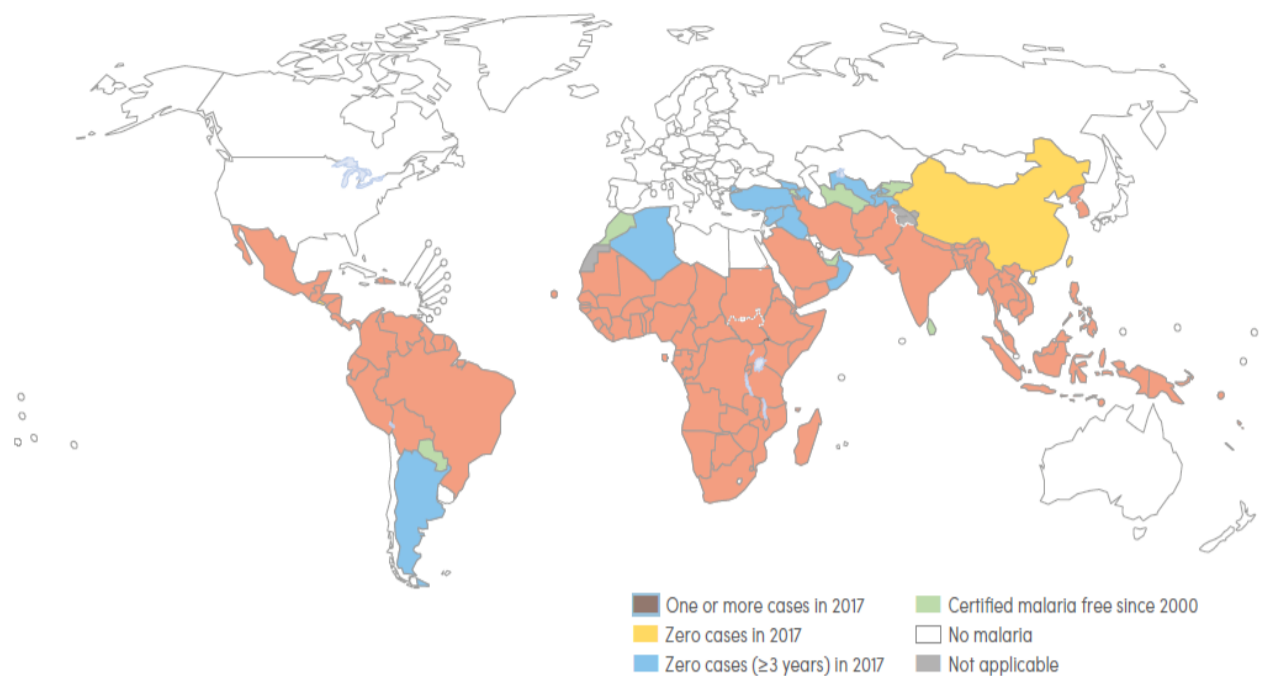
WHO: World Health Organization.

Figure 1. Countries with zero indigenous cases over at least the past 3 consecutive years are considered to be malaria free.
All countries in the WHO European Region reported zero indigenous cases in 2016 and again in 2017. In 2017, both China and El Salvador reported zero indigenous cases (23).

1.2. The Malaria disease.

Malaria is a mosquito (*Anopheles*) borne infectious disease, caused by unicellular parasites of the *Plasmodium* group, which belong to the Eukaryota Domain, Infrakingdom Alveolata, Phylum Apicomplexa, Class Aconoidasida, Order Haemospororida, Family Plasmodiidae, Genus *Plasmodium* (24). *Plasmodium falciparum* infections account for the majority of malaria deaths. The less virulent *P. vivax*, and probably *P. ovale*, also contribute significantly to morbidity (25).

Malaria disease can be categorized as uncomplicated or severe (complicated) (26).

All the clinical symptoms associated with malaria are caused by the blood stage parasites. When the parasite develops in the erythrocyte, numerous proinflammatory substances such as hemozoin pigment associated with *Plasmodium* genomic DNA, GPI (glycosyl-phosphatidylinositol)-anchored proteins and other toxic factors accumulate in the infected red blood cell (26). These factors are released together with infective merozoites into the bloodstream when the infected cells lyse. The hemozoin and GPI anchors stimulate the innate immune response in macrophages and other cells to produce cytokines and other soluble factors, which are perceived as fever and rigors and may in later phases result in the severe pathophysiology associated with malaria (27).

In the uncomplicated malaria the attacks last 6–10 hours. They consist of a cold stage (sensation of cold, trembling), a hot stage (fever, headaches, vomiting; even convulsions in young children); and finally a sweating stage (sweats, return to normal temperature, fatigues) (28). Classically, the attacks occur every second day and every third day depending of the type of parasite infection.

More commonly, the patient presents a combination of the following symptoms: fever, cold, sweats, headaches, nausea and vomiting and/or body aches and general sickness, but may also include weakness, enlarged spleen, mild jaundice, enlargement of the liver, or respiratory symptoms (26).

There are four manifestations of severe malaria caused by *P. falciparum*: placental malaria, cerebral malaria, severe anemia/renal failure, and respiratory distress. Cerebral malaria is indicated when the following symptoms occur: abnormal behavior, impairment of consciousness, convulsions, non-arousable coma, or other neurologic abnormalities (27). Severe anemia occurs due to hemolysis (destruction of red blood cells) and often presents with hemoglobinuria (hemoglobin in the urine). Acute respiratory distress syndrome (ARDS), an inflammatory reaction in the lungs that impairs oxygen exchange, may occur even after the parasite counts have decreased in response to treatment. In severe malaria, there is also often a dysregulation in blood coagulation and low blood pressure. Acute kidney injury may occur in consequence to a malaria pigment/hemozoin overload. Another frequently observed manifestation is metabolic acidosis (excessive acidity in the blood and tissue fluids), often in association with hypoglycemia (26),(29).

Other manifestations of malaria include neurologic defects which may occasionally persist following cerebral malaria, especially in children (28). Such defects include trouble with movements (ataxia), palsies, speech difficulties, deafness, and blindness. Recurrent infections with *P. falciparum* may result in severe anemia, and this occurs especially in young children in tropical Africa with frequent infections that are inadequately treated. Malaria during pregnancy (especially *P. falciparum*) may cause severe disease in the mother, and may lead to stillbirth, premature delivery or delivery of a low-birth-weight baby (26). In *P. vivax* infections, spleen rupture may occur (30). Nephrotic syndrome (a chronic, severe kidney disease) was reported in chronic or repeated infections with *P. malariae* (31). Hyper-reactive malarial splenomegaly (also called "tropical splenomegaly syndrome") occurs infrequently and is attributed to an abnormal immune response to repeated malarial infections. The disease is marked by a very enlarged spleen and liver, abnormal immunologic findings, anemia, and a susceptibility to other infections (such as skin or respiratory infections) (26).

1.3. The *Plasmodium* parasites.

Plasmodium is a genus of unicellular eukaryotes that are obligate parasites of vertebrates and insects. Over 200 species of *Plasmodium* have been described, many of which have been subdivided into subgenera based on parasite morphology and host range (32).

The lineages assigned to the genus *Plasmodium* are characterized by features such as schizogony, production of crystalline pigment, and gametocyte formation within blood cells (32). A limited range of vertebrates, including mammals, birds, and reptiles are susceptible to infection (33). *Anopheles* mosquitoes transmit parasites that infect humans, monkeys, and rodents, whereas *Culex* and *Aedes* mosquitoes predominate in the natural transmission to birds. The vectors of reptilian parasites are largely unknown (32).

The human infection malaria is caused by five *Plasmodium* species: *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi* (Figure 2). *P. falciparum* and *P. vivax* are responsible for most of the reported cases. Infection with *P. falciparum* is the cause of the most lethal forms of malaria (34).
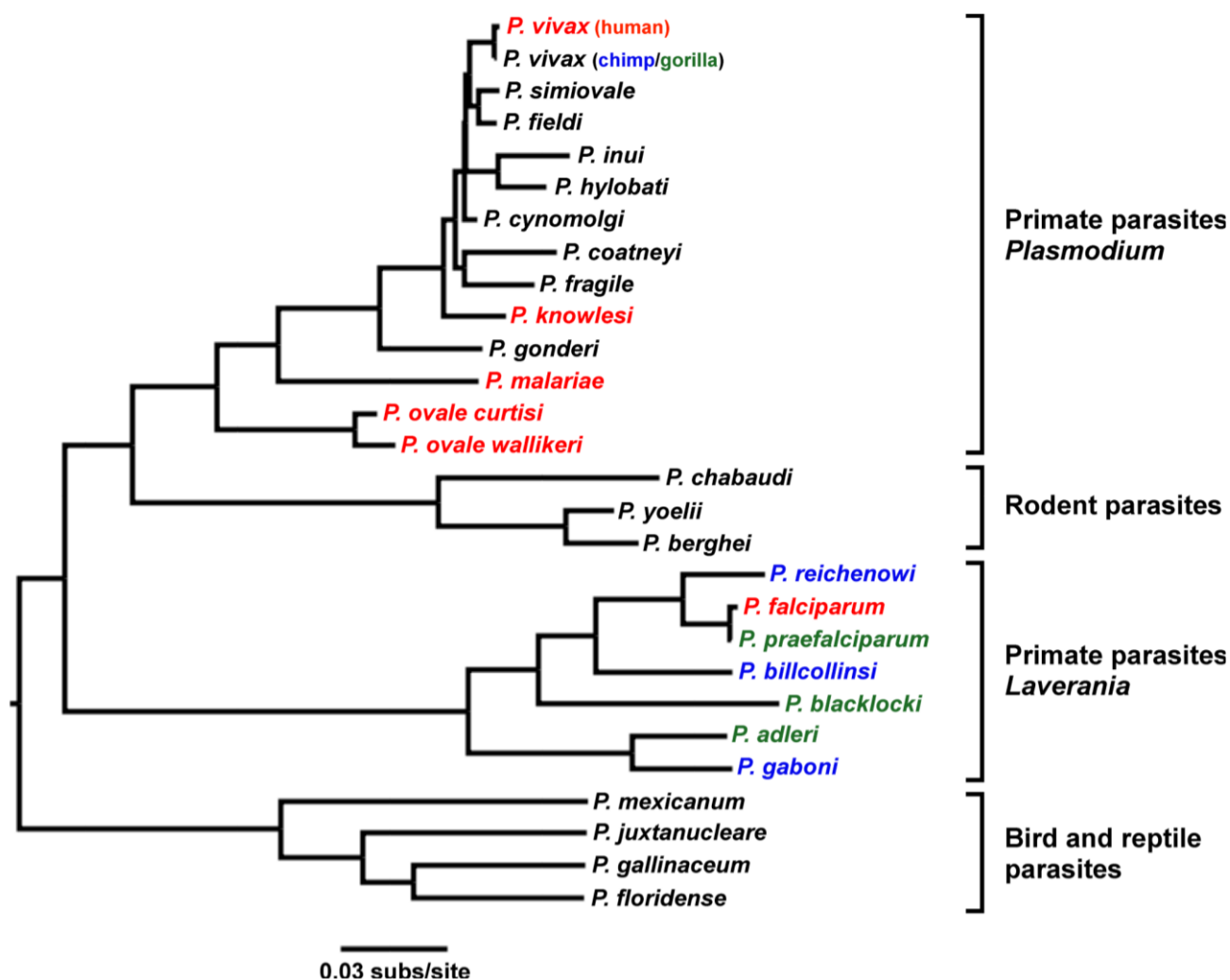
Figure 2. Evolutionary relationships of *Plasmodium* spp.
Colors highlight *Plasmodium* spp. that infect humans (red), chimpanzees (blue) and gorillas (green). Four groups of *Plasmodium* spp. are shown, with subgenus designations indicated for primate parasites. From Loy D.E., et al. 2017 (24).

In a general view, all *Plasmodium* species contain many features that are common to other eukaryotes, and some that are unique to their phylum or genus. The *Plasmodium* genome is separated into 14 chromosomes contained in the nucleus. *Plasmodium* parasites maintain a single copy of their genome through much of the life cycle, doubling the genome only for a brief sexual exchange within the midgut of the insect host (35). Attached to the nucleus is the endoplasmic reticulum (ER), which functions similarly to the ER in other eukaryotes. Proteins are trafficked from the ER to the Golgi apparatus which generally consists of a single membrane-bound compartment in Apicomplexans. From here proteins are trafficked to various cellular compartments or to the parasite or host cell surface (36).

Similar to other apicomplexans, *Plasmodium* species also possess the name-giving apical complex including rhoptries and micronemes which are important for invasion of target host cells. The bulbous rhoptries which contain parasite proteins involved in invading and modifying the host cell once inside (37). Adjacent to the rhoptries are the micronemes that contain parasite proteins required for motility as well as recognition and attachment to host cells (38). Spread throughout the parasite are secretory vesicles called dense granules that contain parasite proteins involved in modifying the membrane that separates the parasite from the host, the parasitophorous vacuole (38).

Plasmodium possesses also two membrane-bound organelles of endosymbiotic origin, the mitochondrion and the apicoplast, both of which play key roles in the parasite's metabolism. *Plasmodium* cells have a single large mitochondrion that coordinates its division with that of the *Plasmodium* cell (39). Like in other eukaryotes, the *Plasmodium* mitochondrion is capable of generating energy in the form of ATP via the citric acid cycle; however, this function is only required for parasite survival in the insect host, and is not needed for growth in red blood cells (39). The apicoplast, is

derived from a secondary endosymbiosis event, in this case the acquisition of a red algae by the *Plasmodium* ancestor (40). The apicoplast is involved in the synthesis of various metabolic precursors, including fatty acids, isoprenoids, iron-sulphur clusters, and components of the heme biosynthesis pathway (41). His main function, however, is the synthesis of isoprenoids and *P. falciparum* can be grown without apicoplast, if isopentenyl pyrophosphate is supplemented to the growth medium.
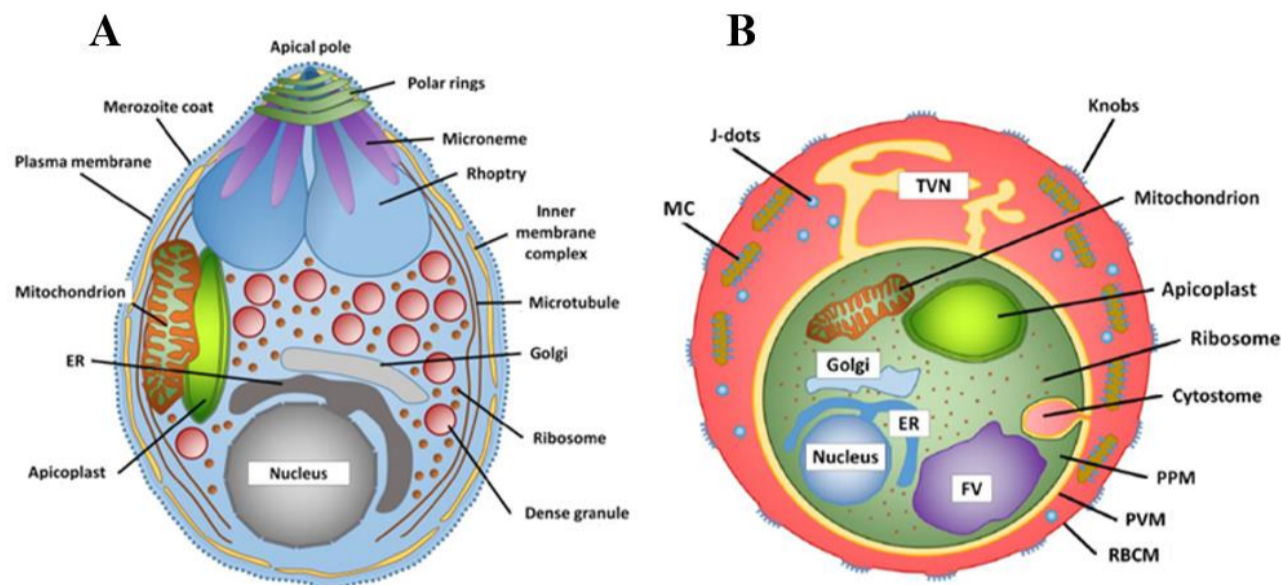
Figure 3.General schema of organelles arrangement in *Plasmodium falciparum* blood stage parasites.
A. The merozoite. **B.** The trophozoite infected red blood cell. ER, endoplasmic reticulum; FV, food vacuole; IMC, inner membrane complex; MC, Maurer's cleft; PPM, parasite plasma membrane; PVM, parasitophorous vacuole membrane; RBCM, red blood cell membrane; TVN, tube vesicular network. Adapted from Flammersfeld A., *et al.* 2018 (42).

1.4. Life cycle.

P. falciparum and *P. vivax* are the most relevant species of *Plasmodium*, since they are the responsible for the majority of the reported (and lethal) cases of human infections. The life cycle of these species are well characterized but each time new aspects come to light (25).

In a general view, the life cycle of *P. falciparum* is highly complex, involving several developmental stages in both the human host and the mosquito vector (figure 4). The malaria parasite is transmitted to the human host when an infected female *Anopheles* mosquito takes a blood meal and simultaneously injects a small number of sporozoites into the skin.

Once in the skin, sporozoites invade small blood vessels and are transported through the bloodstream to the liver. There, the sporozoites are able to pass through liver resident macrophages (Kupffer cells) and invade hepatocytes, where they start to replicate in a specific fashion called schizogony. Subsequently, thousands of merozoites are released into the bloodstream through so-called merosomes made of liver cell membranes which initially protect the non-motile merozoites form phagocytosis. Merozoites then invade red blood cells (43), (44). In the case of *Plasmodium vivax,* the parasite has a dormant stage in the human liver, termed hypnozoites (45). Hypnozoites may remain inside the hepatocytes and not develop into schizonts (46) for months, or even years and the mechanism behind the development into dormancy and the activation is not yet known (47). Nevertheless, activation of hypnozoites leads to a

novel malaria episode (relapse) and hypnozoites become important when transmission is limited to a few months per year in colder areas of the planet.

Once in the bloodstream, merozoites invade erythrocytes and multiply during an approximately 48-h, intraerythrocytic cycle. In this period the parasite progresses through the ring, the trophozoite stage and the schizont stage (schizogony), releasing then 16 to 32 merozoites into the bloodstream by a two-step, coordinated proteolytic rupture of the parasitophorous vacuole and the infected erythrocyte's membrane. Once in the blood stream, merozoites quickly invade new erythrocytes, restarting the intraerythrocytic cycle (48).

During each cycle, a small subset of parasites reprogram from asexual replication and instead produce sexual progeny that differentiate the following cycle into male and female sexual forms, known as gametocytes. A subset of parasites leave the peripheral circulation and enter the extravascular space of the bone marrow, where gametocytes mature and progress through stages I–V over the course of eight to ten days (gametocytogenesis) (46). In stage V, male and female gametocytes re-enter peripheral circulation, in which they become competent for infection to mosquitoes. Once ingested by a mosquito, male and female gametocytes rapidly mature into gametes (gametogenesis). Within the midgut, the male gametocyte divides into up to eight flagellated microgametes (exflagellation), whereas the female gametocyte develops into a single macrogamete. Fertilization of a macrogamete by a microgamete results in the formation of a zygote and develops into an invasive ookinete that penetrates the mosquito gut wall. The ookinete forms an oocyst within which the parasite asexually replicates, forming several thousands of sporozoites (sporogony). Upon oocyst rupture, these sporozoites actively migrate to the salivary glands, invade and then can be transmitted back to the human host during a blood meal (48) (45).
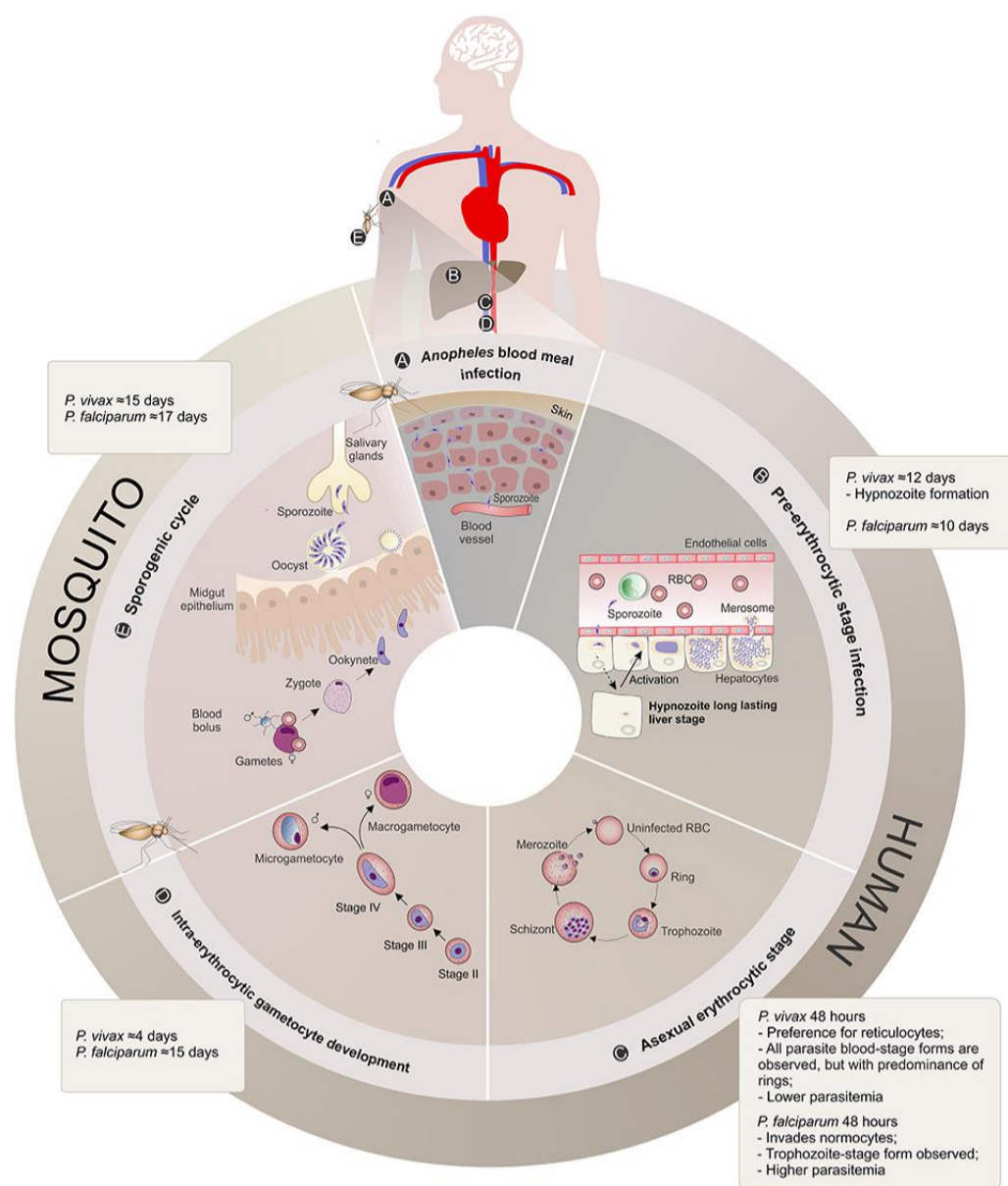
Figure 4. Life cycle of *Plasmodium falciparum /vivax*.
For details, see main text. *Plasmodium vivax* and *Plasmodium falciparum* life cycle comparison. **A**. *Anopheles* blood meal infection. **B.** Pre-erythrocytic stage infection. **C.** Asexual erythrocytic stage. **D.** Intra-erythrocytic gametocyte development. **E.** Mosquito Stages. From Bourgard C., *et al*. 2018 (49).

1.5. *Plasmodium falciparum* pathogenicity: adhesive phenotypes.

The pathogenesis of malaria is incompletely understood. During asexual multiplication and growth in erythrocytes, *P. falciparum* exports numerous proteins targeted to the host cytoplasm and membrane which modify the structure and function of host erythrocytes. The exported proteins then are believed to play key roles in virulence, growth and survival of the parasite (50).

The virulence of *P. falciparum* is mainly attributed to the ability of trophozoite parasitized-erythrocytes to adhere to and sequester in the microvasculature of various organs, including lungs, brain, placenta, and others (51).

The sequestration in deep venules enables parasite survival by avoiding spleen dependent killing and may lead to occlusion of blood flow and local endothelial cell activation (52). This phenomenon is termed as cytoadherence and is mediated by host-parasite molecular interactions provoked by knob-like protrusions on the surface of the infected red blood cell (iRBC). Specific parasite proteins on the surface of iRBC retain these cells to the microvascular endothelium of virtually every organ and tissue, by binding in an specific manner, to a diversity of receptors expressed on different host endothelial cells (27). The proteins of the *P. falciparum* erythrocyte membrane protein 1 (*Pf*EMP1) family mediate this adhesion through specific binding to multiple endothelial cell receptors, including CD36, intercellular

adhesion molecule-1 (ICAM-1), E-selectin, and CD31 (PECAM-1), and to placental chondroitin sulfate A (CSA). Cytoadherence of iRBC is quantitative and commonly seen in trophozoite or schizont infected RBC blood smears from *P. falciparum* infected persons (53).

The expression and interaction of a specific *Pf*EMP1 variant is believed to correlate to the organ-specific syndromes. For example, pregnancy-associated malaria is caused by the expression of the *Pf*EMP1 variant VAR2CSA, which mediates iRBC binding to placental CSA expressed by syncytiotrophoblasts (54). The case of cerebral malaria is induced by the expression of a *Pf*EMP1 variant containing Domain cassette (DC) 8 or DC13 domains (55), which mediate iRBC binding to EPCR (endothelial protein C receptor) on brain endothelium (56). This same *Pf*EMP1 or another variant may also contain DC4 domains, which may enhance the binding of the iRBC to the same brain endothelial cell by binding ICAM-1. In other organs, for example the lung, these same DC8 and DC13 expressing iRBC may contribute to disease but other receptor-DC-binding pairs are proposed to cause organ-specific clinical syndromes, namely respiratory distress (57). Like cerebral malaria, respiratory distress and dyserythropoiesis-associated anemia are organ-specific malaria syndromes that may occur alone or in combination with cerebral malaria. Moreover, adherent iRBCs also activate endothelial cells, leading to proinflammatory and procoagulant responses, reduced barrier function, and impaired vasomotor tonus.

Another adhesion property demonstrated in some *P. falciparum* isolates which has been associated with severe malaria, is the rosette formation. Rosetting is the spontaneous binding of uninfected erythrocytes to erythrocytes infected with mature asexual parasites (58). The molecular basis of rosetting is not well understood. A role for *Pf*EMP1 in rosetting was recently suggested by Rowe and colleges (59) and complement receptor 1 was found to be an associated host counterpart. Also a group of low-molecular-mass proteins called rosettins have been described as potential parasite ligands (60). Another multigene family, the RIFINs (*repetitive interspersed family*) may play a role in rosetting by mediating an *Pf*EMP1-independent vascular sequestration of iRBCs (61). Rosetting is also a virulence feature of *P. falciparum* since it may hinder the blood flow and lead to severe malaria.

Finally, one more variation of an adhesive phenotype, in where infected erythrocytes formed CD36 dependent autoagglutinates in nonimmune serum (62),(platelet-iRBCs association) named clumping and also appears to associate with severe malaria (63).
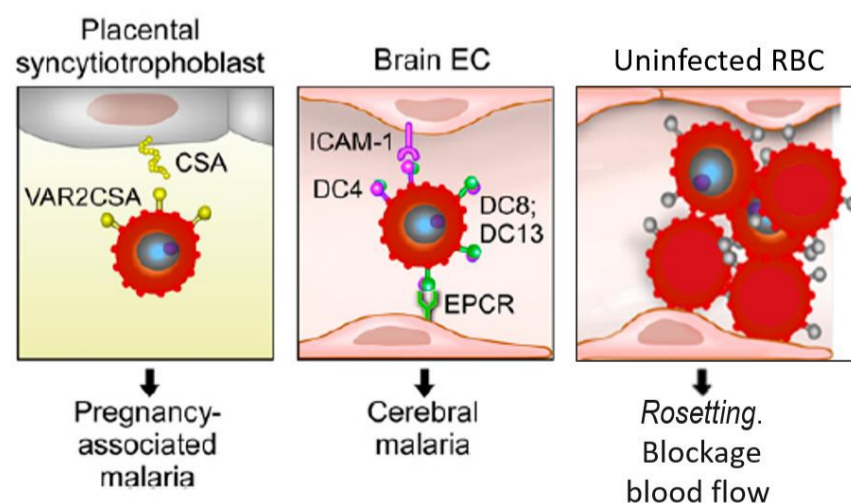
Figure 5. Schema of the organ-specific malaria syndrome by iRBC sequestration.
A. Pregnancy-associated malaria is caused by the interaction of the *Pf*EMP1 variant VAR2CSA and placental CSA receptor. **B.** Cerebral malaria is induced by the interaction of the *Pf*EMP1 variant

containing DC8 or DC13 domains with the EPCR endothelial protein C receptor in brain endothelium. **C.** Rosetting, is the spontaneous binding of uninfected erythrocytes to erythrocytes infected mature parasites. Adapted from Aird WC., *et al*. 2014 (57).

1.6. Antigenic variation in *P. falciparum*.

The expression of proteins on the surface of iRBC described above implies that the parasite becomes recognizable by antibodies. As an adaption, *Plasmodium* has evolved different strategies to escape the immune response and to guarantee the survival inside its vertebrate host. Among the evasion mechanisms, antigenic variation understood as the controlled temporal change of antigens exposed to the host's immune system, is the most intriguing way to evade the immune system (64). The perhaps most studied erythrocyte-surface exposed antigens are members of the *P. falciparum erythrocyte* membrane protein 1 family (*Pf*EMP1). *Pf*EMP1 proteins are involved as a central ligand in cytoadherence processes mentioned above. *Pf*EMP1 protein expression is subject to antigenic variation and of the 50-60 (*var*) genes which encode *Pf*EMP1, only a single allele is expressed per iRBC. Protein expression control occurs on the transcriptional level, meaning that of all *var* loci only one is active in a ring stage parasite. This mode of expression is termed allelic exclusion (65).

Other variant proteins, such as RIFINs expressed during the early trophozoite stage and STEVOR (*subtelomeric open reading frame*) at mature trophozoite stage (66), are also clonally variant and are probably also important in immune evasion (67).

1.6.1. Antigenic diversity developed from multicopy gene families and polymorphic alleles.

1.6.1.1. *Pf*EMP1and *var* gene family.

*Pf*EMP1 proteins are encoded by the large *var* gene family (68), (69), (70). The main part of the molecule is encoded by exon I of *var* genes. The ectodomain is exposed on the erythrocyte surface and is critical for host-cell receptor binding. Exon I encodes 2–5 Duffy-Binding-Like (DBL) domains and a cysteine-rich interdomain region between DBL1 and DBL2. Apart from short conserved sequence motifs in the first DBL domain (DBL1), exon I displays considerable sequence diversity between different *var* genes (55). In comparison, exon II, encodes an intracellular domain and is relatively well conserved between different *var* gene variants and comprises a segment rich in acid amino acids such as glutamate and aspartate (71).

Var genes are distributed throughout all 14 chromosomes (70), (72), (73), (74), (71). Each chromosome end typically contains one, two, or three *var* genes. Many subtelomeric chromosomes regions have two *var* genes arranged in tail-to-tail orientation relative to each other with one or more *rif* genes in between. Centromeric *var* genes can appear singly or in groups that are nearly always tandem arrays (head to tail), containing from three to seven *var* genes. The chromosomal location and transcription orientation of *var* genes can be predicted from its 5´ non-coding region sequence (75). Based upon sequence similarity, the 5´ promoter regions can be defined into four major upstream (Ups) sequence groups, UpsA, UpsB, UpsC, and UpsE. The former UpsD has been grouped with UpsA (76). As stated above, *var* gene expression is stage specific and tightly regulated by *in situ* activation and silencing involving dynamic epigenetic events, and transcriptional switching occurs at a rate of ~2% per generation *in vitro* (69), (62), (59), (77).

1.6.1.2. RIFINs and *rif* (*repetitive interspersed family*) gene family.

RIFIN proteins are products of the largest polymorphic multigene family, comprising approximately 150–200 genes per parasite genome (78). These proteins are expressed on the surface of iRBC, and have a molecular mass of 30-40 KDa. The RIFIN proteins are classified in two groups, the RIFIN A and B. The group RIFIN A consists of proteins which are exported to the iRBC surface. Meanwhile, the other group, RIFIN B, remains inside the parasite (79). This last group was also divided into three subgroups B1, B2 and B3 (78). There is compelling evidence that RIFINs are clonally variant, as different subsets of RIFINs are expressed on the erythrocyte surface by sibling clones derived from an isogenic background (80). The heterogeneity of RIFINs on the erythrocyte surface may arise from a selective repression of transcription of individual *rif* genes (81). The *rif* genes have a two-exon structure, the first exon encodes a predicted signal peptide and the second for polypeptides containing an extracellular domain with conserved cysteine residues and a highly variable region, a transmembrane segment, and a short cytoplasmic tail that is highly conserved (80).
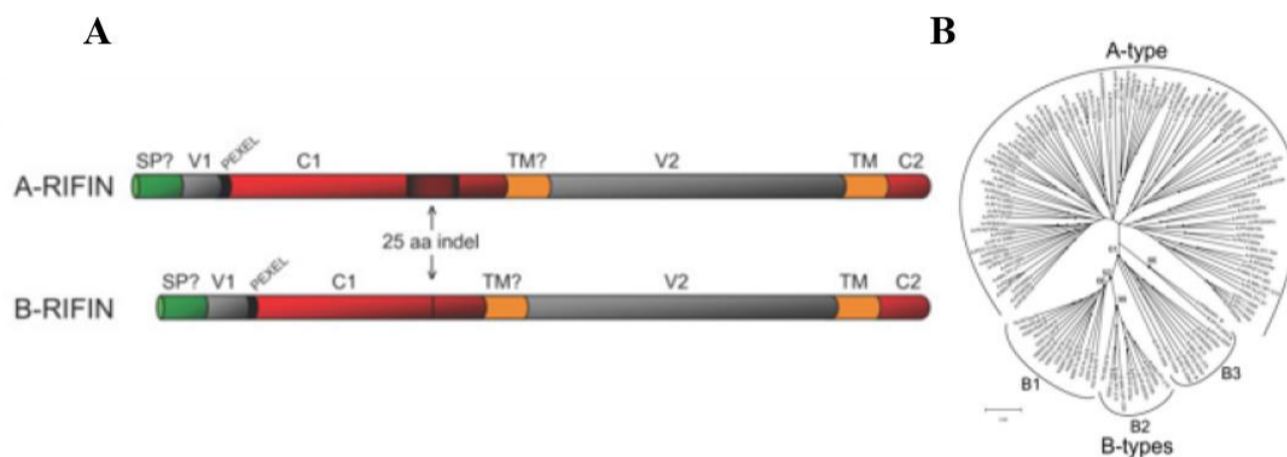
Figure 6. RIFIN proteins and its phylogenetic classification.
A. Graphic illustration of RIFIN proteins. *SP?:* potential signal peptide; *V1*: first variable region; *PEXEL*: *Plasmodium* export element; *C1*: first conserved region; *TM?:* questionable transmembrane region; *V2:* second variable region; *TM:* highly probable transmembrane region; *C2:* second conserved region. **B**. Phylogenetic tree of *rif* cDNA. The tree shows the segregation of A- and B-*rif* genes (gaps considered as complete deletions). The B-*rif* group is further subdivided into B1, B2 and B3 clusters. Adapted from Joannin N., *et al.* 2008 and 2011 (78) (82).

The function of the RIFINs proteins is not clear. It have been suggest that the RIFINs proteins could have a fundamental role in the rosetting phenomena by mediating *Pf*EMP1-independent vascular sequestration of iRBCs (61). Due to their apparent clonally variant expression, they may have a role in immune evasion and in cytoadherence (67) and therefore may be considered as an important virulence factor of *P. falciparum*. The *rif* genes have been shown to undergo expression switching (83). Recently, it was reported that a subset of RIFINs binds to either leucocyte immunoglobulin-like receptor B1 (LILRB1) or leucocyte-associated immunoglobulin-like receptor 1 (LAIR1). LILRB1-binding RIFINs inhibit activation of LILRB1-expressing B cells and natural killer cells. It was suggested that *P. falciparum* has acquired multiple RIFINs to evade the host immune system by targeting immune inhibitory receptors (84).

The few data addressing the transcriptional mode of *rif* genes appear to confirm that these genes are also transcribed in a clonal way as the *var* genes (85) (86). The transcription of all *rif* genes at the same time has never have been reported. Little is known about the mechanisms that regulate *rif* transcription.

Northern blot analysis using a degenerated *rif* probe showed that maximal *rif* gene transcription occurs during the intra-erythrocytic cycle, corresponding with the late ring to early pigmented trophozoite stage in the parasite line Palo Alto (83). At the moment, there are only few reports about the control of transcription of the *rif* gene family. One describes the suppressor elements in a *rif* promoter (81), and another proposes that the *rif* and *var* gene families share the same regulatory factors (87) (88). A later study by Katrin Witmer and colleagues (88) tried to monitor *rif* transcription in the same way as done before by Voss and colleagues using *var* promoters in bicistronic vector constructs (89). There, transfected parasites were selected using one drug, later, these parasites were selected for expression of the other resistance gene under the control of a variant gene promoter. The study by Howitt and colleagues used a monocistronic expression vector and varied only the quantity of blasticidin to modulate the *rif* promoter-driven expression of blasticidin deaminase (87). This required that the plasmidial *rif* promoter must be transcriptionally active all the time, not permitting a perfect "off" state of activation. The second study by Witmer and colleagues did not succeed at all in activating the transgene locus containing a *rif* 5' upstream region. Both groups did not consider differential classification of *rif* genes and the different peaks of transcription post erythrocyte invasion. Recently, in 2018 we proposed a differential control of transcription according to the gene group classification. Based on the data obtained by means of an RNA-seq analysis of a transgenic *P. falciparum* cell line, with contained modulatable *rif* promoter (*PF*3D7_0200700), "on/off" state of activation. We suggested, that while in one of the *rif* groups, the transcription may follow the principle of allelic exclusion, the transcriptional regulation of the other group (*rif A*) may be different. Nevertheless, more studies should be done regarding the transcriptional control of the *rif* genes (90).

1.7. Transcriptional control of variant genes in *Plasmodium falciparum*.

In *P. falciparum,* slow adaptation typically happens at the genetic level. The genetic bases of adaptation have been extensively studied (91), (92). However, the adaptations to pressures of quickly changing environments such as the human-mosquito transition, requires a much faster change which must also be reversible. In *P. falciparum* these responses work like in many other organisms at the transcriptional and posttranscriptional level (93). Clonally variant gene families trigger phenotypic plasticity in *Plasmodium falciparum*, an indispensable process for adaptation and survival in the human host. Switching the expression of variant molecules prolongs the time of infection and generates alternative pathways for the evasion of immune host response (94)(95) .

Clonally variant gene expression in *P. falciparum* was initially described for *var* genes (96). *var* genes are expressed in a mutually exclusive manner, with only one *Pf*EMP1 protein expressed by any individual parasite (94)(71). Mutually exclusive gene expression refers to the ability of an organism to select one member of a large, multicopy gene family for expression while simultaneously silencing all other members of the family (97). Dzikowski and colleagues showed that this process is regulated entirely at the level of transcription, and that protein production and chromosomal context of the genes are not involved. In addition, they identified the DNA elements required for a *var* gene promoter to be recognized and co-regulated along with the rest of the family (97).

The mutually exclusive *var* gene transcription is under the control of epigenetic factors (98). Among these factors, chromatin modifiers (DNA acetylation/methylation), yet unknown transcription factors,

nuclear architecture and noncoding RNAs play a role. These factors can alter the chromatin structure directly by modulating the interactions of proteins with DNA (99).

Table 1. Epigenetic factors coordinate the gene expression of variant genes in *Plasmodium falciparum*.

