Investigating the mechanism of Heterochromatin Protein-1 (HP1) mediated *var* gene regulation in *Plasmodium falciparum*

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

द्वारा / By
ममतारानी डी वी/ Mamatharani D V

पंजीकरण सं. / Registration No.: 20183584

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor: Dr. Krishanpal Karmodiya/
डॉ. कृष्णपाल करमोदिया

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE
2025

Dedicated to Amma, Appa and Jayakrishnan…

CERTIFICATE

Certified that the work incorporated in the thesis entitled "**Investigating the mechanism of Heterochromatin protein-1 (HP1) mediated *var* gene regulation in *Plasmodium falciparum***" Submitted by Mamatharani D V was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

(Supervisor)

Date: 09.05.25

Krishanpal Karmodiya

DECLARATION

Name of Student: Mamatharani D V

Reg. No.: 20183584

Thesis Supervisor(s): Dr. Krishanpal Karmodiya

Department: Biology

Date of joining program: 01.08.2018

Date of Pre-Synopsis Seminar : 13.11.2024

Title of Thesis : **Investigating the mechanism of Heterochromatin protein-1 (HP1) mediated *var* gene regulation in *Plasmodium falciparum***

I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I declare that I have acknowledged collaborative work and discussions wherever such work has been included. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

The work reported in this thesis is the original work done by me under the guidance of Dr. Krishanpal Karmodiya.

Date: 09.05.25

Signature of the student 

ACKNOWLEDGEMENTS

As I am writing this part of my thesis, an endless wave of gratitude and happiness covers me. The time I spent here at IISER Pune has been absolutely joyful.

First and foremost, I would like to express my deepest gratitude towards my thesis supervisor Dr. Krishanpal Karmodiya. Krish has shaped me into an independent person, by opening a door of endless possibilities and letting me dive deep. His patience and kindness have made the hardships bearable. I also appreciate him for allowing me to frankly express my ideas and to provide critical suggestions.

I am grateful to my research advisory committee, Dr. Gayathri Pananghat, Dr. Mridula Nambiar and Prof. Vasudevan Seshadri for their valuable suggestions throughout my PhD. A special thanks to Dr. Gayathri for all the help and inputs with the initial biochemical experiments.

My sincere gratitude towards all the collaborators that contributed to this work. This study wouldn't have taken this shape without their help. I am grateful for the opportunity provided by Prof. Till Voss at Swiss TPH for generating all the transgenic lines used in this study. I would like to acknowledge the contributions of Dr. Igor Niederwieser (Swiss TPH) in designing the CRISPR-Cas9 constructs and in executing the transfection experiments. I appreciate all the new things I learned from Igor and Till.

I am thankful to Prof. Samrat Mukhopadhyay and Sandeep K Rai for all the help with the initial standardization of phase separation assays and experiments done at IISER Mohali. I especially would like to remember the pleasant time I spent at Mohali with Sandeep and friends.

My sincere thanks to Dr. Mahipal Ganji and Prakshi Gaur at IISc Bangalore for the collaboration on the single-molecular DNA tethering experiments. I would like to extend a note of appreciation to all the lab members from the three collaborators (Till Voss's, Samrat's and Mahipal's) for always being helpful and supportive.

I thank Devatrisha Purkayastha for all the help with various experiments included in this thesis. Also, my gratitude to Bhagyashree for the initial harvesting of the ChIP and RNA sequencing samples used in chapter 2.

The plasmids used for recombinant protein expression for this study were a kind gift from Prof. Thomas Pucadyil (IISER Pune). The parasite line 3D7/1G5/DiCre used for PfHP1 mutant line generation was obtained from Michael J. Blackman.

I thank all the undergraduate students (Saakshi, Kalyan, Mohit, Trisha, Vasudha, Srijan, Sharath, Nirav, Radhika, Devika, Varuni, Somdatta and Manas) who worked with me during my PhD. Their enthusiasm and inquisitiveness have fueled the project in the right direction. A special thanks to Kalyan for his contributions to the 'insulator proteins' project.

I am also grateful for the phase separation workshop organized by Prof. Geeta Narlikar. The scientific discussions with her were instrumental to this project.

I am thankful for the support received from the EMBO short-term fellowship for my visit at Swiss TPH. I acknowledge all the funding I received throughout my PhD. I thank IISER Pune for all the infrastructure and support. My PhD fellowship was funded by CSIR and I also received the IISER Pune-IDEAS scholarship. I received support from GRC, Infosys travel award and CSIR travel award, for attending the GRC chromosome dynamics conference (2023).

I am grateful to the various facilities at IISER Pune and the very efficient staff. The NGS facility, mass spectrometry facility, microscopy facility and animal house facility have been important for turning our dreams into a reality.

I would like to use this opportunity to express my appreciation for the diligent staff at the bio department, academic section, accounts department and purchase department. They collectively made my life a little better with their kindness and timely documentation.

I consider myself luckiest to be a part of the Krish lab. I admire the spirit of this lab and will continue to cherish the fond memories with them. The critical suggestions they gave were invaluable to shaping this project. I am thankful to Mukul for the warm welcome he gave when I joined the lab and for his constant motivation. I thank Abhishek for teaching me all the techniques

during my initial days. I will always remember his 'quirkiness'. I am thankful to dear Anjani for her love and warmth. I am especially thankful to Bhagyashree for sticking with me through thick and thin. A special mention to also, Tanwee for always being affectionate. I have never been part of a more wholesome bunch than the current Krish lab members, Trisha, Disha, Saptarshi, Kushankur, Sarthak, Arya, Riya and Neha. I cannot describe them any better than a sudden burst of colorful clouds in the wonderful sky!

I have to mention all the fun we had in the lab parties, especially with Anamika, Ankrisha and Kiyan. Thanks to Anamika and Krish for being such gracious hosts.

I am thankful to the other lab members with whom we shared common lab space. Especially to the SG lab and AKB lab for always being kind and supportive.

I am also grateful to all my friends at IISER Pune, especially my batchmates, Shilpa, Ardhra, Akshay, Keerti, Nikita, Netra, Dhruv and Soumyajit. They have helped me through the hardships of the PhD and made my life here a memorable experience.

I also thank all my friends for always believing in me and accepting me for who I am. Lastly, I thank my family for their unwavering support throughout this journey. I will never forget to eternally love Amma, Appa and Jayakrishnan!

Sincerely,

Mamatharani Deepa Venugopal

CONTENTS

ABSTRACT

Malaria caused by the *Plasmodium falciparum* is driven by alterations in the infected RBC surface by recruiting antigens that can bind to the endothelial receptors (cytoadherence). The clonally variant multigene family of genes known as *var* genes (~60 members) codes for proteins that cause cytoadherence. The *var* gene family shows mutually exclusive expression even though the mechanism of regulation is unclear. Interestingly, *var* genes are marked by a unique set of activation (histone acetylations) and repression (H3K9me3) marks, some of which are exclusive for virulence family genes. PfHP1 (*P. falciparum* Heterochromatin Protein 1), a homolog of HP1, specifically binds to H3K9me3 histone modification. ChIP sequencing of PfHP1 has revealed a restricted pattern of heterochromatin spread across all the *var* introns that is correlated with the intronic transcription from there. These results indicate that PfHP1 shows a reversible spread across the *var* genes that may enable activation of some *var* genes while others are kept suppressed. But the mechanism of PfHP1 dependent spread of heterochromatin is not very well known. In HP1 homologs of higher eukaryotes it is shown that heterochromatinization is facilitated by liquid-liquid phase separation (LLPS). We further explored the property of phase separation of PfHP1 *in vitro* to understand the biochemical basis of heterochromatinization. Our data suggest that the PfHP1 phase separates *in vitro* in an RNA-dependent manner. Higher concentrations of AU-rich non-coding RNA (notably *var* introns are ~ 80 % AT-rich) causes the dissolution of PfHP1 droplets. Our single molecular DNA tethering experiments showed that PfHP1 preferably forms puncta over AT-rich intronic DNA and compacts the DNA. We have also identified point mutations in PfHP1 that disrupts the phase separation *in vitro*. We have generated *P. falciparum* transgenics with these mutations to further study the effect of PfHP1 phase separation *in vivo* and on *var* gene expression. Overexpression of PfHP1 mutants showed reduced chromatin binding on *var* genes. Conditional expression of PfHP1 point mutations showed dispersed nuclear localization as opposed to puncta-like appearance of the wild type protein. In addition some of these point mutations showed growth defects and gametocytogenesis in the parasites. These observations indicate a strong possibility that disruption of phase separating properties of PfHP1 *in vivo* might affect its gene regulatory function. Our study reveals the nature of PfHP1 mediated heterochromatin formation and its role in *var* gene regulation.

SYNOPSIS

Malaria is a mosquito borne disease that is spread across tropical and subtropical regions. The disease kills millions each year, especially in Africa. *Plasmodium falciparum*, which causes the most lethal form of malaria, is the prime focus of this thesis. In the **Chapter 1, Introduction**, the history of malaria, life cycle of *P. falciparum* and major treatment/control of malaria are touched upon.

The main objective of this study is to understand the mechanism of transcriptional control of virulence gene expression. Hence, the basics of transcriptional regulation in *P. falciparum* is talked about in detail. There are clonally variant gene (CVGs) families in *P. falciparum*; amongst those, *var* genes/PfEMP1 (*P. falciparum* erythrocyte membrane protein 1) is of main interest to this study. The *var* genes are responsible for the important pathophysiological process known as cytoadhesion. They showcase antigenic variation where *var* genes show switching to another one upon external cues. At the end of **Chapter 1** we have discussed the discovery and functional aspects of PfEMP1.

In the subsequent Chapters, the main findings are discussed one by one.

In **Chapter 2**, the epigenetic landscape of *var* gene regulation is explored. In this Chapter we are trying to comprehensively understand dynamics of epigenetic modifications at the *var* loci in the light (or darkness) of heterochromatin and the reader protein PfHP1 (*P. falciparum* heterochromatin protein 1). We have harvested the samples for ChIP sequencing and RNA sequencing simultaneously from the 3D7 culture. Also, we raised and validated an antibody against PfHP1 protein. We chose temperature stress as a paradigm to study the effect of environmental factors on epigenetics and gene regulation.

We identify the dynamic change in PfHP1 occupancy in *var* introns, that is correlated to the intronic transcription. PfHP1 could be the central player in the *var* gene regulation. PfHP1 binds to H3K9me3 and is exclusively present on the telomeres and CVGs. Even though it is known that PfHP1 represses *var* genes, the mechanism of gene silencing and heterochromatin formation is

poorly understood. LLPS (liquid-liquid phase separation) is an emerging theme in eukaryotic heterochromatin formation and maintenance. Here, we explore LLPS of PfHP1 as a possible mechanism of gene silencing.

In **Chapter 3,** we test the hypothesis whether PfHP1 can undergo LLPS *in vitro* or not. If yes, then what is the nature of this phase transition and is it a tunable process. The clues from this *in vitro* biochemical study will let us unravel the mechanism of *var* gene silencing.

We have standardized conditions for the PfHP1 purification and droplet formation assay. We have also identified point mutation in the PfHP1 that disrupts the phase separation property.

In **Chapter 4,** we explore the dynamics of PfHP1 mediated DNA compaction. The basic function of HP1 is to condense DNA into highly inaccessible heterochromatin. *P. falciparum* has an unusually AT-rich (~81%) genome. The DNA compaction of such AT-rich genomes is an interesting problem, considering that TA nucleotides have more flexibility that others. Despite that PfHP1 forms stable heterochromatin domains on the high AT rich genome. Also in a recent study, it is shown that highly AT-rich *var* promoters (88%) are the nucleation center for the heterochromatin formation, even though the mechanism is unknown (Pérez-Cantero et al. 2025).

Also we observed in Chapter 3 that PfHP1 forms droplets with highly AT-rich PfDNA. This leads us to the speculation that PfHP1 might have evolved a way to condense the highly AT rich genome of *P. falciparum*, unlike its other eukaryotic counterparts.

We are testing whether PfHP1 will form condensates with DNA in a sequence dependent manner; i.e, by differentiating AT-rich vs. GC-rich DNA sequences. Our observation from Chapter 2 suggests that introns are an interesting paradigm to study the heterochromatin dynamics. Moreover, *var* introns have similar AT-richness as that of *var* promoters. Hence it is possible that *var* introns can also act as nucleation centers similar to *var* promoters.

For this Chapter, we adapted a single molecular DNA tethering strategy where we used a double tethered DNA with a *var* intron at the middle to test how PfHP1 behaves on the PfDNA. We are trying to uncover the molecular basis for the PfHP1 mediated DNA compaction.

We have seen PfHP1 undergoes LLPS and forms co-condensates with DNA in Chapter 3 and 4 respectively. In **Chapter 5,** we are trying to see functional consequences of such unique biochemical and biophysical properties of PfHP1. The point mutations that were identified in Chapter 3 are used here to generate PfHP1 mutants in cell lines. Subsequently we are trying to study the effect of these mutants on chromatin binding and *var* gene expression.

We have utilised three strategies to generate the mutant transgenics;

1) Controlled over-expression lines of PfHP1 mutants: The cell lines will over-express mutant PfHP1 but in the background of wild type PfHP1 from endogenous loci. Helpful to study the dominant negative effect of PfHP1 mutation on *var* gene expression. This approach has a caveat, it will produce the mutant protein in large amounts and may suffer from leaky expression.

2) Gene mutation at the endogenous gene loci of PfHP1: The cells will only express mutant PfHP1 and can study the effect on *var* gene expression in endogenous expression levels of PfHP1. The mutant protein could be lethal to the parasites as PfHP1 is an essential gene. An inducible PfHP1 mutant line will enable us to study the effect of the mutations by expressing PfHP1 for a small window of time.

3) Conditionally expression of GFP tagged mutant PfHP1 from the endogenous loci: Conditional expression of mutant PfHP1 will enable growth of cell lines by only expressing the mutantants upon induction, in case the mutation is toxic to the cells. The GFP tagging will help us to see the changes in localization of PfHP1 upon mutations.

We could successfully generate the controlled over-expression lines (strategy 1) and conditional mutant expressing lines (strategy 3). We observe that these mutations affect PfHP1 localisation and function in the parasites.

In **Chapter 6,** we summarise the major findings of this study and the future directions. There are numerous experiments and hypotheses that revolve around the mechanism of *var* gene switching. Perhaps what is missing is a unifying molecular mechanism that compiles all the existing data

and new findings. In this study, we adopted a comprehensive approach to address the complex and intriguing question of *var* gene regulation.

This study is also one of its kind, because phase separation of proteins is relatively under explored in malaria biology. Due to its rapidly changing dynamics and tunability the phase separation could be a general mechanism adapted by the parasites to help in disease progression.

ABBREVIATIONS

WHO	World Health Organization
RBC	Red Blood Cell
PfHP1	*Plasmodium falciparum* Heterochromatin Protein 1
HP1	Heterochromatin Protein 1
CVG	Clonally Variant Gene
ChIP	Chromatin Immunoprecipitation
LLPS	Liquid-Liquid Phase Separation
RNA	Ribonucleic acid
DNA	Deoxyribonucleic acid
PfDNA	*Plasmodium falciparum* Deoxyribonucleic acid
PfEMP1	*Plasmodium falciparum* Erythrocyte Membrane Protein 1
GFP	Green Fluorescent Protein
IDC	Intraerythrocytic Developmental Cycle
TSS	Transcription Start Site

TTS	Transcription Termination Site
HiC	High-throughput Chromosome Conformation Capture
FISH	Fluorescence In Situ Hybridization
PTM	Post-Translational Modification
Pfmc-2tm	*Plasmodium falciparum* Maurer's cleft-two transmembrane protein
kbp	kilo base pair
bp	base pair
PERC	Perinuclear Repressive Cluster
ncRNA	non-coding RNA
lncRNA	long non-coding RNA
CRISPR-Cas9	Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated protein 9
Pol	Polymerase
TARE	Telomere Associated RNA Element
PCR	Polymerase Chain Reaction
Ni-NTA	Nickel-Nitrilotriacetic Acid

CV	Column Volume
NP-40	Nonidet P-40
DTT	Dithiothreitol
PMSF	Phenylmethanesulfonyl fluoride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
KCl	Potassium Chloride
EDTA	Ethylenediamine tetraacetic acid
EGTA	Ethylene glycol-bis(2-aminoethyl ether)-N,N,N',N'-tetraacetic acid
PVDF	Polyvinylidene Fluoride
PBS	Phosphate Buffered Saline
TEABC	Triethylammonium bicarbonate buffer
NaCl	Sodium Chloride
rpm	Revolutions per minute
tRNA	Transfer ribonucleic acid
hrs	hours

CD	Chromodomain
CSD	Chromo Shadow domain
NTE	N-terminal extension
CTE	C-terminal extension
β-ME	β-Mercaptoethanol
Na_3PO_4	Trisodium phosphate
EMSA	Electrophoretic Mobility Shift Assay
NaOH	Sodium Hydroxide
KOH	Potassium Hydroxide
EtBr	Ethidium Bromide
PEG	Polyethylene Glycol
gRNA	guide RNA
DiCre	Dimerizable Cre recombinase
$MgCl_2$	Magnesium Chloride
K_2HPO_4	Dipotassium hydrogen phosphate

KH_2PO_4	Potassium dihydrogen phosphate
DAPI	4',6-diamidino-2-phenylindole
kDa	kiloDalton
W	Tryptophan
A	Alanine
K	Lysine
E	Glutamate
L	Leucine
D	Aspartate
UTR	Untranslated region
A	Adenine
T	Thymine
G	Guanine
C	Cytosine
U	Uracil

FRAP	Fluorescence Recovery After Photobleaching
NE	Nuclear Extract

Chapter 1: Introduction

1.1 An overview of Malaria

1.1.1 Burden of Malaria

Malaria is a protozoan disease that is mainly spread across the tropical region. Causative agents of human malaria are species of the genus *Plasmodium* that belong to a group of obligate endoparasites called Apicomplexa. *P. falciparum*, being the deadliest in the group, causes the most deaths each year. The emergence of drug resistance in *Plasmodium* species is a major threat to the global efforts to eradicate the disease. A better understanding of the complex life cycle of the parasites and underlying gene regulatory mechanisms could largely help in developing effective strategies to control malaria.

Despite the efforts to control the disease, each year, millions of people get infected. WHO World Malaria Report 2024 emphasizes the same; there have been 263 million cases and 597,000 deaths reported in 2023 due to malaria. Cases from sub-Saharan Africa account for most of the global malaria burden. Since 2020 the number of cases has steadily increased (Fig. 1.1). The contribution from countries such as Pakistan, Nigeria, Madagascar, Ethiopia, and Democratic Republic of Congo has increased during these years. The number of deaths due to malaria has shown a promising trend with a steady decline since 2000 to 2019 (Fig. 1.1). However, in 2020 there has been an increase in the number of deaths mainly due to the disruption caused by COVID-19 pandemic. Since then, deaths have slightly declined (Venkatesan 2025) .

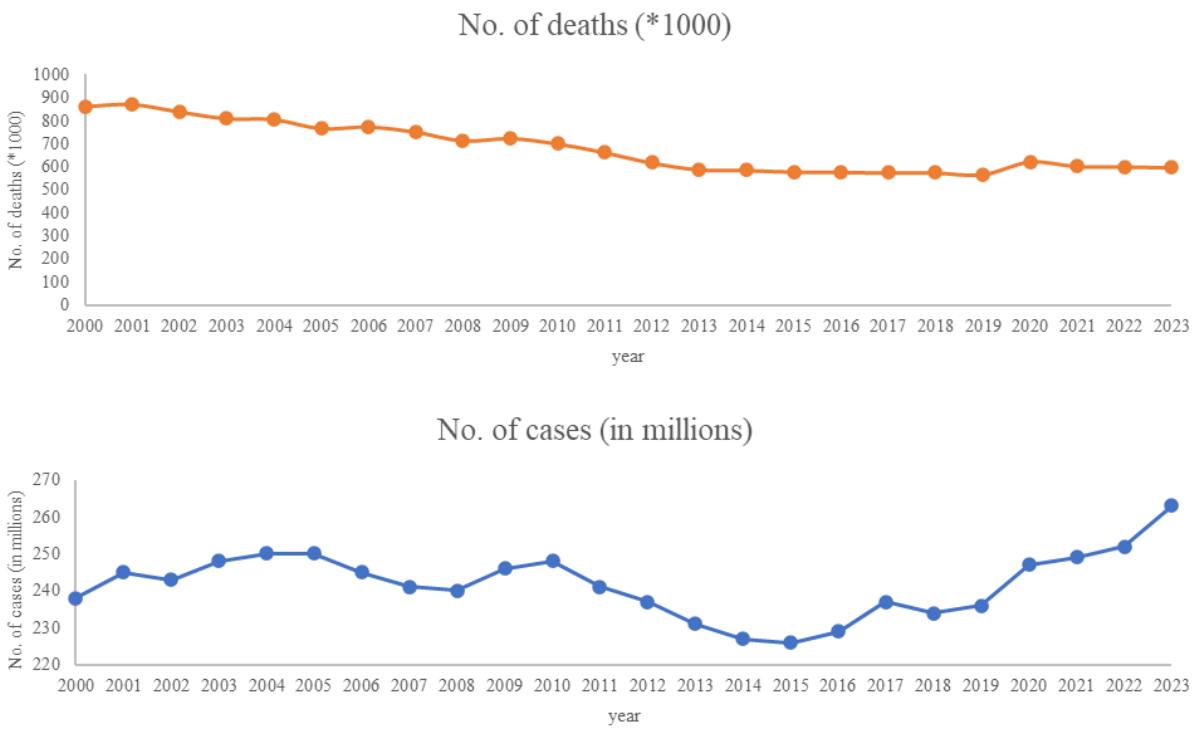


Figure 1.1: Global estimate of malaria cases and deaths since 2000, data plotted from World Malaria Report, 2024.

1.1.2 History of Malaria

Malaria has plagued humanity for centuries, claiming countless lives and leaving a significant mark on human history. According to Chinese medical literature, one of the earliest documentations of malaria was in 2700 BC. In Ancient Greece, it was believed that malaria spread was associated with marshy areas. They propagated the miasmatic theory of disease spread, in which miasma or fumes from swampy water cause malaria. However, the origin of the term 'malaria' is Italian, mal-aria meaning 'bad air' (Hempelmann & Krafts 2013).

With the discovery of bacteria as a cause of infectious disease, the search for the causative agent of malaria intensified. In 1880 Charles Louis Alphonse Laveran discovered malaria parasites by identification of various parasitic forms inside the RBCs from the blood smear of the patients (Fig. 1.2). He was awarded the Nobel Prize in Physiology in 1907 for the discovery of the causative agent of malaria. Later, in 1897, Sir Ronald Ross established that malaria is a vector-

borne disease, the vectors being the mosquitoes. Independently, Giovanni Batista Grassi, in 1898, also identified anopheline mosquitoes to be carrying malaria parasites (Cox 2010). The understanding of the life cycle of the malaria parasites has aided largely in developing strategies to combat the disease.

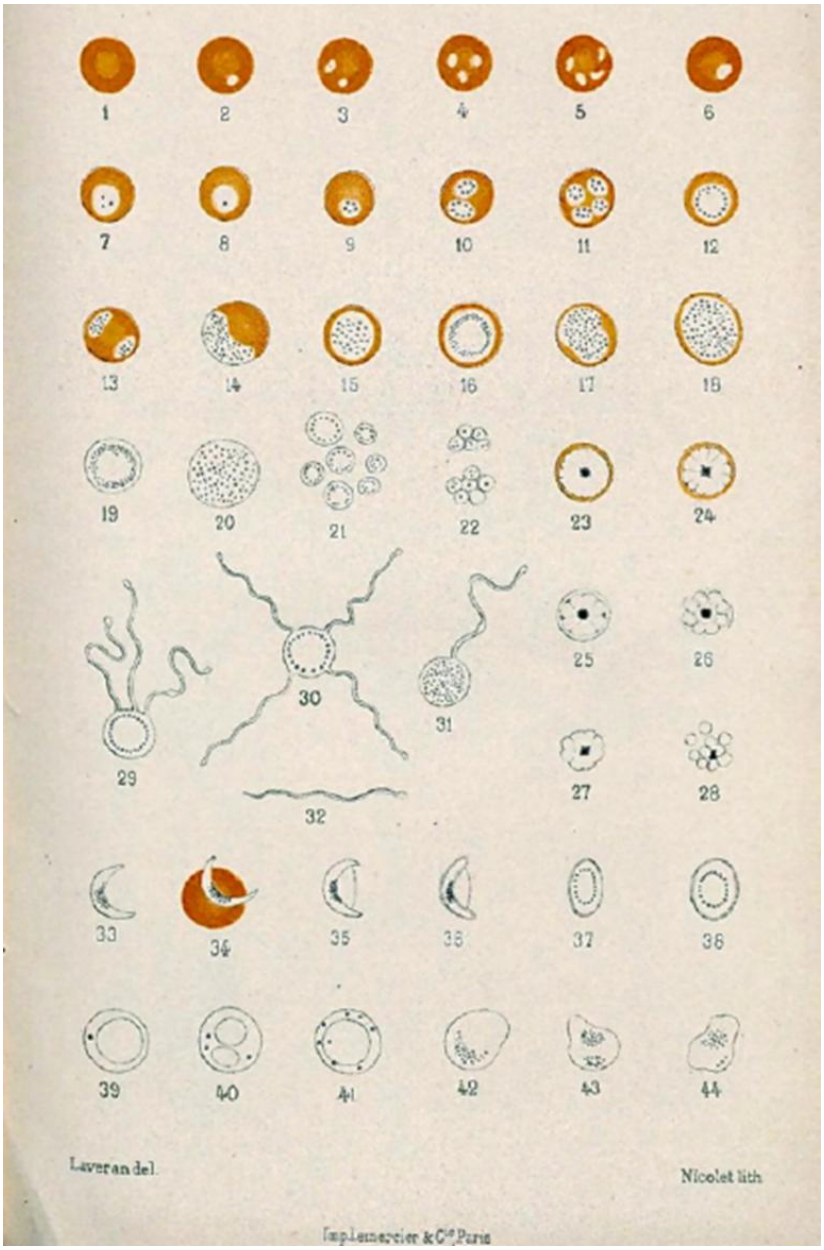


Figure 1.2: Sketches made by Laveran of different stages of malaria parasites, image taken from (Hempelmann & Krafts 2013).

1.1.3 Evolution of malaria in primates

Malaria has been reported in birds, reptiles and mammals including humans. There are five species of *Plasmodium* that infect humans: *Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, Plasmodium malariae*, and *Plasmodium knowlesi.* While *P. falciparum* is the deadliest among the species, *P. vivax* has a wider geographical distribution and is the most prevalent in countries outside sub-Saharan Africa (World Health Organization(*Malaria*)).

The origin of malaria infection in humans has been of great interest for decades. The common ancestors of humans and chimpanzees might have been first infected by malaria. The first study on the identification of *Plasmodium spp.* infecting apes was led by Reichenow in 1920 (Reichenow). Early studies identified three major species that cause malaria in apes: *P. reichenowi, P. rhodaini* and *P. schwetzi* (Brumpt; Sluiter 1912). With the introduction of molecular genetics techniques, rRNA sequencing could be used to study phylogeny. It was identified that *P. reichenowi* and *P. falciparum* are very closely related (Escalante & Ayala 1994). In addition, another parasite that is similar to *P. reichenowi* was identified in two chimpanzees from Gabon named *P. gaboni* (Ollomo et al. 2009). Phylogenetically, *P. falciparum, P. reinchenowi* and *P. gaboni* are closely related and belong to a superclass called Laverania. Whereas, *P. ovale , P. vivax, P. malariae* and *P. knowlesi* all belonged to different clades (Carter & Mendis 2002; Loy et al. 2017; Rich & Ayala 2006). The *P. falciparum* clade is more closely related to avian malaria species than any of the human-infecting species (Loy et al. 2017). The current understanding is that the five human infecting species are phylogenetically independent in their transmission into a human host.

The infection has caused dramatic changes in the human genome (Price 2017). The sickle cell trait that is associated with malaria is one such example (Beet 1946). Mutations in the Duffy antigen receptor on the surface of RBCs confer resistance to *P. vivax i*nfection in the sub-Saharan African population (Zimmerman et al. 1999).