The chromatin structure.	Alter DNA accessibility (100): condense chromatin (heterochromatin) and the relaxed chromatin (euchromatin). The heterochromatin remains supposedly transcriptionally silent while the euchromatin has active transcriptional state. The heterochromatin structural state prevents the access to chromatin remodelers, whereas the access is allowed by the euchromatin state (101). Chromatin remodelers are ATP-dependent proteins that also trigger the post-replicative replacement of canonical histones with histone variants. For example, the SWI2/SNF2 family of protein complex mobilize the nucleosome along the DNA strand by ATP hydrolysis. In *Plasmodium falciparum* there have been identified 11 chromatin-remodeling ATPases of the SWI2/SNF2 family including orthologs of CHD-1, Brahma, ISWI, SNF2L, ELT-1, RAD5, RAD16, and HARP from crown group eukaryotes (102) (103).
Transcription factors (TFs).	*P. falciparum* encodes 27 plant-like transcription factors of the Apetala family, termed ApiAP2 TFs (104). These putative TFs are sequence-specific regulators of gene expression that act directly or indirectly with chromatin remodelers. Some *Pf*AP2 TFs are involved in the epigenetic regulation via heterochromatin formation, genome integrity of subtelomeric region and targeting silenced *var* genes to heterochromatin rich clusters (101). Gene expression data indicated that different ApiAP2 family members were expressed in different stages during parasite development, suggesting that they might be involved in life cycle progression and differentiation processes, as reported for their plant homologues (105). Other TFs identified in *Plasmodium falciparum*, are *Pfmyb1*, involved in intraerythrocytic development and two non-sequence specific TFs that contain the mobility group box (HMGB) motif, involved in the intraerythrocytic development and sexual stages (106).
Non-coding RNAs (ncRNA).	These molecules interact themselves directly associating with the chromatin, and recruiting repressive modify histone complex (107). Recent evidence shows that long ncRNAs are associated with *P. falciparum* centromeres. It was proposed that ncRNAs could be involved in the telomeric silencing, in absence of a functional RNA interfaces (RNAi) system. The suggested mechanism of ncRNAs silencing activities, include a polymerase obstruction and/or the formation of duplex RNA structures to prevent the translation process (108). Overexpression of GC rich ncRNAs in *P. falciparum* led to a break in allelic exclusion (109).
Histone	The acetylation and methylation of histones alters the transcriptional activity of

Modifications (PTMs).	promoters associated to the respective chromatin. The proteins that perform post-transcriptional modifications (PTMs) are chromatin modifiers, the chromatin "readers" and "writers". PTMs have been reported to be involved in the regulation of variant gene expression of *P. falciparum*. The PTMs affect the histones, resulting in the remodeling of the chromatin into either euchromatin (accessible) or heterochromatin (inaccessible). The chromatin "readers" and "writers" cause reversible modifications in the chromatin and promote the recruitment of TFs and RNA polymerase complex. The chromatin "readers" can be classified in several groups based in their conserved domain: 1) Bromodomains, 2) Royal Superfamily proteins, 3) 14-3-3 proteins, 4) Plant Homodomain (PHD) fingers, 5) WD40 Repeat Containing proteins and 6) Chromodomains. These domains recognize and bind at specific histone PTMs. These interactions allow the recruitment of proteins associated with the chromatin(110). In *P. falciparum,* several chromatin readers such as *PfMYST* (a histone acetyltransferase) and *PfHP1* (Heterochromatin 1) have been reported, the latter is associated with subtelomeric and intrachromosomal silenced *var* genes. *Pf*SET1 (histone deacetylase) and *Pf*GCN5 (histone acetyltransferase) contain a single bromodomain (110). The proteins that modify histone PTMs at a specific residue on the histone tail, are denominated chromatin "writers". These modifications modulate the chromatin dynamics, creating an epigenetic mark for a specific effector protein, or for an specific a reader recognition (110).

In the context of the control of *P. falciparum* variant gene expression, the acetylation and methylation of histones are decisive for the three different transcriptional states of *var* genes. An actively transcribed *var* gene in ring stages is associated with H3K4me2, H3K4me3, and H3K9ac in 5'-flanking regions. In the trophozoite/schizont stage the previously active *var* gene is transiently silenced (poised) but maintains the enrichment of H3K4me2. This association may transmit memory of the active *var* gene during cell division. Stably silenced *var* genes (5'-flanking region and exon 1) are enriched in H3K9me3 throughout the asexual life cycle (98). Likewise, an association of the dynamic expression pattern with activation/silencing marks of the *rif* gene family was also reported (86).
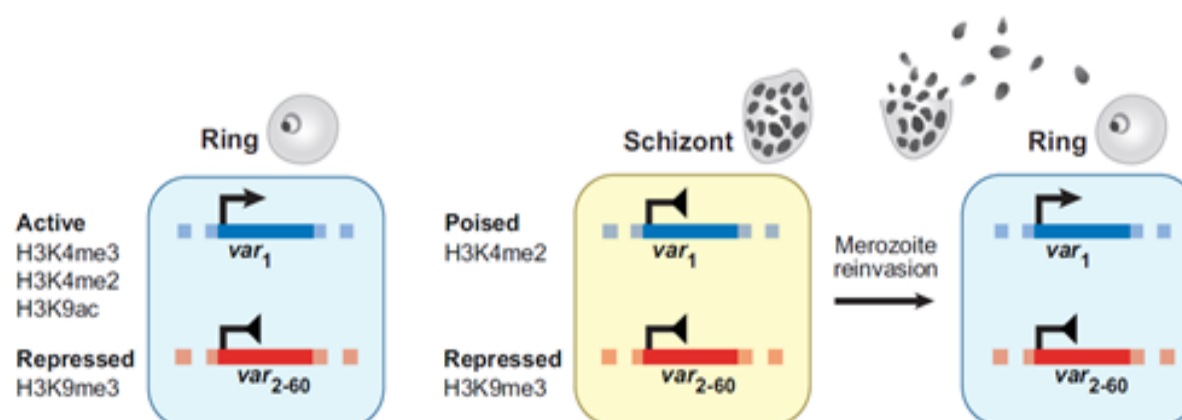

Figure 7. Histone marks linked to *var* gene expression.
Adapted from Scherf A., *et al*. 2008 (98).

34

1.8. *Plasmodium falciparum* histone methylation.

In addition to the acetylation by histone deacetylases (HDACs) (111), (112), the methylation and demethylation via histone lysine methyl transferases (HKMTs) or histone lysine demethylases (HKDMs), are supposed to have critical roles in controlling gene expression in *P. falciparum* (113), (114), (115), (116).

Methylation of histones (a process by which methyl groups are transferred to lysine side chains of histones) can either increase or decrease transcription of genes, depending on which amino acids in the histones are methylated, and how many methyl groups are attached. Methylation may weaken the electrostatic attractions between histone tails and DNA, leading to an increase of transcription, since this allows the DNA to uncoil from nucleosomes, so transcription factors and RNA polymerases can access the DNA. Histones can be methylated on lysine (K) and arginine (R) residues only, but methylation is most commonly observed at lysine residues of histone tails H3 and H4 (117). Lysine is able to be mono- di-, or tri- methylated with a methyl group replacing each hydrogen of its NH_3^+ group (118). There are a total of ten predicted *P. falciparum* HKMTs (*Pf*HKMTs) belonging to the SET domain superfamily (115), (116). As acetylation, histone methylation is a dynamic and reversible (demethylation) process under enzymatic regulation, which catalyzed the placement or removal of mono-, di- and trimethyl groups, positioned into the amino acids residues at the histones N-terminal tails (119). The demethylation play an important role via regulation of steady-state levels of histone methylation (118).

P. falciparum possess two families of lysine demethylases, which are able to demethylate the H3K4 methyl residues. These lysine demethylases employ two different molecular mechanisms of demethylation: the lysine-specific demethylases 1 (LSD1) and Jumonji C (JmjC) - Domain histone demethylases (JHDMs) (120). *P. falciparum* genome encodes at least one LSD1 (*PF*3D7_1211600) and two JHDMs (*PF*3D7_0809900 and *PF*3D7_0602800).

The lysine demethylases Jumonji C –Domain (JHDMs) family, belong to the ion-dependent di-oxygenase superfamily, which possess a metalloenzyme catalytic motif (121). The JHDMs have limited substrate affinities, and often these enzymes target a single lysine residue at the time. As no free electron pair on the nitrogen atom is needed, the JHDMs are capable of removing the methyl group of any of the three states of methylation (119). These enzymes use as cofactor α-ketoglutarate and ion (Fe^{2+}). The demethylation reaction occurs in the presence of oxygen and generates an ion-oxo intermediate (122). The two *P.f* JHDMs, *Pf*JmjC1 and *Pf*JmjC2, seem to have specificity for the H3K9 and even higher for the H3K36, due to the serine-alanine residue substitution in their substrate-binding pocket (115).

The other lysine demethylase present in *P. falciparum* genome, is the Lysine-Specific Demethylase 1 *Pf*LSD1 (*PF*3D7_1211600). This enzyme has not yet been characterized. In a general context, LSD1 is a highly conserved protein (123). LSD1 belongs to the family of flavin-containing oxidases. The amine oxidase-like domain (AOL) is located at its C-terminal end. It is homologous to other FAD-dependent oxidases and consists of two subdomains: a FAD binding domain and a substrate binding domain (124). This enzyme is able to remove methyl groups from histone tails with specificity toward H3K4me1/2 (125), but it is not able to turn over H3K4me3 (126). Due to interactions of the H3 methyl lysine with FAD and the N-terminus of H3 with the anionic pocket, no more than three residues on the N-terminal side of the methyl lysine are able to enter the catalytic center of LSD1 (127). However, LSD1 is also able to demethylate K370 of the non-histone substrate p53 (128). In mammalian cells, the LSD1 enzyme initiates demethylation by generating an imine intermediate from the methylated amino group of the

lysine (Figure 8). LSD1 was found to act in protein complexes, and it seems to change its substrate specificity depending on the associated protein complex. For example, LSD1 which acts specifically with H3K4, possesses a higher affinity toward H3K9me1/2, when interacting with the androgen receptor (AR) (129). In the case of the interaction of LSD1 with CoREST complex, the activity of LSD1 can be modulated by other histone modifications (123). The acetylation occurrence at H3K9 increases the catalytic activity of LSD1 (130). Moreover, LSD1 was also found to act in multiprotein complexes.
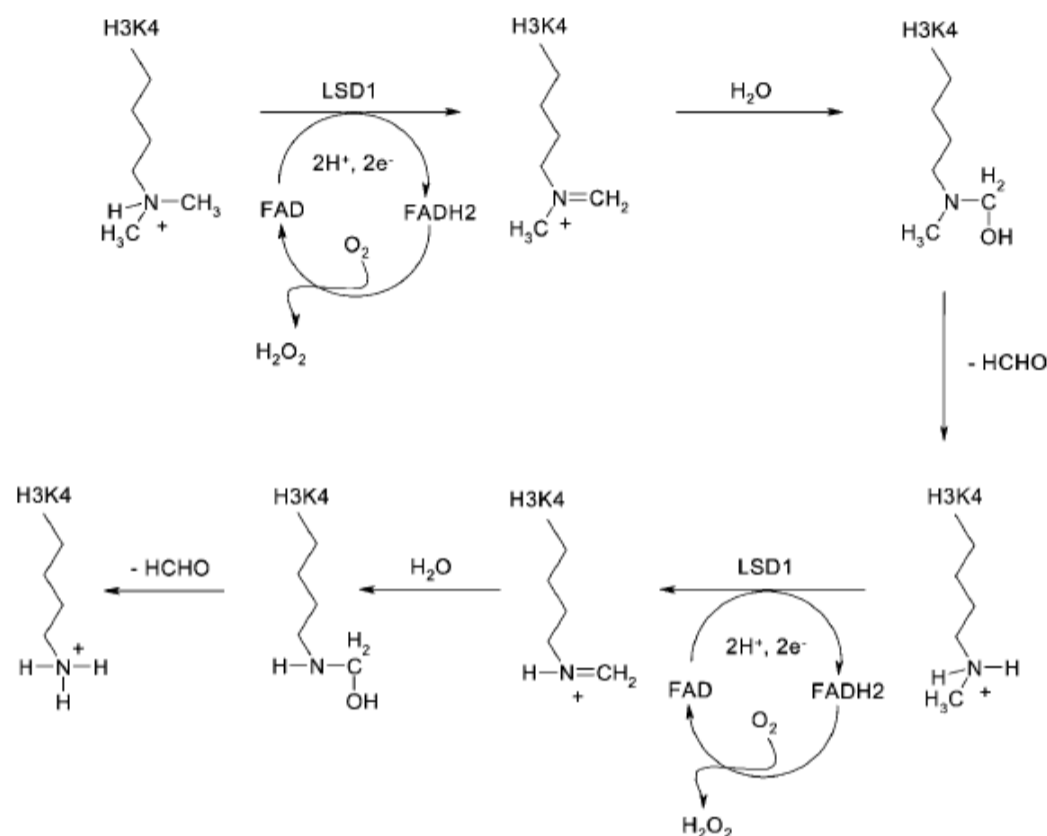
Figure 8. Biochemical steps in the FAD-dependent lysine demethylation by LSD1 (122).

LSD1 is an essential factor in the mammalian biology and has been related to many different specific roles. LSD1 mRNA and protein is highly expressed in undifferentiated human embryonic stem cells and is progressively downregulated during differentiation, suggesting a significant role in development (131). LSD1 is an important regulator of neural stem cell proliferation (132), and neuronal development. Tissue-specific knockouts have shown that LSD1 is required for proper differentiation of endocrine cells of the anterior pituitary (133), for proper function of oocytes (134) and for normal bone marrow function (135). The *C. elegans* homologue of LSD1, *spr-5,* regulates notch signaling (136) (137), and maintains transgenerational epigenetic memory and fertility (138). In yeast *spLsd1/2* and Drosophila *Su (var)3-3* homologs regulate heterochromatic gene silencing (139), (140) and the size of the germline stem cell niche in Drosophila ovary (141). LSD1 has also been reported to have a role in the DNA damage response (142), the circadian cycle (143) and in repression of a mitochondrial metabolism and lipid oxidation energy expenditure program in adipocytes or liver cancer cells (144) (145).

In *P. falciparum*, the highest levels of the *PfLSD1* mRNA have been found in the ookinete and oocyst stage, during the sexual phase of the parasite life cycle, and in the trophozoite stage in the asexual period (146). *PfLSD1* seems not essential for parasite survival (*http://plasmodb.org/plasmo/app/*

36

record/gene/PF3D7_1211600). Volz e colleges, investigated the nuclear locations of 12 *P. falciparum* proteins, between them *Pf*LSD1, which appear to be localized, in an area surrounding and partially overlapping with DAPI staining; the nuclear periphery in trophozoite/schizont stages. *Pf*LSD1 localization overlapped also with *Pf*SET2, *Pf*SET3, and a nucleosome assembly protein. This nucleolar area is characterized by the presence of both histone marks H3K9m3 and H3K9ac. All this suggest a probable dual role, by promoting both gene repression via its H3K4 demethylase activity, and transcriptional activation due its H3K9 demethylase ability, in this stage of the parasite (147). An interaction with the a putative *Pf*SNF2 helicase was also suggested, since this is the *P. falciparum* homologue of the ISWI protein, an important component of the nucleosome remodeling factor complex (148). In apicomplexans SWI/SNF factors are most likely core subunits of large functional complexes that include other chromatin-modifiers, which facilitate the perturbation of chromatin structure *in vitro* in an ATP-dependent manner (103), permitting a change of transcriptional activity of the associated DNA.

Since many recent publications point to reversible chromatin changes, such as histone methylation and acetylation, as important control elements of *P. falciparum* gene silencing and monoallelic activation (98). Also, enzymes that add or remove acetyl and methyl marks in *P. falciparum* have been identified (115), (149), (150). In addition, it has been demonstrated that trimethylation of histone H3 lysine 9 (H3K9me3) is associated to transcriptionally silent *var* genes (113), (151). Upon *var* gene activation, in the 5' flanking region (5'UTR), methylation is replaced by acetylation at lysine 9 of histone H3 (H3K9ac), and histone H3 lysine 4 is di- and trimethylated (H3K4me2/3) (151). Furthermore, an association of the dynamic expression pattern with activation/silencing histone marks of the *rif* gene family was also shown (86). Given its apparent activity as a Histone-lysine modifier, and being not essential for asexual *in vitro* growth. Is possible that *Pf*LSD1 may be involved in variant gene regulation and epigenetic memory maintenance.

Here, we want to elucidate if *Pf*LSD1 influences the transcription or switching of *rif* genes and/or other variant gene families, since in other organisms LSD-1 is a demethylase which acts with differential affinity in the H3K4m1/2 and H3K9m1/2 histone marks, also related to variant gene transcription.

2. OBJECTIVES

2.1. Control of *rif* gene transcription.

2.1.1. General objective.

Provide insights on the mode of transcription of *rif* genes of *P. falciparum.*

2.1.2 Technical procedure.

1. Modify a *rif* locus by single recombination, turning it inducible.

2. Study the transcriptional profile of the *rif* gene family, by activation of a *rif* promoter controlling a resistance gene.

2.2. The role of *Pf*LSD-1, a putative histone modifier, in the transcriptional memory/switching of *rif* and other variant genes.

2.2.1 General objective.

Study the role of LSD-1, a putative histone modifier, in the transcriptional memory/switching of *P. falciparum* variant genes.

2.2.2 Technical procedure.

1. Transform the *P. falciparum lsd-1* locus by single crossover recombination, turning it regulatable.

2. Evaluate if the sudden absence of the *P. falciparum* LSD-1 results in a remarkable phenotype by knockdown of lsd-1 transcripts.

3. Analyze the participation of *Pf*LSD-1 in the transcription of *P. falciparum* variant genes.

4. Study the possible mechanisms of action of *Pf*LSD-1 by protein-protein interactions and chromatin-protein interactions approaches.

Considering that LSD1 is strongly transcribed in sexual stages (ookinetes and oocysts) which cannot easily be observed in *P. falciparum,* we also used the murine model of infection, *P. berghei*, to study the possible role of LSD1 in the sexual development; with the intention to later extrapolate the results to *P. falciparum*. For this we proposed the following procedures:

1. Modify the *P. berghei lsd-1* locus by double cross over recombination, interrupting the *Pblsd1* locus.

2. Evaluate if in the lack of the *P. berghei* LSD-1 results in a notable phenotype.

3. Analyze the participation of *Pb*LSD-1 in the viability and differentiation of the asexual and sexual stages of *P. berghei.*

3. MATERIALS AND METHODS

Here, Materials and Methods not included in section 4.1 are described.

3.1. Polymerase chain reaction (PCR).

Total gDNA was isolated (as described in 4.1) from *P. falciparum* NF54 and *P. berghei ANKA* lines, and used as template to obtain the homology regions for the transfection vector construction. PCRs were performed using a Taq DNA Polymerase from *Sinapse Inc* with the following thermocycling general program: 95 °C, 2 min s; 95 °C, 45 s; AT °C, 40 s; 72 °C, 1 min, over 30 cycles. In table 2, the set of oligos and the optimal annealing temperature (AT) for each set of primers are shown.

Table 2. Oligos used to amplify targets from genomic DNA for the construction of the transfection vectors.

Gene ID	Oligo Fw	Oligo Rv	AT °C	Size (pb)	Construct
PF3D7_0200700	CTGCAGATATAAATTTGT AAAAAACCATGTG	GGATCCTTAATTGTGATACG TATATTATTTAATG	56	1488	pTZ57$_{PF3D7\_0200700}$
PF3D7_0900500	CTGCAGTATTATATTTTTA TATATAATTATTCGTG	GGATCCATTTAATGTGATAC TTATATTATTTTATG	54	1451	pTZ57$_{PF3D7\_0900500}$
PF3D7_1300400	CTGCAGATGTAATATATT ATTATGTTAATATTC	GGATCCTATTGTGATACGTA TATTATTTTATG	54	1457	pTZ57 $_{PF3D7\_1300400}$
PF3D7_0632200	AGATCTCCATTATATTAAT ATATTATTGTTTGCTC	CTGCAGGTTCGTTTAATAAT TTCATAAATTGC	56	1291	p*rif6*GFPHA$_{2A}$-BglmS
PF3D7_0100400	AGATCTCCATTATACTAA TATATTATTGTTTCCTC	CTGCAGCTTCTTCTAATAAT TTTATATATTGGAG	56	1276	p*rifr* GFPHA$_{2A}$-BglmS
PF3D7_1211600.1	AGATCTGAGCAATATTAG TACAGATGAAT	CTGCAGGAGCAACACTTCTT GATAAGTG	58	981	p*Pf*LSD1GFPHAglmS
5´UTR PBANKA_0610100.1	GGGGGTACCGGATTTTGT CACTTGTGTTCGTC	GGGCTCGAGGATAAATGTA TCTACATTCAAGCATAAGCA	56	963	*Pb*LSD1KO
3´UTR PBANKA_0610100.1	GGGAGATCTGCAATACCA TTTGAATCCATGTGT	GGGGAATTCGCTGCCTCCTT ATTTTATGTGTATCC	56	755	

All the amplified fragments and the identity of the sequences was confirmed by semiautomatic sequencing in an Applied Biosystems 7550 sequencer (CEGH/IB-USP).

3.2. Plasmid constructs.

3.2.1. pTZ57 plasmid derivates (see section4.1).

3.2.2. p*rif*GFPHA2A-BglmS plasmid.

The p*rif*GFPHA2A-BglmS plasmid is a derivate of the pTEX150glmS plasmid. The plasmid was constructed introducing an approximately 1.5 Kb DNA fragment of two different *rif* 3' ORFs via Bgl II and Pst I (*PF*3D7_0100400 and *PF*3D7_0632200:1), linked with the GFP reporter gene, through Pst I and Mlu l. Connected to this is the 72pb 2A domain sequence, via Spel and Agel, and the Blasticidin-deaminase (BSD) cassette cloned via Age l and Xho l sites. Finally, the glmS sequence was inserted together with the *P. berghei* hsp90 terminator sequence. As a selection marker the hDHFR gene under the control of the calmodulin promoter was used.

When the single recombination event takes place (mediated by the *rif* homology region), the modified parasites will produce a fused protein, RIFIN-GFP-2A-BSD. The BSD domain is then cleaved and remains in the cytosol while RIFIN-GFP is expected to be trafficked normally. This vector also has the glmS sequence which forms an inactive ribozyme once in RNA form. This means that in the presence of glucosamine (2.5mM) (152) the polyA tail is cleaved off and the transcript becomes unstable. In this sense, modified parasites will be able to modulate the quantity of the *rif*-GFP-2A-BSD transcript in a post-transcriptional manner.
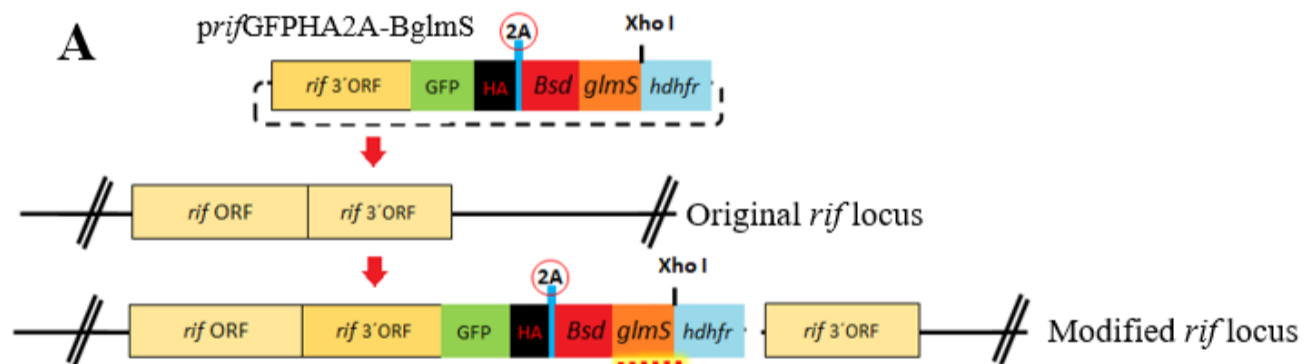
Figure 9. Schematic representation of a single crossover, using the p*rif*GFPHA2A-BglmS vector. See text for details.

3.2.3. p*Pf*LSD1GFPHAglmS plasmid.

The p*Pf*LSD1GFPHA-glmS plasmid is based on the pTEX150glmS plasmid back bone. This plasmid was constructed by introducing ~1Kb DNA fragment of the *Pf*LSD1 3' ORF (*PF*3D7_1211600) via Bgl ll and Pst l, linked with the GFP reporter gene and the influenza virus hemagglutinin (HA) epitope fragment sequence (84 pb), via Pst I and Mlu l, to monitor *Pf*LSD1 protein expression. Finally, the glmS

sequence and the *P. berghei* terminator was cloned. As before, the hDHFR gene under control of the caldmodulin promoter was used as a selection marker.

This vector contains also the glmS sequence, which allows to regulate gene expression in response to the glucosamine metabolite, as explained before.
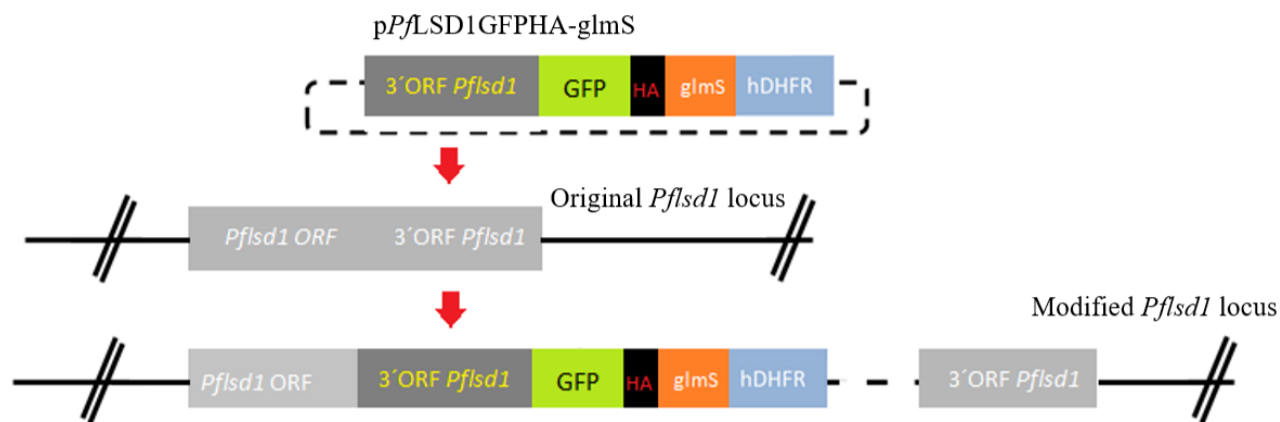
Figure 10. Schematic representation of single crossing over using the p*Pf*LSD1GFPHA-glmS vector. See text for details.

3.2.4. p*Pb*LSD1^{KO} plasmid.

To generate the targeting sequence to knockout LSD1 in *P. berghei*, the *Pblsd1* 5′UTR (963 bp) and 3′UTR (755 bp) were used as homology sequences flanking the hDHFR and mCherry cassettes. This plasmid for knock-out strategies in *P. berghei*, was already available (153) and was kindly provided by professor Daniel Y. Bargieri from the *Laboratório de genética molecular* at the Department of Parasitology/ICB-USP.
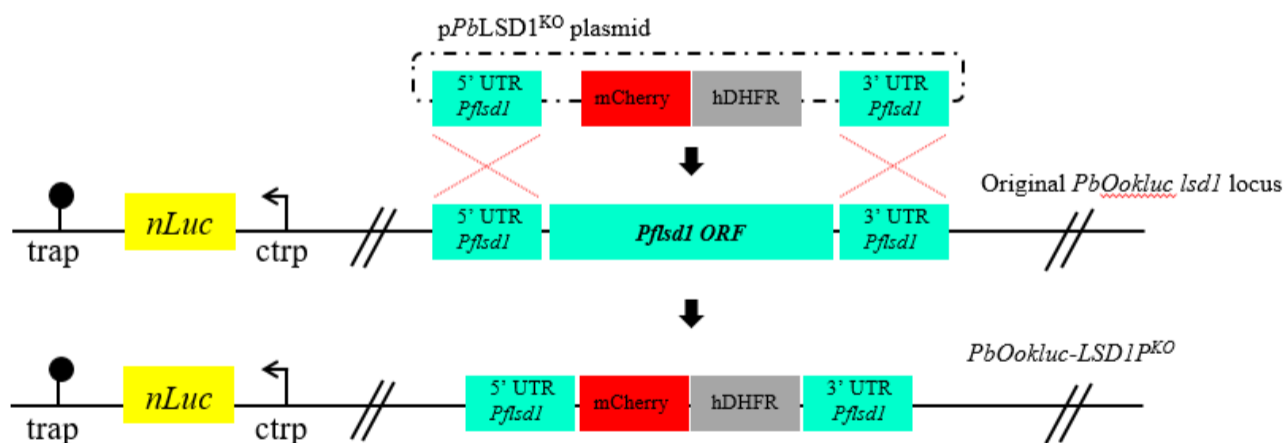
Figure 11. Schematic representation of double crossing over using the pPbLSD1^{KO} plasmid vector, in a *P. berghei-Ookluc* transgenic cell line.
Adapted from (154).

43

*All the constructs were done by cloning the DNA fragments with ligase (*Thermo Fisher Scientific*), in a ratio of 4:1 insert:vector, respectively, and overnight incubation at 4°C. The bacterial transformation was done using $CaCl_2/MnCl_2$ chemocompetent *E. coli* DH10B cells (155).

3.3. *P. falciparum in vitro* culture.

3.3.1. Parasite culture and transfection.

The parasite culture and transfection is described in section 4.1. In the case of parasites transfected with p*R*GFPHA$_{2A}$-BglmS plasmid or p*Pf*LSD1GFPHAglmS, the selection of transfectant parasites was done adding WR99210 at 2.5 nM to the medium, instead of 2.5µg/ml of Blasticidin.

In order to favor the integration via single crossover recombination, transfected parasites were cultivated for 14–20 days without WR99210, and then the drug was added again. Normally, after three cycles locus-integrated parasite lines were obtained. Then, the modified parasites were selected by cloning via limiting dilution (156).

3.3.2. Parasite synchronization.

3.3.2.1. Sorbitol-synchronization of *P. falciparum* infected erythrocytes.

To establish synchrony, the culture was centrifuged at 200g for 5 min, the supernatant discarded, and the pellet (-0.5 ml) resuspended in 10 ml of aqueous 5% D-sorbitol (0.274 M) for 20 min at 37 °C. After an additional centrifugation, an equal volume of complete medium was added to the pellet. Cultures were reestablished by addition of uninfected erythrocytes and complete medium to result in a 5% hematocrit with an appropriate starting parasitemia (generally 0.1%). The medium was changed daily, as previously described (adapted from (157)).

3.3.2.2. Plasmagel purification of *P. falciparum* trophozoite and schizonts stages.

The plasmagel purification is an enrichment of the cultures with late trophozoites and mid-schizont stage parasites (158). This technique is only applicable to K^+ (expressing knobs on parasitized RBC membrane surface) type cultures (159). Late trophozoites and schizonts of K^+ strains sediment slower than non-parasitized or young trophozoite-infected RBC. Briefly, the parasite culture was centrifuged (200g, 5 min) and the pellet (0.5 -1 mL) was then re-suspended in 1.4 volumes of the pellet with complete medium (10% human B^+ plasma in RPMI 1640 and 0.23% sodium bicarbonate), and added slowly 2.4 volumes of the pellet of 6% plasmagel (Voluven 6%, Fresenius-Kabi). The parasite solution was mixed gently and incubated at 37 °C for 30 min. Finally, the supernatant containing older stage parasite RBCs was washed once with RPMI 1640 and parasitemia was evaluated (adapted from (160)).

In order to reach 80% synchronized stages, synchronization was performed in a sequential mode, with one sorbitol lysis when parasites were in ring stage, and approximately 24 h afterwards using the plasmagel procedure.

3.3.3. Parasite culture assays.

3.3.3.1. Parasite growth evaluation in the presence of Blasticidin.

After establishment of a NF54::*rif*6GFPHA2A-BglmS parasite line, the parasite culture at 4% of parasitemia (5% hematocrit), mainly in trophozoite stage, was cultivated in the presence of 2.5µg/ml Blasticidin. The parasitemia was monitored by flow cytometry (Guava EasyCyte Min GE.) using ethidium bromide DNA staining, after 72 h. The percentage of parasitized cells was analyzed in the Guava EasyCyteSoft program.

3.3.3.2. Parasite growth evaluation in the presence of Glucosamine 6 phosphate (GlcN-6P).

Once established the clonal *NF54::rif6GFPHA$_{2A}$-BglmS* and *NF54::pLSD1GFPHA-glmS* parasite lines, Parasite cultures were split into two parallel cultures at 0.5 % parasitemia. The cultures were cultivated in complete medium either 2.5 mM of glucosamine 6 phosphate (*Sigma-Aldrich* GlcN-6P), or not. The parasitemia was monitored by flow cytometry via ethidium bromide DNA staining, from 1 to 4 reinvasion cycles. The percentage of parasitized cells was analyzed in the Guava EasyCyteSoft program.

3.3.3.3. Blasticidin activation/selection assay.

Cultures of clonal the NF54::*rif*6GFPHA$_{2A}$-BglmS parasite line, were split into two parallel cultures and maintained in complete medium with either 2.5µg/ml Blasticidin, or not. After 1 reinvasion cycle, parasites were synchronized at 6 h ring stage and then harvested. Harvested IRBC were treated with 0.1% Saponin for 10min at RT and then pelleted at 12000 g/4 °C for 5min, washed once in 1ml PBS and then resuspended in a final volume of 100µl using TE. Afterwards, 1ml Trizol (*Invitrogen*) was added and the sample was stored at −80 °C until use. Total RNA was prepared following the Trizol protocol provided by the manufacturer (Ambion/Life technologies). Finally, recovered total RNA was dissolved in 20 µl RNAse free water and stored at −80 °C until use.

3.3.3.4. Glucosamine 6 phosphate (GlcN-6P) knockdown assay.

Cultures of clonal *NF54::rif6GFPHA$_{2A}$-BglmS* and *NF54::pLSD1GFPHA-glmS* parasite lines were split into two, and cultivated in complete medium supplemented or not with 2.5mM of glucosamine 6 phosphate (GlcN-6, Sigma-Aldrich). After 1 reinvasion cycle, parasites were synchronized at 6 h ring stage in the case of *NF54::rif6GFPHAHA2A-BglmS* parasite line and 32 h trophozoite stage in the case

of *NF54::PfLSD1GFPHA-glmS,* then harvested. Harvested IRBC were treated as explained before (3.3.3.3).

3.3.3.5. Panning of *P. falciparum* parasites on *CHO-CD36* cells.

Cultivated *P. falciparum* normally express a variety of *var* genes. To establish if knockdown/knockout of factors interferes with *var* gene transcription, it is advantageous to select parasites which express dominantly one *var* gene. This can be done by phenotypic selection (biopanning) over CHO-CD36 cells which leads to the selection of parasites which express *var* gene *PF*3D7_0412400 (161). For biopanning, trophozoite enriched cultures were resuspended in complete medium with plasma. Around 10^7-10^8 parasitized cells were then incubated over confluent *CHO* cell monolayers grown in 25 cm^2 culture flasks for 1 h with gentle agitation every 15 min. Non-adherent infected erythrocytes were washed away three times in RPMI 1640, pH 6.8 by direct aspiration. Bound infected erythrocytes were detached from cells by washing with RPMI 1640, pH 7.4 containing 10% plasma and returned to culture, adjusting the hematocrit to 5%. Cultures were grown to 2–6% parasitaemia before repeating this process. The phenotype was maintained by panning before the GlcN-6P knockdown assay. Adapted from (71) and (161).

3.4. Culture of *CHO* cells.

Stable transfectants of *CHO* expressing the human CD36 endothelial receptor (*CHO-CD36*)(162), were cultured in RPMI 1640 containing 10% FCS, 40 mg/L gentamycin in a 5% CO$_2$ atmosphere at 37°C and replaced each 3-4 days. To maintain CD36 expression, cells were frequently kept for 48 h in the presence of 750 µg/ml Neomycin/G418.

3.5. Real-time PCR with parasite-derived cDNAs.

All real-time PCR assays were done according to section 4.1. As an internal control, the seryl-tRNA ligase (*PF*3D7_0717700) transcript was used. Relative transcript abundance in relation to seryl-tRNA ligase was then calculated by the $2^{-\Delta Ct}$ method (163). Primer3 online software was used to create oligonucleotides for realtime PCR (settings: amplicon size 80-120, Tm 58-62°C, ideally 60°C, GC content 30-70%, (table 3).

Table 3. Real-time PCR oligonucleotides designed in *Primer3* online software.

Target gene	Gene ID	Oligo 5´Fw	Oligo 5´ Rv
rif 6	*PF*3D7_0632200	GTGGAAACCTGGGGCACTTA	ATAGCTTCACCTGCGGCATT
blasticidin deaminase (*bsd*)		TGCAGTTTCGAATGGACAAA	AACACAAAACAATCTGGTGCAT
GFP		TACACGTGCAAGTGCAGCTA	CTGGGTATCTCGCAAAGCAT
Pflsd1	*PF*3D7_1211600	TGCACGTATGCCTGTATCATAA	TCCAAGGGTAGTTTTCTGTGGT

3.6. Southern Analysis.

In order to confirm the appropriate integration occurrence in all our parasite cell lines, was performed a Southern analysis. This analysis was preceded by several PCR tests that indicated the proper integration event and the absence of episomal forms in the cloned parasite lines populations. The Southern analysis was done according to the section 4.1. For both PCRs and Southern analyses, we used material from WT NF54 parasites as negative integration control, and plasmid DNA from the respective constructs for the episome detection control. In table 4, the PCR oligonucleotide sequences used to detect the modified locus and plasmid presence are shown, together with the sequence of the probes used to detect target fragments in the Southern blot analysis.

Table 4. PCR oligos sequences to detect modified locus and episome presence.

Target gene	Gene ID	Oligo 5´Fw	Oligo 5´Rv
rif **6 check**	*PF*3D7_0632200	GTTTTATGTTAAACATATTTGATGTATT	-
rif r **check**	*PF*3D7_0100400	TAGAACACAGAGCCGCAAAA	-
GFP check		-	CTGCACTTGCACGTGTAG
HA check		-	AGCGGCATAATCTGGAACATCGTAC
Plasmid back bone check		GGGTTTCGCCACCTCTGAC	-
Pflsd1 check	*PF*3D7_1211600	GATGATAATAATAATAATAATGGTC	-
Pblsd1 check	PBANKA_0610100	AGCGGCATAATCTGGAACATCGTAC	AGCGGCATAATCTGGAACATCGTAC
hDHFR		-	AGCGACGATGCAGTTTAG
Blasticidine deaminase **probe**		ATGGGAAAAACATTTAACATTTC	AACACAAAACAATCTGGTGCAT
glmS **probe**		CTCGAGTAATTATAGCGCCCGAACTAAGC	GGTACCAGATCATGTGATTTCTCTTTG

3.7. Western blot.

For the detection of recombinant proteins in transgenic parasite lines, whole parasite protein extracts were prepared from saponin-lysed IRBCs as described in (156) Proteins were loaded on standard discontinuous SDS-polyacrylamide gels and transferred to Hybond C membranes (Amersham).

After blocking with 4% skimmed milk in 1xPBS/0.1% Tween20, the target proteins were recognized using one of the following antibodies.

Table 5. Antibodies used in western blot assays.

Epitope	Animal of provenance	specificity	Brand
HA	rabbit	monoclonal	cell signaling
H3	mouse	polyclonal	cell signaling
H4	mouse	polyclonal	cell signaling
PTEX150	mouse	polyclonal	deKoning-Ward T/WEHI-
GFP	mouse	monoclonal	Biomatik

Subsequently an antiMouse/Rabbit IgG-peroxydase antibody (KPL) was used to detect bound antibodies. Blots were exhaustively washed with PBS/Tween between incubations and finally incubated with Western Super signal substrate ECL^{TM} (*GE-healthcare*). Chemoluminescent signals were captured in an ImageQuant (GE) apparatus and intensities were quantified using ImageJ software (NIH).

3.8. Immunofluorescence assay.

To localize the tagged proteins in the organelles of the parasite was used a specific protocol of fixation proposed by (164). Briefly, *NF54::PfLSD1GFPHA-glmS* parasites cultures were synchronized at the desired stages. The cultures were washed once (200g for 5min) in PBS, then fixed with 1 mL of 4% paraformaldehyde and 0.0075% glutaraldehyde in PBS for 30 min, at room temperature in constant agitation. Fixed cells were washed once in PBS and then permeabilized with 1 mL 0.1% Triton X-100/PBS for 10 min, at room temperature in constant agitation. Cells were then washed again in PBS and then treated with 1 mL of ~0.1 mg/ml of sodium borohydride (NaBH4)/PBS for 10 min to reduce any free aldehyde groups. Following another PBS wash, cells were blocked in 5 mL of 3% BSA/PBS for one hour, at room temperature in constant agitation. A primary antibody at 1:500 dilution was added, and allowed to bind for a minimum of 1 h or overnight at 4°C in 3% BSA/PBS in constant agitation. After that, the cells were washed three times in PBS for 10 min each, to remove excess primary antibody. A secondary antibody was added at 1:1000 dilution (in 3% BSA/PBS) and allowed to bind for an hour in the dark, at room temperature in constant agitation. Cells were washed three times in PBS. Finally, the cells were incubated with 40µg/mL of DAPI (4′,6-Diamidine-2′-phenylindole dihydrochloride) during 5 min for nuclear staining and washed once (200g for 5min) in PBS (Adapted from (164)).