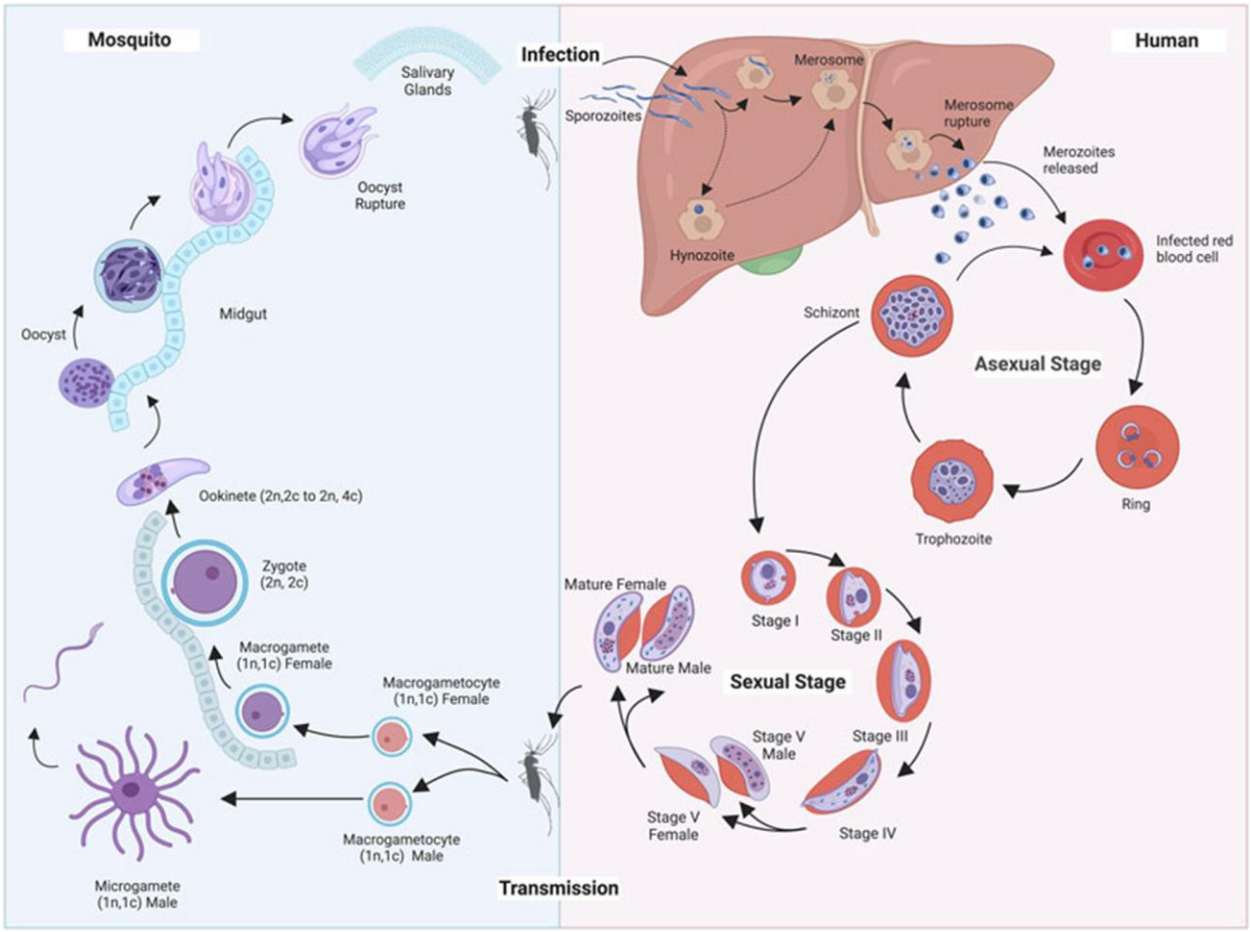


Figure 1.3: Life cycle of *Plasmodium*, alternates between a mosquito and human host, image taken from (Chahine & Le Roch 2022).

1.1.4 Life cycle of P. falciparum

The life cycle of malaria parasites is very complex and is spread across a vertebrate and invertebrate host. *Plasmodium* is transmitted to humans by bite from female *Anopheles* mosquitoes. While the mosquito takes a blood meal, the haploid infective stage of *Plasmodium* known as sporozoites gets into the circulation. Sporozoites divide inside the liver giving rise to schizonts. Schizonts burst out of hepatocytes and release merozoites into the blood. Merozoites invade the RBC, starting a next replication cycle known as intra erythrocytic development cycle (IDC). A typical IDC lasts for 48 hrs in *P. falciparum, P. ovale* and *P. vivax* infections and 72 hrs

in *P. malariae* infection. IDC has three morphologically distinct stages: ring, trophozoite and schizont (Fig. 1.3). At the end of each IDC the infected RBCs burst and release more merozoites for the next round of invasion. Also, a fraction of parasites differentiates into the gametocytes. When a mosquito bites an infected human being, these gametocytes get into the midgut of the mosquito. The male gametocytes undergo exflagellation whereas female gametocytes give rise to macrogametes; these later fertilize to form a zygote. Then it develops into ookinetes and penetrates mosquito midgut to form oocyst. The oocysts self-propagate inside the mosquito gut, haemolymph and salivary glands to produce sporozoites for the next round of infection (Tuteja 2007).

1.1.5 Treatment and control of malaria

1.1.5.1 Vector control

Vector control and anti-malarial drugs are major strategies to combat malaria. Among the available vector management tools, insecticides and environmental management (e.g. bed nets) have controlled the disease burden significantly (Wang et al. 2021b). One notable example is the insecticide DDT (dichloro-diphenyl-trichloro-ethane), which helped during the period following World War II but eventually faced a significant setback due to the emergence of insecticide-resistant mosquitoes (Wang et al. 2019a). Another disadvantage of such strategies is that they cause ecosystem imbalance by un-targetedly killing other insects, including pollinators.

The recently developed genetic control technologies in vector control are sustainable and promising. Gene drives (GD) are one such technique, where GDs are selfish genetic elements that can bias the inheritance in its favor. In this approach, the GDs increase the probability of inheritance of a transgene from 50% (normal mendelian inheritance) to greater than 50%, in some cases even 90% (Fig. 1.4A). Often, fitness related genes are targeted which can reduce their propagation and eventual death of the mosquitoes (Wang et al. 2021b).

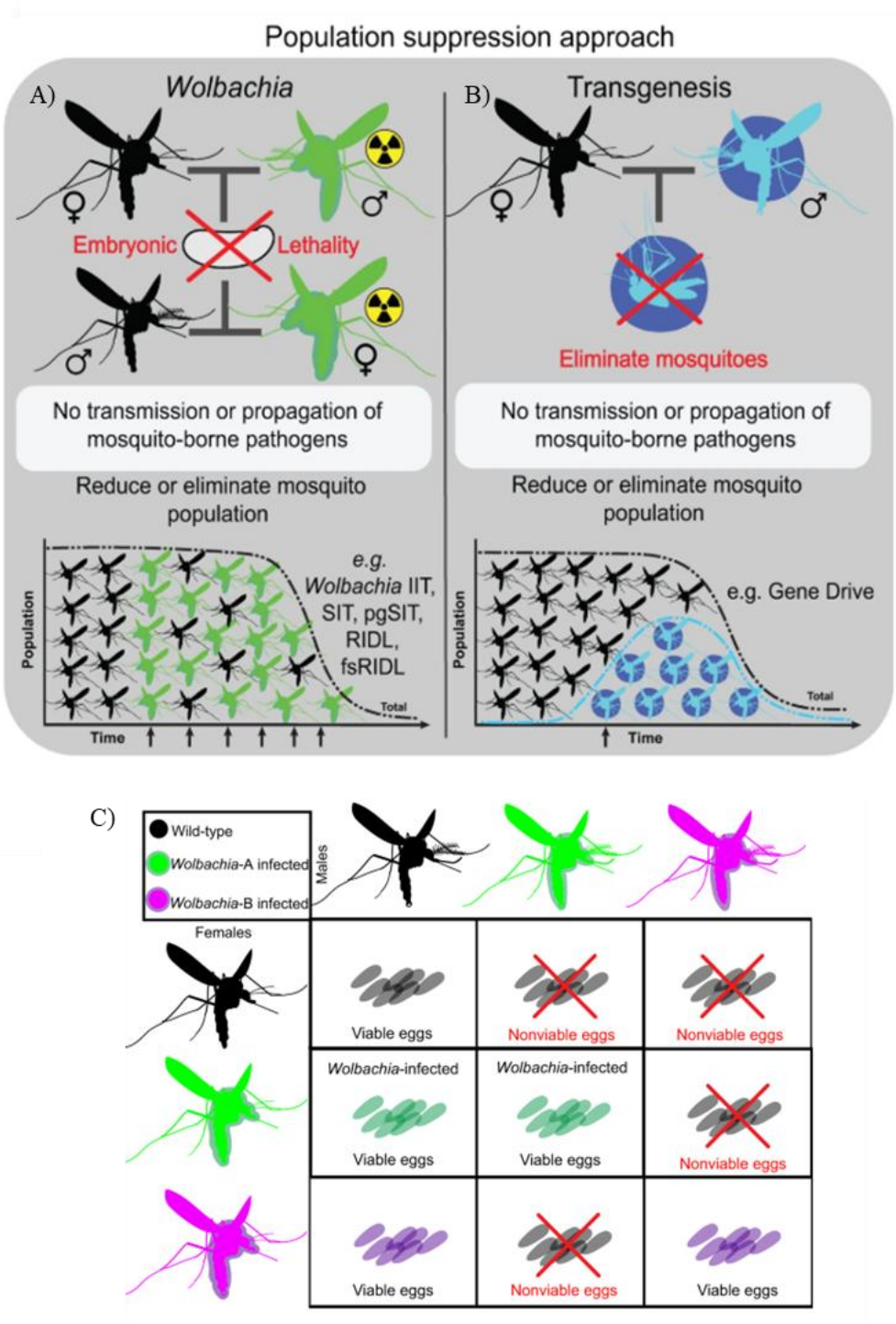


Figure 1.4: Vector control approaches based on A) population suppression by *Wolbachia*, B) Population suppression by GDs, C) Cytoplasmic incompatibility in *Wolbachia* infected mosquitoes; images taken from (Wang et al. 2021b).

Wolbachia-based approaches for mosquito control have also gained popularity in recent years. *Wolbachia* is an intracellular bacterium that infects arthropods. *Wolbachia* can reduce disease transmission in two ways 1) population suppression and 2) population modification. In population suppression strategy, because of the cytoplasmic incompatibility (CI) the *Wolbachia* infected males can only produce viable offspring, if mated with the female with the same *Wolbachia* strain (Fig. 1.4B & C). Hence, for the control of mosquitoes, males with *Wolbachia* are introduced into nature where the population lacks *Wolbachia* (Crawford et al. 2020) . The second strategy, population modification, is based on the bacteria's remarkable ability to reduce the transmission of various pathogens, including *Plasmodium (Bian et al. 2013)*. Hence, field trials are in progress where *Wolbachia* infected mosquitoes are released into the wild. It has been successful for *Aedes* mosquitoes (O'Neill et al. 2018). Recent discovery of presence of *Wolbachia* in *Anopheles* species, opens a new path for malaria control (Wong et al. 2020).

1.1.5.2 Antimalarial compounds to combat malaria

Natural remedies for malaria have been popular since historic times, one notable example being quinine derivatives from the cinchona tree. Later, this would give rise to synthetic drugs like chloroquine and sulfadoxine-pyrimethamine (Alghamdi et al. 2024). The emergence of chloroquine resistance in the early 1950s alarmed the scientific community for innovations to combat malaria. In 1969, a group led by Prof. Youyou Tu started a wide screen for anti-malarial chemicals from traditional herbal medicines. They successfully established artemisinin, a compound derived from Qinghao plant extract, for its antimalarial properties. The introduction of Artemisinin combination therapy (ACT) made a significant breakthrough in reducing malaria-related mortality (Wang et al. 2019a). However, the first case of Artemisinin resistance was reported in 2008 from Cambodia, leading to a considerable setback to global efforts to eradicate malaria (Noedl et al. 2008) (Dondorp et al. 2009) (Fig. 1.5). With the increasing failure of ACTs, new approaches were introduced, including triple Artemisinin combination therapy (TACT), which combines two more partner drugs with the existing ACT (van der Pluijm et al. 2020).

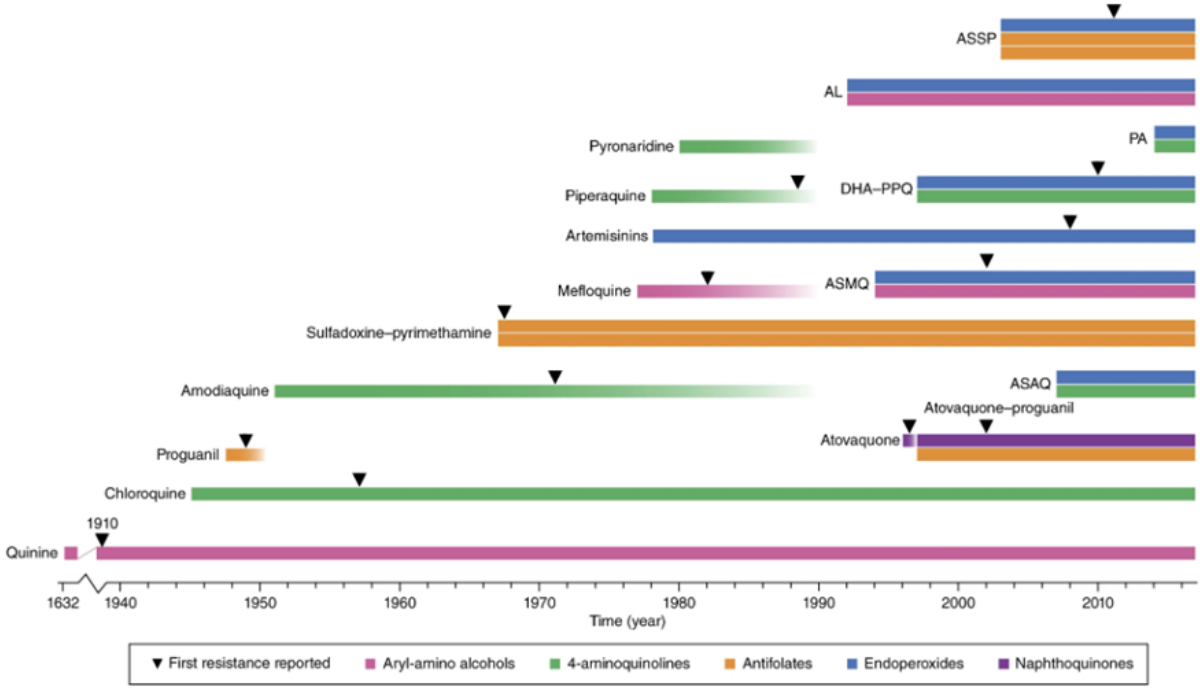


Figure 1.5: Artemisinins (ARTs) were originally administered as standalone treatments, and injectable artesunate remains in use for severe malaria. However, because of their brief presence in the bloodstream and rising resistance, artemisinin-based combination therapies (ACTs) were introduced to treat uncomplicated malaria. Some ACT partner drugs, such as amodiaquine and mefloquine, had already been used as monotherapies and continued to be used individually even after resistance was detected. Piperaquine and pyronaridine, brought into use in China approximately four decades ago as substitutes for chloroquine (CQ), also faced resistance, with piperaquine resistance emerging by the late 1980s (AL, artemether + lumefantrine; ASMQ, artesunate + mefloquine; ASAQ, artesunate + amodiaquine; DHA-PPQ, dihydroartemisinin + piperaquine; PA, artesunate + pyronaridine; ASSP, artesunate + SP); image taken from (Blasco et al. 2017).

1.1.5.3 Vaccination against malaria

Despite the global efforts to develop a vaccine against malaria, we still lack a high efficacy vaccine. Major challenges in vaccine development are the complex life cycle of the parasites, extensive antigenic variation in species like *P. falciparum* and poor understanding of immune reaction against the infection (Crompton et al. 2010). Various studies have shown that *P. falciparum* confers acquired immunity in the host (Ciuca et al. 1934; Doolan et al. 2009).

However, the exact mechanism of naturally acquired immunity is not well understood. Also, even individuals with multiple exposures to the disease don't develop sterile immunity against blood-stages (Crompton et al. 2010).

PfEMP1 (*P. falciparum* erythrocyte membrane protein 1) is a potential candidate for vaccine because studies have shown evidence for acquired humoral immunity against PfEMP1 (Obeng-Adjei et al. 2020; Rambhatla et al. 2019). Recent evidence suggests broadly reactive antibodies are present in patients against the CIDRα domain of PfEMP1 (Reyes et al. 2024). There are other blood-stage antigens against which vaccine development is being attempted, and this includes apical membrane antigen 1 (AMA1), erythrocyte-binding antigen–175 (EBA-175), glutamate-rich protein (GLURP), merozoite surface protein 1 (MSP1), MSP2, MSP3, and serine-repeat antigen 5 (SERA5) (Esen et al. 2009; Hermsen et al. 2007; Sagara et al. 2009;Genton et al. 2002; Horii et al. 2010).

Pre-erythrocytic stages were not targeted for vaccine development because acquiring immunity against the stage couldn't be observed. Later, it was demonstrated that it's possible to induce immunity through multiple sporozoite infections (Roestenberg et al. 2009). In 1985, a British Pharmaceutical company, GlaxoSmithKline started developing a vaccine using circumsporozoite protein (CSP). After several clinical trials in 2021 this vaccine named RTS,S was approved by WHO as the first-ever vaccine against malaria. It showed about 39% efficacy in the phase III clinical trial. Despite the moderate efficacy, it helped reduce the disease burden. It is recommended to use RTS,S to immunize children from the endemic regions. An improved version of RTS,S, called R21is the second vaccine to be prequalified by WHO as of 2023 (Duffy et al. 2024; RTS,S Clinical Trials Partnership 2015; Venkatesan 2025).

Transmission-blocking vaccines (TBV) are also being attempted, where antigens from sexual stages are used to prevent transmission to mosquitoes. This type of vaccine doesn't give immunity to the individual taking it; rather confers immunity against the sexual stages hence blocking the spread of the disease further. Gametocyte proteins such as Pfs48/45 and Pfs23 that belong to a superfamily with six-cysteine domains are mainly targeted in TBVs (Chowdhury et al. 2009; Quakyi et al. 1987).

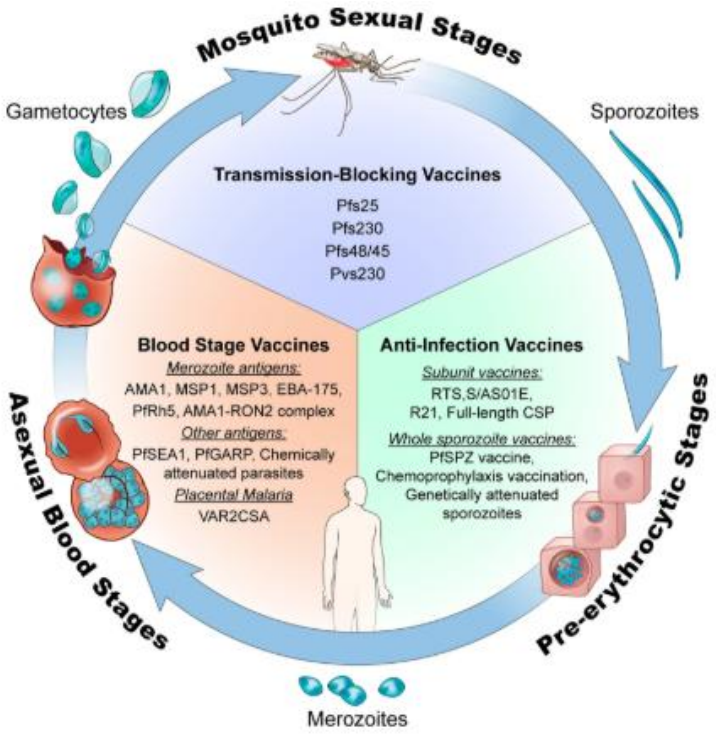


Figure 1.6 : Vaccine candidates that target different life stages of *P. falciparum*, image taken from (Duffy & Patrick Gorres 2020).

1.2 Transcription regulation in *P. falciparum*

1.2.1 Stage-specific transcription

IDC of *P. falciparum* has been well studied for the transcription and the regulatory circuit; considering that IDC shows clinical manifestation of the disease. Overall, 60% of the genes are active during the IDC (Hollin & Le Roch 2020). The phenotypically distinct ring, trophozoite and schizont stages, shows a distinct transcriptome profile. A complex interplay of epigenetic mechanisms gives rise to 'just-in-time' transcription, wherein a gene is only expressed once in the IDC (Fig. 1.7) (Bozdech et al. 2003).

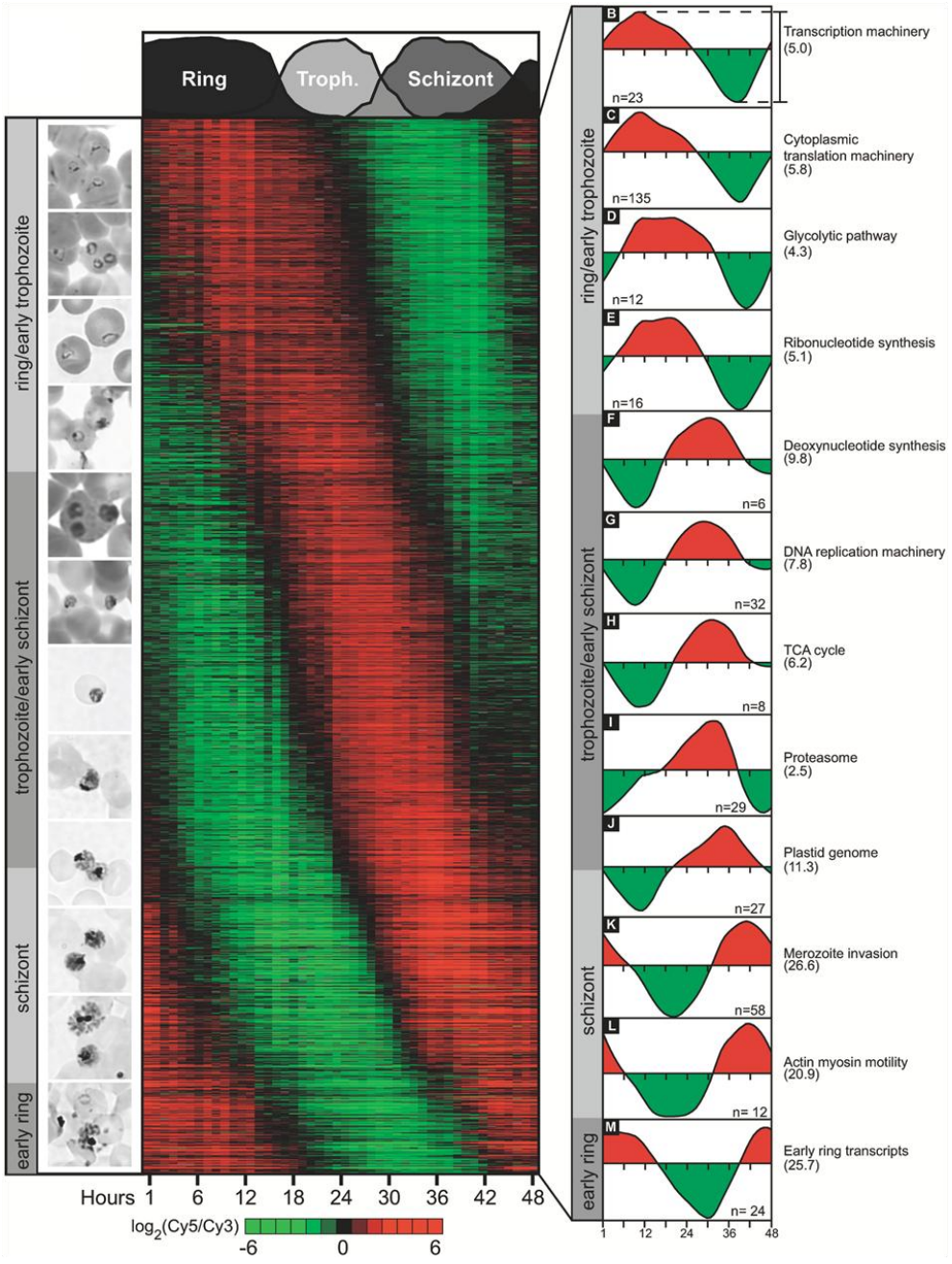


Figure 1.7: 'just-in-time' transcription during the stages of IDC, image taken from (Bozdech et al. 2003).

Ring stage genome is very compact, attributed to the ring shape and volume of the nucleus. Hence the stage shows minimal transcriptional output. Despite this , genes related to antigenic variation

such as *var* are mostly expressed in this stage (Bártfai et al. 2010; Scherf et al. 1998). Trophozoite stage has a more open chromatin and larger nuclear volume. This is positively manifested in the global transcriptional activity (Bozdech et al. 2003). There is a notable escalation in the number of nuclear pores during the stage attributing to the RNA export (Weiner et al. 2011). Schizogony is marked by repacking the genome to a compact stage and hence, global transcription is reduced. Invasion and merozoite surface proteins are highly expressed during this stage (Bozdech et al. 2003; Le Roch et al. 2003).

The gametocytic stages have distinct transcriptome and nuclear architecture (Yeoh et al. 2017) (Bunnik et al. 2018; Kengne-Ouafo et al. 2023). However the switch to the gametocytic stages occurs during the blood stages (Filarsky et al. 2018). The associated regulatory scheme is discussed in section 1.3.4.

1.2.2 Nucleosome landscape in P. falciparum

Nucleosome positioning is believed to be a DNA sequence-driven process, where 10 bp periodicity of AA/TT/AT/TA dinucleotides helps in nucleosome wrapping. Extreme AT richness (81%) of *P. falciparum* genome might hamper the nucleosome formation but MNase digestion of the genome followed by sequencing has indicated that there isn't a bias in nucleosome occupancy between highly AT rich intergenic regions and relative low AT rich coding regions. Also, despite the highly AT-rich genome, the 10 bp, AA/TT periodicity-driven rotational positioning is conserved in *P. falciparum* (Kensche et al. 2016).

Another signature of eukaryotic promoters is nucleosome depleted region (NDR) in the transcription start site (TSS). NDRs are essential for transcription initiation, as they help form the pre-initiation complex. The MNase seq of *P. falciparum* chromatin revealed a similar NDR at the TSS of most of the genes. However a strong positioning of NDRs at +1 is lacking in *P. falciparum* unlike the other model organisms studied.

The canonical histones of *P. falciparum* show a divergence from human histones with identities of only 64%, 68%, 93% and 92% for H2A, H2B, H3 and H4, respectively. This indicates that histone octamers might have evolved to meet the specific needs of the organism. In particular,

H2A and H2B histones wrap around DNA with much weaker interaction, making the *P. falciparum* nucleosomes looser and more open (Silberhorn et al. 2016). Another peculiar feature is the absence of H1 linker histone in *P. falciparum*, and the average linker length is 8 bp; much shorter than higher eukaryotes (Silberhorn et al. 2016).

In addition, there are specific histone variants that occupy highly AT-rich regions. The classic DNA damage associated H2A.X histone variant is absent in *P. falciparum*. The homolog of H2A.Z has been identified along with a new variant that is H2B.Z. The H2A.Z/H2B.Z double histone variant occupies most of the intergenic regions (86 % AT). There is a strong correlation between the AT content of the genome and H2A.Z/H2B.Z occupancy. The CenH3 histone variant that helps form centromeres, wraps the highly AT-rich (97%) DNA around the centromeres (Bartfai et al. 2010; Hoeijmakers et al. 2013; Petter et al. 2013; Hoeijmakers et al. 2012).

Another histone variant, H3.3 occupies relatively GC-rich coding regions and subtelomeric regions (Fraschka et al. 2016). Specifically coding sequence and promoters of active *var* genes. Heterochromatin domains are also formed around low AT-rich (73% AT) coding regions of clonally variant genes (CVGs) and are composed of canonical histones. It is also observed to have a relatively higher nucleosome density in heterochromatin (Hoeijmakers et al. 2012).

This indicates that *P. falciparum* has evolved histone variants that might help in wrapping of the extremely AT-rich DNA. The nucleosomes help in indexing the genome and defining chromatin structure overall (Hoeijmakers et al. 2012).