The parasite fluorescence was analyzed by microscopy, using an Axio Imager M2 microscope (Zeiss, Germany) equipped with an AxioCam HRC digital camera (Zeiss, Germany). The images were analyzed with the AxioVision 4.8 software.

3.9. Pull down assays.

With the purpose to study possible protein-protein interaction, pull down assays were performed using as protein bait a total protein extract of NF54::*Pf*LSD1GFPHA-glmS trophozoite enriched parasite cultures. The pull down was done using the Pierce® Crosslink Immunoprecipitation Kit (*Thermo Fisher Scientific*) flowing the manufacturer's instructions. For the technique approximately 1mg of total protein extract of each sample was used, prepared by saponin-lysis as described in (156). The protein samples were incubated in columns charged with 10 µg of anti-H3 or HA antibodies coupled to protein A/G. After immunoprecipitation, the resulting samples were analyzed by Silver stained-SDS polyacrylamide gels, western blot and mass spectrophotometry analysis (in the case of HA-IP samples).

3.10. Mass spectrophotometry analysis (MS).

To identify differentially precipitated proteins (WT vs *NF54::PfLSD1GFPHA-glmS* protein IP elutions) detected in Silver-stained SDS gels, a mass spectrophotometry analysis was done.

The target proteins were extracted and digested from a Coomassie-blue stained SDS-PAGE gel in the CEFAP/ICB-USP according to their established protocol (165). The proteins were identified using the following database: NCBInr – *(http: //www.ncbi.nlm.nih.gov/protein/)* and Uniprot (*http://www.uniprot.org*/), using the software *Proteome Discovery version 1.4* (*Thermo Scientific* and the SEQUEST search tool (*http://fields.scripps.edu/sequest/).*

3.11. RNA-seq Assay.

All RNA-seq assays were done according to the section 4.1. The cDNA libraries were constructed from samples: grown with or without 2.5 nM WR99210 for 6 reinvasions or grown with or without 2.5 mM GlcN 6P for 2 reinvasions. The TruSeq RNA Sample Preparation Kit v2 low sample (LS) protocol (Illumina Inc., CA) was used, based on the manufacturer's instructions.

After analysis using the CLC workbench, resulting RPKM values of trophozoite-derived cDNA from cultures which were treated or not with WR99210 or GlcN 6-P, were loaded in MS Access and filtered for product ID. The evaluation of significance between FPKM values in "on" and "off" samples was conducted by our collaborator Prof. Paulo Ribolla (UNESP/Botucatu).

3.12. Chromatin Immunoprecipitation Sequencing (ChIP-Seq) assay.

To verify how *Pf*LSD1 interacts with DNA sequences in the genome of *P. falciparum*, a Chromatin-Immunoprecipitation assay followed by DNA sequencing (ChIPseq) was performed.

3.12.1. Chromatin Immunoprecipitation (ChIP).

Trophozoite-enriched cultures of NF54::pLSD1GFPHA-glmS were harvested and crosslinked as described (156). Following this, the material was digested with Micrococcus Nuclease to separate

nucleosomes (1h at 37°C), in concordance with the procedure established by the enzyme provider (Thermo Scientific).

The digested sample was then submitted to antiHA immunoprecipitation using the HA immunoprecipitation kit (*MACS Miltenyi Biotec*) according to the manufacturer's instructions. Subsequently, crosslinking was reversed by incubation with glycin. Then, bound proteins were resolved by SDS gel electrophoresis and silver staining of gels to confirm the presence of the *Pf*LSD1 protein. Additionally, eventually bound DNA was detected by PCR.

3.12.2. gDNA libraries.

From immunoprecipitated material, genomic DNA libraries were prepared for high throughput sequencing using the TruSeq Sample Preparation Kit v2 low sample (LS) protocol (Illumina Inc., CA), based on the manufacturer's protocol. The integrity of the libraries was verified by *Bioanalyzer DNA 100*. The ChIP-seq assay was executed in an Illumina NextSeq500 sequencer following the recommendations of the provider, using mid output flow cells and a total of 10 million reads per sample.

3.13. Animal maintenance and assays.

This part of the work was done in collaboration of the group professor Daniel Y. Bargieri from the *Laboratório de genética molecular* at the ICB-USP. Animals were kept in the isogenic mouse facility (Department of Parasitology, University of São Paulo, Brazil) and used according to the certificate approved by the Institutional Animal Care and Use Committee (CEUA) in the Institute of Biomedical Sciences (ICB) (protocol number 132/2014) and standards established by the Brazilian College of Animal Experimentation guidelines (CONCEA). This part was entirely handled by Prof. Bargieri's students and the following descriptions.

3.13.1. Animals and parasite strains.

BALB/c mice were bred and maintained in the animal facility of the Department of Parasitology at the Institute of Biomedical Sciences, University of São Paulo. The *P. berghei ANKA* recombinant Ookluc line (*PbOokluc*) (154) were stored as frozen stocks in liquid nitrogen or at −80°C. Vial stocks were prepared by mixing 150 µl of parasitized mouse blood with 300 µl of Alsever's solution (*Sigma-Aldrich, A3551*) with 10% glycerol (*Sigma-Aldrich, G5516*). Mice were infected by intraperitoneal injection of 200-µl portions of thawed stocks. The parasitemia was monitored daily by Giemsa stained (*Laborclin, 620529*) thin blood smears counted by direct light microscopy with a 100× oil immersion objective (*Nikon E200*).

3.13.2. *Plasmodium berghei* transfection.

For the transfections, the plasmids were electroporated into synchronized *Pb*Ookluc asexual schizonts using an Amaxa (*Lonza*) Nucleofector electroporator set at program U33 and using a human T cell Nucleofector kit (*Lonza, VPA-1002*) as previously described (166).

For transfection of the *Pb*Ookluc, 5µg pPbLSD1KO plasmid was linearized with *KpnI* and *EcoRI* (FastDigest; Thermo Scientific).

After transfection, transformed parasites (*PbOokluc-LSD1^{KO}*) were selected by the administration of pyrimethamine (*Sigma-Aldrich, 46706*) for 72 h in the drinking water (70 mg/liter) of mice infected with the transfected parasites, starting 24 h after transfection. After selection, the transformed parasites were cloned by limiting dilution.

3.13.3. *Plasmodium berghei* viability assay.

Three weeks old BALB/c mice were divided into two groups of five mice each. The mices were challenged intravenously with 5.10^3 parasitized erythrocytes of *P. berghei* ANKA (control group) and *Pb*Ookluc-LSD1KO (from a donor mice). Then, the parasitemia was followed during 12 days. The parasites were counted by microscopy, each two days, using blood smears.

3.13.4. Conversion assays.

For conversion assays, parasitized mouse blood, with ~0.4 % of gametocytaemia, was obtained by collecting 4µl of blood from the infected mouse tail and added to 80 µl of ookinete medium. The ookinete medium (167) consisted of RPMI 1640 (*Thermo Scientific, 61870*) with 0.025 M HEPES (*Thermo Scientific, 15630080*), penicillin-streptomycin-neomycin (*Sigma-Aldrich, P4083*), hypoxanthine 50 mg/liter (*Sigma-Aldrich, H9636*), and xanthurenic acid 100 µM (*Sigma-Aldrich, D120804*) at a pH of 8.3.

The blood samples were kept at 21°C in an incubator for 24 h. The luciferase activity was determined by measuring the RLU using a microplate reader (*SpectraMax i3; Molecular Devices*) after the addition of 1 volume of the substrate buffer (*Nano-Glo luciferase assay system; Promega*). Each conversion assay sample was also analyzed by counting blood smears, to identify possible phenotypes. Adapted from (154).

4. RESULTS

4.1. *rif* genes and transcription mechanism.

4.1.1. Independent regulation of *P. falciparum rif* gene promoters.

The expression control of *rif* genes, is unclear and several studies reported somewhat contradictory results. Herein, we addressed the mode of *rif* transcription using plasmid vectors with two drug resistance markers. In this construction *rif* 5' upstream regions control the expression of one drug resistance marker (*hdhfr* cassette) while transfectant parasite lines are selected with the second, constitutively expressed marker (*blasticidin deaminase* cassette). When testing three different *rif 5'* upstream regions, we found that one, of *rif PF*3D7_0900500, was impossible to be activated, a second, *rif PF*3D7_1300400, was constitutively expressed while a third region, the *PF*3D7_0200700, showed activity that largely varied with the drug pressure applied in the culture. This last parasite transfected line had also integrated in the genome via single crossing over. When the global transcription of *rif* genes in this modified parasites (NF54:: pTZ57BiPro-*PF*3D7_0200700 *5' rif ups*) was compared via RNA-Seq analysis, in the presence or absence of drugs, we observed that the activation/silencing of all *rif* did not change profoundly between strains. We concluded that there is no crosstalk between *rif* locus or that - unlike *var* promoters - solely the 5' upstream region of *rif* genes is not able to engage in a yet elusive system of allelic exclusion of *rif* gene transcription.

The complete section 4.1 can be consulted in annex 1.

4.1.2. The impact of the transcriptional activation of a single *rif* gene.

4.1.2.1. The strategy p*rif*GFPHA2A-BglmS.

In order to confirm our findings which were based on a single *rif* locus integrated in an ectopic region of the genome, an additional approach was tempted. The novel approach employed the pGFPHA2A-BglmS plasmid to establish a regulatable *rif* locus. This plasmid has also two resistance markers. One resistance marker, the *hdhfr* cassette, is to select the obtained transfected parasite lines. The second marker, the *blasticidin deaminase* cassette, is to selected/promote the expression of the target genes. Of note, the *blasticidin deaminase* gene is only transcribed if a knockin in the target gene occurred which led to the creation of a hybrid gene, containing the target gene, GFP-HA or at least HA, and the blasticidin deaminase, both of which controlled by the target gene promoter. When the single recombination event takes place, due to the *rif* homology region, the modified parasites will produce a fused protein, RIFIN-GFP/HA-2A-BSD. To avoid mislocalization at the moment of exporting proteins to other cell compartments or even into the erythrocyte a 2A peptide domain was included, which promotes the cleavage of the translated protein. In this way, the BSD domain is cleaved and remains in the cytosol while RIFIN-GFP/HA can be exported normally. This construct also has the glmS sequence, which forms a ribozyme once the RNA is synthesized. Then, in the presence of glucosamine (2.5mM) (152), the polyA signal of the corresponding mRNA, is cleaved off and the transcript becomes unstable. In this sense, the modified parasites can be conditionally depleted of the *rif-GFPHA-2a-bsd* transcript.

Two versions of this plasmid, p*rif*6GFPHA2A-BglmS and p*rif*rGFPHA2A-BglmS respectively were created to target two different *rif* genes: *PF*3D7_0632200 (*rif*6) and *PF*3D7_0100400 (*rif*r). Both vectors were transfected as described above, and the positive WR99210 selected parasites were cloned by limiting dilution.

Due to difficulties to identify the correct integration via PCR analysis of the modified parasite lines (the placement of specific oligos in the 5' untranslated region of *rif* genes is virtually impossible due to similarity between rif promoters and/or high AT content), a Southern Blot (SB) analysis was done, using a probe of 156 nt which hybridizes with the glmS sequence. Before blotting, genomic DNA from both parasite lines were digested with Xho I and Spe I. The strategy of the SB assay is shown in Figure 12 A.
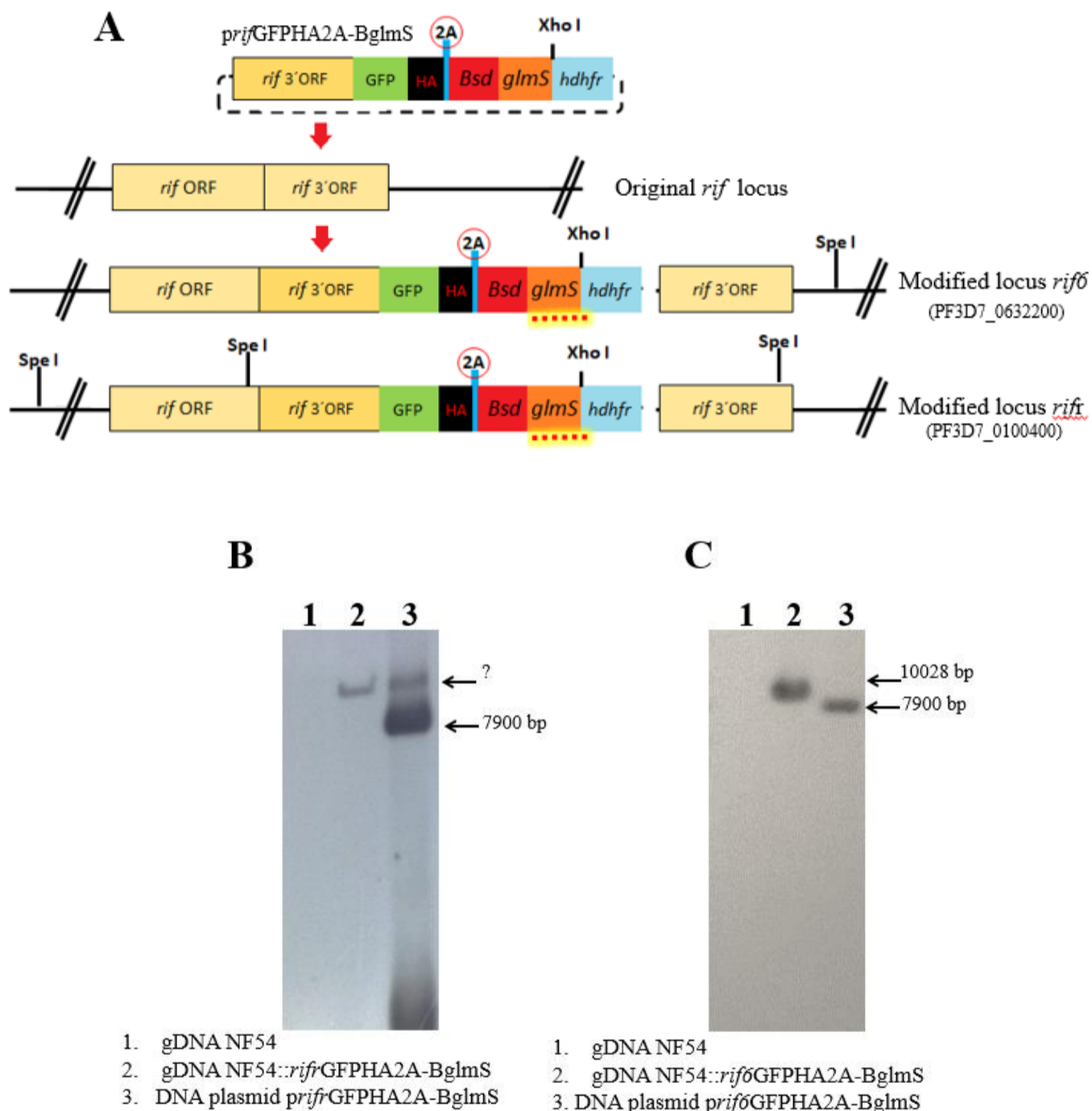
Figure 12. Outline of the transfection p*rif*_GFPHA2A-BglmS plasmids construct.
A. *rif* 3′ ORF regions were interchanged in this plasmid via Bgl II and Pst I. Due to the *rif* 3′ ORF homologous regions with *P. falciparum* genomic sequences, the integration via single crossing-over took place. The *in silico* predicted integration of the p*rif*6GFPHA2A-BglmS (*PF*3D7_0632200) and p*rif*rGFPHA2A-BglmS (*PF*3D7_0100400) in their respective locus is shown. **B**. Approximately 5 µg of total gDNA of NF54 (WT parasite as negative control, lane 1) and the *NF54:: prifrGFPHA2A-BglmS* parasite line, were digested with Xho I and Spe I. In parallel, 25 ng of p*rif*rGFPHA2A-BglmS plasmid DNA (as positive control of the probe recognition and episomal plasmid existence) was loaded. **C**. Approximately 20 µg total gDNA of NF54 and *NF54::prif6GFPHA2A-BglmS*, were digested with Xho I and Spe I restriction enzymes, and in parallel 5 ng of the plasmid p*rif*6GFPHA2A-BglmS.

As shown in figure 12, only the *NF54::The prif6GFPHA2A-BglmS* parasite line appeared to had a proper integration, since the SB analysis coincided with the *in silico* prediction, which is a fragment of approximately 10028pb in the case of *NF54::The prif6GFPHA2A-BglmS* gDNA and 7900pb for plasmid DNA. In the case of the *NF54::prifrGFPHA2A-BglmS* parasite line, a similar signal of approximately 10000-11000 bp is observed, which does not match with the predicted size, 6271pb. Probably integration occurred in another *rif* locus. Due de impossibility to discover in which of the 183 *rif* genes sequences the p*rif*rGFPHA2A-BglmS integrated, we decided not to further analyze this parasite line, and to continue only with the NF54::p*rif6*GFPHA2A-BglmS line.

4.1.2.2. Transcriptional activation/selection of *PF*3D7_0632200 *rif* promoter.

We evaluated the capacity of NF54::*rif6*GFPHA2A-BglmS line to produce blasticidin deaminase, which implies the activation of the *PF*3D7_0632200 *rif* promoter. The parasites were cultured in the presence of 2.5µg/ml of blasticidin as described in section 3. As is shown in the figure 14 A, parasite growth quickly stalled and parasite densities decreased to very low levels. However, the parasitemias never disappeared. Next, we tested if addition of blasticidin led to an increase of *PF*3D7_0632200, *GFP* and *bsd* transcripts. For this, the NF54::*rif6*GFPHA2A-BglmS culture was split in two and treated or not with 2.5µg/ml blasticidin. After 24h the parasites were harvested at ring stage (highest predicted levels of transcription of *PF*3D7_0632200 in the intraerythrocytic cycle) and total RNA were extracted. The levels of mRNA were evaluated by RT-qPCR, using real time oligos for *PF*3D7_0632200 (*rif6*), *bsd* and *GFP*, as indicated in section 3. As an internal control, we used the *P. falciparum seryl tRNA ligase* transcript. The relative quantity of transcripts was computed using the $2^{-\Delta Ct}$ method and the significance of difference between the samples was estimated using Student's T-test (FIG 14B). We also investigated the functionality of the glmS system. We analyzed if the absence of the *PF*3D7_0632200 transcript affected the viability of the parasite and also if the transcript levels of *PF*3D7_0632200 (*rif6*) and *GFP*, were diminished in the presence of GlcN 6P. For this, cultures of the NF54::*rif6*GFPHA2A-BglmS line, were split in two and treated or not with 2.5mM GlcN 6P. After 48h the parasites were harvested at ring stage and total RNAs were extracted (Figure 13C-D).

A
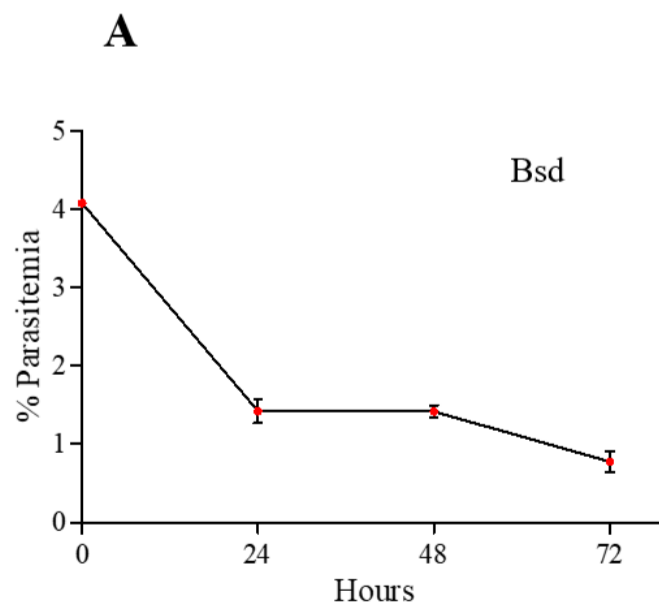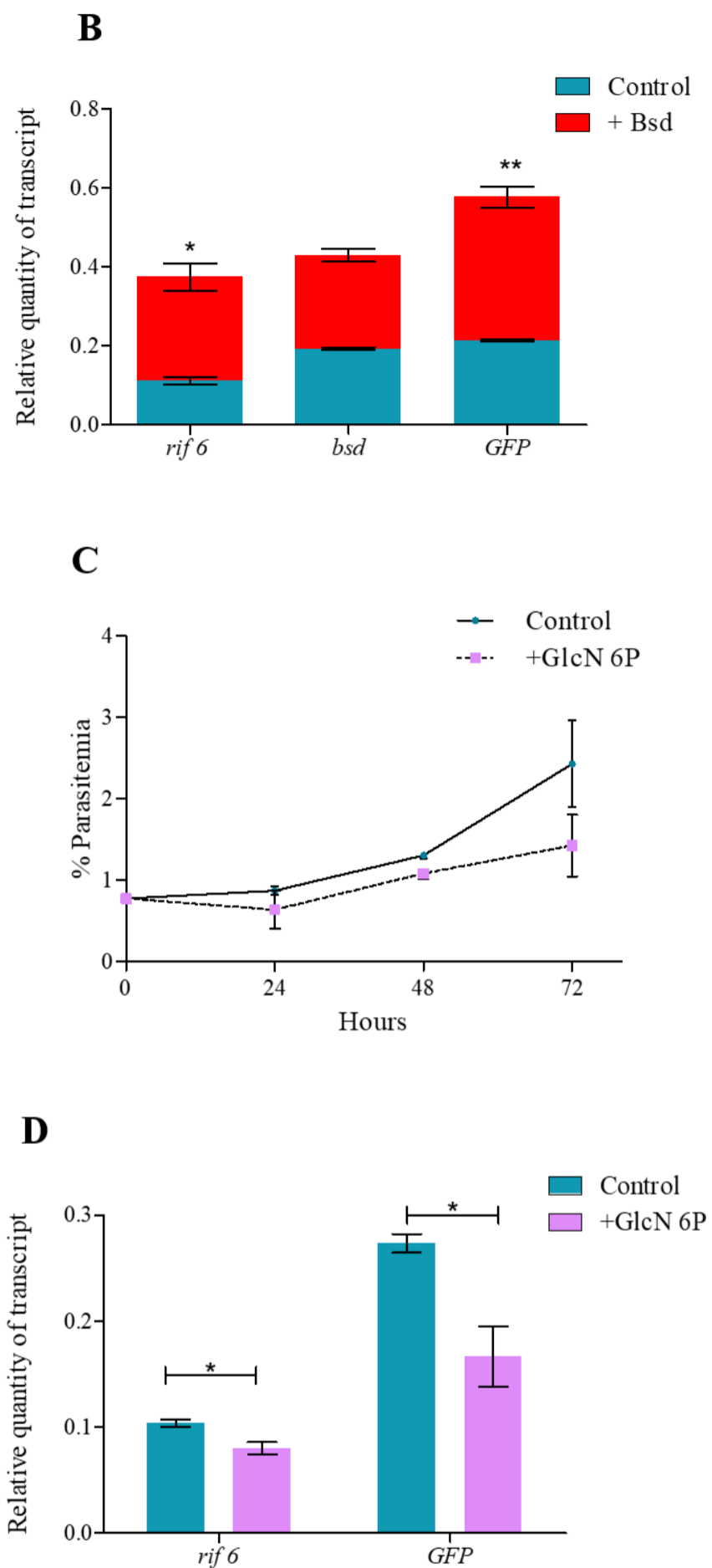
Figure 13. Capacity of NF54::*rif6*GFPHA2A-BglmS parasites to produce blasticidin deaminase transcript.
A-C. Growth analysis of NF54::*rif6*GFPHA2A-BglmS parasite, exposed or not to 2.5 μg/ml blasticidin. A. and 2.5mM of GlcN6P (without blasticidin). **C**. The percentage of parasitemia was determined by cytometric counting of ethidium bromide-stained parasites. **B-D**. Analysis of the transcript abundance of rif 6 (PF3D7_0632200), bsd and GFP was determined by qRT-PCR, after the treatment with 2.5 μg/ml blasticidin B. or 2.5mM GlcN6P. **D**. In both cases untreated parasites were used as controls.

As seen in figure 13 B the levels of transcripts increased significantly when the cultures were submitted to blasticidin treatment. Likewise, GlcN6P-treated parasites (figure 13 D), exhibited a two-fold decrease of transcript levels, which indicates that glmS can be activated as expected. Regarding the viability of

the parasite cultured in presence of the glucosamine metabolite, only a slight difference in the growth pattern was observed during the time frame measured when compared with untreated parasites (figure 13 C). The observed difference in transcript increase by the use of blasticidin was though significant only subtle. Given that fact that the tagged *rif* locus already showed baseline transcriptional activity, we judged that no new information in comparison to the data in 4.1 would be obtained and therefore did not further characterize this parasite line, e.g. by RNAseq to monitor all *rif* transcripts.

4.2. *Plasmodium Specific* Lysine Demethylase 1 (LSD-1).

4.2.1. The role of *P. falciparum* LSD-1, a putative histone modifier, in the transcriptional memory/switching of *rif* and other variant genes.

4.2.1.1. The strategy p*Pf*LSD1GFPHA-glmS.

In order to study the possible role of *Pf*LSD1 in the transcriptional control of *rif* and other variant genes in *P. falciparum*, a knockdown (KD) approach was used. As before for tagged *rif* genes, we employed the glmS system and in the modified parasites the *lsd-1* transcript is then subjected to conditional knockdown dependent on the addition of GlcN6P. Further, we monitored *Pf*LSD1 protein expression using the introduced C-terminal GFP/HA tags.

The transfection plasmid p*Pf*LSD1GFPHA-glmS plasmid was constructed by introduction of a PCR-amplified 981 pb DNA fragment of the 3' end of the *Pflsd1* ORF (*PF*3D7_1211600), introduced via Bgl II and Pst I in the BglII/PstI digested vector p_GFPHAglmS which contains a hDHFR cassette for transfectant selection with WR99210 (figure 15). WR-resistant transgenic parasites were obtained after 3-4 weeks of selection. The integration of the p*Pf*LSD1GFPHA-glmS plasmid was achieved by two WR on/off cycles of 2 weeks of each interval, followed by parasite cloning using the limiting dilution method. The presence of modified DNA in WR-resistant parasites was determined by PCR and SB assays from total parasite genomic DNA (figure 14).
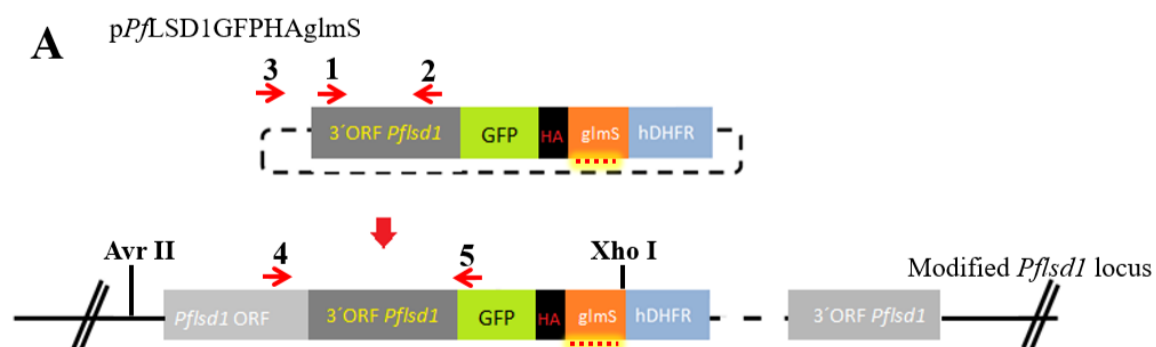
Figure 14. Scheme of the transfection vector p*Pf*LSD1GFPHA-glmS and single crossover recombination.
A. The predicted integration of the p*Pf*LSD1GFPHA-glmS into its respective locus. **B.** PCR analysis detected integration of the pLSD1GFPHA-glmS into the genome using the primer combinations indicated in A (red arrows). To detect episomal plasmid primers 3-5 were used, to confirm DNA integrity and primers performance, primers 4-2 were used, and finally for the detection of the modified locus primers 4-5 were employed. **C.** Southern blot analysis using 20 µg ea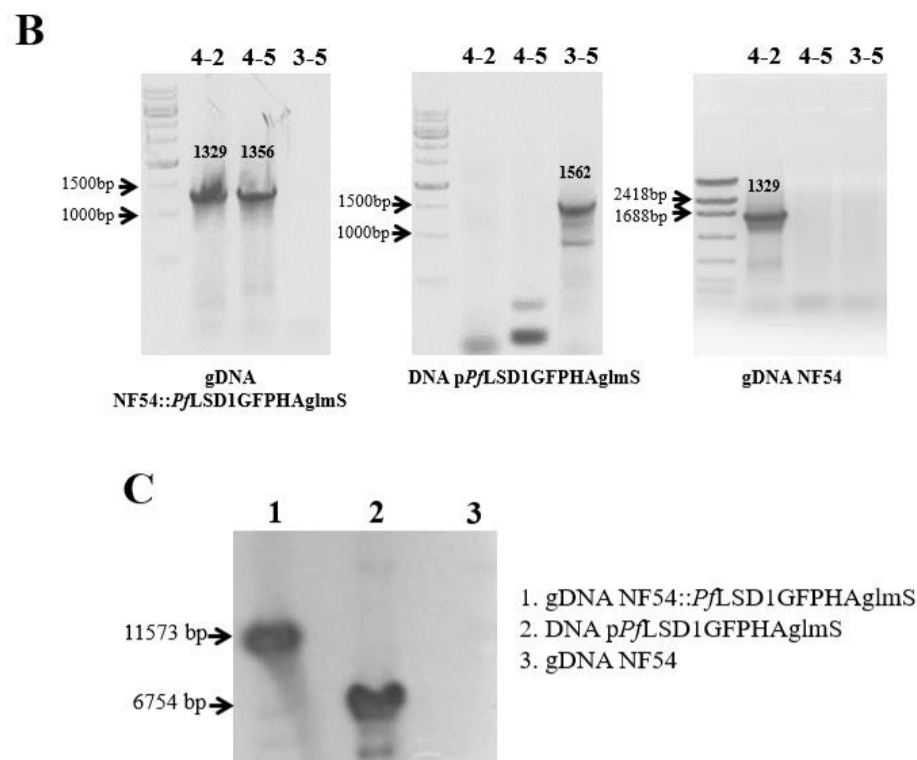ch of total gDNA of the NF54::*Pf*LSD1GFPHA-glmS parasite line and wild-type NF54 were digested with Xho I and Avr II, as well as 25 ng of p*Pf*LSD1GFPHA-glmS plasmid DNA (as positive control of the probe recognition and episomal plasmid existence). As a probe, a 156 nt digUTP labelled fragment representing the glmS sequence was used (see Methods).

As shown in the figure 14 B no residual episomal plasmid was detected in the gDNA samples of the cloned NF54:: *Pf*LSD1GFPHA-glmS parasite line. When we used the primers set 4-5, which hybridize in the *Pflsd1* ORF (Fw primer 4) and in the plasmid GFP tag (Rv primer 5), we were able to detect an amplicon of 1356 bp, which indicated a correct recombination event. As the positive control for episomal plasmid detection, we used DNA from p*Pf*LSD1GFPHA-glmS plasmid. The episomal presence was detected using the 3-5 primer set. The PCR product obtained was a 1562 bp fragment, as shown in the second gel in the figure 14 B. Finally, genomic DNA of NF54 wild type parasites was used as negative control. To confirm these results, we performed a SB analysis. A signal of approximately 11573 bp was detected in the gDNA samples corresponding to the NF54::*Pf*LSD1GFPHA-glmS transgenic parasite line and approximately 6754 bp sign corresponding to the plasmidial DNA of p*Pf*LSD1GFPHA-glmS. The difference of ~4800 bp matches *in silico* predictions.

4.2.1.2. The glmS system in the NF54::*Pf*LSD1GFPHA-glmS parasite line.

After confirmation of the integrity of the NF54::*Pf*LSD1GFPHA-glmS parasite line, we evaluated the performance of the glmS system. In a first approach, we monitored the transcriptional profile of *Pflsd*1in the intraerythrocytic cycle by RT-qPCR. Subsequently, we analyzed the KD efficiency at transcriptional and protein level, via RT-qRT and Western Blot assays, respectively. We also evaluated if the viability of the parasite in the absence of the *Pflsd1* transcript was affected, by following the parasitemia in parasites treated or not with the glmS inductor, 2.5mM of GlcN6P.
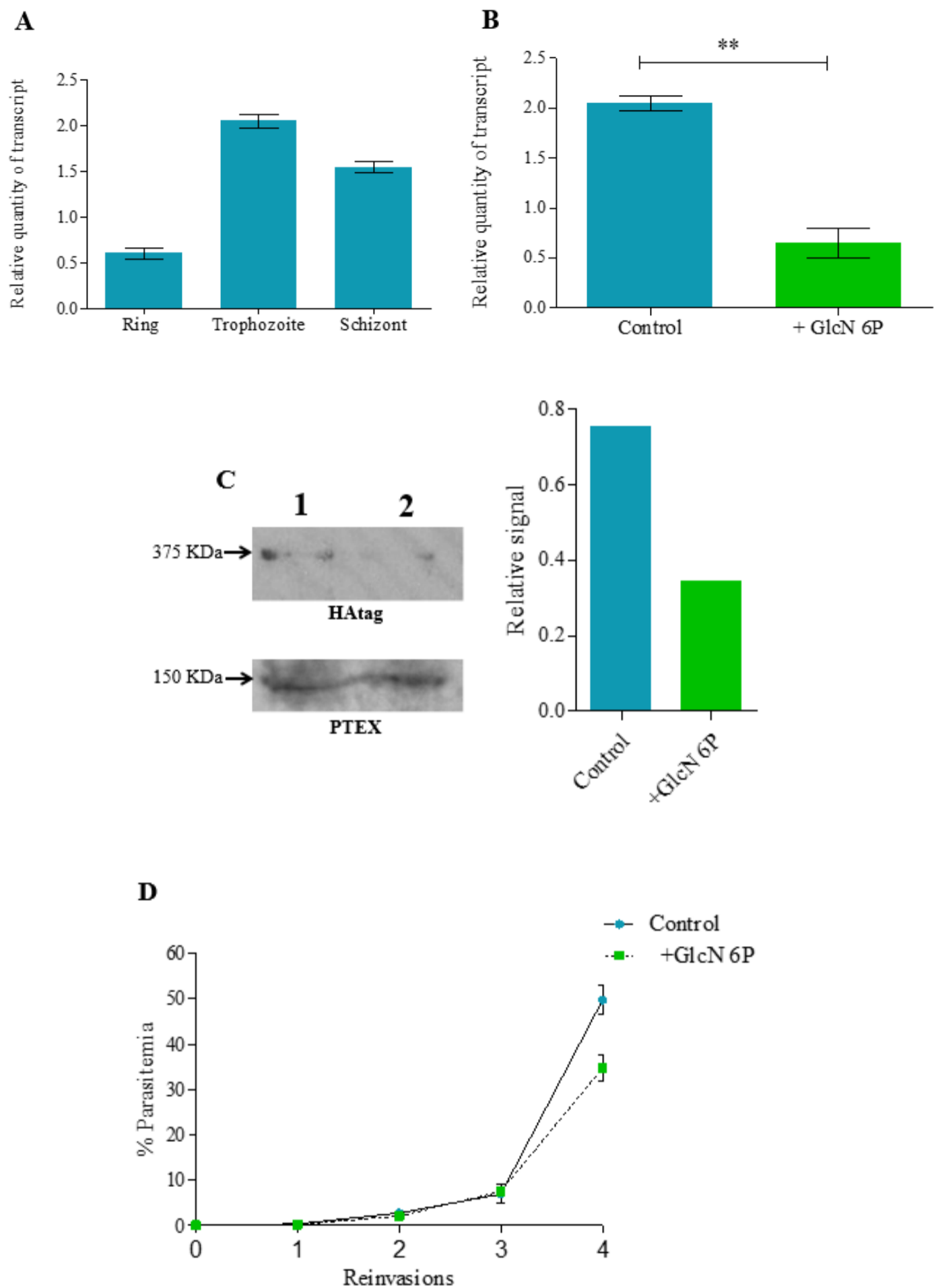
Figure 15. *Pf*lsd1 transcription profile in NF54::*Pf*LSD1GFPHA-glmS and glmS system performance.

A. Total RNA was extracted from synchronized parasite cultures, harvested in ring, trophozoite and schizont stages. The relative quantity of transcripts was determined by RT-qPCR via the $2^{-\Delta Ct}$ method, using the real time primer set of *PF*3D7_1211600.1, described in section 3. A synchronized trophozoite culture of NF54::*Pf*LSD1GFPHA-glmS parasite line was split in two of which one half was treated for 48 h with 2.5mM GlcN6P. In **B,** the relative quantity of *Pf*lsd1 transcripts was analyzed via RT-qPCR. In **C,** *Pf*LSD1 protein expression was evaluated by Western Blot analysis, using an antiHA antibody to detect the *Pf*LSD1 tagged protein (upper part) and an anti-PTEX150 antibody as load control (bottom part). In **D** growth analysis of NF54::*Pf*LSD1GFPHA-glmS parasites in the presence or absence of 2.5mM of GlcN6P is shown and the percentage of parasitemia was determined by cytometric counting of ethidium bromide stained parasites, during 4 reinvasion cycles. To get to measure the last two points of the curve growth was necessary to diluted in a 1/100 DF. This is a graphical representation of one of three independent biological replicates.

Our results coincide with the results from the *plasmodb* data base regarding the *Pflsd1* transcriptional pattern. The highest levels of *Pflsd1* transcripts in the blood stage was registered in the trophozoite stage (Figure 15 A). Therefore, the following experiments were performed in this stage. The parasites treated with of GlcN6P exhibited about 4 times less *Pflsd1* transcripts, when compared with the untreated parasite culture (Figure 15 B). This result was also confirmed at the protein level by densitometry analysis of the 375 KDa detected signal from the Western Blot assay, using *ImageJ* Software, and normalizing to the PTEX150 signal (Figure 15 C). Although there is not a complete deletion of the protein, our results indicated that the glmS system was effective, permitting us to control the abundance of the *Pf*LSD1 protein.

The NF54::*Pf*LSD1GFPHA-glmS line cultivated in the presence of 2.5mM of GlcN6P did not exhibit any meaningful growth difference when compared to the parasites cultured in normal conditions (Figure 15 D), which indicates that *Pf*LSD1 is not essential for the biology of the parasite, at least not in asexual blood stages. This is in accordance with reports by Lubin Jiang and colleagues in (168), who also state that a knockout of *Pf*LSD1 is viable.

4.2.1.3. *P. falciparum* LSD-1 as histone modifier.

P. falciparum LSD1 has not yet been characterized biochemically and little is known about its biological functions. There are no reports about its influence on the transcriptional control, protein-protein and/or protein-DNA interactions.

According to the literature, LSD1 proteins in others organism have an important role in chromatin remodeling and transcriptional regulation (123). It can either repress or activate target genes through interacting with a variety of protein complex and catalyzing demethylation H3K4 or H3K9, depending of the context. In this sense, we wanted to investigate if this protein in any way participates in transcriptional events in *P. falciparum*. We first investigated the possible protein-protein interactions.

An *in silico* analysis of the *Pf*LSD1 amino acid sequences, using *NCBI conserved domain* software, pointed to possible interacting sites of this enzyme (Figure 16).
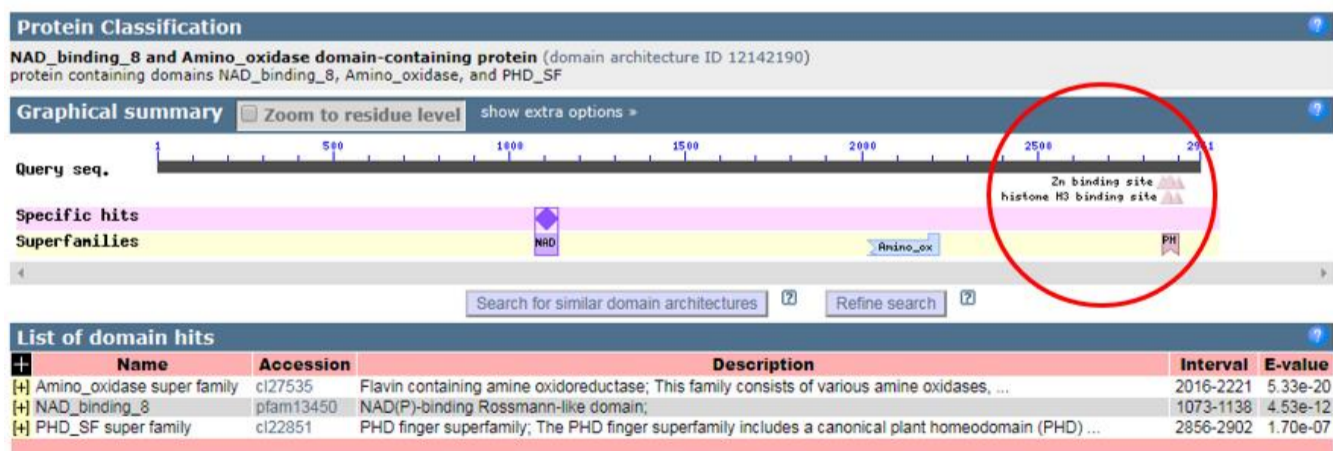
Figure 16. Graphical summary of the conserved domains present in the *Pf*LSD1 protein using NCBI conserved domain software.

The LSD1 proteins belongs to the family of proteins with flavin-containing amine oxidase domains, which is also the case for *Pf*LSD1. This suggests that in *P. falciparum* this protein could also have the

similar role and a H3K4/9 demethylation activity. The fact that *Pf*LSD1 has a putative PHD domain which can recognize the unmodified and modified histone H3 tail, is in accordance with these suppositions. This protein also seems to have a specific site to bind histone 3. The histone H3 binding site lies in its PHD domain, next to the Zn binding site, which is involved in DNA/RNA binding. Taken together, these data suggest that *Pf*LSD1 may be involved in control of gene transcription and chromatin dynamics. It has been indicated that the proteins which show these characteristics, function as epigenome readers. This class of proteins can control the gene expression through molecular recruitment of multi-protein complexes of chromatin regulators and transcription factors (169).

4.2.1.4. *P. falciparum* **LSD-1 interact with the histone 3.**

Since the *Pf*LSD1 seems to have a specific site to bind histone 3 in its PHD domain, we checked for a possible co-localization with histone 3. For this, an immunofluorescence analysis was performed using different intraerythrocytic stages (figure 17).
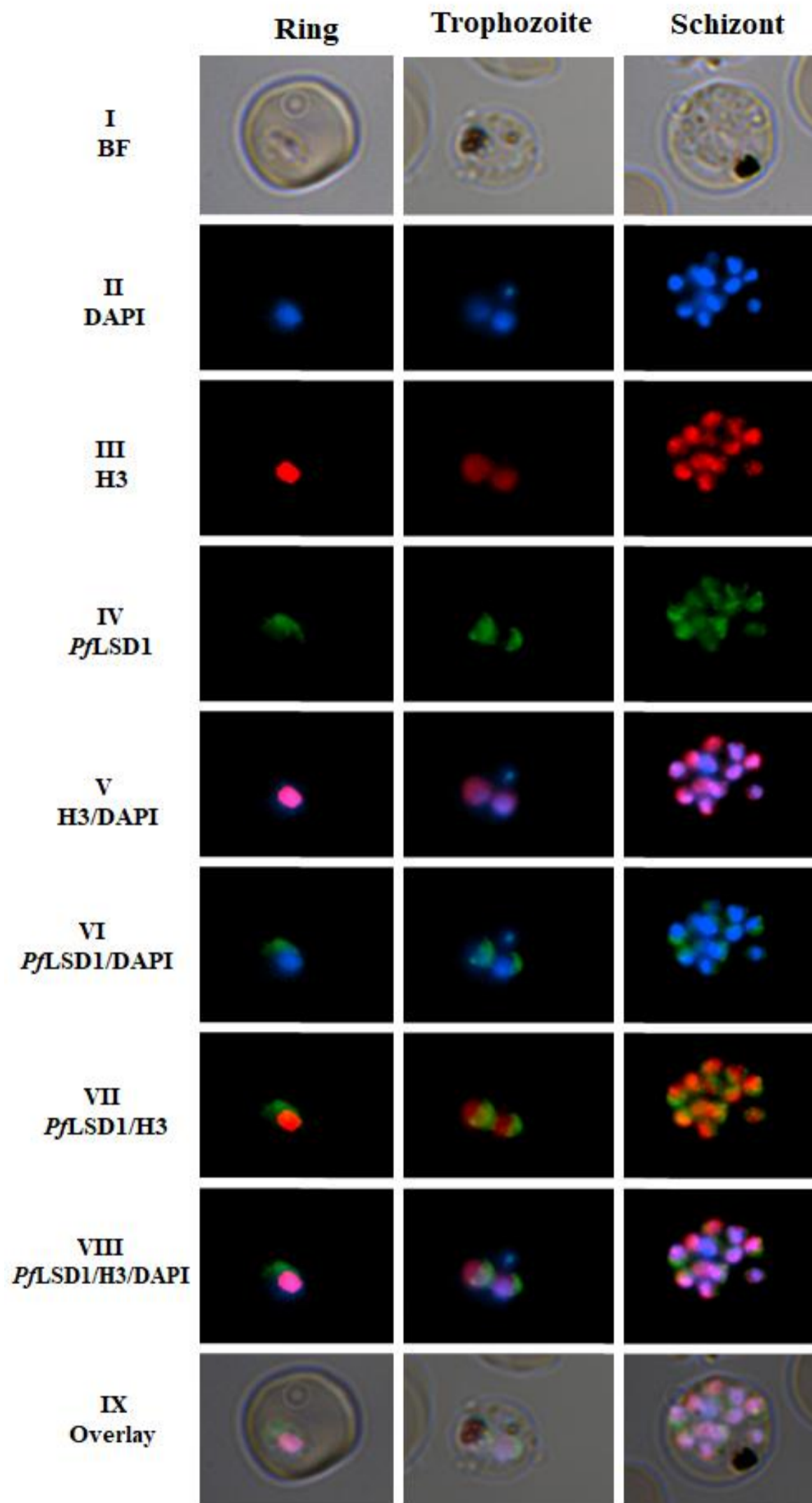
Figure 17. Immunofluorescence microscopy of PfLSD1 and histone H3.
Synchronized parasite cultures were harvested in ring, trophozoite and schizont stages. Cells were fixed as described before. For the localization of histone 3 was used a rabbit anti-H3 ab as primary antibody and goat anti-rabbit Alexa Fluor 568 (*Life technologies*) conjugate dye as secondary antibody. For the

GFP-HA tagged *Pf*LSD1, the GFP signal was improved using a mouse anti-GFP as primary ab and a goat anti-mouse IgG Alexa Fluor 488 (*Life technologies*) conjugate dye as secondary antibody. **BF** bright field, **DAPI** nuclear staining, **H3** immuno-localized histone 3 and **GFP** immune-localized tagged *Pf*LSD1.

*Pf*LSD1 localizes mostly around and partially overlapping with the DAPI stained portion of the nucleus in all intraerytrocytic parasite stages (row VI, *Pf*LSD1/DAPI). The same localization pattern is seen when *Pf*LSD1 is overlaid with the H3 localization (rows VII and VIII, *Pf*LSD1/H3 and *Pf*LSD1/H3/DAPI). Of note, the H3 signal (in red) overlaps fully with nuclear DAPI signal (row V, H3/DAPI). We conclude that *Pf*LSD1 is, as expected, localized in part of the nucleus and its peripheries. It is also present in all intraerythrocytic forms in accordance with RT-qPCR analyses (figure 15 A) which showed that the *Pf*LSD1 is expressed throughout the intraerytrocytic cycle.

Next, we set out to identify *Pf*LSD1 interacting factors. For this, we performed a differential pull down assay as explained in section 3. We used an anti-HA antibody to separate our GFP/HA tagged *Pf*LSD1together with proteins that were strongly interacting with it (figure 18).
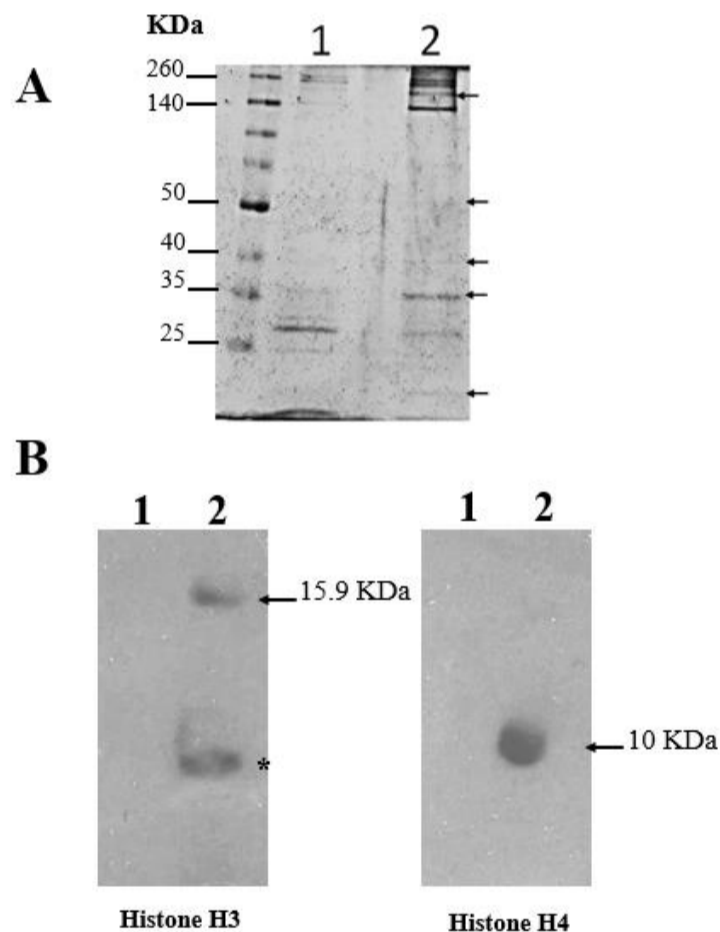
Figure 18. Protein pull-down reveals a number of proteins species reacting with *Pf*LSD1.
Cultures of trophozoite stage of NF54::*Pf*LSD1GFPHA-glmS and NF54 (WT control) parasites were harvested and total protein was extracted .1mg of total protein extract of each samples was loaded into the anti-HA coupled protein A/G column. Eluted proteins were analyzed by **A**. Silver stained SDS gel, in lane **1** NF54 (WT control) immunoprecipitated and eluted proteins, in lane **2** NF54::*Pf*LSD1GFPHA-glmS IPed and eluted proteins. In **B,** detection of IPed proteins using antiH3 and antiH4 antibodies. While no histones could be detected in NF54 immunoprecipitated extracts (lane 1 both pictures), H3 and H4 were readily detected by specific antibodies for these proteins in a wild type protein extract (lane 2 both pictures).

As shown in the silver-stained SDS gel (figure 18 A) a different pattern of protein bands was observed, when we compared protein elutions obtained from NF54::*Pf*LSD1GFPHA-glmS and from NF54 (WT control). The differential bands obtained from NF54::*Pf*LSD1GFPHA-glmS IP protein elutions (arrows in the figure 18 A), were subsequently extracted from the gel and sent to mass spectrometry analysis. The results of this analysis are shown in Table 6.

Table 6. Several proteins involved in epigenetic processes are identified in *Pf*LSD1 co-precipitations.
In this table, *Pf*LSD1-GFP-HA-co-precipitating/co-eluting proteins are listed. These proteins are possibly involved in epigenetic processes. The values of score represent the quantity and quality of the result of the identified peptides. Coverage, the percentage of the protein sequence covered by identified peptides and the # PSM's the number of peptide spectrum matches. The number of PSM's is the total number of identified peptide spectra matched for the protein. For the full list, see appendix 2.

Accession	Gene ID	Description	Score	Coverage	# PSMs	MW [KDa]
Q8IIV1	*PF*3D7_1105100	Histone H2B	40.83	12.82	11	13.0
Q8I457	*PF*3D7_0504400	ATP-dependent helicase	28.36	8.87	10	87.5
U5P4F6	*PF*3D7_0617800	Histone H2A	10.49	13.01	5	13.1
C6KSV0	*PF*3D7_0610400	Histone H3	3.52	23.53	2	15.5
C6KT18	*PF*3D7_0617800	Histone H2A	12.10	22.40	5	15.5
O00914	*PF*3D7_1104200	*Pf*SNF2L: chromatin remodeling protein	3.07	1.05	2	166.8
O00882	*PF*3D7_1105000	Histone H4	4.59	15.19	2	9.1
Q8IAX8	*PF*3D7_0814200	DNA/RNA-binding protein Alba 1	10.19	18.15	4	44.2

Among the high number of proteins found in the MS analysis, a number of proteins, listed in table 6, may hint to a possible function of *Pf*LSD1 in epigenetic processes. In accordance with the presence of a histone 3 binding domain inside the PHD domain in *Pf*LSD1, histone 3 and 4 were among the proteins which co-precipitated with *Pf*LSD1. This became also evident when eluted proteins were analyzed in Western blots, testing with histone 3 and 4 specific antibodies (figure 18). Of note, no histones were detected when immunoeluting NF54 extracts without tagged *Pf*LSD1 (figure 18 B). In addition to the histones found, different nucleolar proteins with chromatin remodeling proprieties were also detected these include ATP-dependent helicases. We also detected *Pf*SNF2L which is the *P. falciparum* homologue of the Helicase SNF (Sucrose Non-Fermentable). SNF is part of the SWI/SNF complex, a nucleosome remodeling complex found in eukaryotes (170). This protein complex determine the way DNA is packaged. In Apicomplexan SWI/SNF factors are characterized by fusion to PHD fingers (148).

In order to further confirm an interaction between histones and *Pf*LSD1, the pulldown assay was repeated with the difference that antiH3 antibody was used to isolate coprecipitating proteins. After SDS gel electrophoresis and transfer, blotted proteins were detected using the antiHA antibody. As shown in figure 19, a huge protein in the expected size of the *Pf*LSD1-GFP-HA fusion was detected in the transfectant line, but not in wildtype NF54 extract pulldowns/elutions.
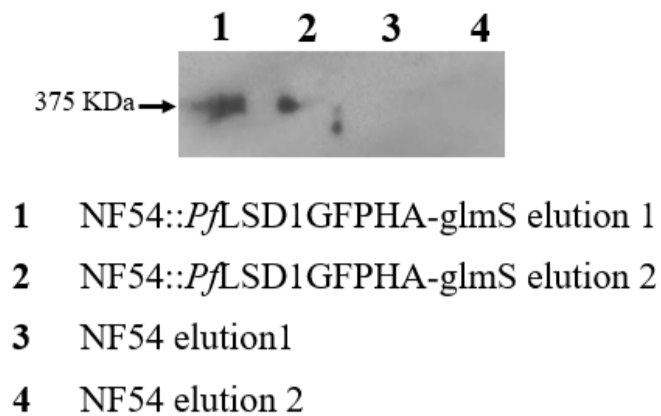
Figure 19. Pull down of H3 associated proteins reinforce interaction with *Pf*LSD1.
trophozoites of NF54::*Pf*LSD1-GFP-HA-glmS and NF54 (WT control) parasite cultures were harvested and total protein was extracted from each culture. 1mg of the total protein extract of each sample was loaded onto the anti-H3 coupled protein A/G column. The protein elutions were analyzed by western blot, using an anti-HA antibody to detect the*Pf*LSD1-GFP-HA tagged protein. Lane **1-2** NF54::*Pf*LSD1-GFP-HA-glmS IP protein elutions 1 and 2, and lane **3-4** NF54 (WT control) IP protein elution 1 and 2.

Taken together (*in silico* predicted H3 binding site, IFA of *Pf*LSD1-H3 co-localization, pull down of *Pf*LSD1 and pull down of H3), it appears that *P. falciparum* LSD1 interacts with the histone 3. Together with the characteristic domains present in PfLSD1, this points to a role as a histone modifier.

4.2.1.5. Influence of a *Pf*LSD1 knockdown on *var* gene transcription.

To elucidate whether *Pf*LSD1 exerts any influence on variant gene expression, we evaluated the transcriptional profile of the *var* genes under depletion of *Pf*LSD1. First, we selected a particular *var* gene transcriptional profile by biopanning/cytoadherence selection of NF54::*Pf*LSD1GFPHA-glmS over CHO-CD36 cells which leads to the dominant expression of *var* gene *PF*3D7_0412400 (*Pf*D0615c). Afterwards, the NF54::*Pf*LSD1-GF-PHA-glmS – line dominantly expressed *var* *PF*3D7_0412400. This culture was then submitted to the presence or absence of GlcN6P for 48 h. Subsequently, RNA was purified form untreated and treated cultures in ring stage and the *var* transcription profile was determined by RT-qPCR (figure 20).

Figure 20. No influence of a *Pf*LSD1 knockdown on *var* genes transcription.

Cytoadherence selected NF54::*Pf*LSD1-GFPHA-glmS-parasites (expressing *var* gene *PF*3D7_0412400) were synchronized in ring stage, split in two, and were treated or not with 2.5mM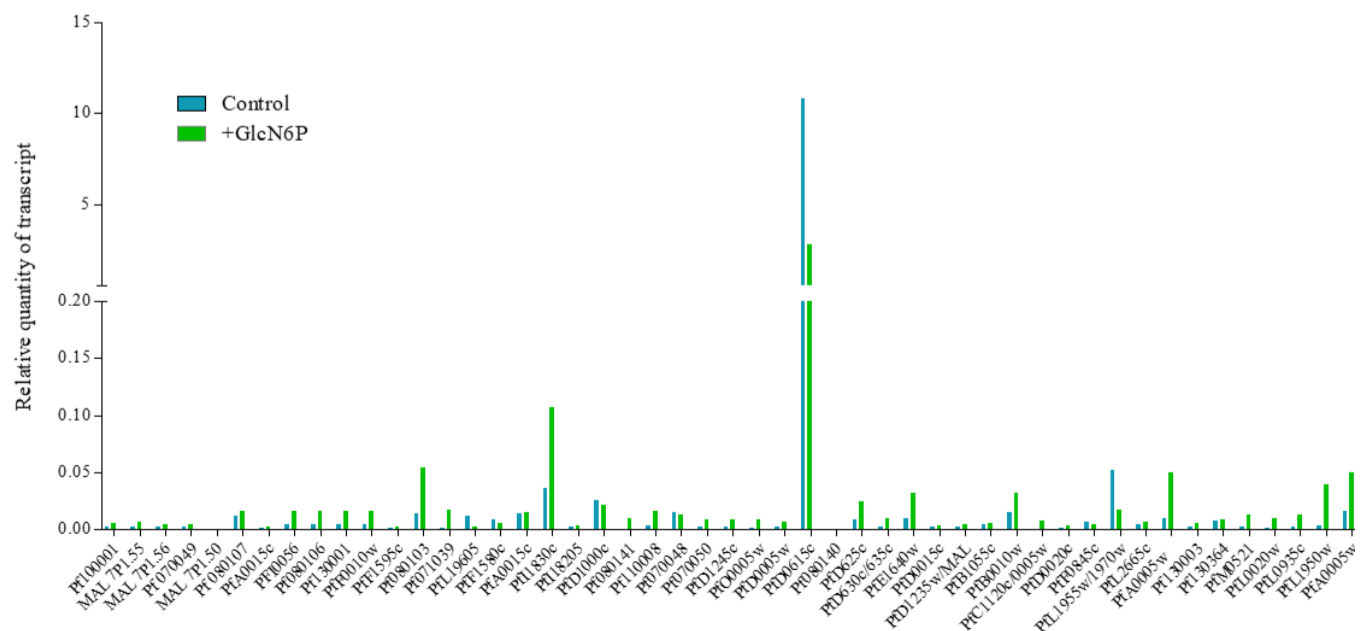 of GlcN6P during 48 h. The cultures were harvested and total RNA was extracted and the relative quantity of all *var* gene transcripts was analyzed via RT-qPCR using primers specific for each of the 57 different *var* genes identified in the *P. falciparum* 3D7 genome, described in (161). Relative transcript quantities are shown.

After one cycle of *Pf*LSD1 knockdown, we did not observe any significant variations in the *var* gene transcriptional profile, when we compared the glucosamine-treated parasites with the untreated parasites. Perhaps, the *Pf*LSD1 protein is not involved in the control of *var* gene regulation, or the partial knockdown is not sufficient to reveal fundamental changes in *var* transcription.

4.2.2 Possible involvement of *Pf*LSD1 in the regulation of sexual development genes.

In order to elucidate the possible participation of *Pf*LSD1 in the regulation of other genes/gene families, we performed an evaluation via RNA-seq analysis.
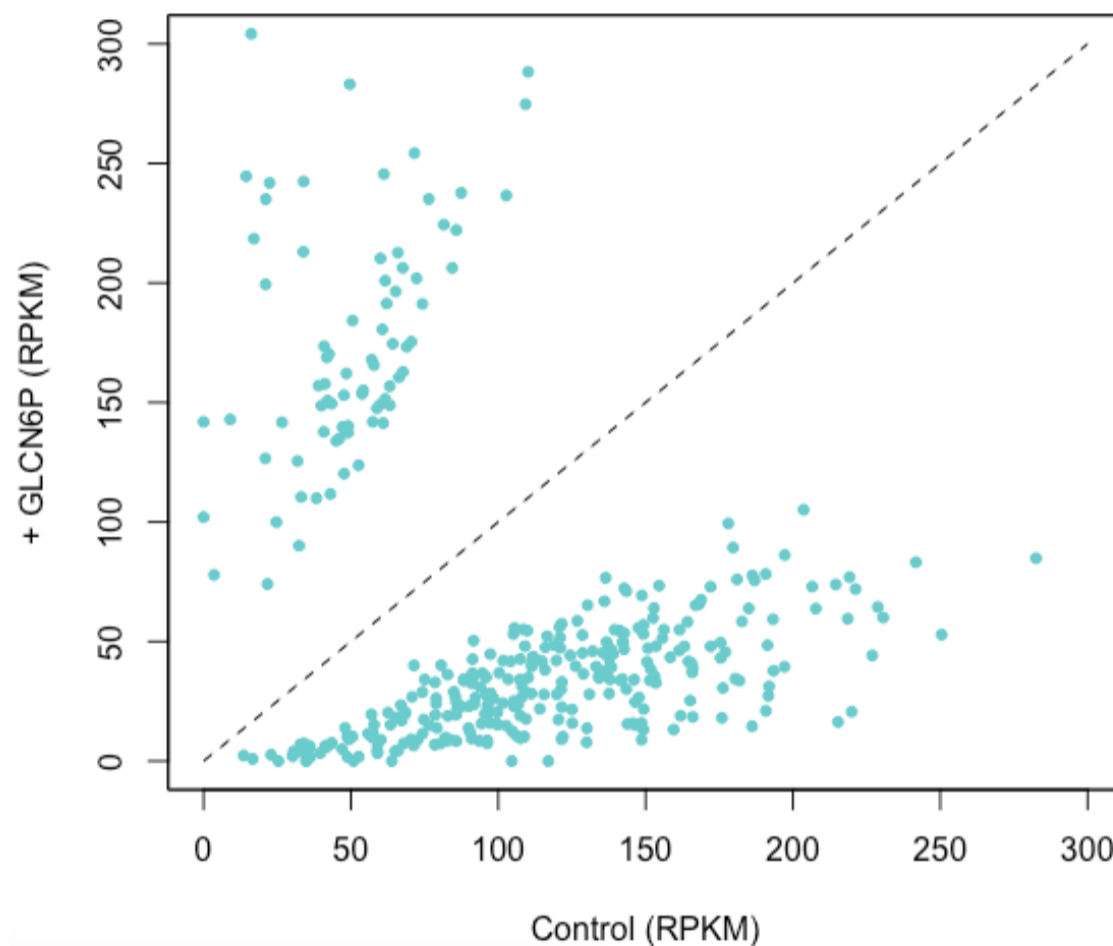
Figure 21. *Pf*LSD1 in the regulation of other gene families.
Expression RPKM values of the transcripts in GlcN 6P treated vs. untreated NF54::*Pf*LSD1-GFP-HA-glmS-parasites was determined by RNA-Seq analysis. Two independent biological replicates (p-value < 0.05). The cultures were synchronized in trophozoites stage, split in two, and was treated or not with 2.5mM of GlcN6P during 48 h. The samples were analyzed in biological duplicates. After that, the cultures were harvested and total RNA was extracted. The cDNA libraries were made as explained in section 3. The data replicates were analyzed using the Bonferroni test.

From the total data set generated in the experiment, we arbitrarily established a cutoff of 2 or -2 of the fold of change to select the significantly different expressed genes. See annex 3.

In figure 21 are shown the total of genes which had their transcript quantities significantly changed when the *Pf*LSD1 protein was depleted (control vs GlcN6P treated).
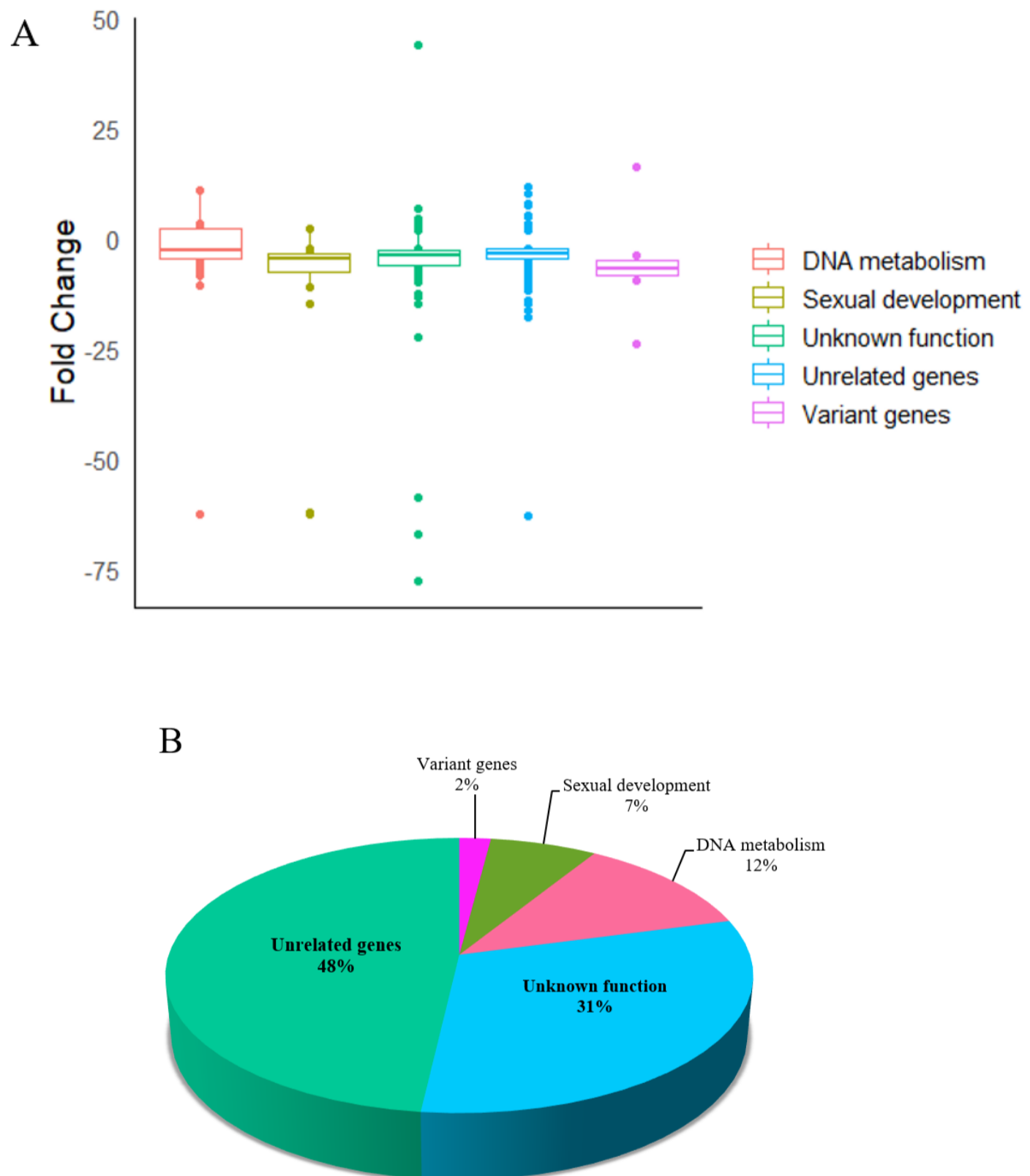
Figure 22. *Pf*LSD1 is related to the epigenetic regulation of the sexual stage development genes.
A. Functional categorization of the differential 367 expressed genes. DNA metabolism, Sexual development, Unrelated genes and Unknown function genes. For the full list, see annex 3. **B.** Percentage of genes present in each category.

With the total amount of genes which showed significant changes in their transcript quantities, we generated different categories based on their Gene Ontology predicted function (figure 22). We found two big groups of genes which have related functions, a group of genes which are related with the DNA metabolism (12 %) and a group of genes that are related to the sexual development stages (7 %). The genes which have no related functions, were grouped in the "unrelated genes" category (48 %), and the

genes described as *conserved protein, unknown function*, were classified as genes with unknown function (31 %), and in a small percentage, the variant genes group with a 2%.
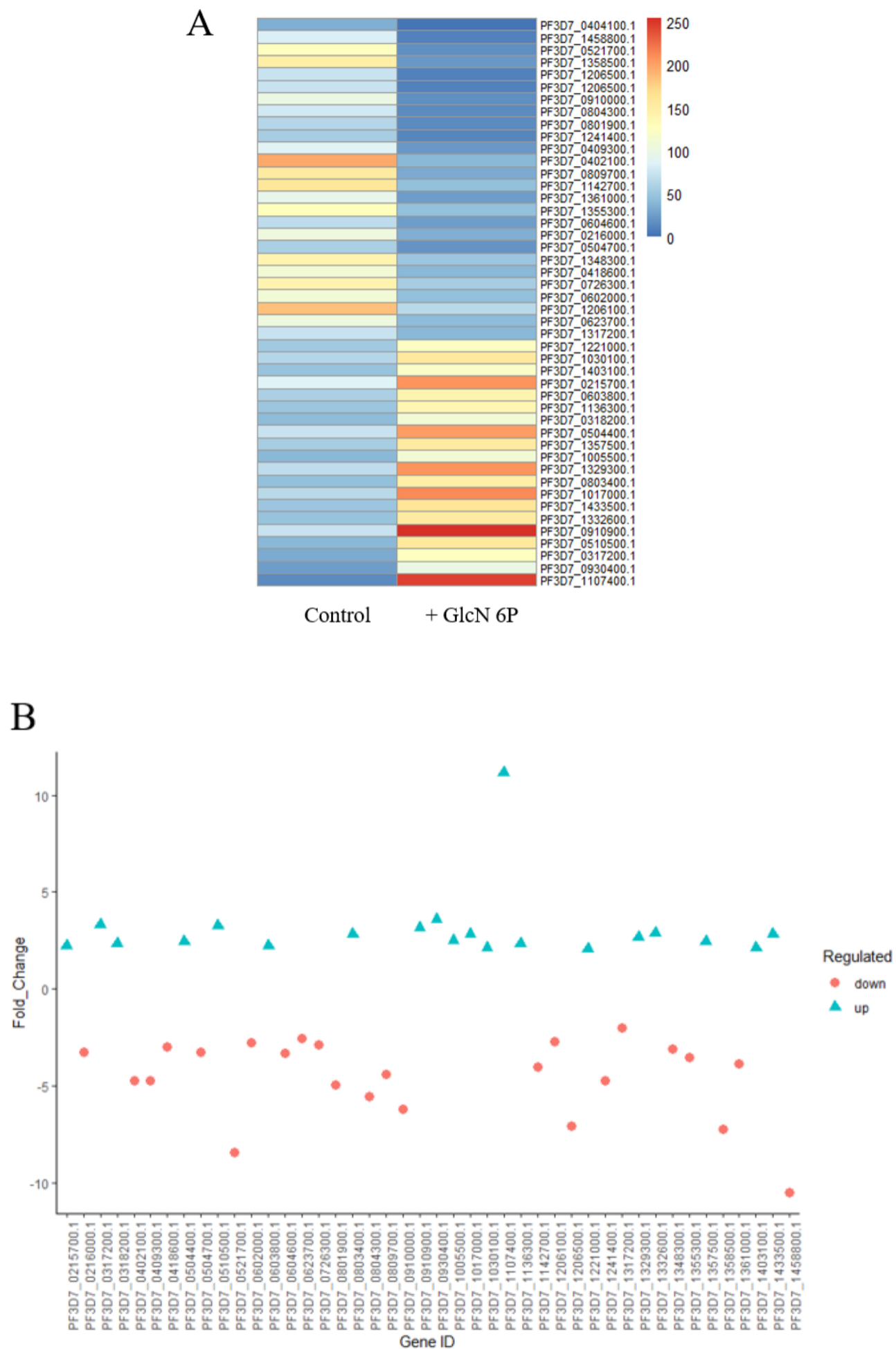
Figure 23.Influence of the *Pf*LSD1 presence in DNA metabolism genes regulation.
Genes with significant transcript quantity change under *Pf*LSD1 knockdown, classified in the DNA metabolism group. **A.** Expression levels of transcripts from the DNA metabolism related genes. Average of the RPKM expression values from each treatment (control and GlcN 6P treatment), from the two

independent biological replicates. **B.** Fold of change average from the two independent biological replicates. Each point shows a value of the fold of change of each DNA metabolism-classified gene.

A total of 46 genes related with the DNA metabolism were found to present a significant (more that 2 or less than -2 of fold of change, with a p value of ≥ 0.05) fold change when we compare the two conditions, in GlcN 6P treated vs untreated NF54::*Pf*LSD1-GFP-HA-glmS-parasites. In this category, we found many important genes related to epigenetic control. For example, *pfset4* (PF3D7_0910000.1) was proposed to encode a putative H3K4 methyltransferase, *ap2-sp2* (PF3D7_0404100.1) an ApiAP2 transcriptional factor, *lsd2* (PF3D7_0801900.1) that encodes a putative second lysine-specific histone demethylase (homolog of *Pf*LSD1), *pfset6* (PF3D7_1355300.1) which encode for another histone-lysine N-methyltransferase, *ap2-g3* (PF3D7_1317200.1), another ApiAP2 transcription factor, *pfset10* (PF3D7_1221000.1) histone-lysine N-methyltransferase, H3 lysine-4 specific, and two other genes that encode for putative methyltransferases (PF3D7_1142700.1 and PF3D7_0409300.1). A high number of DNA helicase encoding genes, ATP-dependent helicases and zinc finger encoding genes appeared also with different transcript quantities, as well as other genes that are somehow related with DNA remodeling and conformation.
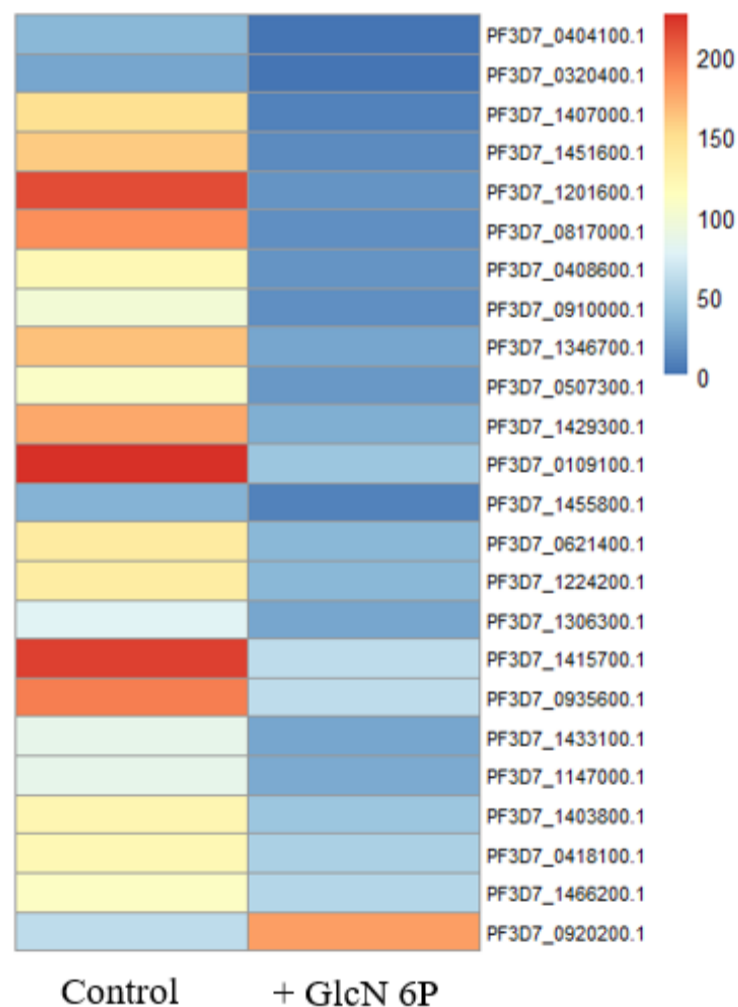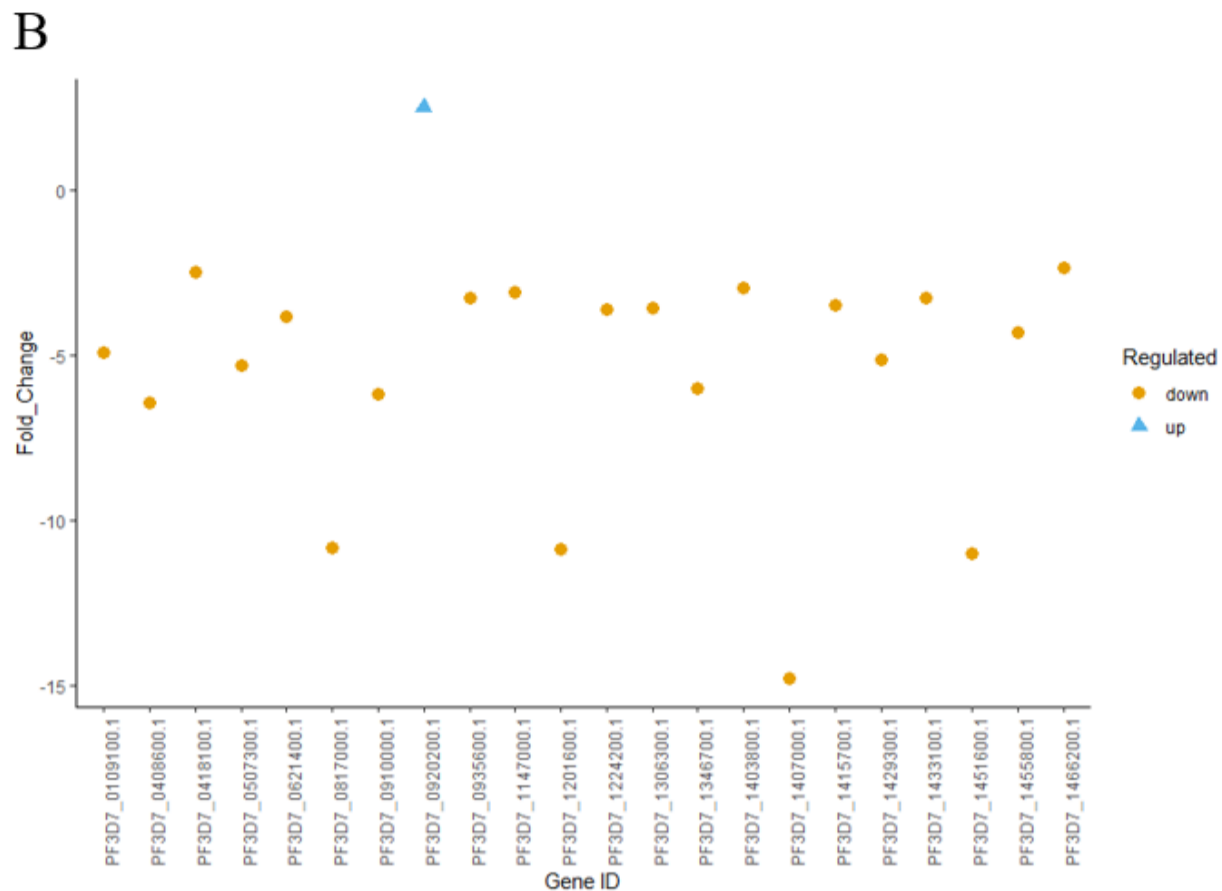
Figure 24. Influence of the *Pf*LSD1 presence in sexual development genes expression.
Genes with significant transcript quantity change under *Pf*LSD1 knockdown, classified in the group of genes putatively involved in sexual development. **A.** Expression levels of transcripts from the sexual development related genes. Average of the RPKM expression values from each treatment (control and GlcN 6P treatment), from the two independent biological replicates. **B.** Fold of change average from the two independent biological replicates. Each point shows a value of the fold change of each sexual development-classified gene.