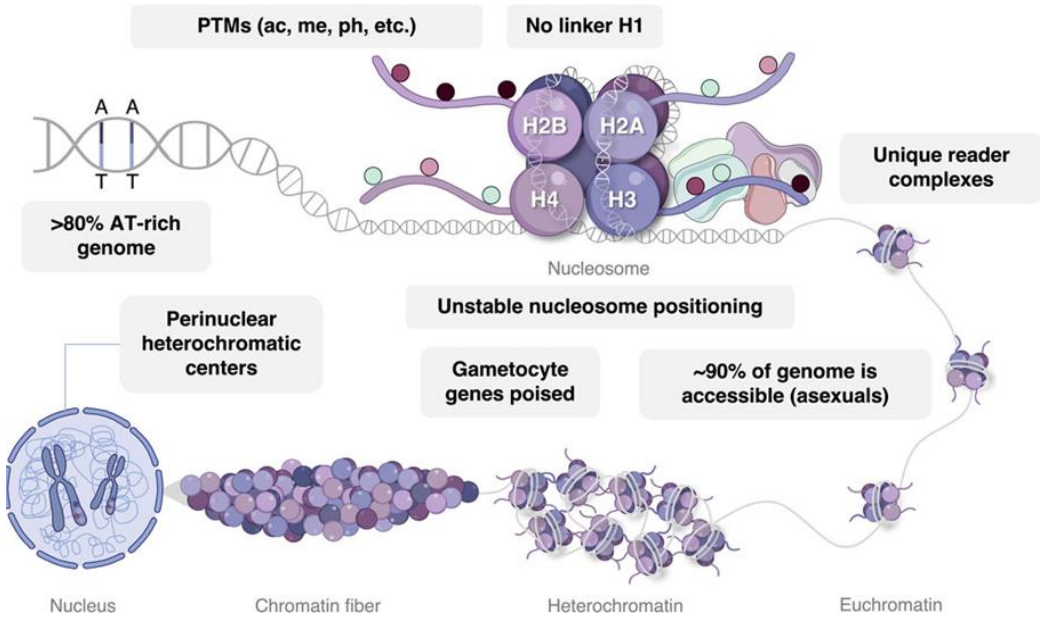


Figure 1.8: Epigenetic regulation of transcription in *P. falciparum*, image taken from (Connacher et al. 2022).

1.2.3 Chromatin organization and nuclear architecture

The nucleosomes are organised into condensed forms along with other proteins called chromatin. Heterochromatin is the transcriptionally repressed part of the genome that is tightly packed. The term was first coined by Emil Heitz when he observed darkly stained regions in chromosomes while preparing cytological samples (Heitz 1928). In higher eukaryotes, heterochromatin is further categorised into facultative and constitutive heterochromatin. As the names suggest, facultative heterochromatin genes are kept silent based on the cell signals, whereas constitutive heterochromatin is always repressed and contains mostly repetitive DNA that is gene-poor. The H3K9me3 histone modification usually flags the constitutive heterochromatin, and heterochromatin protein 1 (HP1) helps in tight packaging. H3K27me3 marks are enriched in facultative heterochromatin (Morrison & Thakur 2021).

In *P. falciparum*, even though H3K9me3-associated heterochromatin was discovered, facultative heterochromatin and associated H3K27me3 could never be found. Interestingly, the heterochromatic regions harboured many genes, particularly those of the CVG family (Flueck et

al. 2009; Lopez-Rubio et al. 2009). The gene regulation of the CVGs are addressed in detail in the antigenic variation section 1.2.

Like its higher eukaryotic counterparts, *P. falciparum* heterochromatin was bound by PfHP1 (*P. falciparum* HP1). The heterochromatin is mostly positioned at the telomeric and subtelomeric regions except for a few internal islands of heterochromatin in some chromosomes (Flueck et al. 2009; Lopez-Rubio et al. 2009).

Euchromatin is the region of chromatin that is less condensed, gene-rich, and actively transcribed. In *P. falciparum*, most of the genes are open and euchromatic. Over 90% of the genome is accessible during the asexual stages. The classic marks of activation such as H3K9ac and H3K4me3, are present in the *P. falciparum* euchromatin (Bártfai et al. 2010).

Heterochromatin has the ability to spread in a sequence-independent manner. One classic example is position effect variegation (PEV) in *Drosophila melanogaster,* where the rearrangement of euchromatic genes in proximity to heterochromatin leads to the inactivation of those genes due to the heterochromatin spread (Muller 1930). In *P. falciparum*, heterochromatin is tightly regulated, and the mechanism of formation and spread has not been explored well.

Using HiC sequencing along with FISH (Fluorescence in-situ hybridisation) studies, it is found that the heterochromatin forms a repressive cluster around the nuclear periphery and the centromeres form a cluster at the radially opposite end (Ay et al. 2014; Bunnik et al. 2018; Freitas-Junior et al. 2005; Lopez-Rubio et al. 2009). Also *Plasmodium spp* lack a topologically associated domain (TAD) probably due to the small genome size and minimal proteome for higher order genome organisation (Bunnik et al. 2018;Ay et al. 2014).

1.2.4 Histone modifications and epigenetic regulators

The N-terminal tail of histones often harbors post-translational modifications (PTMs) that have gene regulatory functions and decide chromatin states. The methylation, acetylation, phosphorylation, sumoylation and ubiquitination are the most studied histone PTMs. In *P.*

falciparum, using the proteomics approach, 232 histone PTMs were identified. In addition to canonical PTMs there were 88 PTMs identified exclusively in *P. falciparum* (Saraf et al. 2016).

Heterochromatin in *P. falciparum* is marked by H3K9me3 and bound by heterochromatinizing protein PfHP1 (Flueck et al. 2009; Pérez-Toledo et al. 2009). However, the facultative heterochromatin and its corresponding modification H3K27me3 are not robust in *P. falciparum.* Hence, the existence of a global repressive mark in *P. falciparum* is still under investigation. One such potential mark could be H3K36me2, as it shows a global negative correlation with the gene expression (Karmodiya et al. 2015). DNA methylations are also associated with gene repression in mammalians, but the presence of DNA methylation has been long debated in *Plasmodium*. The general consensus is that there are very low levels of 5mC present in *Plasmodium*, far less in levels as compared to mammals and birds. The hallmark of transcription elongation in mammalian systems, H3K36me3, has a different function in the parasites, where it enriches in the heterochromatin domain along with H3K9me3 (Jiang et al. 2013). Activating histone modifications such as H3K4me3 and H3K9ac are present in the promoter region upstream of the TSS (Transcription start site) of the active genes (Cui et al. 2008; Gupta et al. 2017; Salcedo-Amaya et al. 2009). In addition, H3K14ac, H4K8ac, and H4K20me1 have been shown to have activating roles (Cui et al. 2008; Gupta et al. 2017; Salcedo-Amaya et al. 2009; von Grüning et al. 2022). The H3K18ac and H3K27ac, along with the histone variant H2A.Z, dynamically mark gene promoters, colocalising with transcription factor AP2-I and BDP1 (Bromodomain protein 1) (Tang et al. 2020). Arginine methylation H3R17 and H3R40 are also abundant in blood stages.

Histone methyltransferases (HMT) are the enzymes that deposit methylation marks on histones. In *P. falciparum* ten HMTs are identified that have enzymatic domain SET (Su (var)3-9, Enhancer of Zeste and Trithorax). Product specificities of most of the PfSET proteins are ambiguous. PfSET8 has been shown to have methyltransferase activity towards H4K20 (Chen et al. 2016; Cui et al. 2008). PfSET10 is among one of the better-studied HMTs. It is shown to be associated with antigenic variation genes and H3K4 methylation (Volz et al. 2012). However, contrasting results have been reported where PfSET10 was not essential for regulating antigenic variation (Wyss et al. 2024; Ngwa et al. 2021). Another independent study has shown that PfSET10 knock-out results in the loss of H3K18 methylation but not H3K4 methylation (Musabyimana et al.

2024). PfSET7 methylate H3K4 and H3K9 with a preference for pre-existing H3K14ac on the substrate (Chen et al. 2016). PfSET2/ PfSETvs have also been discussed in the context of antigenic variation (Jiang et al. 2013). Knock-out of SET2 affects the levels of H3K36me3. Other studied HMTs are SET3, which methylates H3K9 and PfSET4/PfSET5, which methylates H3K64 (Jabeena et al. 2021).

The **histone demethylases** majorly removes methyl functional groups from the histone tails. They are sub-categorised into two; LSD1 (Lysine specific demethylase 1) and Jumonji (Jmjc) demethylases. The LSD1 class usually enzymatically removes mon/di methyl groups whereas, PfJmjc1, PfJmjc2 and PfJmjc3 are for removing di/tri methyl groups. Current understanding of these enzymes is incomplete.

Histone acetyltransferases (HATs) function by acetylating lysine residues on histone proteins, leading to chromatin remodeling and transcriptional activation. Histone acetylation is a highly dynamic process, hence antagonising enzymes HATs and HDACs (**Histone deacetylases**) play an important role. PfGCN5 (*P. falciparum* general control non-depressible 5) and PfMYST are two of the well studied HATs that acetylate H3 and H4 respectively. PfGCN5 *in-vitro* acetylates H3K9 as well as H3K14 (Fan et al. 2004). It's role in stress regulation including artemisinin drug resistance. Proteolytic cleavage of the enzyme is imperative to its activity (Bhowmick et al. 2020; Rawat et al. 2021a). PfMYST has a role in antigenic variation, cell cycle progression and DNA damage response (Miao et al. 2010).

There are five PfHDACs identified, three belonging to class I & II along with two sirtuins ; Sir2A and Sir2B (Kanyal et al. 2018). PfHDAC1 has a role in asexual proliferation and artemisinin stress response genes (Kanyal et al. 2024). PfSir2A and PfSir2B, both have a role in regulation of antigenic variation (Freitas-Junior et al. 2005; Tonkin et al. 2009). While most of these histone modifying enzymes are under-studied; demonstration of chemical inhibition of the enzymes makes them an amenable therapeutic target (Reyser et al. 2024).

There is relatively paucity of specific **transcription factors (TF)** in *Plasmodium spp.* However, there is a family of transcription factors identified that show similarity to Apetala2 (AP2) TFs

from the plants. In *P. falciparum,* 27 of these are present and collectively called ApiAP2 (Apicomplexan AP2) (Balaji et al. 2005).

AP2 TFs bind to specific DNA motifs and are temporally expressed at different stages of IDC and gametocytogenesis following the pattern of just in time transcription. The AP2 TFs have varied target genes and functions. The co-immunoprecipitation followed by LC-MS/MS finds that AP2-HC is associated with heterochromatin (Carrington et al. 2021). AP2-EXP, another factor, has a role in expression of antigenic variation genes (Martins et al. 2017). AP2-I was implicated in regulation of invasion genes and is present on these gene promoters (Santos et al. 2017). Recently, a putative TF, AP2-P was shown to have a role in host cell remodelling, antigenic variation, egress, invasion and gametocytogenesis (Subudhi et al. 2023). Unlike other AP2 proteins, PfSIP2 has a role in heterochromatin formation and genome integrity rather than gene regulation. It binds to SPE2 motifs on subtelomeric regions along with PfHP1 (Flueck et al. 2010). In a genome wide screen to identify heterochromatin associated TFs, 8 ApiAP2 factors were identified (Shang et al. 2022).

Gametocytogenesis has long been established to be an essential part of the *Plasmodium* life cycle. But the molecular mechanism of gametocytogenesis was unknown, until in the previous decade it was discovered that ApiAP2 factors are at the core of this regulation. AP2-G was responsible for the commitment to gametocytogenesis along with AP2-G2 which is part of a cascade (Kafsack et al. 2014; Sinha et al. 2014). Interestingly, AP2-G was one of the PfHP1 targets and was heterochromatinized throughout IDC (Brancucci et al. 2014). Only during the sexual conversion, another protein, GDV1 (gametocyte development 1) causes the eviction of PfHP1 from AP2-G loci and hence triggers its activation. The expression of GDV1 is regulated by its own antisense RNA (Filarsky et al. 2018). *P. falciparum* has a huge repertoire of ncRNAs (non-coding RNAs), that are shown to have gene regulatory roles; most studied in the context of antigenic variation.

1.3 Antigenic variation

1.3.1 History

Bacteria, fungi and many other pathogens including *Plasmodium spp.* have evolved an immune evasion strategy known as antigenic variation where the pathogens express variant antigens in a temporal fashion to fool the host immune system. The first indication of antigenic variation in *Plasmodium* species was provided by Cox et al, in 1959. It is stated that the relapse of *P. berghei* in the mouse could be because of the antigenic variation ability of the parasites (Cox 1959). The first evidence for the antigenic variation is provided by Brown and Brown in 1965. They hypothesize that chronicity of *Plasmodium* species could be because they are i) poorly immunogenic ii) intracellular and hence not accessible. Experiments with the Nuri strain of *P. knowlesi*, a highly virulent strain found in rhesus monkeys, proved that the above hypotheses are wrong, as antigenic variation is the main reason for chronicity. K1, a frozen stabilate of the Nuri strain and its relapse stabilates K1A , K1B, K19, and K21 were used in the schizont infected cell agglutination (SICA) test (Brown & Brown 1965). In 1938, Eaton reported that RBCs with schizonts agglutinate in presence of immune sera from monkeys immune to *P. knowlesi* (Eaton 1938). Antigens on the surface of RBCs from a specific stabilate reacted only with antisera from the same stabilate. This shows that during each relapse of the infection different antigens are expressed on an infected RBC surface. And the immune response is elicited against the antigens by the production of specific antibodies (Brown & Brown 1965). Later another report showed that antigenic variation in *P. knowlesi* was absent in splenectomized monkeys. This shows the possible role of the host microenvironment on SICA virulent gene expression in parasites (Brown et al. 1968). Stage specificity of variant antigenic variation was also first shown in *P. knowlesi*. An IFA test showed that only schizont infected erythrocytes showed antigenic variation and not merozoites. The important implication of this experiment is that merozoites can't escape immune clearance by this mechanism (Hommel & David 1981). In 1982, McLean et al. showed antigenic variation in *P. chabaudi*. Immune serum from mice infected with *P. chabaudi* is transferred to mice infected with the recrudescent population and the original population. The immune serum delayed patient parasitemia reaching 2% in the original infecting population as compared to the recrudescent population. This shows that the recrudescent population has undergone antigenic

variation (McLean et al. 1982). *P. fragile* also shows antigenic variation like *P. knowlesi* and *P. chabuadi*. Surface immunofluorescence test on the RBC showed no fluorescence in splenectomized animals infected with *P. fragile* and showed fluorescence in non-splenectomized animals (Handunnetti et al. 1987).

The first report on the antigenic variation of *P. falciparum* came in 1983. Reactivity of sera taken at different stages of infection was analyzed against the original and recrudescent population of Indo-china strain of *P. falciparum*. The effect of spleen was also tested, using splenectomized squirrel monkey. *P. falciparum* from the splenectomized animals was transferred to normal animals, resulting in antigenic switching (Hommel et al. 1983). Leech et al. 1984 identified 285 kDa and 260 kDa proteins on the surface of erythrocytes infected with two different strains of *P. falciparum*. These proteins got immunoprecipitated with immune sera with a striking similarity to proteins coded by PfEMP1. It is also indicated in the paper that these antigens are of parasitic origin rather than host origin (Leech et al. 1984). Three papers published in the same issue of the Cell in 1995 showed that *var* genes are responsible for the antigenic variation on the surface of RBCs infected with *P. falciparum*. The *var* genes is the collective term used for the family of genes encoding large proteins (200-350 kDa) that are antigenic in nature (Baruch et al. 1995; Smith et al. 1995a,b; Su et al. 1995).

1.3.2 Cytoadherence in P. falciparum

Molecular interactions on the surface antigens of RBC are important for pathogenesis. They tend to modify the surface of RBCs by transporting antigens on to them. There are three kinds of adhesive processes that happen because of these antigens; cytoadherence to endothelial cells, agglutination, and rosetting (Fig. 1.9). Cytoadherence is the adhesion of the infected RBCs to the other type of cells such as vascular endothelial cells, platelets, and dendritic cells. Agglutination is the clumping of infected RBCs and rosetting is clumping of uninfected RBCs around a single infected RBC. Besides, during pregnancy, the clumps are formed in the placenta due to the adhesion of RBCs on syncytiotrophoblast (Rowe et al. 2009).

The infection of the RBCs leads to very high rigidity and deformation which is recognized by the spleen and cleared. But cytoadherence to the endothelial cells keeps iRBCs (infected RBCs) away

from peripheral circulation and hidden from the spleen. But there is a trade-off, even though parasites could escape splenic clearance, it is now more exposed to the immune system. So, it raises a question of whether this is the only advantage of the expression of antigen on the iRBC surface. In their review Saul, 1999, has suggested that chronicity of the infection is very important for the propagation of malaria to a new host. Hence, keeping parasitemia on check is very important, otherwise, all RBCs will be destroyed and eventually, the host will die before propagating malaria to a new host. Thus, the antibody response is the mechanism that keeps check on parasitemia. The switching rate of the antigen will be correlative to the antibody responses generated; hence, it will take care of the problem of wiping out all parasites (Saul 1999).

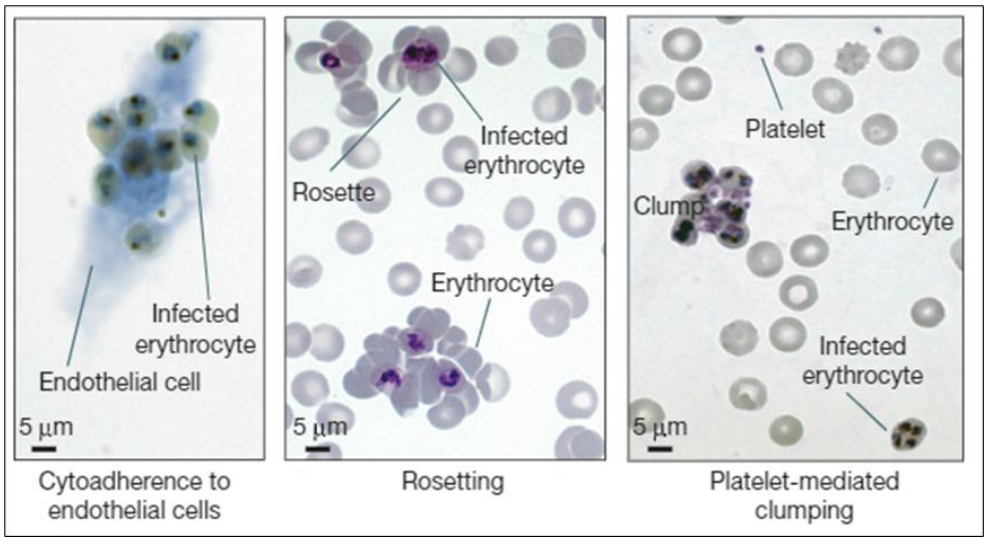


Figure 1.9: Adhesive interactions between RBCs infected with *P. falciparum* and different types of cells (image taken from Rowe et al. 2009).

1.3.3 Antigenic variation in P. falciparum

There are about 60 *var* genes identified in the *P. falciparum* genome. Most of the *var* genes are present at the sub-telomeric region of the chromosome, they are present on 13 out of 14 chromosomes. Six distinct telomere adjacent repeats (TARE) are present *ups*tream to the *var* genes. Each *var* gene has two exons and one intron. Exon 1 is variable; intron is conserved and exon 2 is semi-conserved in all *var* genes. Exon 1 codes for mostly extracellular domains that interact with other proteins, mostly receptors. It codes for two to five Duffy binding like (DBL)

domain and cysteine-rich interdomain region (CIDR) that is interspaced between DBL1 and DBL2. Scherf et al. in 1998 showed the differential binding of PfEMP1 to the receptors in Saimiri brain endothelial cells (SBEC) expressing three different receptors CD36, ICAM1 and CSA. SBEC cells expressing a particular receptor type will put selection pressure on parasite cells to express the *var* gene that binds to the receptor. Four generations of parasites were selected this way and isolated. Then again panned over CHO (Chinese Hamster Ovary) cells specifically expressing CSA and transgenics of CD36, ICAM-1. FCR3-CSA showed binding to CHO-CSA, similarly for others. Hence proved that each protein coded by *var* genes have specificity to the protein interaction. To check how many *var* genes are expressed at time rings and trophozoites were isolated after panning assay on CHO-CSA. Ring stages showed expression of 14 *var* genes at RNA level whereas in trophozoites only one *var* gene was expressed. At the protein level, only one *var* gene is expressed at a time and all others are repressed, this is called the mutually exclusive expression of *var* genes. The above results reiterated the fact that *var* genes are regulated at the transcription level (Scherf et al. 1998).

Other than *var* genes there are CVG families that show antigenic variation. *Rifins* and *stevors* are a prime example, which codes for transmembrane proteins. They are present on Maurer's cleft, parasitophorous vacuoles, and merozoites in addition to erythrocytes (Scherf et al. 2008). But the clear function is not yet known. Rosetting functions of *rifins* and *stevors* have been reported (Niang et al. 2014).

Chapter 2: Epigenetic landscape on *var* genes

2.1 Introduction

2.1.1 Singular expression of var genes

Monoallelic expression is a vital mechanism in mammals where in a diploid genome only one allele is expressed and other/s are kept suppressed. There are several multicopy genes that show allelic exclusion/mutually exclusive expression; prime examples being allelic exclusion of immunoglobulins in B cells, X chromosome inactivation and olfactory receptor genes (Vettermann & Schlissel 2010; Lee 2011; Monahan & Lomvardas 2015). *P. falciparum var* genes are a multigene family with 60 members, it shows a similar trait of mutually exclusive expression where a single *var* gene is expressed at a time and others are simultaneously repressed. While parallels can be drawn between monoallelic expression in mammalian cells and mutually exclusive expression of *var* genes, distinct genomic location of *var* genes and its role in pathophysiology of the parasite makes it a complex multilayered problem (Scherf et al. 1998). In addition, *var* genes show switching during the course of infection as a part of the bet-hedging strategy adapted by the parasites (Stubbs et al. 2005).

The *var* genes have a role in cytoadherence (as discussed in section 1.3.2), this makes them susceptible to detection by host immunity. The tug-of-war between the *var* genes and host immunity gave rise to a large repertoire of *var* genes by recombination and mutation. However, switching to a different variant upon an immune pressure is quick and effective, as it takes time by the host immunity for the recognition of antigen every time. The clinical data suggests a fast-switching rate for *var* genes (Gatton et al. 2003).

Regulation of mutually exclusive expression of *var* genes is largely transcriptional and by various epigenetic mechanisms (summarized in Table 2.1 and discussed in following sections). The dynamic *var* loci is regulated by reversible histone modifications. A large number of epigenetic players are identified but it is still unclear how this complex mechanism is executed (Deitsch &

Dzikowski 2017; Diffendall & Scherf 2024). Still, dots need to be connected and gaps need to be filled to understand the longstanding question of 'how *var* gene switching is regulated?'.

2.1.2 The var gene structure and sequence

The *var* genes are arranged on the sub-telomeric ends of the chromosome along with other CVGs such as *rifin, stevor, and Pfmc-2tm*. A few *var* genes are also positioned on the internal islands of CVGs in chromosome 4, 6, 7, 8, and 12 (Fig. 2.1A). All *var* genes have the same basic structure where there is a large exon 1 and ~1 kbp intron and a relatively shorter exon 2. Exon 1 codes for the hypervariable region that interacts with the endothelial receptors and exon 2 that is highly conserved in the gene family codes for the transmembrane domain (Fig. 2.1A). The *var* genes have a characteristic organization with two promoters; first positioned 1 kbp upstream of transcription start site and another exists in the intron. The cooperative interaction between the 5' flanking region and intron of a *var* gene is important for its repression (Frank et al. 2006).

Based on the sequence analysis of *var* upstream region the *var* genes are classified into three upsA, upsB and upsC and two intermediate groups; upsA/B and upsB/C. The orientation of upsA is usually pointing towards the telomere and the location is sub-telomeric, For upsB the orientation is away from telomere and the location is sub-telomeric. UpsC class has centrally located *var* genes on the chromosome. The *var* gene associated with placental malaria var2csa has a specialised promoter, hence it is classified as upsE separately (Kraemer et al. 2007; Scherf et al. 2008).

The 5' flanking regions of *var* genes contains promoters. The episomal expression of reporters under upsC promoters are demonstrated, indicating that UpsC *var* promoter could alter the chromatin environment and express the gene. Also, promoters can mediated the nucleation and spreading of silenced chromatin with the help of gene silencing machinery (Voss et al. 2006).

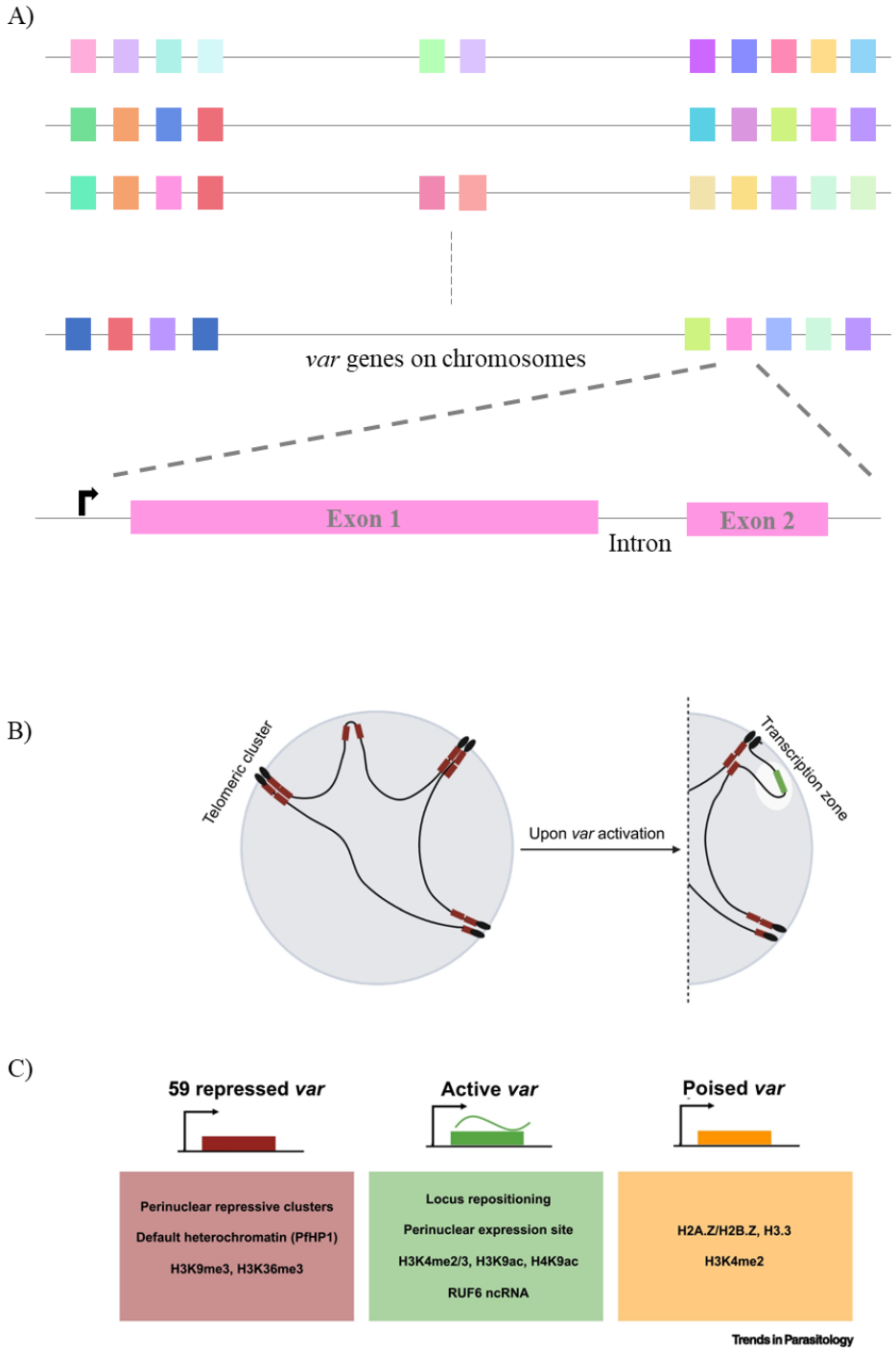


Figure 2.1: A) *var* gene organization on the chromosomes and gene structure B) Spatial genome organization of *var* genes C) Identified epigenetic signatures of *var* gene regulation (B and C are reproduced from the Diffendall and Scherf, 2024).

2.1.3 Heterochromatin in var gene silencing

The first report of involvement of heterochromatin in *var* gene regulation is by Duraisingh et al. 2005 and Freitas-Junior et al. 2005. It was found that PfSir2A (a histone deacetylase) to be a part of heterochromatin and regulates *var* genes by keeping all but one under suppression. (Duraisingh et al. 2005). Later H3K9me3 was identified to be present on inactive *var* loci (Chookajorn et al. 2007; Lopez-Rubio et al. 2007). ChIP-on-chip experiment with H3K9me3 antibody has revealed that heterochromatin is restricted to telomeric, sub-telomeric and some internal islands on the chromosome. In other words, heterochromatin has an exclusive enrichment on clonally variant genes including *var*. In addition, PfSir2A knockout parasites that showed derepression of *var* genes also had a striking absence of H3K9me3 from the loci. The *var* genes that had H3K9me3 enrichment, spatially clustered together to form a perinuclear repressive cluster (PERC) (Lopez-Rubio et al. 2009)(Fig. 2.1B).