RNAseq analysis also revealed 24 genes that appear differentially regulated during the knockdown of *Pf*LSD1, and these are putatively involved in the sexual development or even essential for the survival of the parasite in the mosquito host.

Table 7. Summary of the sexual development genes differentially regulated in the absence of *Pf*LSD1 protein.
The predicted product for each differentially regulated gene and its respective symbol is shown. The stage where maximum transcript abundance is shown together with the average fold change under knockdown of PfLSD1.

Gene ID	Protein product description	Symbol	Stage of expression	Fold of change
PF3D7_1407000	LCCL domain-containing protein	CCp3	Ookinete	-14.78
PF3D7_0109100	LCCL domain-containing protein	CCp5	Gametocyte V	-4.91
PF3D7_1451600	LCCL domain-containing protein	LAP5	Ookinete	-11.01
PF3D7_0408600	Sporozoite invasion-associated protein 1	SIAP1	Sporozoite	-6.44
PF3D7_0507300	Subtilisin-like ookinete protein SOPT	SOPT	Ookinete	-5.28

PF3D7_1466200	Early gametocyte enriched phosphoprotein EGXP	EGXP	Ookinete	-2.34
PF3D7_1201600	NIMA related kinase 3	NEK3	Gametocyte V	-10.85
PF3D7_1224200	BRO1 domain-containing protein, putative	-	Gametocyte V	-3.62
PF3D7_1415700	Serine palmitoyltransferase, putative	SPT	Gametocyte II	-3.49
PF3D7_0621400	*Pf*77 protein	ALV7	Ookinete	-3.81
PF3D7_1147000	Sporozoite and liver stage asparagine-rich protein	SLARP	Sporozoite	-3.07
PF3D7_0817000	NEDD8-activating enzyme E1 catalytic subunit, putative	UBA3	Gametocyte V	-10.84
PF3D7_1346700	6-cysteine protein	P48/45	Gametocyte V	-5.98
PF3D7_0935600	Gametocytogenesis-implicated protein▲	GIG	Gametocyte II	-3.26
PF3D7_1433100	HID1 domain-containing protein, putative	-	Gametocyte V	-3.26
PF3D7_0920200	CS domain protein, putative	-	Oocyst	2.49
PF3D7_1429300	CPW-WPC family protein	-	Gametocyte V	-5.14
PF3D7_1403800	Nuclear formin-like protein MISFIT, putative	MISFIT	Gametocyte V	-2.96
PF3D7_0418100	Protein SOC1, putative	SOC1	Oocyst	-2.49
PF3D7_1317200	AP2 domain transcription factor AP2-G3, putative*▲	AP2-G3	Gametocyte V	-2.03
PF3D7_1306300	SAM dependent methyltransferase, putative*	-	Gametocyte V	-3.54
PF3D7_1455800	LCCL domain-containing protein	CCp2	Gametocyte V	-4.30
PF3D7_0404100	AP2 domain transcription factor AP2-SP2, putative*▲	AP2-SP2	Oocyst	-62.50
PF3D7_0320400	Oocyst capsule protein Cap380	Cap380	Oocyst	-62.07

▲ Also highly express in the blood stage.

* Also classified in the DNA metabolism category.
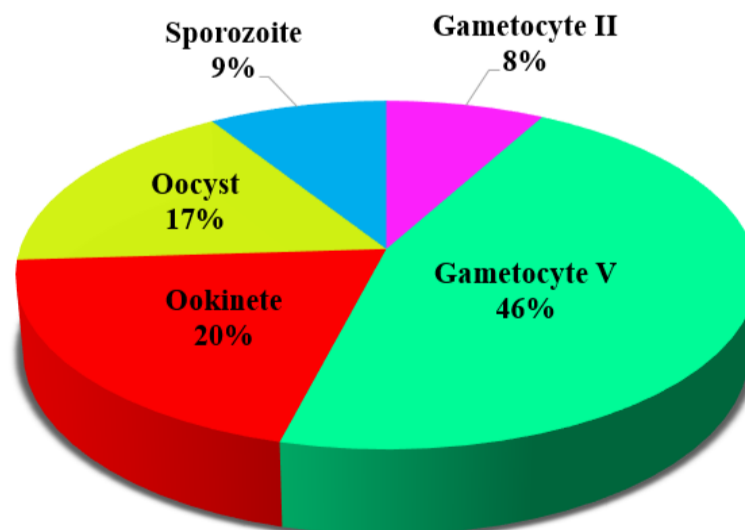
Figure 25. Percentage of genes related with sexual development from each sexual stage in *Plasmodium falciparum*.

4.2.3 The possible influence of *Plasmodium berghei* LSD1 in the asexual and sexual development of the parasite live cycle.

Given the immediate impossibility to analyze the importance of PfLSD1 in sexual or mosquito stage, we opted to analyze the function of LSD1 in the rodent parasite *P. berghei* where it is possible to monitor parasites *in vitro* at least until ookinete stage. Of note, LSD1 is widely conserved among apicomplexa (171) and an *in silico* comparison of *Pf*LSD1 and *Pb*LSD1 showed that both alleles share a similar transcription pattern, being highly expressed in sexual forms (ookinetes/oocysts) of the parasites (www.plasmodb.org). Also, both *lsd1* genes are in synteny. Moreover, both LSD1 proteins are huge proteins, 350KDa for *Pf*LSD1 and 325KDa for PbLSD1, are predicted to have nucleolar localization and common enzymatic domains, as shown in figure 24.

Figure 26. *Plasmodium* LSD1 Proteins.
A. Graphical summary of the conserved domains present in the *Pf*LSD1and PbLSD1 proteins, using *NCBI conserved domain* software. **B.** Amino acid alignment of the Zn finger and histone 3 binding site present in the PHD domain of both LSD1 enzymes using the *Clustal Omega < Multiple Sequence*

Alignment function. In the alignment, an <u>*</u> (asterisk) indicates positions which have a single, fully conserved residue. "<u>:</u>" indicates conservation between aminoacids of highly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix. A "<u>.</u>" indicates conservation between amino acids with less similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix. A classical PHD2_CHD_II domain (cd15532) was used as reference for the conserved site alignments.

As shown in figure 25, *Pb*LSD1 also has the flavin-containing amine oxidoreductases domain, grouping it in the flavin-containing oxidase super family, as all LSD1 proteins. In general, both proteins share a similar domain architecture. *P. falciparum* and *P. berghei* LSD1 protein might exert the same functions, regarding a possible H3K4/9 demethylation activity. Furthermore, there is 100% percentage identity of the Zn finger site, and 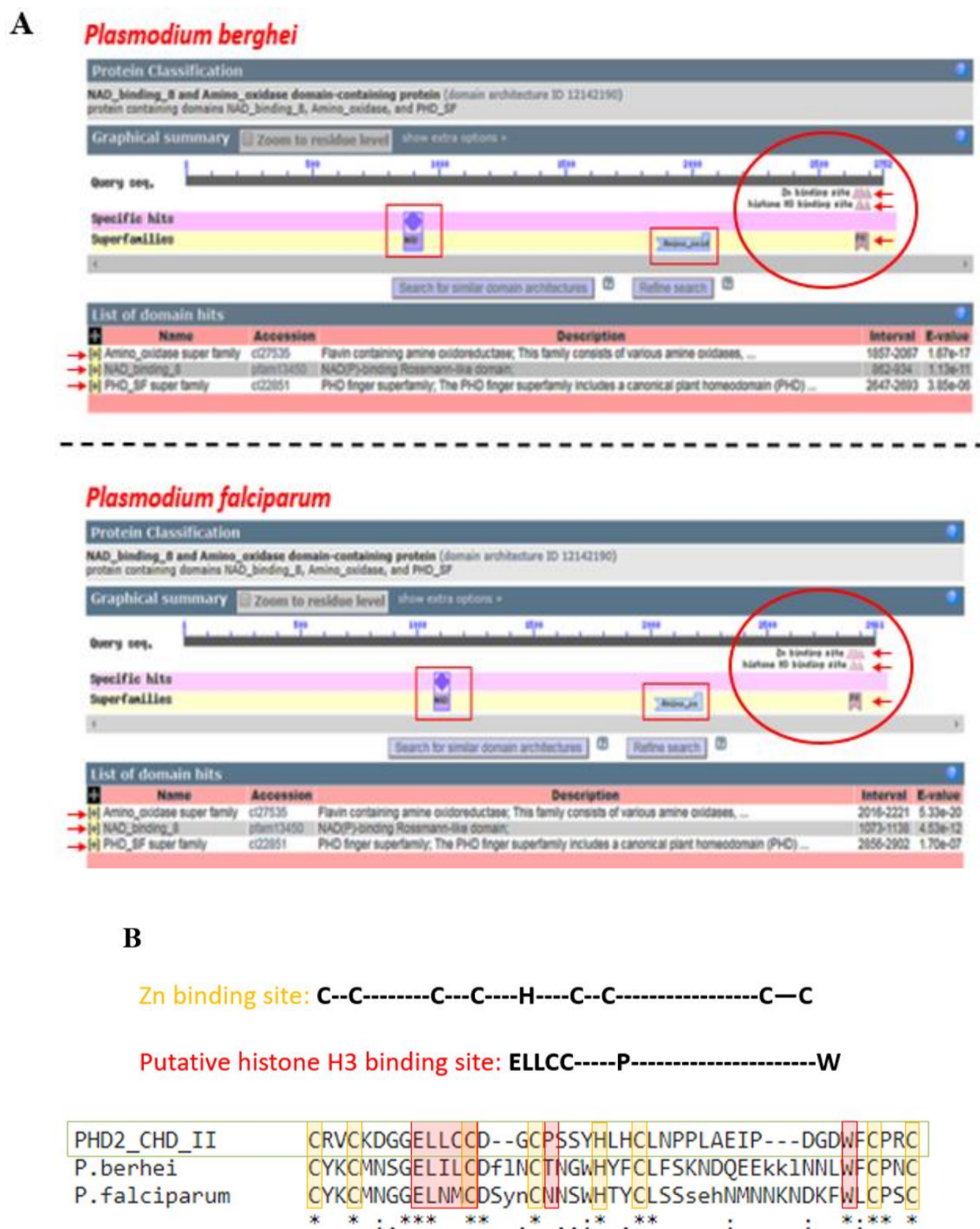about 70% in the H3 binding site, present in their PHD domains. This suggests that *Pf*LSD1and PbLSD1 may share similar functions, probably most accentuated in ookinete/oocyst stage parasites. Based on this, we decided to analyze the *Pb*LSD1 constructing a transgenic *P. berghei lsd1* knockout parasite line.

4.2.3.1 The strategy *PbOokluc-LSD1^{KO}*.

The PbLSD1^{KO} was done in the background of the *PbOokluc* parasite line and developed in collaboration with Professor Daniel Y. Bargieri from the *Laboratório de Genética Molecular* at the Department of Parasitology/ICB-USP. The PbOokluc parasite line contains a nanoluciferase under control of the ookinete specific promoter of the CTRP gene (154). We constructed the pPbLSD1KO transfection vector, which contains the *Pblsd1* 5′UTR (963 bp) and 3′UTR (755 bp) as homology sequences flanking the hDHFR and mCherry cassettes, as described in methods.

Transfected parasites were positively selected by the administration of pyrimethamine for 72 h. The drug was administered in the drinking water of mice infected with the transfected parasites 24 h after the transfection. After selection, the transformed parasites were cloned by limiting dilution.

To confirm appropriate integration, a PCR analysis was performed, using gDNA from transfected and cloned parasites (Figure 27).
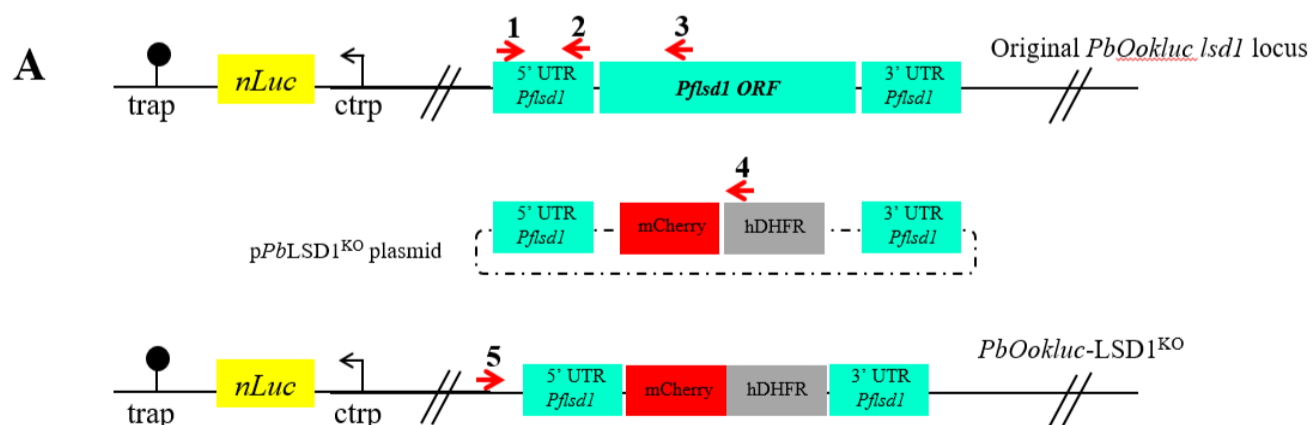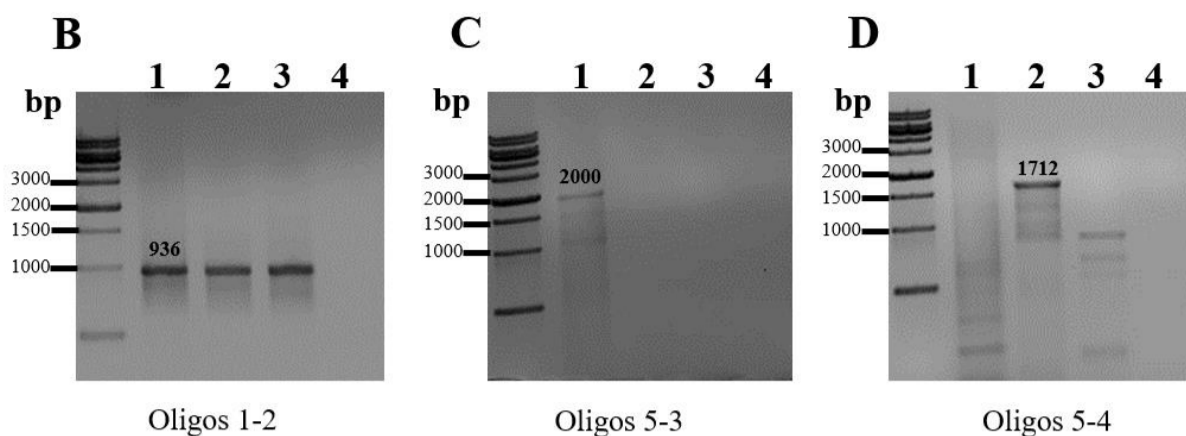
Figure 27. Outline of the production of a *Pb*LSD1 knockout parasite line.
A. Integration into the *P. berghei* ANKA Ookluc strain was performed using two homology regions flanking a mCherry/hDHFR cassette. PCR analysis to detect the *PbOokluc-lsd1* integration was conducted using the oligonucleotide primers indicated as red arrows. **B**. Oligos 1-2 was used to confirm the quality of genomic DNA which was analyzed. **C.** Oligos 5 and 3 were used to detect the native *Pb*LSD1 locus, while **D.** Oligos 4 and 5 detected integrated/recombined locus. The following DNAs were analyzed: lane **1** gDNA from *PbOokluc* parasites, lane **2** gDNA from *PbOokluc-lsd1*[KO] parasites, and **3** DNA from the *PbOockluc-lsd1*[KO] vector, and lane **4** water (negative PCR control).

In Fig 27 A, a 936 bp PCR product was detected in all DNA samples, indicating the integrity of analyzed samples. In Fig 27 B, the wildtype, unrecombined LSD1 locus was only detected in *PbOokluc* parasites, resulting in the expected 2000 bp amplicon. In Fig 27 C, the expected 1712pb PCR product was observed only in recombinant parasites (*PbOockluc-lsd1*[KO]). A smaller sized fragment is observed with the same primer pair also in transfection plasmid DNA, probably result of false priming or excessive template.

4.2.3.2 The *Pb*LSD1 knockout does not influence the viability of *P. berghei* blood stage forms.

To analyze the possible influence of the PbLSD1 on the viability of blood stage parasites, Balb/c mice were challenged with $5*10^3$ *P. berghei ANKA* (control group) and PbOokluc-LSD1KO (from donor mice) parasites. Then, the parasitemia and the gametocytemia were followed by thin blood smears in 2-day intervals (Figure 28).
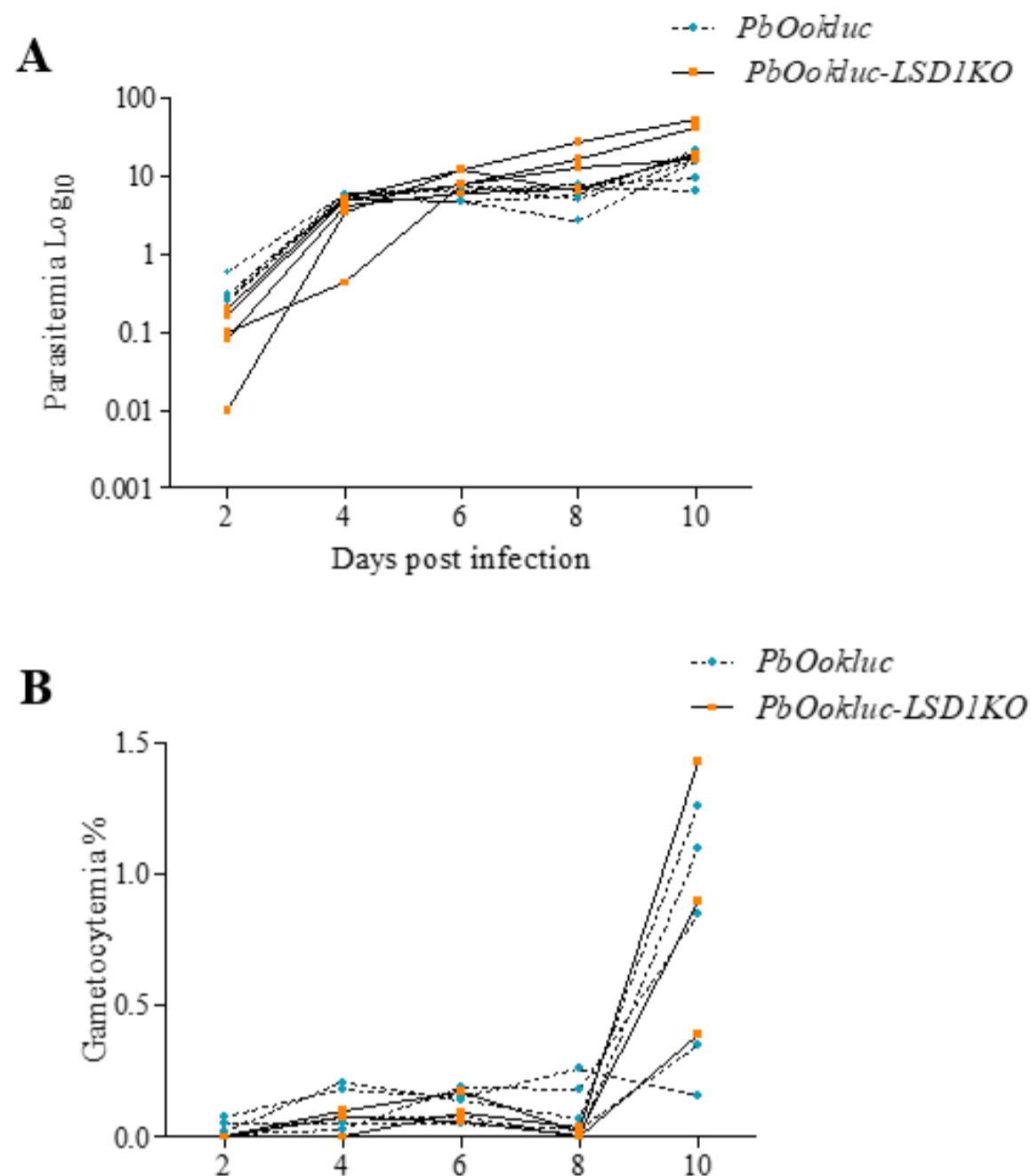
Figure 28. Growth analysis of *PbOokluc-LSD1^{KO}* Parasites.
A. Two mice groups, of five animals each, were challenged intravenously *P. berghei ANKA* parasitized erythrocytes or *PbOokluc-LSD1^{KO}*. The total parasitemia was followed during 12 days by thin blood smears. **B** The percentage of gametocytemia was monitored via thin blood smear counting.

Similar as observed before in the *P. falciparum Pf*LSD1-KD parasites (NF54::*Pf*LSD1GFPHA-glmS) growth analysis (figure 15 D), *P. berghei* wild type and *PbOokluc-LSD1^{KO}* parasites showed no significant difference in the growth and gametocytemia patterns (figure 28 A and B). This indicates that the absence of the LSD1 protein in *P. falciparum* and *P. berghei* did not affect the parasite development, at least not in the blood stages. This points to the view that LSD1 is not necessary in asexual stages even in the presence of the host immune response.

4.2.3.4 The *Pb*LSD1 influence on the sexual development of *P. berghei*.

Given that LSD1 is mostly expressed in sexual stages, we evaluated the influence of a PbLSD1 knockout in the development to sexual stages. For this, we analyzed the fertilization and sexual differentiation capacities in *PbOokluc-LSD1^{KO}* parasites in *in vitro* approaches (figure 29), taking advantage of the reporter protein nanoluciferase under the control of the ookinete *crtp* promoter which accurately indicates transformation into ookinetes.
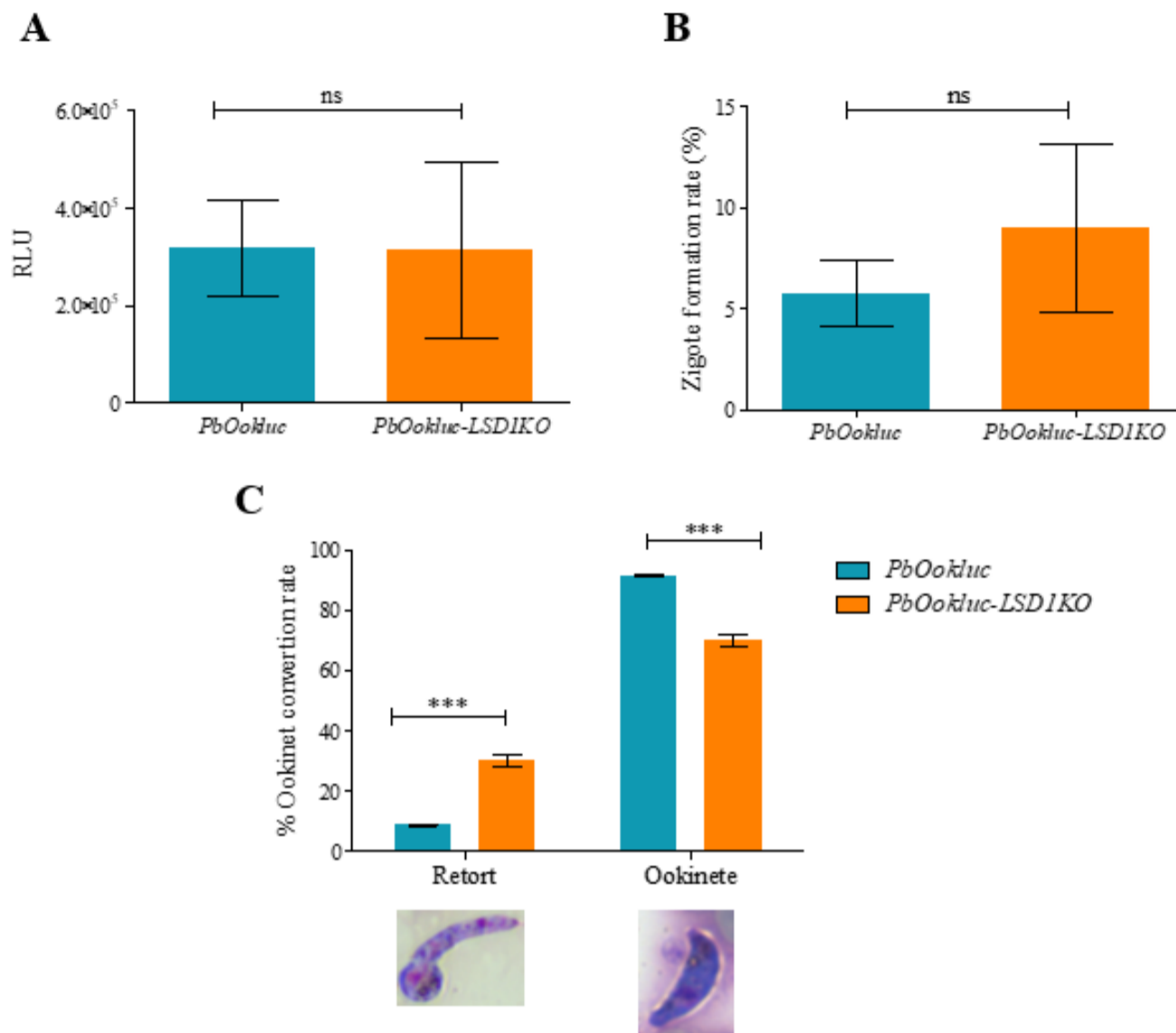
Figure 29. *In vitro* analyses of sexual differentiation in *PbOokluc-LSD1*[KO] parasites.
A. nLuc activity in conversion assay of PbOokluc and PbOokluc-LSD1[KO] parasites. For this, 4µl of blood from the infected mouse, with ~0.4 % of gametocytemia, was added to 80 µl of ookinete medium. After 24 h incubation at 21 °C, nLuc activity was determined. The nLuc activity is expressed as relative luminescence units (RLU). Each sample of the conversion assay was morphologically analyzed by thin blood smear and the percentage of zygote **B**. mature **C.** and immature (retort) ookinete forms, were counted. The significance of differences between the samples was calculated using the Two-way ANOVA test. "ns" means no significant difference, and *** highly significant difference with a P value of < 0.0001. A graphical representation of four biological replicates is shown.

As shown in figure 29 A, no significant differences between the *PbOokluc* and *PbOokluc-LSD1*[KO] conversion into zygotes could be detected. Thus, the fertilization capacity of this parasites is not affected by the absence of the LSD1 protein (figure 29 B). However, when we analyzed the percentage of mature ookinete forms, we found that a certain percentage of the converted ookinetes didn't reached the mature forms, as expected after 24 hours of maturation. These parasites remain in an immature intermediary ookinete form, the so-called retort forms. Our data showed about 30% of retort forms in the *PbOokluc-LSD1*[KO] conversion assay while only 8% in the *PbOokluc* parasites were observed (Figure 29 C).

These data suggest that *Plasmodium* LSD1 enzyme may exert an important role in later sexual stages which in the *P. berghei* context cannot yet be approached in Brazil due to the absence of a competent mosquito vector, such as *Anopheles stephensi*.

5. DISCUSSION

5.1 *rif* genes and their mode of transcription: Independent regulation of *P. falciparum rif* gene promoters and the impact of the transcriptional activation of a single *rif* gene

In our first approach we tried to elucidate if *rif* gene transcription followed the allelic exclusion model as happens for *var* genes, using a similar approach as was successfully done for this variant gene family (90). From three different constructs, only one presented a variable transcription activity, which is a pre-requisite of the corresponding RIFIN to function in antigenic variation. For the other two constructs, which also did not show integration at genomic sites, we found either a permanently activated or silenced status. Notably, all tested 5′-upstream regions have several, almost completely identical regions in the genome but only one of the *rif* genes controlled by these identical regions seemed transcriptionally active to the same degree as the transgene, strongly pointing to epigenetic modifications for activation or silencing rather than sequence requirements.

It is probable that the chromosomal context surrounding *rif* 5′ups-regions may play a pivotal role in *rif* expression. All of the three 5′-ups regions were designed in the same way and contained almost 1500 nt upstream of the corresponding *rif* ATG, and all three upstream regions contained either terminator regions or head-to-head promoter regions of the adjacent gene locus possibly including heterochromatin boundary regions. However, only two 5′-ups were functional in providing a sufficient number of transcripts and only the apparently integrated allele showed a tunable behavior. In this part of the work we showed that transcription of the *rif* gene family, specifically of the RIFIN A subset, is probably not controlled by a mechanism related to allelic exclusion as is valid for *var* genes.

However, not all *rif* genes are transcribed, so there still must be some kind of controlled activation and repression. It also still remains elusive if the B subfamily of *rif* genes is differentially regulated, given the fact that we found only one B-type *rif* gene activated in our cultures. Interestingly, B type RIFINs are expected to remain inside the parasite, while RIFIN A type proteins become exposed. Regarding the function of RIFINs, a recent study of Saito *et al* 2017 (84), showed that A-type RIFINs interact with LAIR1 receptors on certain lymphocytes, leading to a downregulation of their function. In this sense, it could be hypothesized that the more A type RIFINs are expressed, the better are chances for successful survival of the parasites by more effectively downregulating the immune response. Conversely, the recent detection of so called "public antibodies", recognizing many different *P. falciparum* strains, may target A type RIFINs, since these antibodies show a LAIR1 motif inserted in the Fv region (172), which possibly block LAIR1 receptor-RIFIN interaction.

Since our first study had the disadvantage of ectopic insertion creating a functional (tuneable), nevertheless artificial locus, we opted for an approach which would provide higher chances of correct integration into a genomic *rif* locus. Initially, three different *rif* genes, which were chosen either for their RIFIN type (A and B) or for a functional property (rosetting ligand, PF3D7_0100400 (*rif*r)). Instead of highly A/T rich promoter regions, we chose *rif* coding regions for integration via homology-driven recombination. While two of the three showed either difficulties in obtaining integration at all or integration at the expected locus, we focused on the parasite line NF54::p*rif6*GFPHA2A-BglmS (PF3D7_0632200 (*rif 6*) a RIFIN A encoding gene). When we evaluated the capacity of this parasite line to sustain growth in the presence of blasticidin, which should select for parasites actively transcribing from the *PF*3D7_0632200 *rif* promoter, very few parasites were recovered and no expansion of parasitemias was observed. Nevertheless, the culture also did not completely die, which is the case for untransfected parasites. Also, a slight increase in transcript abundance (or selection of blasticidin deaminase producing parasites) was observed when submitting parasites to blasticidin

treatment. This may be interpreted that the activity of the controlling *rif* promoter in the conditions of our *in vitro* culture is not sufficient. When monitoring the relative *rif, gfp* and *bsd* transcript abundance, we observed that the highest values never reached the transcript abundance of the internal control gene (t-seryl ligase). In a previous work, much higher values were found in wildtype parasites (86). It is possible that different *rif* loci have different maximum transcript quantities.

Due to the very low biomass of parasites recovered after blasticidin treatment, our initial plan to perform genome wide *rif* transcript analysis in these parasites could not be performed. Although there are novel technologies which may permit single cell transcriptomic analysis, it may be expected that the view about how *rif* gene transcription is globally controlled may not be altered.

The difficulty to modulate *rif* transcription was also present in other studies. For example, in 2012 Witmer and colleagues (88) tested the 5′-ups region of *rif* *PF*3D7_1300400 (the same gene of our p*rif*rGFPHA2A-BglmS construct). In their study, in one construct the *rif* promoter portion could be activated, but adjacent head-to-head *var ups* couldn't be activated at all. In the second trial, a slightly modified construct, both promoters could be selected for active transcription, but the level of induction of the *rif* 5′-*ups* in their study was not significantly different, similar to our attempt with the NF54::p*rif*6GFPHA2A-BglmS parasite line. This indicates that a larger context surrounding *rif* promoters may influence their activity. The mechanism how *rif* gene transcription is controlled remains still elusive. An influence of nuclear architecture and chromatin positioning on transcriptional activity as occurs with *var* genes (173) may play a role but has still to be tested.

5.2 The role of Lysine-specific histone demethylase 1 (LSD1) in malaria parasites.

LSD1 was the first protein reported to exhibit histone demethylase activity (125). LSD1 enzymatic activity is widely conserved among the different biological kingdoms. Classically, this enzyme removes specifically the methyl groups H3K4me1/2 and H3K9me1/2 from the histone H3 N-terminal (123). LSD1 protein has been described in Apicomplexa with the classical mechanisms of action (147).

The LSD1 gene (*PF*3D7_1211600) has been successfully annotated in *P. falciparum* (110). A recent study mentioned the presence of another *Pf*LSD gene (*PF*3D7_0801900), called *Pf*LSD2 (supplementary table 1 (168)), and suggested that the putative PfLSD2 interacted with AP2-G, important for gametocyte induction (174). In our RNAseq results, PfLSD2 appeared differentially regulated by knockdown of PfLSD1. Classically, LSD1 and LSD2 possess the same function (175) however LSD2 have a PHD domain contrary to LSD1 (176). Interestingly, *Pf*LSD1 possess a PHD domain (*https://plasmodb.org/plasmo/app/record/gene/PF3D7_1211600#category:sequence*) while the PHD domain in PfLSD2 was predicted only with a very low probability (174), when joining conserved sequences of the entire polypeptide, but not by simple prediction in NCBIs conserved domain server (compare *Pf*LSD2 (*https://plasmodb.org/plasmo/app/record/gene/PF3D7_0801900#category:sequence*) (figure30). According to the literature the presence of this PHD domain, is the basic structural characteristic which differentiates LSD2 from LSD1 proteins (123).
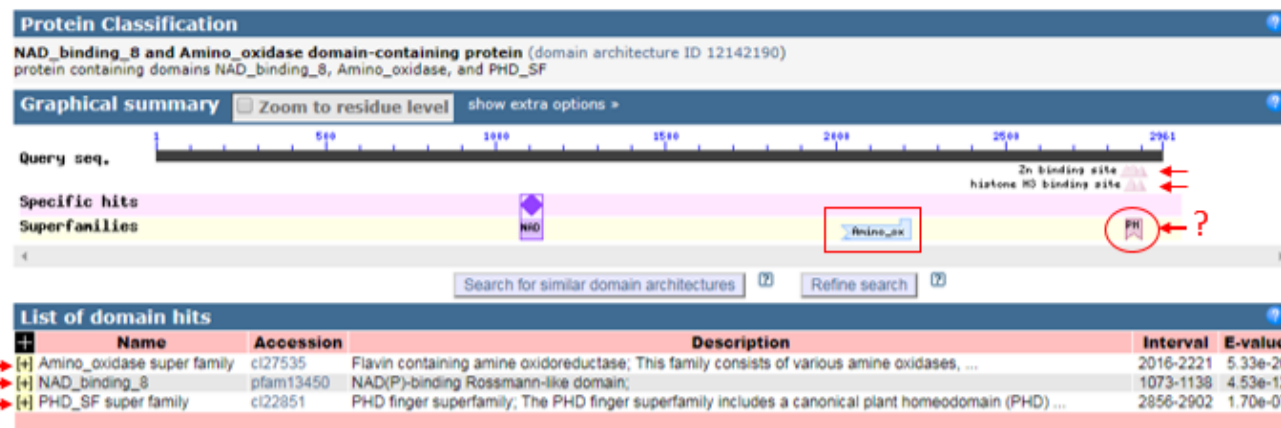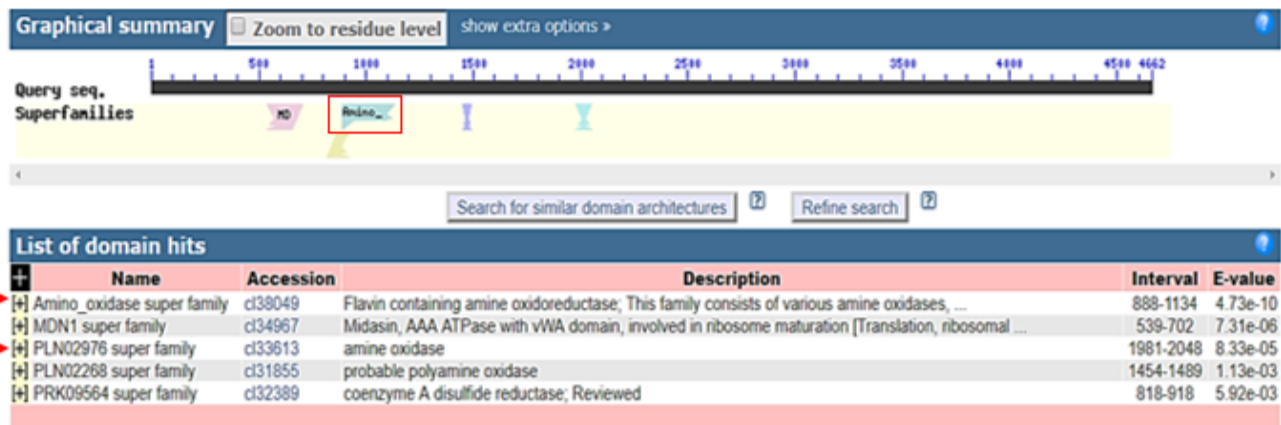
Figure 30. Graphical summary of the conserved domains present in the *Pf*LSD1and *Pf*LSD2 proteins.
Using the *NCBI conserved domain* tool, no PHD domain is encountered in the full length PfLSD2 polypeptide.

As described above, LSD1 and LSD2 have a redundant role. Both enzymes specifically demethylate H3K4me1/2 and H3K9me1/2 from the histone H3 (175). It is in accordance with our results which show no significant alteration of the phenotype and the *var* gene transcriptional profile in *Pf*LSD1-KD parasites during the intraerythrocytic cycle. Similarly, *Pb*LSD1-KO parasites did not exhibit any phenotype change compared to wild type parasites in the blood stage phase. Possibly, LSD2 protein exerts the function of LSD1 when LSD1 is absent.

In other organisms, the LSD1 proteins directly interact with histone 3 to demethylate lysine rests (177). Although the magnification is obviously not sufficient, we observed a partial co-localization of *Pf*LSD1 with the histone 3 at the nucleus area (figure 17). A higher resolution, for example by transmission electron microscopy and immuno-gold staining of both H3 and LSD1 may show better spatial proximity. Likewise, we observed that the *Pf*LSD1 protein was also localized in the area surrounding the nucleus. Remarkably, the periphery of the nucleus can be a transcriptionally active chromatin region (147). The Alan Cowman group reported in 2010 the co-localization of *Pf*LSD1 with several proteins of the chromatin remodeling protein complex such as *Pf*SET10 (PF3D7_1221000) and *Pf*SET4 (PF3D7_0910000) H3K4 specific histone-lysine N-methyltransferases (115) and *Pf*SNF2, a helicase and ISWI homolog protein (147). They showed that these proteins are distributed within the same area in speckles. These data are consistent with our data showing a robust, precipitable *Pf*LSD1-Histone 3 interaction.

A pull down assay followed by a MS of the CoIP-proteins was performed to address the *Pf*LSD1-protein association profile. The MS analysis (table 6 appendix 2), allowed us to detect the presence of five

histones (Histone H2B, Histone H2A, Histone H3, Histone H2A and Histone H4) belonging to the nucleosome core. Furthermore, we also detected several ATP-dependent helicases which are nucleolar proteins with chromatin remodeling proprieties. This strongly indicates that *Pf*LSD1 interacts with complex network of proteins. This protein complex probably remodels the chromatin, participating in the modulation of gene expression or promoter activity. Even more interesting, as predicted by Alan Cowman's group, we found the presence of *Pf*SNF2L protein among the co-immunoprecipitated proteins. *Pf*SNF2L is the *P. falciparum* homolog of the SNF protein, which is part of ISW/SNF chromatin remodeling protein complex. Apicomplexan ISWI/SNF proteins can interact with PHD-finger proteins (148). PHD-finger proteins are involved in chromatin-mediated transcription regulation and it has been shown to bind to all nucleosomal histones (178). Briefly, ISWI/SNF-PHD-finger protein belong to a large functional protein complexes that include other chromatin-modifiers, such as acetylases, and methylases enzymes, affecting the chromatin structure in an ATP-dependent manner (103). The ISWI/SNF complex has been proposed to recognize H3K4 methylated tails and mediate transcriptional activation, permitting access and recognition of the H3K4 demethylases (179). Furthermore, *Pf*SNF2 (180), the NAP nucleosome assembly protein and CHD1 (181) have been identified to be part of this transcription/chromatin remodelling network.