Even though initial studies believed that PfSir2A to be driving the heterochromatinization, in 2009 two different independent studies showed that PfHP1 (ortholog of heterochromatin protein 1) to be specifically binding to H3K9me3 but no other histone methylations (Flueck et al. 2009; Pérez-Toledo et al. 2009). Also, PfHP1 was shown to be a part of PERC placing the protein at the center of the *var* gene regulome(Flueck et al. 2009). Later PfHP1 was found to be an essential gene for asexual proliferation but did not show any dramatic changes throughout the IDC as one would expect. However conditional depletion of PfHP1 resulted in de-repression of 52 out of 60 *var* genes (Brancucci et al. 2014). Controlled human malaria infection (CHMI) showed heterochromatin signature is low in expressing *var* genes from the patient samples (Pickford et al. 2021).

Phosphorylation of HP1 homologs has shown to be instrumental to its function (Canzio et al. 2014). But PfHP1 phosphorylation was important for neither asexual proliferation nor its nuclear localization. However, whether it has a role in *var* gene regulation remains elusive (Bui et al. 2019).

PfSET2 and PfHda2 were also important in the mutually exclusive expression of *var* genes. PfSET2 dependent H3K36me3 is present on the entire gene body and TSS of all silenced *var*

genes. The knockout of PfSET2 leads to the expression of all *var* genes simultaneously. The depletion of PfHda2 also leads to the universal loss of transcriptional repression of *var* genes along with the ApiAP2 transcription factor that helps in the sexual conversion of asexual stages (Coleman et al. 2014; Jiang et al. 2013).

2.1.4 Activating marks on var gene

The *var* promoters are shown to be instrumental for its activation (Voss et al. 2006). The active *var* gene is a part of much more permissive chromatin that is repositioned from the heterochromatin to the perinuclear expression site (Duraisingh et al. 2005). The first modification that was shown to be present on active *var* promoters was H4ac and was present only on var2csa that is active. PfSir2A acts antagonistically by only occupying *var* genes that are inactive. Also, PfSir2 can do direct deacetylation of H4ac (Freitas-Junior et al. 2005). H3K9ac and H3K4me2/3 were also found to be active *var* promoters (Lopez-Rubio et al. 2007) (Fig. 2.1B). Newer modifications are added to the list; H4K8ac was shown to be present on active *var* promoters (Gupta et al. 2017). Enzymes that deposit these histone marks on the *var* genes also have been identified, PfMYST; an acetyltransferase for H4 and PfSET10; a methyltransferase for H3K4 (Miao et al. 2010; Volz et al. 2012). Nonetheless, PfSET10s role in *var* regulation was disproved (Ngwa et al. 2021; Wyss et al. 2024).

In addition to histone modifications, histone variants are also shown to have a role in singular *var* choice. PfH2A.Z, PfH2B.Z and PfH3.3 variants were enriched on the active *var* promoters (Fraschka et al. 2016; Petter et al. 2011, 2013). Furthermore active histone modifications H3K9ac, H3K4me3 and H2A.Z showed a perfect colocalization at a global level (Bártfai et al. 2010).

Protein	Classification	Effect on var gene expression
H3K9me3	Histone modification	Keeps var genes in repressed state, exclusively present on the CVGs
PfHP1	Reader protein	Depletion of PfHP1 leads to de-repression var genes
PfSir2A	Histone deacetylase	Depletion of Sir2A leads to loss of H3K9me3 from var loci
PfSir2B	Histone deacetylase	Knock out leads to activation of specific var repertoire
PfSETvs/PfSET2	Histone methyltransferase	Methyl transferase for H3K36me3, knockout leads to dysregulation of var expression
H3K36me3	Histone modification	Present on silenced var genes along with the H3K9me3
PfHda2	Histone deacetylase	Essential for repression of all CVGs
H3K9ac	Histone modification	Present on active var promoters
H3K4me3	Histone modification	Present on active var promoters
H4K8ac	Histone modification	Present on active var promoters
H2A.Z	Histone variant	Part of transcriptionally permissive var loci, may be part of the epigenetic memory
H2B.Z	Histone variant	Part of transcriptionally permissive var loci, may be part of the epigenetic memory
H3.3	Histone variant	Part of transcriptionally permissive var loci, may be part of the epigenetic memory
RUF6	ncRNA	Associated to specific repertoire of active var gene
Antisense var intron	ncRNA	Can act in trans for var activation
Sense var intron	ncRNA	Associate with var activity not clear

Table 2.1: Summary of epigenetic factors that affect antigenic variation.

2.1.5 Intronic transcription mediated regulation

All *var* genes have the same basic structure where there is a large exon1 and ~1 kbp intron and a relatively shorter exon 2. Interestingly, intron has an AT-rich central region and a bidirectional promoter activity, this gives rise to sterile transcripts. While *var* full length transcripts show mutually exclusion expression, the intronic promoters of many var genes are active at all times.

RNA immunoprecipitation showed that sense and antisense non-coding RNAs (ncRNA) from the bidirectional promoter bind to the chromatin (Epp et al. 2009). They showed that this antisense ncRNA gets localised to special perinuclear foci and gets incorporated into the chromatin. Furthermore, the other studies show that these antisense lncRNAs (long non-coding RNA) are important for *var* activation. Antisense lncRNA are only transcribed from active var genes, unlike the sense that is transcribed from all loci. A *var* gene can be activated by its own antisense lncRNA, even in trans. A silenced endogenous *var* gene gets activated upon transcription of its antisense in an episome (Amit-Avraham et al. 2015). However there are other studies that showed there is no overall correlation between the sense or antisense originated from a particular *var* gene to its full length transcript (Ralph et al. 2005).

The deletion of the intronic promoter in an episomal construct disrupts the *var* gene silencing. This indicates the importance of the intron region and sterile transcripts on *var* gene silencing (Calderwood et al. 2003; Deitsch et al. 2001). Nevertheless, the context of chromatin is absent in these episomal constructs that use reporter genes, hence using CRISPR-Cas9 deletion an intronless *var* gene was created. Neither activation nor silencing of the *var* gene was affected but it did result in more switching rate (Bryant et al. 2017). It is possible that because of the sequence similarity between the *var* genes the intronic transcripts might act in trans and compensate for each other even if one intron was knocked out.

2.1.6 Other ncRNAs

RUF6 are a family of GC-rich ncRNAs that belong to a family called RNA of unknown function. It is unique to a genome like *P. falciparum* that is ~80% AT rich to have GC-rich ncRNAs. RUF6

is located along with the upsC *var* genes and shows promoter features of RNA pol III transcribing genes (Diffendall & Scherf 2024).

The RUF6 ncRNAs are generally associated with *var* gene activation. The interactome of this RNA includes RNA pol II subunits, nucleosome assembly proteins, and DEAD box helicase 5 (DDX5). The *var* genes harbors plenty of G quadruplex structures, DDX5 proposed to be resolving the secondary structures and hence helps in transcription initiation (Diffendall et al. 2023).

Another class of lncRNAs are TARE (Telomere associated RNA elements), which is suggested to have a role in heterochromatin maintenance and gene regulation (Simantov et al. 2022).

2.1.7 Genomic organization and var gene expression

Telomeric arrangement and *var* gene organization were initially studied by the FISH experiment. Very high recombination frequency was observed for *var* genes when two different strains of *P. falciparum* were crossed; even higher than usual meiotic recombination frequency. FISH experiments on rep20 in an adjacent region to the *var* gene in telomeres showed clustering of telomeres near to the nuclear periphery. The haploid genome of *P. falciparum* has 14 chromosomes; it is expected that 14 dots will appear in the FISH experiment if chromosome ends are dispersed. Instead, 4 to 7 chromosome ends cluster together to give rise to a bouquet-like structure. The clustering of telomeres will help in gene conversion events of *var* genes that are located sub-telemetrically (Freitas-Junior et al. 2000). It was widely believed that the telomere position effect (TPE) is responsible for *var* genes repression. But this will not be true for *var* genes that are located in the center of the chromosome. Sub-telomeric *var* gene var2csa was colocalized with the telomere cluster when it was inactive but it moved from the telomere cluster when activated. But even in the activated state, the var2csa was located near the nuclear periphery. This was later called the perinuclear expression site (PES). Electron microscopy of the *Plasmodium* nucleus showed a higher electron density near the nuclear periphery corresponding to the heterochromatin and lesser electron density at the center corresponding to the euchromatin. But in all nuclei, there is a small gap in the periphery that is devoid of heterochromatin. This area could be utilized by active *var* genes for their transcription. But it is very unlikely that this is the

only mechanism that keeps a *var* active or repressed. There are *var* genes that are only 10 kbp apart but have different transcriptional states. There will be additional mechanisms that help in the regulation of *var* gene expression (Ralph et al. 2005). It has been later shown that this re-localization is carried by perinuclear F actin. Tethering of F actin on to the intron of *var* genes is critical in this process (Zhang et al. 2011). Recent Hi-C studies on *P. falciparum* and related species showed the domain like structure (DLS) formation around *var* genes. But other *Plasmodium* species that lack *var* genes or antigenic variation failed to form such structures. Interestingly *P. knowlesi* that shows antigenic switching of SICA family of genes also has DLS in the genome (Bunnik et al. 2019).

2.1.8 Environmental factors in var gene regulation

Epigenetic modifications are highly susceptible to environmental changes. As an intracellular parasite, the *P. falciparum* life cycle is intimately connected to host life conditions. One major clinical manifestation of malaria is cyclic fevers. Temperature stress (41°C) has shown to upregulate *var* genes in a few studies (Jabeena et al. 2021; Oakley et al. 2007; Rawat et al. 2021b; Tabassum et al. 2021). PfSir2A and PfSir2B get down regulated under febrile temperatures in a Hsp90 dependent manner. This in turn leads to derepression of *var* genes (Tabassum et al. 2021).

Host diet is also shown to have effect on sexual conversion and virulence in malaria cases (Brancucci et al. 2017; Mancio-Silva et al. 2017). The principal methyl donor SAM (S-Adenosylmethionine) is vulnerable to dietary changes as it is part of a complex metabolic pathway. Changes in lipid metabolism resulted in changes in virulence gene expression. Genetic modulation of SAM led to similar changes in *var* gene expression (Schneider et al. 2023).

A clinical study on asymptomatic malaria in the dry season showed a decreased RNA pol III transcription and hence low RUF6 ncRNAs. As discussed in section 2.1.6, *var* genes are positively regulated by RUF6, hence diminished *var* gene levels were observed (Diffendall et al. 2024).

2.2 Scope of this Chapter

Epigenetics is shown to be at the core of *var* gene regulation. There are numerous experiments and hypotheses that revolve around the mechanism of *var* gene switching. Perhaps what is missing is a unifying molecular mechanism that compiles all the existing data and new findings.

In this Chapter we are trying to comprehensively understand dynamics of epigenetic modifications at the *var* loci in the light (or darkness) of heterochromatin. We have harvested the samples for ChIP sequencing and RNA sequencing simultaneously from the 3D7 culture. Also, we raised and validated an antibody against PfHP1 protein. We chose temperature stress as a paradigm to study the effect of environmental factors on epigenetics and gene regulation.

2.3 Materials and Methods

2.3.1 Cloning, bacterial expression and purification of recombinant PfHP1 protein

Full length PfHP1 is cloned in the pET28a+ bacterial expression vector with His tag. The vector was transformed in Bl21 (DE3) *Escherichia coli.* The next day colonies were picked to expand the culture in LB (Luria Broth) under kanamycin antibiotic selection. At 0.6 OD (optical density) at 600nm the culture was induced for protein expression using 0.5 mM IPTG (isopropyl-1-thio-β-d-galactopyranoside). The induction was done overnight at 18°C. The pellet was spun down and stored at -80°C until the protein purification.

The cell pellet was resuspended in sonication buffer (see table 2.2 for the buffer recipe; 40 ml sonication buffer for the pellet from 1L culture) and sonicated for 40 minutes at 60% amplitude, 2 sec ON and 5 sec OFF cycles in a probe sonicator (Vibra-cell, Sonics and Material Inc.). The lysate was spun at 12000 rpm, 4°C, 20 min. The supernatant was allowed to bind to Ni-NTA beads for 2 hours in rotation at 4°C. The unbound lysate was removed by gravity flow. The beads were washed with 5 CV (column volume) of wash buffer I, II, and III. The PfHP1 was eluted in 1 CV of elution buffer; a total of 5 elutions were taken. The protein was estimated using

spectrophotometric measurement or a nanodrop. The protein was stored at 4°C for short term and -80°C for long term storage. All the buffer recipes are given in the table.

2.3.2 Antibody generation

The purified PfHP1 protein was used as an antigen for the antibody generation. The antibodies were raised in The National Facility for Gene Function in Health and Disease, IISER Pune. The New Zealand White rabbit was immunised with 300 µg of protein with complete Freund's adjuvant (CFA). The antigen-adjuvant emulsion was made by mixing equal volumes of the protein and CFA using a three-way syringe. Before the first injection collected pre-immune sera. For the subsequent booster shots 200 µg protein was injected along with Incomplete Freund's adjuvant. After the first booster shot the bleed was collected every 14 days. The immune sera was separated from the whole blood by centrifugation.

2.3.3 Western blot

For the antibody validation *P. falciparum* nuclear lysate was prepared following the given protocol. Briefly the parasites were enriched by saponin (0.015%) lysis of RBCs. The pellet was first lysed by a cytoplasmic buffer (10mM HEPES pH-7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 0.65% NP40, 1 mM DTT, 2X PMSF) and then by a high salt nuclear extraction buffer (20 mM HEPES pH 7.9, 400 mM NaCl, 1 mM EDTA, 1 mM EGTA and 1 mM DTT). The nuclear extract (25 µg) was used for western blotting. After the transfer, PVDF was blocked using 5% milk and probed with primary antibody (3rd bleed of PfHP1 antibody in 1:200 dilution) overnight. Next day, probed with anti-rabbit HRP (Horseradish peroxidase) antibody in 1:5000 dilution. Gave three washes with TBST (1X Tris-buffered saline, 0.1% Tween 20) and developed the blot using Clarity max ECL substrate (Bio Rad).