We propose that *Pf*LSD1 participates in the regulation of the gene transcription and acts as epigenomic controller. This regulation occurs through the interaction of *Pf*LSD1 with the histone H3, via its H3 binding site. The enzyme probably recognizes and removes the mono- or di-methyl groups of theirs lysine 4 residues, with the help of *Pf*SNF2 and others ATP-dependent helicases. The demethylation allow the subsequent acetylation event (182). As in other organisms, this complex is probably composed of some acetyltransferase and methyltransferase proteins (123). However, we did not detect these enzymes in our co-immunoprecipitation assay. We can hypothesize that some acetylases and methyltransferases do not or weakly interact directly with *Pf*LSD1.

Interestingly, the DNA/RNA-binding protein Alba1, *Pf*Alba 1, was detected between the Co-IP proteins. The presence of Alba family proteins in eukaryotes is limited to plants and protozoan parasites (183). In *P. falciparum*, *Pf*Alba1 is involved in the mRNA homeostasis and translational repression (184). It has been reported that *Pf*Alba1 is associated with a large number of transcripts and is essential for growth and proliferation during intraerythrocytic developmental stages of parasites. *Pf*Alba1 is also a post-transcriptional gene regulator during asexual blood stages of *P. falciparum* (184), (185).The interactions between *Pf*LSD1 and *Pf*Alba1, reported in our MS-data, suggest that both proteins could belong to a protein complex participating in the post-transcriptional gene regulation.

We also performed another pull down assay to confirm the *Pf*LSD1-Histone 3 interaction. In this pull-down assay we used as bait *Pf*Histone 3. In figure 19, we observed the presence of a high molecular weight protein (more than 260 KDa), among the co-inmunoprecipitated proteins. This protein was detected by an Anti-HA mAb, in a western blot analysis. Our result strongly indicated the presence of the *Pf*LSD1-GFPHA-tagged protein. Taken together, our data shows that *Pf*LSD1 directly interacts with Histone 3 and probably participates in the H3K4me1/2 and/or H3K9me1/2 demethylation processes.

We then performed a comparative RNA-Seq analysis to assess an influence of *Pf*LSD1 on global transcription, evaluating the transcriptional response of *P. falciparum* parasites in the presence of *Pf*LSD1 protein and under knockdown conditions. Our data, show that 367 genes are altered when we compare the two conditions (figure 21). The putative function of each gene was classified within the

following categories: DNA metabolism, variant genes, sexual development genes, genes without evident related function with LSD1 and genes with unknown function (figure 22 A). We observe that 2 % of the variant genes present difference in the relative transcript levels (figure 22 B). These results are consistent with our previous RT-qPCR data (figure 20) showing that the variant gene expression is not altered when less PfLSD1 is present. We hypothesize that *Pf*LSD1 does not participate in the transcriptional memory/switching of *P. falciparum* variant genes. Nevertheless, *Pf*LSD2, which possess a redundant deacetylase function, may exert LSD1's roles in its absence.

The *PfSET10* gene (PF3D7_1221000.1), appeared slightly upregulated with a fold of change of 2.08, in the absence of *Pf*LSD1. The *Pf*SET10 protein is a histone 3 lysine 4 methyltransferase that localizes exclusively at the perinuclear area of *P. falciparum*. *Pf*SET10 protein maintains the transcriptionally permissive chromatin environment and is supposed to retain memory for heritable transmission of epigenetic information during parasite division. *Pf*SET10 is also tightly related to *var* gene expression dynamics (186). On the other hand *pfset4* and *6*, which encode other putative H3K4 methyltransferases (115), (187), appear to be down regulated, in the same condition, with a fold of change of -6 and -4 respectively. The absence of *Pf*LSD1 affect the transcription pattern of *pfset4*, 6 and *10* genes. Our results are in accordance with the hypothesis of Alan Cowman's group (147), who point to a functional link between *Pf*LSD1 and the *Pf*SET4/10 proteins.

Our comparative RNA-Seq analysis results also show the remarkable downregulation of the transcriptional factor *ap2-sp2* and also of *ap2-g3* with a fold of change of -62.5 and -2 respectively, in absence of *Pf*LSD1. Both genes belong to the ApiAP2 family of putative transcription factors. Of note, PfAP2-G or GDV, master regulators of gametocyte induction (188) were not influenced. Recently, AP2-Sp2 TF was shown to be involved in sporozoite formation in the oocyst. The AP2-Sp2 knockout led to a block of development in the sporoblast stage (ap2-sp2 sporogony phenotype) (189). On the other hand, AP2-G3 has been identified as a third gametocyte-related ApiAP2 factor (190). It is hypothesized, that the AP2-G3 TF, senses environmental signals like homocysteine and activates AP2-G, which in turn initiates the differentiation process by activating gametocyte specific genes (191), (105).

The RNA-Seq data also show 24 other genes related with the sexual development stages of the parasite (table 7). The percentage of frequency of these genes for each sexual stage is, 8% of gametocyte II, 46% of gametocyte V, 20% of ookinete, 17% of oocyst and 9% of sporozoite. All our data point out, that *Pf*LSD1 probably contribute to the epigenetic regulation of gene families related with the sexual development. It could be possible that LSD1 protein activate/repress the expression of various genes such as: TFs, HMT, KDM, HDAC, which participated in the regulation of the sexual development stages. These results clearly call for the testing of the transfectants in mosquito stage.

As a substitute, we investigated the role of LSD1 in *P. berghei* to study its role in early sexual stages as explained above. It was proposed that the *Pb*LSD1 protein have the same demethylation activity. Both genes *pflsd1* and *pblsd1*, seem to have their highest levels of transcription during the sexual differentiation (146).

We performed several *in vitro* analyses of some the sexual stages of *P. berghei* (*PbOokluc-LSD1^{KO}*) parasites (figure 29). In the *in vitro* fecundation assay (figure 29 A), we didn't observe any significant difference between the control (WT) and the *PbOokluc-LSD1^{KO}* parasites. It probably means that until the zygote formation (figure 29 B), passing through the gametocyte development (figure 28B) and activation, there is no critical influence of the *Pb*LSD1 protein. Nevertheless, when we evaluated the

ookinete maturation phenotype (figure 29 C), we observed a significant difference in the maturation state of the *PbOokluc-LSD1^{KO}* parasites. We observe an elevated percentage of *PbOokluc-LSD1^{KO}* parasites that remained in an intermediary ookinete form, the retort parasite form, when we compared to the wildtype parasites. The ookinete is the parasite form which invades the mosquito midgut and is an important stage for genetic recombination. In 2006, Anna Raibaud and colleagues (192), studied the relative transcript levels of several potential ookinete genes in zygote and ookinete *P. berghei* parasites. They found that the majority of the gene transcripts were detected only in the ookinete, indicating a significant change in the transcription repertoire upon zygote to ookinete development (192). Our data also show that *Pb*LSD1 may participate in the control of gene expression. *Pb*LSD1 is probably involved in the PTMs that govern transcriptional control, especially in the sexual stage. This is deduced from the observation that an incomplete maturation of ookinetes occurred, when at least one of the *Pb*LSD proteins was absent. *Pb*LSD1 probably acts as a key post-transcriptional regulator in the genomic reprogramming of ookinete forms. We propose that *Pb*LSD1 protein plays its main role in this stage. It is possible that a more radical phenotype might be reached in the absence of both *Pb*LSD enzymes. However, these *in vitro* results should be confirmed *in vivo,* in mosquitoes, where parasites face complex environmental pressures. In this case, LSD1 may then activate signaling pathways that are translated into cellular responses, which involves changes in the development of the zygote to the ookinete, to become competent to cross the midgut cells (193). Again, further analyses in the mosquito need to be done to evaluate the biologic influence of microenvironment on the capacity to differentiate of *PbOokluc-LSD1^{KO}*. Furthermore, our RNA-Seq analysis exhibited that many downregulated genes (table 7) in the absence of the *Pf*LSD1 that should be involved in oocyst and sporozoite formation. In resume, it is quite possible that LSD1 knockout parasites are transmission-impaired.

6. CONCLUTIONS

6.1 rif genes and transcription mechanism

1.   The transcription of the *rif* gene family, is probably not controlled by a mechanism related to a strict mono-allelic exclusion as is valid for *var* genes.

2.   We found only one B-type *rif* gene activated in our cultures turning it possible that the B subfamily of *rif* genes is differentially regulated from the A *rif* genes subfamily.

3.   It could be possible that only the *5′ upstream* region of *rif* genes is not able to engage allelic exclusion system of *rif* gene transcription.

6.2 Plasmodium Specific Lysine Demethylase 1 (LSD-1).

1. The *P. falciparum* genome possess a LSD1 gene, which encodes for a high molecular weight protein (350 KDa), localized at the nucleus and periphery area.

2. *Pf*LSD1 does not seem to participate in the transcriptional memory/switching of *P. falciparum* variant genes, but appears to act as an epigenetic controller of many genes including sexual stage specific genes.

3. *P. berghei* parasites possess a LSD1 gene which can be deleted. *PbOokluc-LSD1^{KO}* parasites presented a significant percentage of ookinete immature forms (retort forms), hinting to the hypothesis that *Pb*LSD1 has a central role in the parasite's sexual development.

7. REFERENCES

1. Reiter P. From Shakespeare to Defoe: Malaria in England in the Little Ice Age. Emerging Infectious Diseases. 2000.

2. Tan SY, Sung H. Carlos Juan Finlay (1833-1915): of mosquitoes and yellow fever. Singapore Med J. 2008;

3. Cox FE. History of the discovery of the malaria parasites and their vectors. Parasites and Vectors. 2010.

4. Hardenbergh WA. Preventive Medicine in World War II. Vol. Volume II. 1955. 179–232 p.

5. Clarke J. The Effects of Pandemics on History. In: Armies of Pestilence. 2004. p. 102.

6. JP B. Encyclopedia of Pestilence, Pandemics, and Plagues. In: MALARIA AND MODERN MILITARY HISTORY. 2008. p. 383.

7. M H. Malaria: Poverty, Race, and Public Health in the United States. Johns Hopkins University Press, editor. 2001. 256 p.

8. Meremikwu MM, Odigwe CC, Akudo Nwagbara B, Udoh EE. Antipyretic measures for treating fever in malaria. Cochrane Database Syst Rev. 2012;

9. Piotr Kajfasz. Malaria prevention. Int Marit Health. 2009;vol 60(No 1-2):67–70.

10. Mehlhorn H ed. Disease Control, Methods. In: Springer, editor. Encyclopedia of Parasitology. 3rd ed. 2008. p. 362–66.

11. Asante KP. Malaria vaccine : WHO position paper – January 2016 Abstracts of references provided in the position paper and GRADE tables. Lancet. 2016;

12. Kaufman TS, Rúveda EA. The quest for quinine: Those who won the battles and those who won the war. Angewandte Chemie - International Edition. 2005.

13. Kyle RA, Shampe MA. Discoverers of quinine. J Am Med Assoc. 1974;

14. Achan J, Talisuna AO, Erhart A, Yeka A, Tibenderana JK, Baliraine FN, et al. Quinine, an old anti-malarial drug in a modern world: Role in the treatment of malaria. Malaria Journal. 2011.

15. Sinha S, Medhi B, Sehgal R. Challenges of drug-resistant malaria. Parasite. 2014;

16. World Health Organization, WHO WHO. World Malaria Report. World Malaria Report 2011. 2012.

17. Ashley EA, Dhorda M, Fairhurst RM, Amaratunga C, Lim P, Suon S, et al. Spread of Artemisinin Resistance in Plasmodium falciparum Malaria . N Engl J Med. 2014;

18. Wongsrichanalai C, Meshnick SR. Declining artesunate-mefloquine efficacy against falciparum malaria on the Cambodia-Thailand border. Emerging Infectious Diseases. 2008.

19. Nussenzweig RS, Vanderberg J, Most H, Orton C. Protective immunity produced by the injection

of X-irradiated sporozoites of plasmodium berghei. Nature. 1967.

20.   Vanderberg JP. Reflections on early malaria vaccine studies, the first successful human malaria vaccination, and beyond. Vaccine. 2009.

21.   Han L, Hudgens MG, Emch ME, Juliano JJ, Keeler C, Martinson F, et al. RTS,S/AS01 Malaria Vaccine Efficacy is Not Modified by Seasonal Precipitation: Results from a Phase 3 Randomized Controlled Trial in Malawi. Sci Rep. 2017;

22.   World Health Organization. Malaria vaccine pilot launched in Malawi [Internet]. Country first of three in Africa to roll out landmark vaccine. 2019. Available from: https://www.who.int/news-room/detail/23-04-2019-malaria-vaccine-pilot-launched-in-malawi

23.   World Health Organisation. World Malaria Report. World Malaria Report. 2018.

24.   Loy DE, Liu W, Li Y, Learn GH, Plenderleith LJ, Sundararaman SA, et al. Out of Africa: origins and evolution of the human malaria parasites Plasmodium falciparum and Plasmodium vivax. International Journal for Parasitology. 2017.

25.   Soulard V, Bosson-Vanga H, Lorthiois A, Roucher C, Franetich JF, Zanghi G, et al. Plasmodium falciparum full life cycle and Plasmodium ovale liver stages in humanized mice. Nat Commun. 2015;

26.   Bartoloni A, Zammarchi L. CLINICAL ASPECTS OF UNCOMPLICATED AND SEVERE MALARIA. Mediterr J Hematol Infect Dis. 2012;

27.   Hora R, Kapoor P, Thind KK, Mishra PC. Cerebral malaria – clinical manifestations and pathogenesis. Metabolic Brain Disease. 2016.

28.   Warrell DA. Clinical features of malaria. In: Essential Malariology, Fourth Edition. 2017.

29.   World Health Organization. Treatment of Severe Malaria. In: Guidelines for the treatment of malaria. 2015.

30.   D.K. K, V. S, N. S, S.K. K, S.V. K, A. D. Plasmodium vivax malaria. Emerg Infect Dis. 2005;

31.   Magill AJ, Ryan ET, Solomon T, Hill DR. Hunter's Tropical Medicine and Emerging Infectious Disease: Ninth Edition. Hunter's Tropical Medicine and Emerging Infectious Disease: Ninth Edition. 2012.

32.   Aravind L, Iyer LM, Wellems TE, Miller LH. Plasmodium Biology: Genomic Gleanings. Cell. 2003.

33.   GARNHAM PCC. Malaria Parasites and other Haemosporidia. 5, Alfred Street, Oxford: Blackwell Scientific Publications Ltd.; 1966. 1114 p.

34.   Pérez-Toledo K, Rojas-Meza AP, Mancio-Silva L, Hernández-Cuevas NA, Delgadillo DM, Vargas M, et al. Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. Nucleic Acids Res.

2009;

35. Obado SO, Glover L, Deitsch KW. The nuclear envelope and gene organization in parasitic protozoa: Specializations associated with disease. Mol Biochem Parasitol. 2016;

36. Jimenez-Ruiz E, Morlon-Guyot J, Daher W, Meissner M. Vacuolar protein sorting mechanisms in apicomplexan parasites. Mol Biochem Parasitol. 2016;

37. Counihan NA, Kalanon M, Coppel RL, De Koning-Ward TF. Plasmodium rhoptry proteins: Why order is important. Trends in Parasitology. 2013.

38. Kemp LE, Yamamoto M, Soldati-Favre D. Subversion of host cellular functions by the apicomplexan parasites. FEMS Microbiol Rev. 2013;

39. Sheiner L, Vaidya AB, McFadden GI. The metabolic roles of the endosymbiotic organelles of Toxoplasma and Plasmodium spp. Current Opinion in Microbiology. 2013.

40. McFadden GI, Yeh E. The apicoplast: now you see it, now you don't. International Journal for Parasitology. 2017.

41. van Dooren GG, Striepen B. The Algal Past and Parasite Present of the Apicoplast. Annu Rev Microbiol. 2013;

42. Flammersfeld A, Lang C, Flieger A, Pradel G. Phospholipases during membrane dynamics in malaria parasites. International Journal of Medical Microbiology. 2018.

43. Amino R, Giovannini D, Thiberge S, Gueirard P, Boisson B, Dubremetz JF, et al. Host Cell Traversal Is Important for Progression of the Malaria Parasite through the Dermis to the Liver. Cell Host Microbe. 2008;

44. Sturm A, Amino R, Van De Sand C, Regen T, Retzlaff S, Rennenberg A, et al. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. Science (80-). 2006;

45. Nilsson SK, Childs LM, Buckee C, Marti M. Targeting Human Transmission Biology for Malaria Elimination. PLoS Pathogens. 2015.

46. Mueller I, Galinski MR, Baird JK, Carlton JM, Kochar DK, Alonso PL, et al. Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite. The Lancet Infectious Diseases. 2009.

47. Hulden L, Hulden L. Activation of the hypnozoite: A part of Plasmodium vivax life cycle and survival. Malar J. 2011;

48. Warncke JD, Vakonakis I, Beck H-P. Plasmodium Helical Interspersed Subtelomeric (PHIST) Proteins, at the Center of Host Cell Remodeling. Microbiol Mol Biol Rev. 2016;

49. Bourgard C, Albrecht L, Kayano ACA V., Sunnerhagen P, Costa FTM. Plasmodium vivax Biology: Insights Provided by Genomics, Transcriptomics and Proteomics. Front Cell Infect Microbiol. 2018;

50.   Hiller NL, Bhattacharjee S, Van Ooij C, Liolios K, Harrison T, Lopez-Estraño C, et al. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science (80-). 2004;

51.   Rowe JA, Claessens A, Corrigan RA, Arman M. Adhesion of Plasmodium falciparum-infected erythrocytes to human cells: Molecular mechanisms and therapeutic implications. Expert Reviews in Molecular Medicine. 2009.

52.   Kim H, Higgins S, Liles WC, Kain KC. Endothelial activation and dysregulation in malaria: A potential target for novel therapeutics. Curr Opin Hematol. 2011;

53.   Weatherall DJ, Miller LH, Baruch DI, Marsh K, Doumbo OK, Casals-Pascual C, et al. Malaria and the red cell. Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program. 2002.

54.   Hromatka BS, Ngeleza S, Adibi JJ, Niles RK, Tshefu AK, Fisher SJ. Histopathologies, Immunolocalization, and a Glycan Binding Screen Provide Insights into Plasmodium falciparum Interactions with the Human Placenta. Biol Reprod. 2013;

55.   Rask TS, Hansen DA, Theander TG, Pedersen AG, Lavstsen T. Plasmodium falciparum erythrocyte membrane protein 1 diversity in seven genomes - divide and conquer. PLoS Comput Biol. 2010;

56.   Turner L, Lavstsen T, Berger SS, Wang CW, Petersen JEV, Avril M, et al. Severe malaria is associated with parasite binding to endothelial protein C receptor. Nature. 2013;

57.   Aird WC, Mosnier LO, Fairhurst RM. Plasmodium falciparum picks (on) EPCR. Blood. 2014;

58.   Rowe A, Obeiro J, Newbold CI, Marsh K. Plasmodium falciparum rosetting is associated with malaria severity in Kenya. Infect Immun. 1995;

59.   Alexandra Rowe J, Moulds JM, Newbold CI, Miller LH. P. falciparum rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. Nature. 1997;

60.   Helmby H, Cavelier L, Pettersson U, Wahlgren M. Rosetting Plasmodium falciparum-infected erythrocytes express unique strain-specific antigens on their surface. Infect Immun. 1993;

61.   Goel S, Palmkvist M, Moll K, Joannin N, Lara P, R Akhouri R, et al. RIFINs are adhesins implicated in severe Plasmodium falciparum malaria. Nat Med. 2015;

62.   Roberts DJ, Craig AG, Berendt AR, Pinches R, Nash G, Marsh K, et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature. 1992;

63.   Pain A, Ferguson DJP, Kai O, Urban BC, Lowe B, Marsh K, et al. Platelet-mediated clumping of Plasmodium falciparum-infected erythrocytes is a common adhesive phenotype and is associated with severe malaria. Proc Natl Acad Sci. 2002;

64.   Gomes PS, Bhardwaj J, Rivera-Correa J, Freire-De-Lima CG, Morrot A. Immune escape strategies of malaria parasites. Frontiers in Microbiology. 2016.

65.   Belachew EB. Immune Response and Evasion Mechanisms of Plasmodium falciparum Parasites. Journal of Immunology Research. 2018.

66.   Casares S, Richie TL. Immune evasion by malaria parasites: a challenge for vaccine development. Current Opinion in Immunology. 2009.

67.   Dinko B, Pradel G. Immune evasion by <i>Plasmodium falciparum</i> parasites: converting a host protection mechanism for the parasite′s benefit. Adv Infect Dis. 2016;

68.   Baruch DI, Pasloske BL, Singh HB, Bi X, Ma XC, Feldman M, et al. Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. Cell. 1995;

69.   Smith JD, Chitnis CE, Craig AG, Roberts DJ, Hudson-Taylor DE, Peterson DS, et al. Switches in expression of plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell. 1995;

70.   Su X zhuan, Heatwole VM, Wertheimer SP, Guinet F, Herrfeldt JA, Peterson DS, et al. The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of plasmodium falciparum-infected erythrocytes. Cell. 1995;

71.   Scherf A, Hernandez-Rivas R, Buffet P, Bottius E, Benatar C, Pouvelle B, et al. Antigenic variation in malaria: In situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. EMBO J. 1998;

72.   Rubio JP, Thompson JK, Cowman AF. The var genes of Plasmodium falciparum are located in the subtelomeric region of most chromosomes. EMBO J. 1996;

73.   Fischer K, Horrocks P, Preuss M, Wiesner J, Wünsch S, Camargo AA, et al. Expression of var genes located within polymorphic subtelomeric domains of Plasmodium falciparum chromosomes. Mol Cell Biol. 1997;

74.   Thompson JK, Rubio JP, Caruana S, Brockman A, Wickham ME, Cowman AF. The chromosomal organization of the Plasmodium falciparum var gene family is conserved. Mol Biochem Parasitol. 1997;

75.   Voss TS, Thompson JK, Waterkeyn J, Felger I, Weiss N, Cowman AF, et al. Genomic distribution and functional characterisation of two distinct and conserved Plasmodium falciparum var gene 5' flanking sequences. Mol Biochem Parasitol. 2000;

76.   Kyes SA, Kraemer SM, Smith JD. Antigenic Variation in Plasmodium falciparum : Gene Organization and Regulation of the var Multigene Family . Eukaryot Cell. 2007;

77.   Chen Q, Barragan A, Fernandez V, Sundström A, Schlichtherle M, Sahlén A, et al. Identification of Plasmodium falciparum Erythrocyte Membrane Protein 1 (PfEMP1) as the Rosetting Ligand of the Malaria Parasite P. falciparum . J Exp Med. 2002;

78.   Joannin N, Abhiman S, Sonnhammer EL, Wahlgren M. Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. BMC Genomics. 2008;

79.   Petter M, Bonow I, Klinkert MQ. Diverse expression patterns of subgroups of the rif multigene family during Plasmodium falciparum gametocytogenesis. PLoS One. 2008;

80.   Fernandez V. Small, Clonally Variant Antigens Expressed on the Surface of the Plasmodium falciparum-infected Erythrocyte Are Encoded by the rif Gene Family and Are the Target of Human Immune Responses. J Exp Med. 2002;

81.   Tham WH, Payne PD, Brown G V., Rogerson SJ. Identification of basic transcriptional elements required for rif gene transcription. Int J Parasitol. 2007;

82.   Joannin N, Kallberg Y, Wahlgren M, Persson B. RSpred, a set of Hidden Markov Models to detect and classify the RIFIN and STEVOR proteins of Plasmodium falciparum. BMC Genomics. 2011;

83.   Kyes SA, Rowe JA, Kriek N, Newbold CI. Rifins: A second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. Proc Natl Acad Sci. 2002;

84.   Saito F, Hirayasu K, Satoh T, Wang CW, Lusingu J, Arimori T, et al. Immune evasion of Plasmodium falciparum by RIFIN via inhibitory receptors. Nature. 2017;

85.   Wang CW, Magistrado PA, Nielsen MA, Theander TG, Lavstsen T. Preferential transcription of conserved rif genes in two phenotypically distinct Plasmodium falciparum parasite lines. Int J Parasitol. 2009;

86.   Cabral FJ, Wunderlich G. Transcriptional memory and switching in the Plasmodium falciparum rif gene family. Mol Biochem Parasitol. 2009;

87.   Howitt CA, Willnski D, Llinás M, Templeton TJ, Dzlkowski R, Deitsch KW. Clonally variant gene families in Plasmodium falciparum share a common activation factor. Mol Microbiol. 2009;

88.   Witmer K, Schmid CD, Brancucci NMB, Luah YH, Preiser PR, Bozdech Z, et al. Analysis of subtelomeric virulence gene families in Plasmodium falciparum by comparative transcriptional profiling. Mol Microbiol. 2012;

89.   Voss TS, Healer J, Marty AJ, Duffy MF, Thompson JK, Beeson JG, et al. A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. Nature. 2006;

90.   Araujo RBD, Silva TME, Kaiser CS, Leite GF, Alonso Di, Ribolla PEM, et al. Independent regulation of Plasmodium falciparum rif gene promoters. Sci Rep. 2018;

91.   Mu J, Myers RA, Jiang H, Liu S, Ricklefs S, Waisberg M, et al. Plasmodium falciparum genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. Nat Genet. 2010;

92.   Ochola LI, Tetteh KKA, Stewart LB, Riitho V, Marsh K, Conway DJ. Allele frequency-based and polymorphism-versus-divergence indices of balancing selection in a new filtered set of polymorphic genes in Plasmodium falciparum. Mol Biol Evol. 2010;

93. Rovira-Graells N, Gupta AP, Planet E, Crowley VM, Mok S, De Pouplana LR, et al. Transcriptional variation in the malaria parasite Plasmodium falciparum. Genome Res. 2012;

94. Leech JH. Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. J Exp Med. 2004;

95. Lopez-Rubio JJ, Mancio-Silva L, Scherf A. Genome-wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites. Cell Host Microbe. 2009;

96. Scherf A, Rivière L, Lopez-Rubio JJ. SnapShot: var Gene Expression in the Malaria Parasite. Cell. 2008;

97. Dzikowski R, Frank M, Deitsch K. Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. PLoS Pathog. 2006;

98. Scherf A, Lopez-Rubio JJ, Rivière L. Antigenic Variation in Plasmodium falciparum. Annu Rev Microbiol. 2008;

99. Goldberg AD, Allis CD, Bernstein E. Epigenetics: A Landscape Takes Shape. Cell. 2007.

100. Kouzarides T. SnapShot: Histone-Modifying Enzymes. Cell. 2007;

101. Croken MM, Nardelli SC, Kim K. Chromatin modifications, epigenetics, and how protozoan parasites regulate their lives. Trends in Parasitology. 2012.

102. Eisen JA, Sweder KS, Hanawalt PC. Evolution of the SNF2 family of proteins: Subfamilies with distinct sequences and functions. Nucleic Acids Res. 1995;

103. Ji D Der, Arnot DE. A Plasmodium falciparum homologue of the ATPase subunit of a multi-protein complex involved in chromatin remodelling for transcription. Mol Biochem Parasitol. 1997;

104. Warrenfeltz S, Basenko EY, Crouch K, Harb OS, Kissinger JC, Roos DS, et al. EuPathDB: The eukaryotic pathogen genomics database resource. In: Methods in Molecular Biology. 2018.

105. Jeninga M, Quinn J, Petter M. ApiAP2 Transcription Factors in Apicomplexan Parasites. Pathogens. 2019;

106. Rodríguez MA, Gomez C, Esther Ramirez M, Calixto-Galvez M, Medel O. Regulation of gene expression in protozoa parasites. Journal of Biomedicine and Biotechnology. 2010.

107. Hoeijmakers WAM, Stunnenberg HG, Bártfai R. Placing the Plasmodium falciparum epigenome on the map. Trends in Parasitology. 2012.

108. Merrick CJ, Duraisingh MT. Epigenetics in Plasmodium: What do we really know? Eukaryot Cell. 2010;

109. Guizetti J, Barcons-Simon A, Scherf A. Trans-acting GC-rich non-coding RNA at var expression site modulates gene counting in malaria parasite. Nucleic Acids Res. 2016;

110. Cui L, Miao J. Chromatin-Mediated Epigenetic Regulation in the Malaria Parasite Plasmodium falciparum. Eukaryot Cell. 2010;

111. Tonkin CJ, Carret CK, Duraisingh MT, Voss TS, Ralph SA, Hommel M, et al. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. PLoS Biol. 2009;

112. Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. Cell. 2005;

113. Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, Hartl DL, et al. Epigenetic memory at malaria virulence genes. Proc Natl Acad Sci U S A. 2007;

114. Salcedo-Amaya AM, Van Driel MA, Alako BT, Trelle MB, Van Den Elzen AMG, Cohen AM, et al. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proc Natl Acad Sci U S A. 2009;

115. Cui L, Fan Q, Cui L, Miao J. Histone lysine methyltransferases and demethylases in Plasmodium falciparum. Int J Parasitol. 2008;

116. Aravind L, Abhiman S, Iyer LM. Natural history of the eukaryotic chromatin protein methylation system. In: Progress in Molecular Biology and Translational Science. 2011.

117. Wang Y, Jia S. Degrees make all the difference: The multifunctionality of histone H4 lysine 20 methylation. Epigenetics. 2009;

118. Zhang Y, Reinberg D. Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails. Genes and Development. 2001.

119. Hou H, Yu H. Structural insights into histone lysine demethylation. Current Opinion in Structural Biology. 2010.

120. Shi Y, Whetstine JR. Dynamic Regulation of Histone Lysine Methylation by Demethylases. Molecular Cell. 2007.

121. Tian X, Fang J. Current perspectives on histone demethylases. Acta Biochimica et Biophysica Sinica. 2007.

122. Spannhoff A, Hauser AT, Heinke R, Sippl W, Jung M. The emerging therapeutic potential of histone methyltransferase and demethylase inhibitors. ChemMedChem. 2009.

123. Maiques-Diaz A, Somervaille TCP. LSD1: Biologic roles and therapeutic targeting. Epigenomics. 2016.

124. Anand R, Marmorstein R. Structure and mechanism of lysine-specific demethylase enzymes. Journal of Biological Chemistry. 2007.

125. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, et al. Histone demethylation

mediated by the nuclear amine oxidase homolog LSD1. Cell. 2004;

126. Schurter BT, Koh SS, Chen D, Bunick GJ, Harp JM, Hanson BL, et al. Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. Biochemistry. 2001;

127. Yang M, Culhane JC, Szewczuk LM, Jalili P, Ball HL, Machius M, et al. Structural basis for the inhibition of the LSD1 histone demethylase by the antidepressant trans-2-phenylcyclopropylamine. Biochemistry. 2007;

128. Huang J, Sengupta R, Espejo AB, Lee MG, Dorsey JA, Richter M, et al. p53 is regulated by the lysine demethylase LSD1. Nature. 2007;

129. Metzger E, Wissmann M, Yin N, Müller JM, Schneider R, Peters AHFM, et al. LSD1 demethylates repressive histone marks to promote androgen-receptor- dependent transcription. Nature. 2005;

130. Lee MG, Wynder C, Bochar DA, Hakimi M-A, Cooch N, Shiekhattar R. Functional Interplay between Histone Demethylase and Deacetylase Enzymes. Mol Cell Biol. 2006;

131. Foster CT, Dovey OM, Lezina L, Luo JL, Gant TW, Barlev N, et al. Lysine-Specific Demethylase 1 Regulates the Embryonic Transcriptome and CoREST Stability. Mol Cell Biol. 2010;

132. Sun G, Alzayady K, Stewart R, Ye P, Yang S, Li W, et al. Histone Demethylase LSD1 Regulates Neural Stem Cell Proliferation. Mol Cell Biol. 2010;

133. Wang J, Scully K, Zhu X, Cai L, Zhang J, Prefontaine GG, et al. Opposing LSD1 complexes function in developmental gene activation and repression programmes. Nature. 2007;

134. Kim J, Singh AK, Takata Y, Lin K, Shen J, Lu Y, et al. LSD1 is essential for oocyte meiotic progression by regulating CDC25B expression in mice. Nat Commun. 2015;

135. Kerenyi MA, Shao Z, Hsu Y-J, Guo G, Luc S, O'Brien K, et al. Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation. Elife. 2013;

136. Jarriault S, Greenwald I. Suppressors of the egg-laying defective phenotype of sel-12 presenilin mutants implicate the CoREST corepressor complex in LIN-12/Notch signaling in C. elegans. Genes Dev. 2002;

137. Eimer S, Lakowski B, Donhauser R, Baumeister R. Loss of spr-5 bypasses the requirement for the C.elegans presenilin sel-12 by derepressing hop-1. EMBO J. 2002;

138. Katz DJ, Edwards TM, Reinke V, Kelly WG. A C. elegans LSD1 Demethylase Contributes to Germline Immortality by Reprogramming Epigenetic Memory. Cell. 2009;

139. Lan F, Zaratiegui M, Villén J, Vaughn MW, Verdel A, Huarte M, et al. S. pombe LSD1 Homologs Regulate Heterochromatin Propagation and Euchromatic Gene Transcription. Mol Cell. 2007;

140. Rudolph T, Yonezawa M, Lein S, Heidrich K, Kubicek S, Schäfer C, et al. Heterochromatin Formation in Drosophila Is Initiated through Active Removal of H3K4 Methylation by the LSD1 Homolog SU(VAR)3-3. Mol Cell. 2007;

141. Eliazer S, Palacios V, Wang Z, Kollipara RK, Kittler R, Buszczak M. Lsd1 Restricts the Number of Germline Stem Cells by Regulating Multiple Targets in Escort Cells. PLoS Genet. 2014;

142. Mosammaparast N, Kim H, Laurent B, Zhao Y, Lim HJ, Majid MC, et al. The histone demethylase LSD1/KDM1A promotes the DNA damage response. J Cell Biol. 2013;

143. Nam HJ, Boo K, Kim D, Han DH, Choe HK, Kim CR, et al. Phosphorylation of LSD1 by PKCα Is Crucial for Circadian Rhythmicity and Phase Resetting. Mol Cell. 2014;

144. Hino S, Sakamoto A, Nagaoka K, Anan K, Wang Y, Mimasu S, et al. FAD-dependent lysine-specific demethylase-1 regulates cellular energy expenditure. Nat Commun. 2012;

145. Sakamoto A, Hino S, Nagaoka K, Anan K, Takase R, Matsumori H, et al. Lysine demethylase LSD1 coordinates glycolytic and mitochondrial metabolism in hepatocellular carcinoma cells. Cancer Res. 2015;

146. Aurrecoechea C et al. PlasmoDB: a functional genomic database for malaria parasites. [Internet]. Nucleic Acids Res. 2008. Available from: https://plasmodb.org/plasmo/

147. Volz J, Carvalho TG, Ralph SA, Gilson P, Thompson J, Tonkin CJ, et al. Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in Plasmodium falciparum. Int J Parasitol. 2010;

148. Iyer LM, Anantharaman V, Wolf MY, Aravind L. Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. International Journal for Parasitology. 2008.

149. Merrick CJ, Duraisingh MT. Plasmodium falciparum Sir2: an Unusual Sirtuin with Dual Histone Deacetylase and ADP-Ribosyltransferase Activity . Eukaryot Cell. 2007;

150. Miao J, Fan Q, Cui L, Li J, Li J, Cui L. The malaria parasite Plasmodium falciparum histones: Organization, expression, and acetylation. Gene. 2006;

151. Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R, Scherf A. 5′ flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol Microbiol. 2007;

152. Prommana P, Uthaipibull C, Wongsombat C, Kamchonwongpaisan S, Yuthavong Y, Knuepfer E, et al. Inducible Knockdown of Plasmodium Gene Expression Using the glmS Ribozyme. PLoS One. 2013;

153. Bargieri DY, Thiberge S, Tay CL, Carey AF, Rantz A, Hischen F, et al. Plasmodium Merozoite TRAP Family Protein Is Essential for Vacuole Membrane Disruption and Gamete Egress from Erythrocytes. Cell Host Microbe. 2016;

154. Calit J, Dobrescu I, Gaitán XA, Borges MH, Ramos MS, Eastman RT, et al. Screening the Pathogen Box for Molecules Active against Plasmodium Sexual Stages Using a New Nanoluciferase-Based Transgenic Line of P. berghei Identifies Transmission-Blocking Compounds . Antimicrob Agents Chemother. 2018;

155. Sambrook J, Russell DW. Molecular Cloning - Sambrook & Russel - Vol. 1, 2, 3. Cold Spring Harbor Laboratory Press. 2001.

156. Moll K, Ljungström I, Perlmann H, Scherf A, Wahlgren M. Methods in Malaria Research: Fifth Edition. Malaria Research and Reference Reagent Resource Center (MR4). 2008.

157. Lambros C, Vanderberg JP. Synchronization of Plasmodium falciparum erythrocytic stages in culture. J Parasitol. 1979;

158. Jensen JB. Concentration from continuous culture of erythrocytes infected with trophozoites and schizonts of Plasmodium falciparum. Am J Trop Med Hyg. 1978;

159. Pasvol G, Wilson RJM, Brown J. Separation of viable schizont-infected red cells of plasmodium falciparum from human blood. Ann Trop Med Parasitol. 1978;

160. Lelièvre J, Berry A, Benoit-Vical F. An alternative method for Plasmodium culture synchronization. Exp Parasitol. 2005;

161. Gölnitz U, Albrecht L, Wunderlich G. Var transcription profiling of Plasmodium falciparum 3D7: Assignment of cytoadherent phenotypes to dominant transcripts. Malar J. 2008;

162. Hasler T, Albrecht GR, Van Schravendijk MR, Aguiar JC, Morehead KE, Pasloske BL, et al. An improved microassay for Plasmodium falciparum cytoadherence using stable transformants of Chinese hamster ovary cells expressing CD36 or intercellular adhesion molecule-1. Am J Trop Med Hyg. 1993;

163. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods. 2001;

164. Tonkin CJ, Van Dooren GG, Spurck TP, Struck NS, Good RT, Handman E, et al. Localization of organellar proteins in Plasmodium falciparum using a novel set of transfection vectors and a new immunofluorescence fixation method. Mol Biochem Parasitol. 2004;

165. Centro de Facilidades para a Pesquisa (CEFAP). BIOMASS – MASS SPECTROMETRY AND PROTEOME RESEARCH. Available from: http://cefap.icb.usp.br/biomass-easy-nlc-ltq-orbitrap-velos/

166. Janse CJ, Franke-Fayard B, Mair GR, Ramesar J, Thiel C, Engelmann S, et al. High efficiency transfection of Plasmodium berghei facilitates novel selection procedures. Mol Biochem Parasitol. 2006;

167. Blagborough AM, Delves MJ, Ramakrishnan C, Lal K, Butcher G, Sinden RE. Assessing transmission blockade in plasmodium spp. Methods Mol Biol. 2013;

168. Jiang L, Mu J, Zhang Q, Ni T, Srinivasan P, Rayavara K, et al. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. Nature. 2013;

169. NCBI Conserved Domains [Internet]. Available from: https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

170. Pazin MJ, Kadonaga JT. SWI2/SNF2 and related proteins: ATP-driven motors that disrupt protein- DNA interactions? Cell. 1997.

171. Hakimi MA, Deitsch KW. Epigenetics in Apicomplexa: control of gene expression during cell cycle progression, differentiation and antigenic variation. Current Opinion in Microbiology. 2007.

172. Tan J, Pieper K, Piccoli L, Abdi A, Foglierini M, Geiger R, et al. A LAIR1 insertion generates broadly reactive antibodies against malaria variant antigens. Nature. 2016;

173. Ralph SA, Scheidig-Benatar C, Scherf A. Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. Proc Natl Acad Sci U S A. 2005;

174. Poran A, Nötzel C, Aly O, Mencia-Trinchant N, Harris CT, Guzman ML, et al. Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. Nature. 2017;

175. Binda C, Valente S, Romanenghi M, Pilotto S, Cirilli R, Karytinos A, et al. Biochemical, structural, and biological evaluation of tranylcypromine derivatives as inhibitors of histone demethylases LSD1 and LSD2. J Am Chem Soc. 2010;

176. Fang R, Chen F, Dong Z, Hu D, Barbera AJ, Clark EA, et al. LSD2/KDM1B and Its Cofactor NPAC/GLYR1 Endow a Structural and Molecular Model for Regulation of H3K4 Demethylation. Mol Cell. 2013;

177. Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y. Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell. 2005;

178. Eberharter A, Vetter I, Ferreira R, Becker PB. ACF1 improves the effectiveness of nucleosome mobilization by ISWI through PHD-histone contacts. EMBO J. 2004;

179. Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N, Morillon A, Weise C, et al. Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. Mol Cell. 2003;

180. Suthram S, Sittler T, Ideker T. The Plasmodium protein network diverges from those of other eukaryotes. Nature. 2005;

181. Hall JA, Georgel PT. MINIREVIEW / MINISYNTHE CHD proteins : a diverse family with strong ties 1. Biochem Cell Biol. 2007;

182. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell. 2004;

183. Goyal M, Banerjee C, Nag S, Bandyopadhyay U. The Alba protein family: Structure and function. Biochimica et Biophysica Acta - Proteins and Proteomics. 2016.

184. Bunnik EM, Le Roch KG. PfAlba1: Master regulator of translation in the malaria parasite. Genome Biol. 2015;

185. Vembar SS, Macpherson CR, Sismeiro O, Coppée JY, Scherf A. The PfAlba1 RNA-binding protein is an important regulator of translational timing in Plasmodium falciparum blood stages. Genome Biol. 2015;

186. Volz JC, Bártfai R, Petter M, Langer C, Josling GA, Tsuboi T, et al. PfSET10, a Plasmodium falciparum Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division. Cell Host Microbe. 2012;

187. LaCount DJ, Vignali M, Chettier R, Phansalkar A, Bell R, Hesselberth JR, Schoenfeld LW, Ota I, Sahasrabudhe S, Kurschner C, Fields S HR. A protein interaction network of the malaria parasite Plasmodium falciparum. Nature. 2005;438:103–7.

188. Kafsack BFC, Rovira-Graells N, Clark TG, Bancells C, Crowley VM, Campino SG, et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. Nature. 2014;

189. Modrzynska K, Pfander C, Chappell L, Yu L, Suarez C, Dundas K, et al. A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle. Cell Host Microbe. 2017;

190. Zhang C, Li Z, Cui H, Jiang Y, Yang Z, Wang X, et al. Systematic CRISPR-Cas9-mediated modifications of plasmodium yoelii ApiAP2 genes reveal functional insights into parasite development. MBio. 2017;

191. Dharia N V., Bright AT, Westenberger SJ, Barnes SW, Batalov S, Kuhen K, et al. Whole-genome sequencing and microarray analysis of ex vivo Plasmodium vivax reveal selective pressure on putative drug resistance genes. Proc Natl Acad Sci U S A. 2010;

192. Raibaud A, Brahimi K, Roth CW, Brey PT, Faust DM. Differential gene expression in the ookinete stage of the malaria parasite Plasmodium berghei. Mol Biochem Parasitol. 2006;

193. Kappe SHI, Gardner MJ, Brown SM, Ross J, Matuschewski K, Ribeiro JM, et al. Exploring the transcriptome of the malaria sporozoite stage. Proc Natl Acad Sci U S A. 2001;

8. APPENDICES

8.1. APPENDIX 1: Independent regulation of *Plasmodium falciparum rif* gene promoters

SCIENTIFIC REP🞮RTS

Independent regulation of *Plasmodium falciparum rif* gene promoters

Rosana Beatriz Duque Araujo[1], Tatiane Macedo Silva[1], Charlotte Sophie Kaiser[2], Gabriela Fernandes Leite[1], Diego Alonso[3], Paulo Eduardo Martins Ribolla[3] & Gerhard Wunderlich[1]

All *Plasmodium* species express variant antigens which may mediate immune escape in the vertebrate host. In *Plasmodium falciparum*, the *rif* gene family encodes variant antigens which are partly exposed on the infected red blood cell surface and may function as virulence factors. Not all *rif* genes are expressed at the same time and it is unclear what controls *rif* gene expression. In this work, we addressed global *rif* transcription using plasmid vectors with two drug resistance markers, one controlled by a *rif* 5′ upstream region and the second by a constitutively active promoter. After spontaneous integration into the genome of one construct, we observed that the resistance marker controlled by the *rif* 5′ upstream region was expressed dependent on the applied drug pressure. Then, the global transcription of *rif* genes in these transfectants was compared in the presence or absence of drugs. The relative transcript quantities of all *rif* loci did not change profoundly between strains grown with or without drug. We conclude that either there is no crosstalk between *rif* loci or that the elusive system of allelic exclusion of *rif* gene transcription is not controlled by their 5′ upstream region alone.

The infection with one of the five *Plasmodium* species which cause malaria in humans still is a challenge for the public health predominantly in underdeveloped countries. *Plasmodium falciparum* alone is still responsible for 445000 deaths per year, mostly children under five years or pregnant women[1]. Part of the virulence exerted by *Plasmodium* is caused by the presence of variant antigens expressed on the surface of the host's infected red blood cells. Members of the best-characterized family of variant antigens, the *P. falciparum* erythrocyte membrane protein 1 (PfEMP1), play a central role in immune evasion. PfEMP1 are encoded by approximately 50–60 different *var* genes[2] which are highly recombinogenic[3,4], possibly through specific three-dimensional DNA conformations near breakpoints[5], and possess a modular structure[6]. In order to successfully evade immune mechanisms exerted by the human host, *var* gene transcription is tightly controlled in a way that normally only one or two *var* genes are expressed. A number of factors are involved in this control and these include not only sequences in the 5′ upstream regions of *var* genes[7,8], including untranslated ORFs[9,10], the pairing of *var* promoters and *var* introns[11], but also specific DNA/chromatin binding factors[12] and the involvement of several chromatin modifiers (reviewed in[13]). Recently, the participation of non-coding GC-rich RNAs in *var* transcription control was shown[14]. The current model of *var* transcription regulation also suggests a specific subnuclear site in which *var* transcription occurs and to which *var* loci translocate in order to be transcribed. However, the exact factors and dynamics which orchestrate this process and license one *var* locus for transcription while excluding all other *var* loci are largely unknown. Still more elusive is what determines that a *var* locus and its associated histone modifications switch from an active to a silent state or vice versa.

Another major variant gene family which is not only found in human or primate *Plasmodium* species but also in murine species is the *pir* (*Plasmodium* interspersed repeat) gene family[15], and the biological function of gene products from this family is not well understood. Recent results indicate that their encoded proteins may function at different points of the parasite-host interface[16]. If results from *Plasmodium chabaudi* can be extrapolated to all *Plasmodium* species, *pir* transcription seems to be reset during mosquito passage and reinfection[17] and earlier evidence pointed to transcriptional diversity[18] and quick switching in several models[19,20]. A recent study implied a specific PIR protein of *P. falciparum* (termed RIFINs, repetitive interspersed family[21]) as a factor involved in the

[1]Department of Parasitology, Institute of Biomedical Sciences, University of São Paulo, Av. Prof. Lineu Prestes, 1374, São Paulo - SP, 05508000, Brazil. [2]Institute of Animal Physiology, Schloßplatz 8, Westfälische Wilhelms Universität, Münster, Germany. [3]Department of Parasitology, IBB/IBTEC, State University of São Paulo, Botucatu, São Paulo, Brazil. Correspondence and requests for materials should be addressed to G.W. (email: gwunder@usp.br)
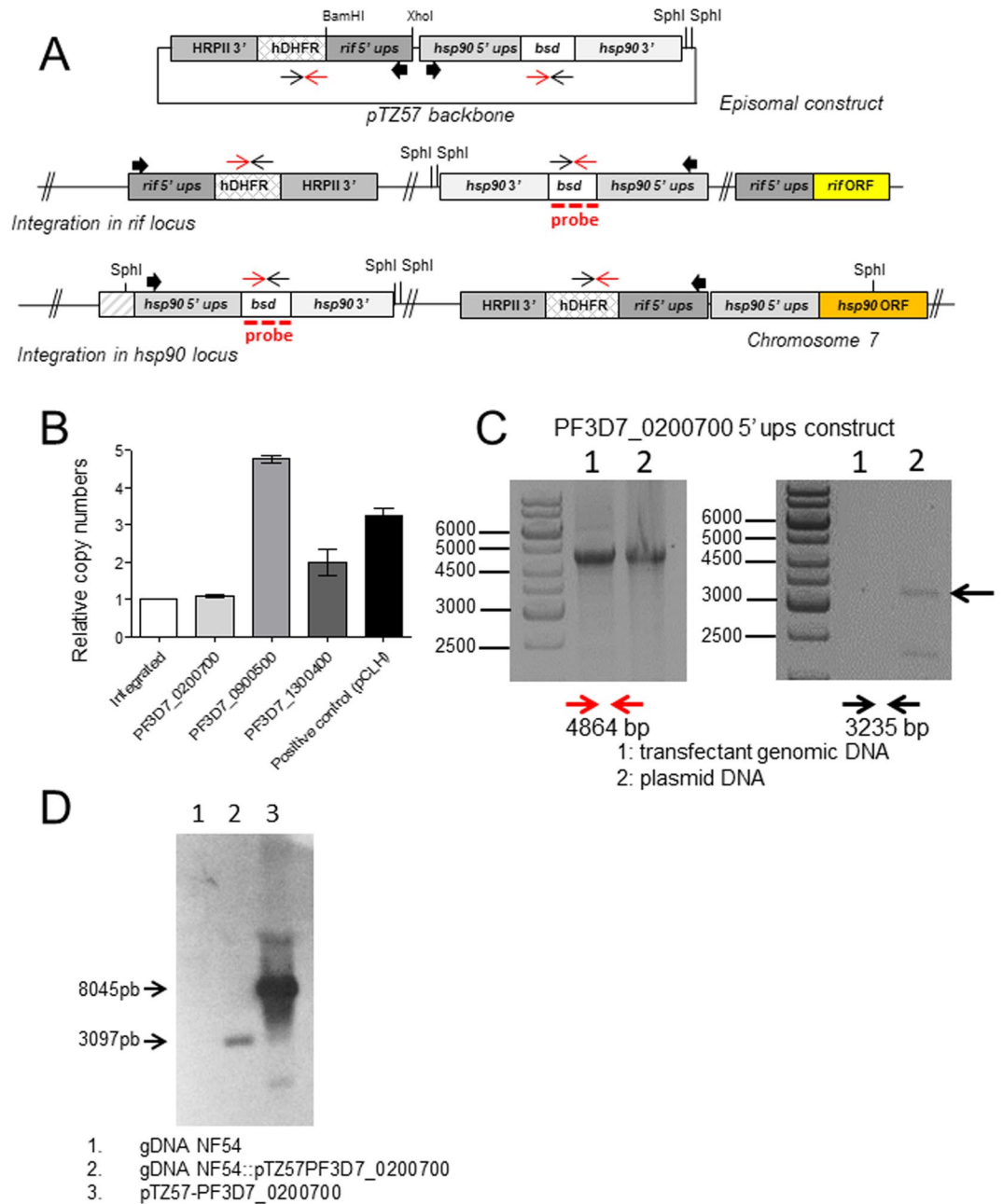
Figure 1. Only the PF3D7_0200700 construct shows single integration in the hsp90 locus. In (**A**) the outline of the transfection plasmid constructs is shown. *rif* 5′ ups regions were interchanged in this plasmid by ligation of XhoI/BamHI fragments representing the *rif* 5′ ups sequence. See Material and Methods for oligonucleotide sequences used for amplification of *rif* 5′ ups. Due to identity with *P. falciparum* genomic sequences, Integration may occur in 4 possible sites: hsp90 5′ ups and 3′ downstream regions, in the hrp3 3′ downstream region and in *rif* 5′ ups regions. Due to the size of identical stretches, we focused on the possible integration events in hsp90 5′ ups and *rif* 5′ ups. Bold arrows indicate the expected direction of transcription in the 5′-ups regions. (**B**) Copy number comparison of different transfected parasite lines with gDNA from the integrated NF54::RESA-GFP strain[54], an episomal pCLH NF54 strain, and the parasite lines after transfection with plasmids containing different 5′ *rif* ups. The primer performance of hDHFR primers and for the t-seryl RNA synthetase were tested and judged identical (less than 0.5 Ct difference on the same substrate in qPCR using gDNA from the NF54::RESA-GFP strain). Then, copy numbers of the hDHFR locus in relation to the genomic t-seryl RNA synthetase were calculated. In (**C**) results from PCRs using long-range polymerases provide indirect evidence that the PF3D7_0200700 5′ *rif* ups was integrated in the genome. On the left, PCR results showing amplification with forward *hDHFR* and forward *bsd* primers (in red) over the backbone of the transfection plasmid (positive for the untransfected plasmid (2) as well as the gDNA from the PF3D7_0200700 5′ *rif* ups transfected strain). On the right, amplification products from PCRs using reverse *hDHFR* and reverse *bsd* oligos. This amplification is only possible when the plasmid is in the episomal form (see scheme for episomal construct and integrated locus in **A**). No amplification product is seen for the gDNA from the strain with the PF3D7_0200700 5′ *rif* ups. See Supplementary Figure 2 for results with the other *rif* ups constructs. In (**D**), digestion with SpHI of NF54

genomic DNA, transfectant line genomic DNA and transfected plasmid DNA and subsequent Southern blot analysis with a probe consisting of a digoxigenin-labeled bsd fragment. Note that only in the case of integration in the hsp90 locus a 3097 bp fragment is formed, while linearized or concatemerized plasmids will result in 8045 bp fragments.

pathogenic process of erythrocyte rosetting[22]. Another study revealed that a specific motif in RIFINs promoted binding to leucocyte immunoglobulin-like receptor B1 (LILRB1) or leucocyte-associated immunoglobulin-like receptor 1, thereby inhibiting activation of B-cells and natural killer cells which express the LILRB1 receptor[23]. This turns evident that at least some RIFINs can be understood as virulence factors. Importantly, VIR proteins of *Plasmodium vivax* also appear to mediate cytoadherence and participate in pathogenic processes and probably immune evasion[24,25]. In *P. falciparum*, the vast majority of *rif* genes are localized adjacent to *var* genes, often in a tandem organization. In version 36 of PlasmoDB, there are 221 genes in the 3D7 strain genome which encode PIR proteins. Of these, 158 are full-length RIFINs, 27 are truncated or defect RIFINs, and the remaining are STEVOR (**s**ub**te**lomeric **v**ariant **o**pen **r**eading frame) or truncated or defect STEVOR. RIFINs can be categorized into two major groups of *rif* genes and RIFINs: A and B. These groups differ by a short conserved 25mer peptide sequence in the first half of the protein which is present only in the 97 A-group RIFINs in the *P. falciparum* strain 3D7 genome[26]. While the A-group RIFINs seem to be exported to the infected red blood cell (IRBC) surface, B-group RIFINs are believed to remain associated with the vesicular network in the IRBC (Maurer's clefts). It is still unclear what controls *rif* transcription, a recent study pointed to the transcription factor AP2-SP[27] which somehow seems to influence transcription in blood stage *P. falciparum*[28]. While earlier studies indicated that *rif* and *var* gene transcription may be controlled by similar factors[29], no clear-cut allelic exclusion mechanism could be detected in other studies which specifically addressed *rif* transcription or switching[20,30]. A drawback in the study from Howitt and colleagues[29] was that the tested *rif* promoter controlled the unique drug resistance marker blasticidin deaminase in their construct, turning a considerable baseline activity of the promoter essential in order to obtain transfectant parasite lines. This did not permit a genuinely "switched-off" state of the *rif* 5′ ups-controlled transgene. In another study, the activity profile of a 5′-*rif* upstream region appeared more related to *var* genes[31] and no significantly different regulation could be discerned upon activation. Also, no "crosstalk" – understood as the influence of the activity of one promoter on the activity of remaining promoters - between 5′ upstream regions such as occurs in *var* gene regulation was detectable. In contrast, in the study by Goel and colleagues, phenotypic selection procedures pointed to the expression of a single *rif* gene in parasites with PfEMP1-independent rosetting[22], supporting the view that allelic exclusion and crosstalk may occur. In order to settle the question if there are allelic exclusion and crosstalk between *rif* 5′ ups regions, we used three different *rif* 5′ upstream regions in bicistronic transfection plasmids. Previously, this approach was successfully applied on *var* promoters[7,32]. After transfection, we divided transfected parasite lines and in one culture we selected the growth of parasites which actively transcribed a drug resistance marker (human dihydrofolate reductase). Afterwards, we compared *rif* transcripts between pyrimethamine-derivate sensitive and resistant lines by RNAseq.

Results

The expression mode of *rif* genes and the proteins they encode, RIFINs, is unclear. In order to test if *rif* genes are expressed in a similarly coordinated way as *var* genes[33], we created bicistronic plasmids similar to those used by Voss[7] or Witmer[31] and colleagues. In these, one resistance marker gene was controlled by the constitutive plasmodial heat shock protein 86/90 promoter and the other by potentially inducible *rif* 5′ upstream sequences. We chose three *rif* upstream sequences (ups) based on previous observations using wild-type parasites of the 3D7 lineage where *rif* transcripts from these 5′ upstream regions were detected[20]. This assured that functional promoters were being used. The size (~1500 nt) of the inserted putative promoter sequence was chosen observing the distance to adjacent ORFs (Supplemental Fig. 1). In Fig. 1, an outline of the plasmid constructs is shown. Plasmids were transfected and blasticidin-resistant parasite lines were readily established. It is possible that transcription of episomal 5′ upstream sequences is different from genomic loci, due to an increase in the number of circulating episomes (for example[31]). When testing the copy number of the hDHFR gene, we observed that the transfectant line containing the PF3D7_0200700 5′ *rif* ups showed only one copy. In contrast, the hDHFR gene in a freshly transfected parasite line containing the *Photinus* luciferase encoding plasmid pCLH (a derivate of pDC10[34] where the chloramphenicol acetyltransferase was changed for a luciferase coding sequence from pGL2 (Promega)) appeared in approximately three copies per cell (selection using 2.5 nM WR99210, Fig. 1). Also, the transfectant lines containing the PF3D7_0900500 or the PF3D7_1300400 5′ *rif* ups showed five or two hDHFR copies per parasite genome, respectively (Fig. 1). The appearance of only one copy hints to the integration of the bicistronic plasmid. To verify this, we used two PCR amplifications to specifically detect integrated or episomal forms of plasmids in all transfectant lines containing *rif* 5′ ups. As shown in Fig. 1, no amplicon was detected using an oligo pair which amplifies over the putative breakpoint upon single crossover integration of the construct carrying PF3D7_0200700 5′ *rif* ups. On the other hand, an amplicon representing the plasmid backbone both in integrated or episomal forms was detected in both transfectants genomic DNA and purified plasmid controls. The transfectants with the PF3D7_0900500 or the PF3D7_1300400 5′ *rif* ups showed each several amplicons, one of which consistent with episomal forms of the plasmids (Supplementary Fig. 2). Since the transfection plasmid backbone contains four sequence regions which may provoke single crossover recombination into the genome (the hrp3 terminator, the hsp86/90 5′ ups and the hsp86/90 3′ region plus the *rif* 5′ ups region), we conducted a Southern blot analysis using a digoxigenin-labeled bsd-gene fragment as a probe. The observed pattern indicated that recombination occurred at the hsp86/90 locus and not at the PF3D7_0200700 *rif* 5′ locus (Fig. 1).
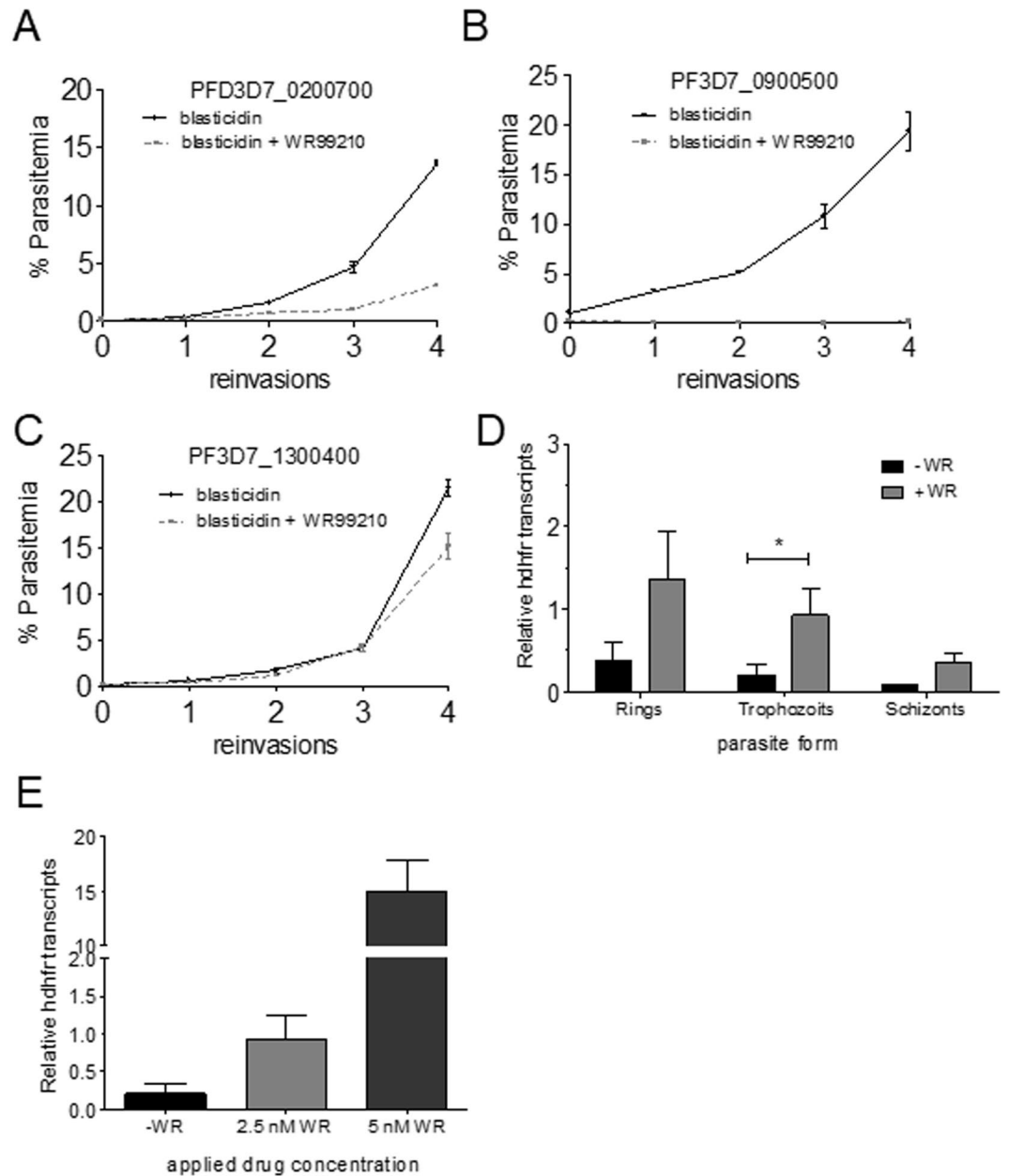
Figure 2. Differential growth of strains containing different 5′ rif ups controlling hDHFR mediated resistance. In (**A**–**C**), transfected NF54 parasite strains with plasmids containing the indicated 5′ *rif* ups controlling hDFHR transcription, established using blasticidin at 2.5 μM as described in methods, were synchronized with plasmagel flotation and sorbitol treatment and submitted to additional WR99210 treatment (2.5 nM) or not. Parasitemias were observed every 48 hours briefly after reinvasion. In (**D**) the strain with the PF3D7_0200700 5′ *rif* ups construct was analyzed for hDHFR transcript abundance dependent on the blood stage form. This experiment was done in three biological replicates and data are shown. In (**E**) the same trophozoite stage parasites as in (**D**) were submitted to higher concentrations of WR99210 and the hDHFR transcript was measured by RT-qPCR. Error bars in all graphs show standard deviation. Statistical differences between relative transcript quantity values were calculated using the two-way ANOVA test (*p < 0.05).

We then tested if transfectant parasites were able to survive in the presence of the second drug WR99210. A prerequisite for survival is the sufficient transcription from the cloned 5′ *rif* ups. To test this, the cultures of the three parasite lines were split and one half of each culture was cultivated in the presence of 2.5 μg/ml blasticidin and 2.5 nM WR99210, while the other half was solely grown in the presence of 2.5 μg/ml blasticidin. The parasite line with PF3D7_0200700 5′ *rif* ups grew slowly for the first few cycles and then proliferated normally, in accordance with a selection of parasites which showed upregulation of the hDHFR controlling PF3D7_0200700 5′ *rif* ups (Fig. 2). In contrast to this, the parasite line containing the PF3D7_0900500 5′ *rif* ups was unable to grow in the presence of WR99210 meaning that no parasites were present which had activated this locus. Three experiments over ten growth cycles were tried with this parasite line and parasites never became resistant to WR99210.
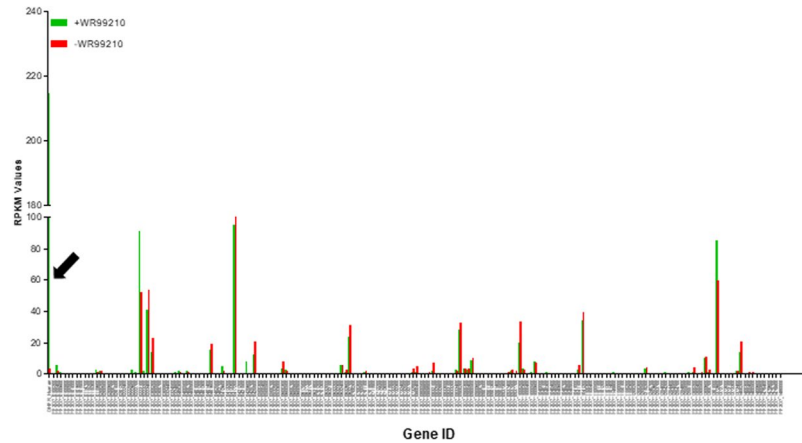
Figure 3. RNAseq shows no substantial differences between *rif* transcription profiles upon presence or absence of activity from a specific 5′ *rif* ups (PF3D7_0200700). A: RNAseq was conducted as described with RNA from parasites grown for four reinvasions in the presence of 2.5 nM WR99210 or in the absence of this drug. In red, the RPKM values for all *rif* loci with the presence of transcripts in parasites grown without WR99210 and in green in the presence of 2.5 nM WR99210. The arrow depicts the strongly different RPKM values for the hDHFR transcript in these cultures. Note that the difference in relative hDHFR transcript quantity is higher than measured by RT-qPCR (see Fig. 2D,E). For individual RPKM data see Supplementary Table 1.

This means that the cloned PF3D7_0900500 5′ *rif* ups is either not sufficiently functional or completely silenced. The third construct containing episomes with the PF3D7_1300400 5′ *rif* ups controlling hDHFR readily grew in the presence of WR99210 and no striking difference could be discerned between blasticidin/WR99210 and blasticidin treated parasites (Fig. 2). Taken together, from the three constructs, only the integrated construct with PF3D7_0200700 5′ *rif* ups showed a dynamic that was expected for a transcriptionally variable member of the multigene family which is functional in blood stage parasites.

We then quantified the relative transcript quantity in the lineage with PF3D7_0200700 5′ *rif* ups in parasites "on" and "off" WR99210 drug pressure during the intraerythrocytic cycle. As shown in Fig. 2D and E, the relative transcript quantity difference between parasites under WR drug treatment was higher by a factor ranging from 1:4 to 1:40 using reverse transcription-qPCR in three independent experiments. The highest and significant differences were observed in trophozoite stage parasites. In order to monitor if the *rif* promotor activity could be further increased, we analyzed the steady-state transcript quantities in this parasite lineage grown under 5 nM WR99210 instead of 2.5 nM. As shown in Fig. 2E, the relative transcript quantities strongly increased under these conditions, indicating that the promoter activity may be modulated in a wide range. Importantly, the cultures that were grown in 5 nM WR99210 also did not increase the relative copy number of the artificial hDHFR, reinforcing the view that the 5′ *rif* ups PF3D7_0200700 had integrated into the genome (Supplementary Fig. 3).

The transcriptional activity of 5′ ups regions of *var* genes is strictly regulated resulting normally in the expression of one unique gene[33] and silencing of all other loci, meaning that *var* loci are in crosstalk. Accordingly, it was observed that genomic *var* loci can be silenced by artificial activation of an episomal *var* 5′ upstream region[7,32]. To detect if the artificially activated *rif* locus had any influence on *rif* transcripts from other genomic loci, RNAseq was performed using a paired sample from the PF3D7_0200700 5′ *rif* ups-construct containing trophozoites which were submitted or not to WR99210 treatment. As shown in Fig. 3, parasites of this transgenic lineage showed a number of transcripts from genomic *rif* loci that appeared in higher RPKM numbers (>20 RPKM). In parasites that were not grown in the presence of WR99210, few hDHFR transcripts from the modified PF3D7_0200700 locus were detected (RPKM ~3), corroborating previous qPCR results. In contrast, higher RPKM values for hDHFR as the most abundantly detected *rif* transcripts were observed in parasites grown in the presence of WR99210. This indicates that the hDHFR-controlling *rif* promoter can be either active or inactive, pre-requisites of a variant gene promoter. Also, the RPKM values of the modified PF3D7_0200700 were in the same range as other simultaneously active *rif* promoters (e.g., PF3D7_1372600). When the relative *rif* transcript quantities from other loci were compared between parasites grown under WR99210 pressure or not, no profound changes were found. This can be interpreted that the activity of the modified PF3D7_0200700 *rif* promoter did not influence transcription from other loci. Such a result is in contrast with similar experiments using *var* promoters where the activity of a *var* 5′ ups led to a substantial silencing of the remaining genomic *var* loci[7]. When analyzing the RPKM values of all transcripts (Supplementary Table 1), a number of genes appeared differentially expressed between the two samples, and the observation that invasion-related genes were detected with higher RPKM values indicates that the sample cultivated without WR99210 was slightly advanced in the erythrocytic cycle (Supplementary Fig. 4).

It is possible that subgroups of *rif* upstream regions are regulated independently, meaning that RIFIN A and/or RIFIN B are not subject to allelic exclusion. When focusing only on *rif* transcripts which were detected in larger quantities in RNAseq (cutoff RPKMs > 20), we observed that only one *rif* B type transcript was detected in slightly elevated levels in both treated or untreated parasites, while several *rif* A-type transcripts were present (Fig. 4). This

5′ ups	ID	% identity	Matches	Score	RPKM	
PF3D7_0200700	PF3D7_0223100	99%	1497/1514	2936	91.36	51.95
	PF3D7_1373000	98%	1497/1520	2926	0.00	0.00
	PF3D7_0732200	99%	1491/1512	2912	0.35	0.00
	PF3D7_0101600	98%	1491/1515	2900	0.35	0.00
PF3D7_1300400	PF3D7_0600500	99%	1506/1511	2702	2.40	1.76
	PF3D7_0937500	99%	1506/1511	2702	0.00	0.00
	PF3D7_1150300	99%	1506/1511	2702	0.00	0.00
	PF3D7_0425700	99%	1502/1514	2960	0.00	0.00
	PF3D7_0100400	98%	1502/1528	2924	5.50	1.76

Table 1. Genomic *rif* 5′-ups identical to cloned *rif* 5′-ups (≥98% identity) are all but one silenced in RNAseq experiments. In the column "RPKM", the values on the left refer to RPKM values of the WR99210 selected parasites, on the right, RPKM values of the WR99210-untreated samples are shown (without selection for hDHFR expression). Given RPKM values refer to transcripts from the *rif* genes preceded by the almost identical *rif* 5′ ups informed in the "ID" column.
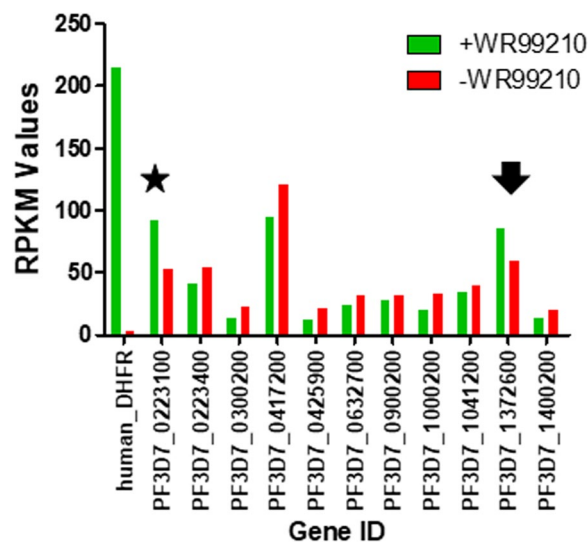

Figure 4. Rif loci with higher (any value above >20) RPKM values were mostly from loci with A-type rif genes. Results from Fig. 3 were filtered for higher RPKM values and are shown. Note that the locus with an almost identical 5′ *rif* ups (asterisk) on the same chromosome as the transfected 5′ *rif* ups controlling hDHFR (bars on the left) is also transcriptionally active independently of activation of the transgene 5′ *rif* ups PF3D7_0200700. The arrow indicates the only significantly transcribed type B *rif*.

reinforces that "A"-grouped *rif* 5′ ups, encoding antigens which are potentially associated with the infected red blood cell surface, are not influenced by an artificially activated "A" type *rif* 5′-ups.

Discussion

The expression control of variant gene families is an intriguing phenomenon in many human parasites ranging from *Trypanosoma brucei* (reviewed in[35]) and *Giardia lamblia*[36] to *Plasmodium*. Each parasite seems to have developed different molecular mechanisms to ensure that antigenic repertoires are not prematurely exhausted. In *Giardia* trophozoites, an RNAi-based mechanism enables the translation of a single variant surface antigen per parasite[37]. In *Trypanosoma brucei*, selective variant surface antigen (VSG) production is controlled at the transcriptional level and also through genetic recombination. Single *vsg* genes are transcribed by RNA Polymerase 1 at a specific subnuclear expression site[38] and the activation of *vsg* expression sites is associated with specific chromatin modifications (reviewed in[39]). *Var* gene expression in *P. falciparum* is tightly controlled at the transcriptional level in ring-stage parasites[40] and all but one or two sites remain silenced. Silencing of sites is associated with a number of specific chromatin modifications, mainly at histone H3 which - when trimethylated at lysine 9 - recruits heterochromatin protein 1[12], initially perceived as a landmark of silenced chromatin in different cell types[41]. A still not answered question is why not all *var* sites are silenced and the reason for this possibly lies in the concentration of ncRNAs transcribed from a number of GC rich regions present in the nucleus which somehow seem to coordinate gene counting[14]. Overproduction of members of these ncRNAs lead to the simultaneous transcription of *var* genes – an effect which is also observed when histone deacetylases SIR2A and SIR2B are repressed

or deleted[42]. Here, we tried to interfere with the transcription of members of the *rif* gene family using an approach that was successfully applied to monitor *var* gene transcription. Interestingly, from three different constructs, only one presented a variable transcription activity, which may be considered as a pre-requisite of the corresponding RIFIN to function in antigenic variation. For the other two constructs, which also did not show integration at genomic sites, we found either a permanently activated or silenced status. Notably, all tested 5′-upstream regions have several, almost completely identical regions in the genome but only one of the *rif* genes controlled by these identical regions seemed transcriptionally active to the same degree as the transgene (not considering the 5′-ups *rif* construct with PF3D7_0900500 which was always silenced (Table 1)). Interestingly, the only identical locus with considerable amounts of transcripts from the genomic locus appeared for the differentially controlled PF3D7_0200700 *rif* 5′-ups construct. Intriguingly, the original *rif* locus PF3D7_0200700 from which the 5′-ups region was cloned had no detectable transcripts at all (Supplementary Table 1), while the *rif* locus lying on the opposite end of chromosome 2 showed transcripts in the RPKM value range of 91 and 51 depending on the presence or absence of transcripts from the artificially integrated PF3D7_0200700 *rif* locus (Table 1).

This may be interpreted that absolute sequence may not be associated with promoter activity or silencing. It seems that epigenetic marking of chromatin surrounding *rif* promoters plus factors associating to them have a decisive role in 5′-ups activation and silencing, similar to what is found for *var* genes. How this regulation is achieved at the molecular level is still elusive. In the study of Howitt and colleagues, a still unknown - but limited in number- factor which provides transcriptional activation of *rif* and *var* transcripts was suggested[29]. In the present RNAseq analysis only one *var* locus showed significantly altered transcript levels (data not shown). This *var* locus – PF3D7_0400200 – is annotated as a pseudogene and consists only of exon 2. PlasmoDB predicts also the presence of the *var* intron, which itself is also a promoter. It is unclear if RNA from this site is involved in *rif* transcription regulation. Our data do not support co-regulation of *rif* and *var* genes. Of note, the material analyzed in our study may also not permit any conclusion about *rif-var* co-regulation, since RNAs from middle/late trophozoite stage parasites were employed. Normally, *var* transcription has ceased in trophozoites 20 h post reinvasion[43]. It must also be reinforced that the shown RNAseq results are unique for this specific experiment. It may be expected that the *rif* genes which appear transcribed in the current samples change over time due to switching and that in another experiment different, dominant *rif* transcripts are detected. Additionally, it is unclear if the observed *rif* transcripts occurred in the same parasites or if they originated from different parasites, concomitantly with the activated PF3D7_0200700 locus in all parasites. Also, it appears that the untreated sample was slightly advanced in the erythrocytic cycle since a number of schizont related genes such as MSPs showed more transcripts in the sample not treated with WR99210 compared to the treated sample (Supplementary Table 1, spreadsheet "2", and "3", Supplementary Fig. 4). Notwithstanding, the main result is still valid and other *rif* gene transcripts, besides the artificially activated PF3D7_0200700 locus, are not influenced.

Based on our data, it may be postulated the chromosomal context surrounding *rif* 5′ups-regions may play a pivotal role in *rif* expression. All of the three 5′-ups regions were designed in the same way and contained almost 1500 nt upstream of the corresponding *rif* ATG, and all three upstream regions contained either terminator regions or head-to-head promoter regions of the adjacent gene locus possibly including heterochromatin boundary regions (Supplementary Fig. 1). Nonetheless, only two 5′-ups were functional in providing a sufficient number of transcripts and only the apparently integrated allele showed a tunable behavior. In the study by Witmer and colleagues[31], the 5′-ups region of *rif* PF3D7_1300400 was tested. This region is somehow special in that it lies in a head-to-head position to a *var* upsA promoter region. In their study, in one construct the *rif* promoter portion could be activated while the adjacent *var* ups was completely silenced. In a second, slightly modified construct, both promoters could be selected for active transcription. The degree of induction of the *rif* 5′-ups in their study was not significantly different, indicating that transcription was possibly partly induced even without drug selection for transcription from this promoter. In our study, we used 1465 nt of the same 5′ *rif* ups and also no differential regulation was observed (construct 3 with PF3D7_1300400 5′ *rif* ups). Given the ambiguous results for both *rif* 5′ ups PF3D7_0900500 and PF3D7_1300400 which are either not integrated or may have been rearranged during the transfection procedure, no further conclusions can be drawn why constructs containing these 5′ ups led to the observed results of permanent silencing or activity, respectively.

Here, the copy number of plasmids encountered in parasite lines bearing most probably episomal forms of *rif* promoted hDHFR was partially lower than in the related study from Howitt[29] or Witmer[31] and colleagues. This may be due to the fact that we used low concentrations of blasticidin for selection. Howitt and colleagues used 2 or 10 µg/ml blasticidin and copy numbers at 2 µg/ml were in the same range or slightly higher than in our episomal constructs. Witmer and colleagues also used 2.5 µg/ml blasticidin but their plasmid backbones contained the TARE/rep20 element, known to improve plasmid segregation[44]. When we compared results of the hDHFR marker transcript in the transgenic PF3D7_0200700 parasite line growing or not in the presence of WR99210, we found smaller discrepancies between the "on" and "off" in qPCR than in RNAseq. The reason for this may lie in the differing range of linearity of both techniques, although this was not specifically tested for.

As a central result, we showed that transcription of the *rif* gene family, specifically of the RIFIN A subset, is most probably not controlled by a mechanism related to allelic exclusion as is valid for *var* genes and that almost 100% sequence identity in *rif* 5′ ups regions is not sufficient to predict promoter co-activation or silencing. However, not all *rif* genes are transcribed meaning that there still must be some kind of controlled activation and repression. Considering the results from Guizetti and colleagues[14], it may be postulated that ncRNAs are also involved in the activation of *rif* loci, with the difference that these ncRNAs allow for activity from multiple loci – similar to what is seen for *var* genes when certain ncRNA from GC rich regions are overexpressed[14]. From the study of Guizetti and colleagues, it is not clear what influence the *var* relevant ncRNAs had on *rif* gene transcription although these authors speculated that a co-regulation could happen. Importantly, there are ncRNAs for most *rif* loci with yet elusive function[45]. A specific transcription machinery in the periphery of the nucleus was postulated for *var* genes[46–48]. To date, it is unclear if *rif* genes are also transcribed from the same machinery. This

could be elucidated by RNA FISH or other adequate methods to monitor nuclear substructures, and the created parasite line may be useful for this kind of experiment. It also still remains elusive if the B subfamily of *rif* genes is differentially regulated, given the fact that we found only one B-type *rif* gene activated in our cultures.

Material and Methods

Parasite culture and transfection.   Parasites (strain NF54) were cultured under biological level 2 conditions at 5% hematocrit in human B+ blood supplemented with 0.5% Albumax 1 (Invitrogen) or 10% human B+ plasma in RPMI and 0.23% sodium bicarbonate under a 90% N_2, 5% CO_2, 5% O_2 atmosphere or in candle jars as described earlier[49]. Human blood and plasma were obtained from the local blood bank and ethical clearance for using this blood for this research was granted by the Ethics Committee of the Institute of Biomedical Sciences at the University of São Paulo (No. 842/2016). The medium was changed daily or every two days when parasitemias were low (<0.5%). Parasitemias were checked by Giemsa-stained thin blood smears. For transfection, the protocol suggested by Hasenkamp *et al*.[50] was used. Essentially, 150 µl of cytomix-washed fresh red blood cells were electroporated in a BioRad Gene Pulser at 310 V, 960 µF with 40 µg purified plasmid and later mixed with $2*10^7$ plasmagel-purified[51] schizont stage parasites. On day 2 after transfection, Blasticidin (Sigma) was added at 2.5 µg/ml to the medium. At day 6 after addition of the drug, no more viable parasites were visible and the medium was changed every two days until the appearance of parasites on days 17–25 post-transfection.

Plasmid constructs.   The bi-cistronic plasmid used here is based on the pTZ57 (Thermo/Fermentas) backbone, where a cassette containing the hsp86 promoter fragment (BglII-NcoI) from pPF86[52] was inserted. The blasticidin-deaminase coding sequence was excised from pBMNL106P-PpLuciBlast[53] (NcoI-SalI) and inserted into the vector. Then, the hsp86 terminator sequence was inserted from pPF86 via SalI and BamHI. In this plasmid, the hDHFR resistance cassette from pRESA-GFP-HA[54] was inserted via EcoRI/EcoRV inserting in EcoRI/SmaI. This plasmid was used to exchange the Calmodulin promoter for a 5′ upstream region from three *rif* genes which were previously shown to be transcribed in wild-type 3D7 cultures[20]. The putative *rif* promoter regions were amplified using oligonucleotides (forward/reverse) **ctcgag**atataaatttgtaaaaaccatgtg/**ggatcc**ttaattgtgatacgtatat-tatttaatg (precedes PF3D7_0200700, *rif* A type), **ctcgag**tattatattttttatatataattattcgtg/**ggatcc**atttaatgtgatacttatattat-ttatg (precedes PF3D7_0900500, *rif* B type), and **ctcgag**atgtaatatattattatgttaatattc/**ggatcc**tattgtgatacgtatattattttatg (precedes either PF3D7_1300400, PF3D7_0600500, PF3D7_0937500, PF3D7_1150300, all *rif* A type). The amplified fragments were 1400 to 1500 nt and the identity of the sequences was confirmed by semiautomatic sequencing in an Applied Biosystems 7550 sequencer. An outline of the different *rif* 5′ upstream region containing plasmid is given in Fig. 1.

Selection of WR99210 resistant parasite lines and total RNA preparation.   After outgrowth of parasites resistant to 2.5 µg/ml Blasticidin, parasites were split into two parallel cultures and cultivated in RPMI/Albumax supplemented with either 2.5 µg/ml Blasticidin/2.5 nM or 5 nM WR99210 or solely 2.5 µg/ml Blasticidin. After 4 reinvasion cycles, parasites were synchronized by plasmagel floatation[51] and subsequent Sorbitol lysis[55] and then harvested after another reinvasion. harvested in ring, trophozoite or schizont stage forms. Harvested IRBC were treated with 0.1% Saponin for 10 min at RT and then pelleted at 12000 g/4 °C for 5 min, washed once in 1 ml PBS and then resuspended in a volume of 100 µl using TE. Afterwards, 1 ml Trizol (Invitrogen) was added and the sample was stored at −80 °C until use. Total RNA was prepared following the Trizol protocol provided by the manufacturer. Final total RNA was dissolved in 20 µl RNAse free water and stored at −80 °C until use.

Real-time PCR with parasite-derived cDNA and PCR test for integration.   Total RNA was converted to cDNA using the previously published protocol[56] using 5× Hotfire Pol SYBR Mix (Solis Biodyne Inc.) on an Eppendorf realplex2 thermocycler. In order to monitor the hDHFR transcription controlled by *rif* upstream regions, the following oligonucleotides were employed: forward 5′-ctggttctccattcctgagaag/reverse 5′-ttgtggaggttccttgagttct. These oligos were predicted using Primer3 online software[57] using the same settings as used for the design of *var* oligos and the internal control oligonucleotide pair for the amplification of seryl tRNA ligase (PF3D7_0717700)[58]. Relative transcript quantities were then calculated by the $2^{-\Delta Ct}$ method[59] using the seryl tRNA ligase transcript as endogenous control. The performance of the hDHFR oligonucleotide pair does not differ from seryl t-RNA ligase specific oligos (data not shown). To test for integration of plasmids, genomic DNAs (gDNA) were prepared from parasite cultures using the Promega genomic DNA preparation kit and aliquots of the gDNAs were tested combining forward and reverse real-time PCR oligos for hDHFR (see above) and blasticidin deaminase (forward: 5′-tgcagtttcgaatggacaaa, reverse: 5′-aacacaaaacaatctggtgcat). PCRs were performed using Thermo/Invitrogen Elongase with the following thermocycling program: 94 °C, 40 s; 54 °C, 40 s; 65 °C, 4 min, over 30 cycles.

RNAseq with parasite-derived cDNA.   Total RNA was isolated from the harvested trophozoite-stage parasites with Trizol reagent following the manufacturer's instruction. The concentration of the isolated RNA was quantified using a Nanodrop device (Thermo Scientific, USA) and the integrity of the RNA was measured by a 2100 Bioanalyzer (Agilent Technologies, CA). Paired-end sequencing cDNA libraries were constructed from two samples (grown with or without 2.5 nM WR99210 for 6 reinvasions) using a TruSeq RNA Sample Preparation Kit v2 low sample (LS) protocol (Illumina Inc., CA), based on the manufacturer's instructions. RNAseq was conducted in an Illumina NextSeq500 sequencer following the recommendations of the provider using mid output flow cells and a total of 10 million reads per sample. CLC Genomics Workbench 7.01 platform was used to remove the adapter and assess reads quality from the raw reads. The same platform was used to subject reads to the reference *P. falciparum* genome available in the Ensemble Genome database (Release 26) and to generate gene reads count table used in the following pipelines. Two packages built in Bioconductor[60] were used for further

analysis: edgeR[61] and limma[62]. The first was used to filter and normalize the data set, while linear modeling and empirical Bayes to assess differential expression were performed with the limma package. Finally, the RPKM values of trophozoite-derived cDNA from cultures which were treated with WR99210 or not were loaded in MS Access and filtered for product ID containing "*rif*", and the corresponding RPKM values for each *rif* gene were plotted against their ID.

Southern Analysis. Genomic DNA was isolated from the WT (NF54) and NF54::pTZ57$_{PF3D7\_0200700}$ parasites, using the Wizard® Genomic DNA Purification Kit (Promega). 5μg of each gDNA and 25ηg of plasmid DNA (pDNA) from the pTZ57$_{PF3D7\_0200700}$ construct, were digested using SpHI (Thermo Fisher Scientific). The probe was amplified using standard PCR conditions with digoxigenin-dUTP from DIG High Prime DNA Labeling and Detection Starter Kit I (Roche Diagnostics), using the Blasticidin deaminase cassette from the pTZ57$_{PF3D7\_0200700}$ construct as a template. The oligonucleotide primers used were 5′- atgggaaaaacatttaacatttc-3′ and 5′-aacacaaaacaatctggtgcat-3′. The Southern procedure was performed following the protocol provided by the manufacturer of the labeling kit (Roche), using Hybond N membranes (Amersham/GE Healthcare) and a hybridization temperature of 44 °C. All washings steps were performed at room temperature.

References

1. WHO: World Malaria Report. 94 (www.who.int/malaria/publications/world-malaria-report-2017/en/) (2017).
2. Su, X. Z. *et al*. The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. *Cell* **82**, 89–100 (1995).
3. Freitas-Junior, L. H. *et al*. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature* **407**, 1018–1022 (2000).
4. Claessens, A. *et al*. Generation of Antigenic Diversity in Plasmodium falciparum by Structured Rearrangement of Var Genes During Mitosis. *PLoS Genet*. https://doi.org/10.1371/journal.pgen.1004812 (2014).
5. Sander, A. F. *et al*. DNA secondary structures are associated with recombination in major Plasmodium falciparum variable surface antigen gene families. *Nucleic Acids Res*. **42**, 2270–2281 (2014).
6. Rask, T. S., Hansen, D. A., Theander, T. G., Gorm Pedersen, A. & Lavstsen, T. Plasmodium falciparum erythrocyte membrane protein 1 diversity in seven genomes–divide and conquer. *PLoS Comput. Biol*. **6**, (2010).
7. Voss, T. S. *et al*. A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. *Nature* **439**, 1004–8 (2006).
8. Brancucci, N. M. B., Witmer, K., Schmid, C. D., Flueck, C. & Voss, T. S. Identification of a cis-acting DNA-protein interaction implicated in singular var gene choice in Plasmodium falciparum. *Cell. Microbiol*. **14**, 1836–1848 (2012).
9. Amulic, B., Salanti, A., Lavstsen, T., Nielsen, M. A. & Deitsch, K. W. An upstream open reading frame controls translation of var2csa, a gene implicated in placental malaria. *PLoS Pathog*. **5**, e1000256 (2009).
10. Brancucci, N. M. B., Witmer, K., Schmid, C. & Voss, T. S. A var gene upstream element controls protein synthesis at the level of translation initiation in Plasmodium falciparum. *PLoS One* **9**, e100183 (2014).
11. Frank, M. *et al*. Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite Plasmodium falciparum. *J Biol Chem* **281**, 9942–9952 (2006).
12. Pérez-Toledo, K. *et al*. Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. *Nucleic Acids Res*. **37**, 2596–606 (2009).
13. Guizetti, J. & Scherf, A. Silence, activate, poise and switch! Mechanisms of antigenic variation in Plasmodium falciparum. *Cell. Microbiol*. **15**, 718–26 (2013).
14. Guizetti, J., Barcons-Simon, A. & Scherf, A. Trans-acting GC-rich non-coding RNA at *var* expression site modulates gene counting in malaria parasite. *Nucleic Acids Res*. gkw664 https://doi.org/10.1093/nar/gkw664 (2016).
15. Janssen, C. S., Phillips, R. S., Turner, C. M. R. & Barrett, M. P. Plasmodium interspersed repeats: the major multigene superfamily of malaria parasites. *Nucleic Acids Res*. **32**, 5712–20 (2004).
16. Yam, X. Y. *et al*. Characterization of the Plasmodium Interspersed Repeats (PIR) proteins of Plasmodium chabaudi indicates functional diversity. *Sci. Rep*. **6**, 23449 (2016).
17. Spence, P. J. *et al*. Vector transmission regulates immune control of Plasmodium virulence. *Nature* **498**, 228–231 (2013).
18. Ebbinghaus, P. & Krucken, J. Characterization and tissue-specific expression patterns of the Plasmodium chabaudi cir multigene family. *Malar J* **10**, 272 (2011).
19. Cunningham, D. *et al*. Rapid changes in transcription profiles of the Plasmodium yoelii yir multigene family in clonal populations: lack of epigenetic memory? *PLoS One* **4**, e4285 (2009).
20. Cabral, F. J. J. & Wunderlich, G. Transcriptional memory and switching in the Plasmodium falciparum rif gene family. *Mol Biochem Parasitol* **168**, 186–190 (2009).
21. Cheng, Q. *et al*. stevor and rif are Plasmodium falciparum multicopy gene families which potentially encode variant antigens. *Mol Biochem Parasitol* **97**, 161–176 (1998).
22. Goel, S. *et al*. RIFINs are adhesins implicated in severe Plasmodium falciparum malaria. *Nat. Med*. **21**, 314–17 (2015).
23. Saito, F. *et al*. Immune evasion of Plasmodium falciparum by RIFIN via inhibitory receptors. *Nature* **552**, 1–49 (2017).
24. Bernabeu, M. *et al*. Functional analysis of Plasmodium vivax VIR proteins reveals different subcellular localizations and cytoadherence to the ICAM-1 endothelial receptor. *Cell Microbiol* **14**, 386–400 (2012).
25. Carvalho, B. O. *et al*. On the cytoadhesion of Plasmodium vivax-infected erythrocytes. *J. Infect. Dis*. **202**, 638–647 (2010).
26. Joannin, N., Abhiman, S., Sonnhammer, E. L. & Wahlgren, M. Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. *BMC Genomics* **9**, 19 (2008).
27. Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T. & Kaneko, I. Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Mol. Microbiol*. **75**, 854–63 (2010).
28. Martins, R. M. *et al*. An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite Plasmodium falciparum. *Sci. Rep*. **7**, 14042 (2017).
29. Howitt, C. A. *et al*. Clonally variant gene families in Plasmodium falciparum share a common activation factor. *Mol Microbiol* **73**, 1171–1185 (2009).
30. Wang, C. W., Magistrado, P. A., Nielsen, M. A., Theander, T. G. & Lavstsen, T. Preferential transcription of conserved rif genes in two phenotypically distinct Plasmodium falciparum parasite lines. *Int J Parasitol* **39**, 655–664 (2009).
31. Witmer, K. *et al*. Analysis of subtelomeric virulence gene families in Plasmodium falciparum by comparative transcriptional profiling. *Mol Microbiol* **84**, 243–259 (2012).
32. Dzikowski, R., Frank, M. & Deitsch, K. Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog* **2**, e22 (2006).
33. Scherf, A. *et al*. 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 (1998).

34. Fidock, D. A., Nomura, T. & Wellems, T. E. Cycloguanil and its parent compound proguanil demonstrate distinct activities against Plasmodium falciparum malaria parasites transformed with human dihydrofolate reductase. *Mol. Pharmacol.* **54**, 1140–7 (1998).
35. Horn, D. & McCulloch, R. Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. *Curr Opin Microbiol* **13**, 700–705 (2010).
36. Prucca, C. G. & Lujan, H. D. Antigenic variation in Giardia lamblia. *Cell. Microbiol.* **11**, 1706–15 (2009).
37. Prucca, C. G. *et al.* Antigenic variation in Giardia lamblia is regulated by RNA interference. *Nature* **456**, 750–4 (2008).
38. Navarro, M. & Gull, K. A pol I transcriptional body associated with VSG mono-allelic expression in Trypanosoma brucei. *Nature* **414**, 759–763 (2001).
39. Figueiredo, L. M., Cross, G. A. & Janzen, C. J. Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat Rev Microbiol* **7**, 504–513 (2009).
40. Kyes, S. *et al.* Plasmodium falciparum var gene expression is developmentally controlled at the level of RNA polymerase II-mediated transcription initiation. *Mol. Microbiol.* **63**, 1237–47 (2007).
41. Zeng, W., Ball, A. R. & Yokomori, K. HP1: heterochromatin binding proteins working the genome. *Epigenetics* **5**, 287–92 (2010).
42. Tonkin, C. J. *et al.* Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. *PLoS Biol.* **7**, e84 (2009).
43. Schieck, E., Pfahler, J. M., Sanchez, C. P. & Lanzer, M. Nuclear run-on analysis of var gene expression in Plasmodium falciparum. *Mol. Biochem. Parasitol.* **153**, 207–12 (2007).
44. O'Donnell, R. A. *et al.* A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes. *EMBO J.* **21**, 1231–9 (2002).
45. Raabe, C. A. *et al.* A global view of the nonprotein-coding transcriptome in Plasmodium falciparum. *Nucleic Acids Res.* **38**, 608–17 (2010).
46. Guizetti, J., Martins, R. M., Guadagnini, S., Claes, A. & Scherf, A. Nuclear Pores and Perinuclear Expression Sites of var and Ribosomal DNA Genes Correspond to Physically Distinct Regions in Plasmodium falciparum. *Eukaryot. Cell* **12**, 697–702 (2013).
47. Ralph, S. A., Scheidig-Benatar, C. & Scherf, A. Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. in. *Proc Natl Acad Sci USA* **102**, 5414–5419 (2005).
48. Issar, N., Ralph, S. A., Mancio-Silva, L., Keeling, C. & Scherf, A. Differential sub-nuclear localisation of repressive and activating histone methyl modifications in P. falciparum. *Microbes Infect.* **11**, 403–407 (2009).
49. Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science (80-.).* **193**, 673–675 (1976).
50. Hasenkamp, S., Russell, K. T. & Horrocks, P. Comparison of the absolute and relative efficiencies of electroporation-based transfection protocols for Plasmodium falciparum. *Malar. J.* **11**, 210 (2012).
51. Lelievre, J., Berry, A. & Benoit-Vical, F. An alternative method for Plasmodium culture synchronization. *Exp Parasitol* **109**, 195–197 (2005).
52. Militello, K. T. & Wirth, D. F. A new reporter gene for transient transfection of Plasmodium falciparum. *Parasitol. Res.* **89**, 154–7 (2003).
53. Banaszynski, L. A., Chen, L.-C., Maynard-Smith, L. A., Ooi, A. G. L. & Wandless, T. J. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* **126**, 995–1004 (2006).
54. de Azevedo, M. F. *et al.* Systematic Analysis of FKBP Inducible Degradation Domain Tagging Strategies for the Human Malaria Parasite Plasmodium falciparum. *PLoS One* **7**, e40981 (2012).
55. Lambros, C. & Vanderberg, J. P. Synchronization of Plasmodium falciparum erythrocytic stages in culture. *J Parasitol* **65**, 418–420 (1979).
56. Golnitz, U., Albrecht, L. & Wunderlich, G. Var transcription profiling of Plasmodium falciparum 3D7: assignment of cytoadherent phenotypes to dominant transcripts. *Malar J* **7**, 14 (2008).
57. Untergasser, A. *et al.* Primer3–new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
58. Salanti, A. *et al.* Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Mol. Microbiol.* **49**, 179–91 (2003).
59. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2-\Delta\Delta CT$ Method. *Methods* **25**, 402–408 (2001).
60. Huber, W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).
61. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
62. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).

Acknowledgements

Author Contributions

R.B.D.A., T.M.S., C.S.K. produced qPCR data, G.F.L. and G.W. constructed transfection plasmids and produced transfectants, D.A. and P.M.R. performed RNAseq and its analysis, G.W. conceived the project and wrote the paper. All authors read and approved the manuscript and its data.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-27646-0.

Competing Interests: The authors declare no competing interests.

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

Supplementary Figures

Independent regulation of *Plasmodium falciparum rif* gene promoters

Rosana Beatriz Duque Araujo, Tatiane Macedo Silva, Charlotte Sophie Kaiser, Gabriela Fernandes Leite, Diego Alonso, Paulo Eduardo Martins Ribolla and Gerhard Wunderlich
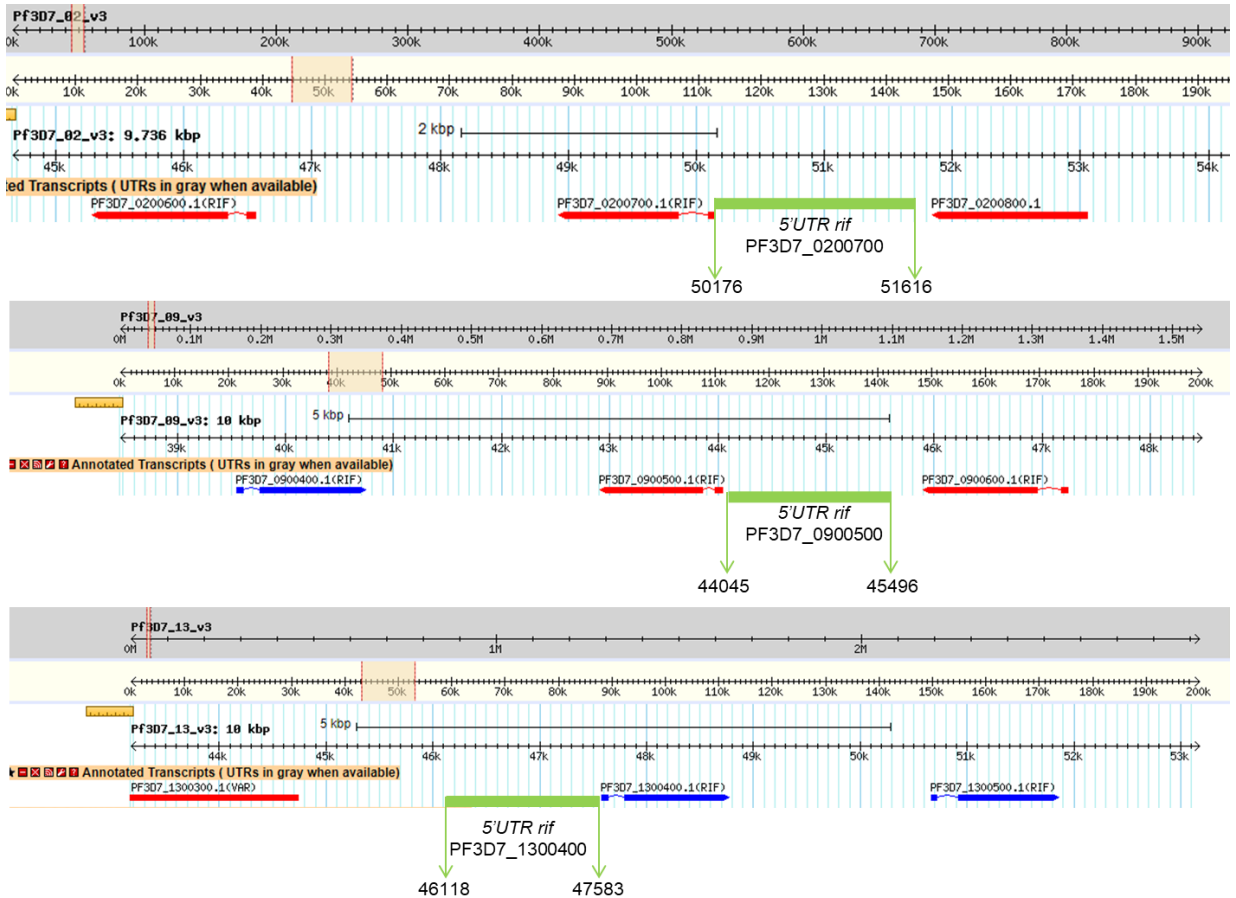
Screenshots from PlasmoDB V.36 showing the genomic context of 5' ups rif regions that were cloned in the three expression vectors used in the experiments.
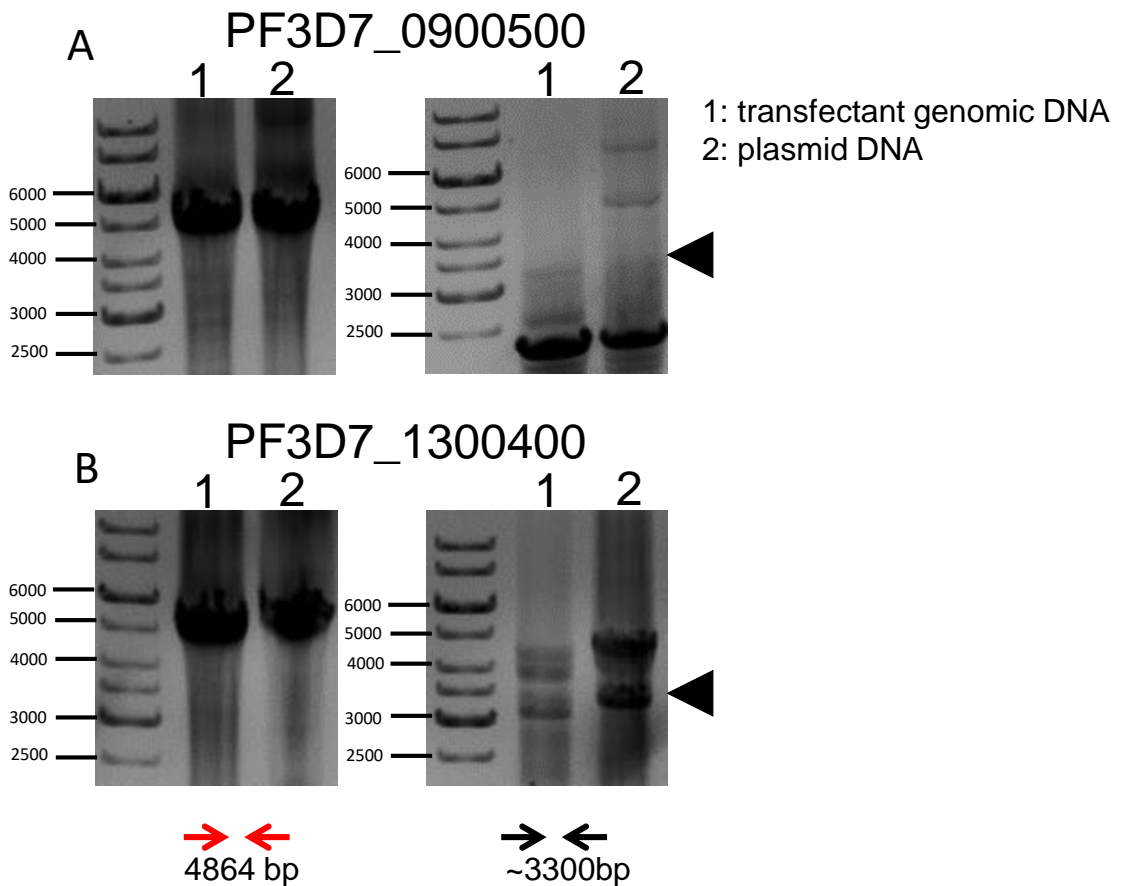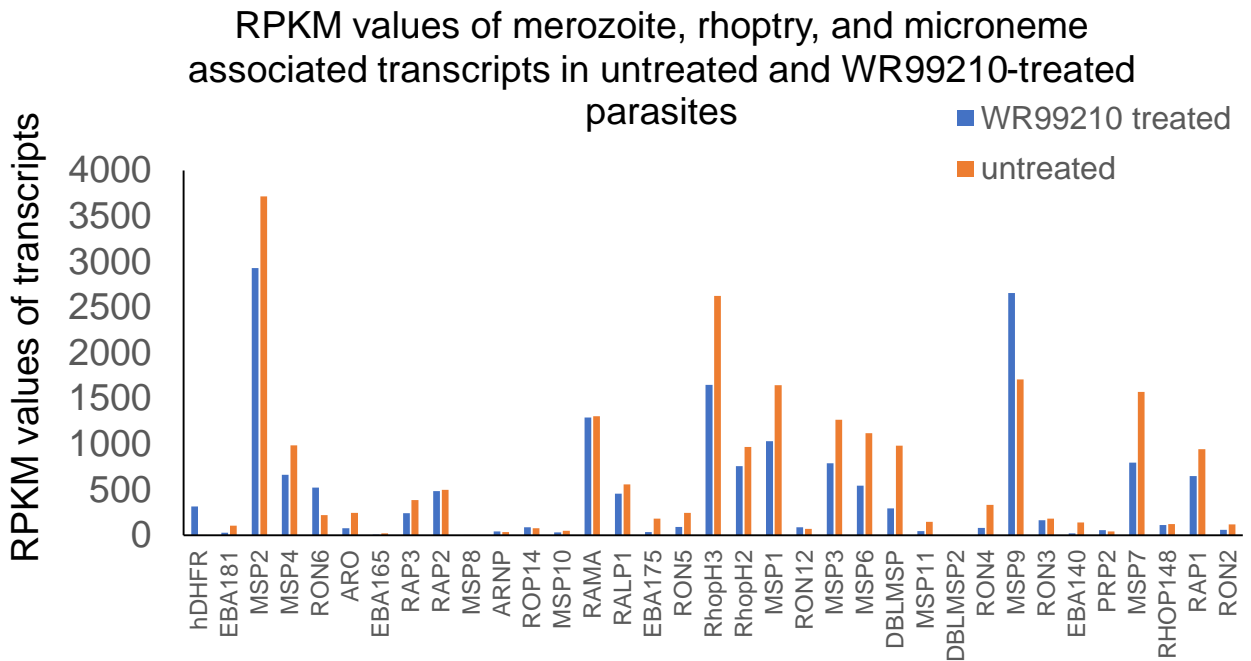
Results from PCRs using long-range polymerases show that constructs containing 5' *rif* ups PF3D7_0900500 (in **A**) and PF3D7_1300400 (in **B)** showed episomal forms in contrast to construct with PF3D7_0200700 (Figure 1). Gel pictures on the left, PCR results showing amplification with forward *hDHFR* and forward *bsd* primers (in red) over the backbone of the transfection plasmid (positive for the untransfected plasmid (2) as well as the gDNA from the 5' *rif* ups transfected strain (1)). On the right, amplification products from PCRs using reverse *hDHFR* and reverse *bsd* oligos. This amplification is only possible when the plasmid is in the episomal form (see scheme for episomal construct and integrated locus in Fig. 1 A). Probably due to the high A/T amount, amplification products over two *Plasmodium* promoter regions (see map in Figure 1) show several amplicons for the "reverse" oligo pair and the black triangle depicts the expected 3300 bp product.

gDNA
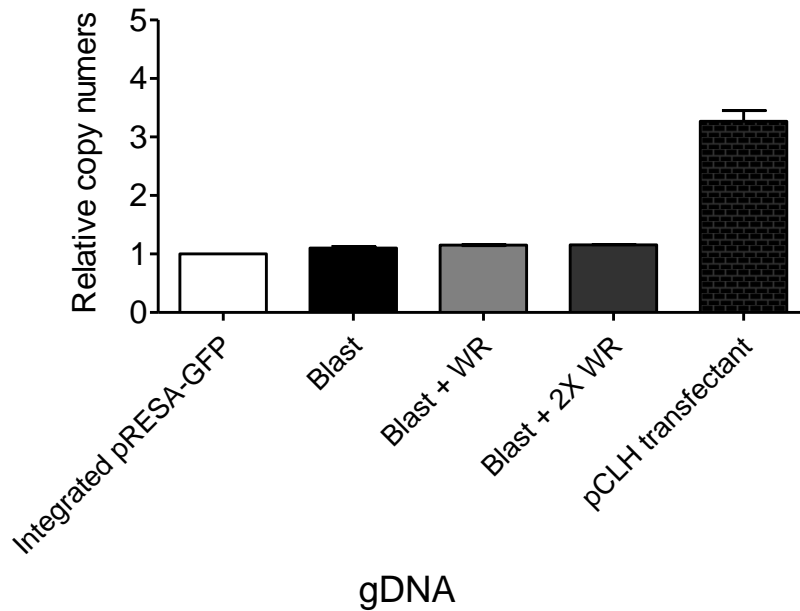
The copy number of the hDHFR locus in the PF3D7_0200700 transfectant line remains stable independent of increased drug pressure, indicating integration. As a template for qPCR using hDHFR and t-seryl synthetase oligos, genomic DNAs from indicated cultures were tested and the relative copy numbers of the hDHFR versus the genomic t-seryl synthetase locus were calculated and plotted. Note that increased presence of WR99210 did not lead to increased hDHFR copy numbers ("Blast + 2x WR" sample). WR99210 was applied in all transfectants at 2.5 nM with the exception of the "Blast + 2xWR" sample where 5 nM was used.

RPKM values of merozoite, rhoptry, and microneme associated transcripts in untreated and WR99210-treated parasites

RNAseq results (RPKM values) from the parasite lineage NF54::pTZ57PF3D7_0200700 were filtered for transcripts encoding merozoite, rhoptry, and microneme proteins and plotted. Note that the untreated parasites showed higher RPKM values in typical merozoite associated factors such as MSP1 and MSP2 than the WR99210 treated parasites, indicating that untreated parasites were recovered at a slightly more advanced stage in the intraerythrocytic cycle. The hDHFR transcript from the artificially integrated rif 5' ups is shown on the left.

The data were extracted from Supplementary table 1 from which the corresponding PlasmoDB IDs can be obtained.

Full-length gel from Figure 1 as photographed
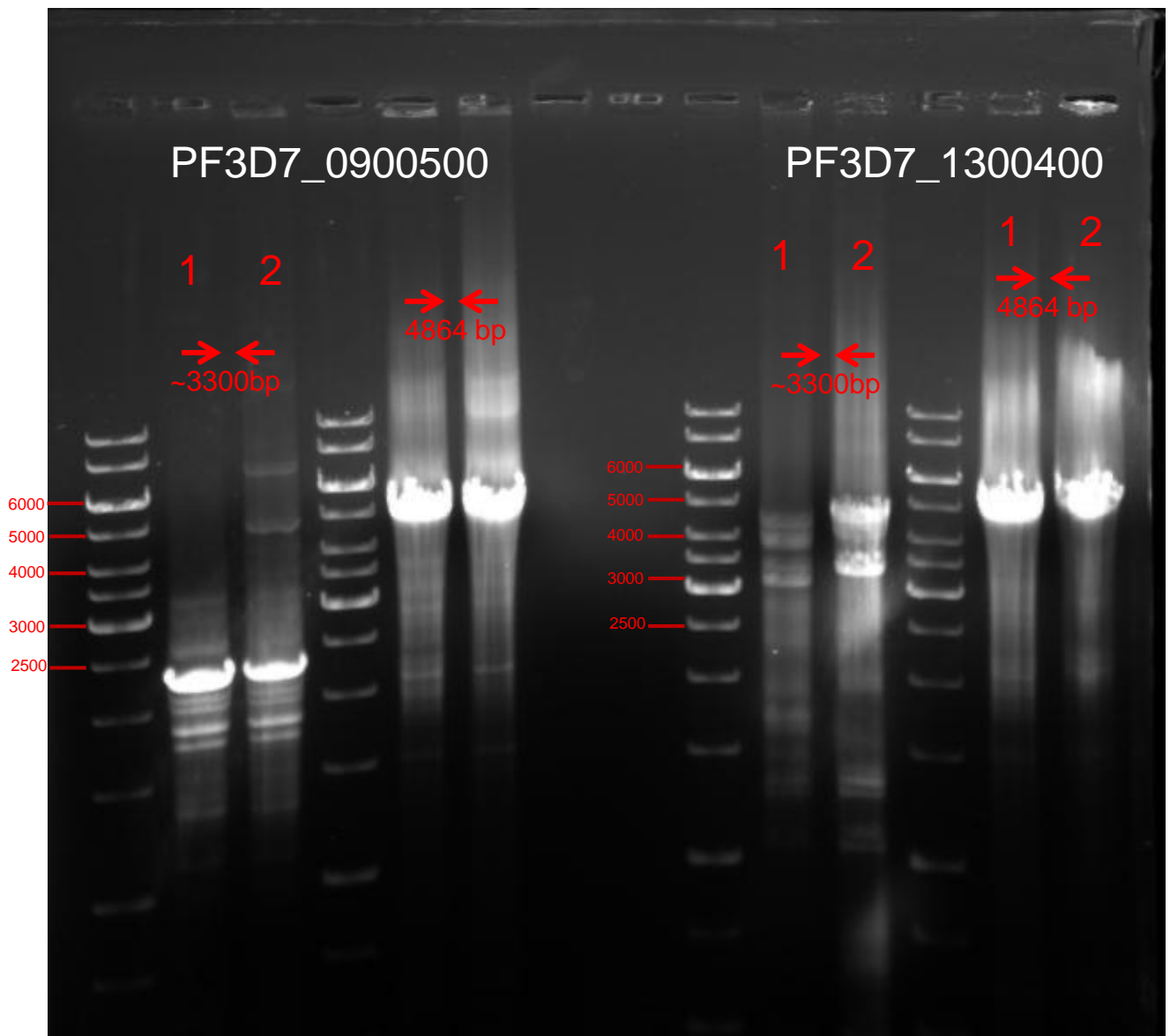
1: transfectant genomic DNA
2: plasmid DNA
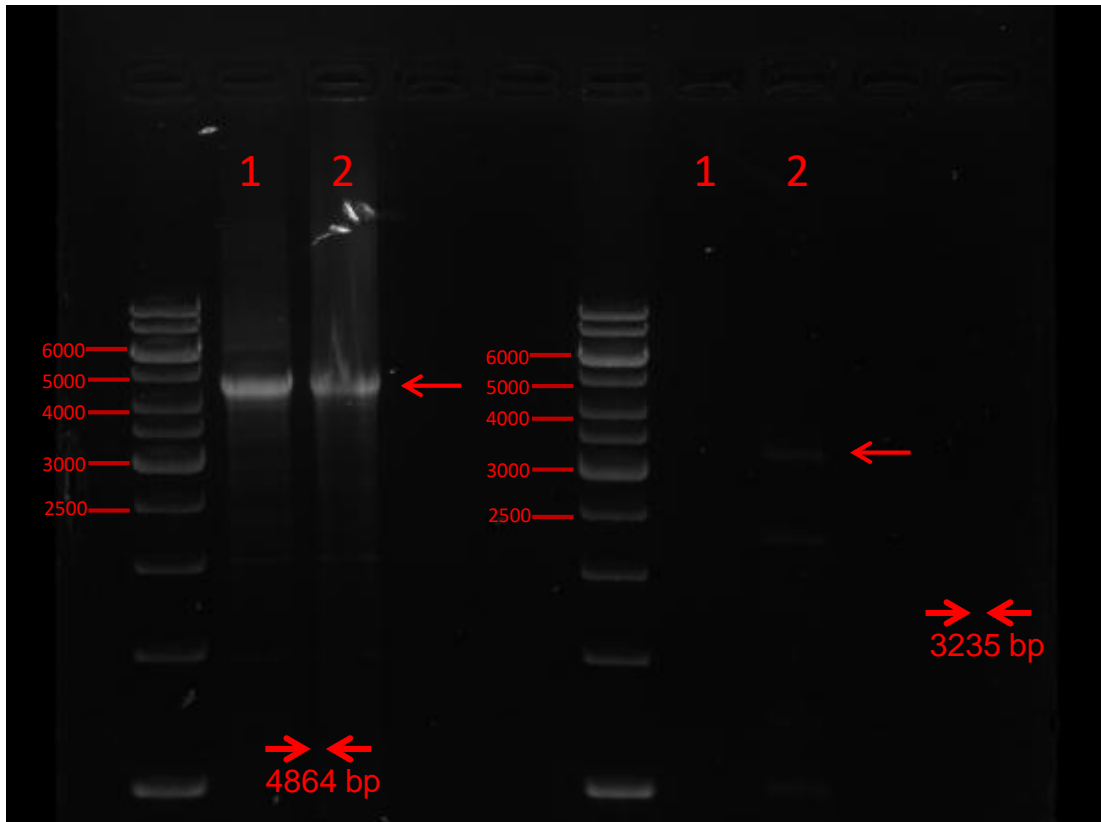
1% TAE agarose gel from Figure 1 using 1 Kb ladder
from Thermo as molecular weight standard. The oligo pairs
employed herein result in the indicated fragment sizes (see
legend of Fig 1 for details).

Original Gel photograph from gel shown in Supplementary Figure 2 - Integration check PCR results

PF3D7_0900500 and PF3D7_1300400 5' ups constructs

1: transfectant genomic DNA
2: plasmid DNA

1% TAE agarose gel from Supplementary Figure 1 using 1 Kb ladder from Thermo as molecular weight standard. The oligo pairs employed herein result in the indicated fragment sizes (see legend of Supplementary Fig 1 for details).

8.2. APPENDIX 2: *Pf*LSD1 Mass Spectrometry data set.

https://rduquepflsd1msandrnaseqappendices.blogspot.com/

8.3. APPENDIX 3: *Pf*LSD1 RNA-Seq data set.

https://rduquepflsd1msandrnaseqappendices.blogspot.com/