Buffer	Recipe
Sonication buffer	HEPES pH 8.0 : 20mM KCl : 300mM imidazole: 10mM Triton X100: 0.5% PMSF : 1mM Glycerol : 10% Lysozyme: 1 mg/ml β-mercaptoethanol (β-me): 2mM
Wash buffer 1	pH 8.0 HEPES: 20mM KCl: 1M Glycerol: 10% Imidazole: 10mM β-me: 2mM
Wash buffer 2	pH 8.0 HEPES: 20mM KCl: 300mM Glycerol: 10% Imidazole: 30mM β-me: 2mM
Wash buffer 3	pH 8.0 HEPES: 20mM KCl: 300mM Glycerol: 10% Imidazole: 50mM β-me: 2mM
Elution buffer	pH 8.0 HEPES: 20mM KCl: 300mM glycerol: 10% Imidazole: 500mM β-me: 2mM

Table 2.2: Buffer recipes for the PfHP1-his protein purification

2.3.4 Immunoprecipitation followed by mass spectrometry

For the Immunoprecipitation (IP), the protocol from Thermo Fisher Scientific Dynabeads protein A was followed. Briefly, 50µl Dynabead was used per pulldown reaction. Two washes were given to the beads with a wash buffer (1X PBS + 0.02% Tween-20). The PfHP1 antibody (from bleed 3) was used and 10µl antibody (or pre-immune sera as a control) was diluted in 190µl buffer. Kept for binding with the Dynabeads at room temperature for 1 hr in rotation. Separate the beads using a magnetic stand and give one wash with the bead wash buffer. Prepares the nuclear extract as detailed above. A total of 850 ng of NE was added to the dynabead-antibody complex and incubated overnight at 4°C, in rotation. Next day separated the beads and gave three washes each

with IP wash buffer 1 (25mM Tris-Cl pH 7.9, 5mM $MgCl_2$, 10% glycerol, 0.1% NP-40, 0.3mM DTT) and IP wash buffer 2(25mM Tris pH 7.9, 5mM $MgCl_2$, 10% glycerol, 100mM KCl, 0.1% IGEPAL, 0.3mM DTT). Then eluted with 30µl elution buffer (50mM Glycine pH 2.5). Later neutralized by adding 30µl Tris-Cl pH 8.8.

For the mass spectrometry the elute was trypsinized and desalted using the standard protocol. Samples were reduced using 10mM DTT made in 50mM TEABC and incubated at 60°C in a thermomixer with shaking for 45 minutes. For alkylation used 10mM Iodoacetamide (IAA) in 50mM TEABC. Incubated in the dark for 15 min at room temperature. Trypsinized the alkylated samples with 1µg trypsin (Promega) and incubated at 37°C overnight in shaking (800 rpm). 0.1% Formic acid was added to stop trypsinization by acidification. Incubated at room temperature for 15 min and spun at 14000 rpm and collected the supernatant. Desalting was performed using self-assembled columns using C18 disks. Load the samples to the column and let it bind. Gave one wash using 100µl of 0.1% formic acid and eluted using 50µl of 60% acetonitrile in 0.1% formic acid. The samples were vacuum dried at 30°C. Mass spectrometry was performed in SCIEX QTOF and results were analyzed in protein pilot software.

2.3.5 Parasite culture

The *P. falciparum* line 3D7 was maintained in 2% hematocrit in RPMI 1640 10.4g/liter, HEPES 25 mM, AlbuMAX II 0.5% (wt/vol), hypoxanthine 100 µM, gentamicin 12.5 µg/ml, sodium bicarbonate 1.77 mM (just before use by adding 3 ml of 5% (wt/vol) sodium bicarbonate to 100 ml of medium). The parasites were maintained at 37 °C in an atmosphere with 5% carbon dioxide, and rest atmospheric. For the temperature stress experiment the 18 hr ring parasites were incubated at 40 °C for 6 hours then harvested for ChIP and RNA sequencing. Parasites at normal condition were harvested as a control. The parasites for the experiments were synchronized by 5% sorbitol and 63% percoll gradient cyclically to achieve a tight synchrony.

Buffer	Recipe
Swelling buffer	25 mM Tris pH 7.9 1.5 mM $MgCl_2$ 10mM KCl 0.1% NP40 1mM DTT 0.5 mM PMSF 1x PIC (Roche)
Sonication buffer	50 mM Tris-Cl pH7.9 140 mM NaCl 1mM EDTA 1% Triton X-100 0.1% Sodium deoxycholate 1.0% SDS 0.5 mM PMSF 1x PIC (Roche)
ChIP dilution buffer	0.01% SDS 1.1% Triton X-100 1.2 mM EDTA 16.7mM Tris-Cl pH 8.0 167 mM NaCl
Low salt buffer	0.1% SDS 1% Triton X-100 2 mM EDTA 20 mM Tris-Cl pH 8.0 150 mM NaCl
High salt buffer	0.1% SDS 1% Triton X-100 2 mM EDTA 20 mM Tris-Cl pH 8.0 500 mM NaCl
LiCl buffer	0.25 M LiCl 1% NP40 1% Sodium deoxycholate 1 mM EDTA 10 mM Tris-Cl pH 8.0
TE buffer	10 mM Tris-Cl pH 7.5 1 mM EDTA
Elution buffer	1% SDS 1.1 M $NaHCO_3$

Table 2.3: Buffer recipes for the ChIP protocol

2.3.6 Chromatin immunoprecipitation (ChIP)

To the parasite culture, add 1% formaldehyde dropwise (16% formaldehyde, methanol-free, Invitrogen). Incubate in the dark, at room temperature and slow rocking for 10 minutes. Quench the formaldehyde by adding 150 mM glycine from a 2M stock solution. Incubate at room temperature and slow rocking for 10 minutes. Spin at 4^0C, 6000 rpm for 5 minutes. Discard the supernatant and wash it three times with chilled 1X PBS (filtered). Store the pellet (with no residual liquid) at -80^0C until further proceeding.

To the harvested *P. falciparum* parasites, add 6 times of swelling buffer (Table 2.3 for the recipe). Keep on ice for 10 minutes and Dounce 100 times for the ring stage and 75 times for troph and schizont stage. Spin at 5000 rpm for 5 minutes at 4^0C. Discard the supernatant and add a sonication buffer 8 times to the packed pellet. Sonication conditions on the disruptor (diagenode) are 5 times 20 cycles, 10 seconds ON/OFF. Use a 1.5 ml tube for sonication and keep the samples in a slush of ice and water. Change the slushy after each 20-cycle iteration of the sonications. If the machine is Covaris M220, the sonication conditions are; average incident power: 10.1 W, peak incident power: 75.0 W, Duty factor: 13.4%, cycle/burst: 200, Duration: 720 seconds. After the sonication, spin at 13000 rpm, 15 minutes at 4^0C.

To estimate the chromatin amount and fragment size, take 20 µl of the supernatant and the remaining freeze down and keep it at -80^0C. Add 20 µl of 5M NaCl (0.3M final conc.) and 1 µl of RNase (10mg/ml). Make the volume 300 µl with sonication buffer. Incubate at 65^0C O/N on a thermomixer. Next day, add 20 µg proteinase K, 20 µl of Tris pH 7.9 (1M) and 10 µl of EDTA (0.5M). Incubate at 42^0C for 1hr. Perform DNA extraction using Phenol: chloroform: isoamyl alcohol. Add 20 µl of glycogen (2mg/ml) for better pellet visibility while precipitation with isopropanol. Estimate the amount using nanodrop reading of the DNA. For size estimation, run the entire sample on a 1% agarose gel at 160 V for 10 minutes. The fragment size should be between 200-500 bp. Over-sonication may result in dissociation of protein from the chromatin. Bigger fragment sizes are not advisable because it may result in erroneous results.

Pulldown of chromatin: Take 20-25 µg of chromatin (for *P. falciparum*, samples can go as low as 10-20 µg if the samples are limited). Make the volume up to 1.0 ml with a chromatin dilution buffer. The pre-clearing step is optional if using protein A dynabead (Invitrogen). For pre-clearing, add 10 µl beads to the chromatin lysate for 1 hour, 4^0C, and 10 rpm end-to-end rotation. Separate the supernatant using a magnetic stand and add antibody (Primary antibodies used were α-H3K9me3 (Invitrogen, PA5-31910), α-H3K9ac (Millipore, 06-942), α-H3K79me2 (Sigma, 04-835), α-H4K8ac (Sigma, 07-328), α -RNA Pol II Serine-2-phosphorylation antibody (Abcam, ab5095), α-PfHP1 (in-house), α-PfSir2A (in-house), α-H3K36me3 (Invitrogen, MA524687)). At this stage the 10% input can be removed just before the addition of the antibody. Incubate overnight 4^0C and 10 rpm. Next morning, add 40 µl of saturated beads to the lysate. To prepare saturated beads, incubate the beads with 0.4 mg/ml tRNA and 0.5 mg/ml BSA for 3 hrs, 4^0C, and 10 rpm. The lysate is incubated with the beads for 4 hrs, 4^0C, and 10 rpm. Separate the beads by a magnetic stand and wash thrice with the chilled low salt buffer, high salt buffer, LiCl buffer and TE buffer (See the Table 2.3 for the recipe). Incubate for 10 minutes with each buffer at 4^0C and 10 rpm. Finally, elute the chromatin using 150 µl elution buffer at 37^0C for 5 minutes. Repeat the step once more to get a total of 300 µl elute. As explained above, de-cross linking is performed, followed by DNA extraction. Estimate the DNA amount with Qubit DNA high-sensitivity kit. Proceed with the library preparation.

2.3.7 ChIP library preparation and sequencing

The sequencing library was prepared using the NEB Ultra II DNA kit (E7645), following the manufacturer's instructions. The prepared library was estimated for the average size using Agilent bioanalyzer and for the amount using Qubit fluorometer. A final amount of 1.3 pM library was loaded from the multiplexed pool of all libraries. Paired end sequencing was performed using Illumina 300 cycle mid output kit. The NextSeq 550 (IISER Pune, NGS facility) was used for the sequencing.

2.3.8 RNA isolation

The samples were spun down and harvested using 0.015% saponin. The cell pellet was resuspended in Trizol and snap frozen and stored -80^0C until further proceeding.

The RNA isolation was done on the day of library preparation to minimize freeze thaw cycles. The samples were thawed and added 0.2 volumes of chloroform and mixed thoroughly by inverting. Incubate at room temperature for 15 minutes, followed by a spin at 12000 rpm, 15 minutes, 4^0C. Separate the top layer and add 0.8 volume of isopropanol (may add glycogen during this step to increase the visibility of the pellet). Invert mix and incubate at room temperature for 30 minutes. Spin down at 12000 rpm, 15 minutes, 4^0C. Observe a white pellet and wash the pellet with 70% ethanol. Dry the pellet at room temperature to evaporate and residual ethanol. Resuspend the pellet in 20 µl nuclease free water. The RNA quality was assessed using RNA high sensitivity nano chip for Agilent bioanalyzer and the RNA was quantified using Qubit RNA high sensitivity kit.

2.3.9 RNA library preparation and sequencing

The isolated RNA was subjected to polyA enrichment using NEB polyA mRNA magnetic isolation module (E7490). The library preparation was done using the NEB Ultra II RNA kit (E7770) following the manufacturer's instructions. The library quality was tested using a DNA high sensitivity kit in Agilent bioanalyzer. The library amount was estimated using Qubit DNA high sensitivity kit. A final pool of 1.7 pM library was loaded on a 75 cycle High output illumina sequencing kit. The single end sequencing was performed using NextSeq 550 from Illumina at IISER Pune NGS (Next generation sequencing) facility.

2.3.10 ChIP sequencing data analysis

Trimming and quality control (QC) of the raw data was done using Trim Galore. The trimmed data was aligned against the *P. falciparum* reference genome (PlasmoDB_v53) using Bowtie2. The aligned files were sorted using samtools which gave output files in bam format. The peaks were called using MACS2, using the input files as control. The normalized bedgraph files were loaded on IGV (Integrative Genomes Browser) to visualize the enrichment pattern. For the mean tag density and gene body profile seqMINER was used. The heatmaps were generated using Morpheus, Broad Institute.

2.3.11 RNA sequencing data analysis

Trimming and QC of the raw data was done using trim galore. The trimmed data was aligned against the *P. falciparum* reference genome (PlasmoDB_v53) using HISAT2. The aligned files were sorted using samtools which gave output files in bam format. The tpm (transcripts per million) value for genes were obtained from StringTie analysis. The mean tag density for the heat maps were obtained from seqMINER software. The heatmaps were generated using Morpheus, Broad Institute.

2.4 Results

2.4.1 Raising PfHP1 antibody and validation of antibody

To gain insights into PfHP1 mediated gene regulation in *P. falciparum* it is important to have ample tools to perform ChIP sequencing or immunofluorescence imaging. Both of which are antibody based. In the absence of a commercial antibody against PfHP1, we decided to raise a polyclonal antibody in-house. PfHP1 was purified using his-tag based affinity purification (Fig. 2.2A). The purified full-length protein was used for the antibody generation. Anti-serum obtained from the immunized rabbits could pick up PfHP1 bands from the nuclear extracts prepared from 3D7 (Fig. 2.2C). However, pre-immune sera collected from the rabbit did not pick up any bands (Fig. 2.2B). An IP followed by mass spectrometry further confirmed that the antibody is indeed specific to PfHP1. The pulldown gave PfHP1 to be top hit after subtracting the proteins pulled down by the pre-immune sera (*Appendix* Table 1 & 2). We have successfully raised and validated the PfHP1 antibody that shall be used in the further experiments required in this study.

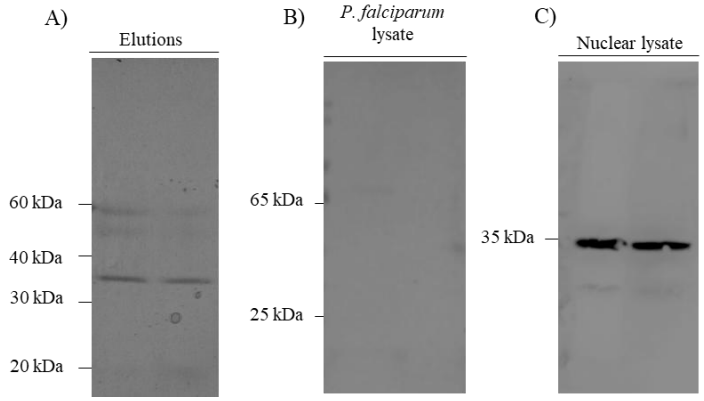


Figure 2.2: A) Protein purification of PfHP1-His, image from the SDS-PAGE stained with Coomassie Brilliant Blue, B) Western blot of *P. falciparum* protein lysate probed with pre-immune sera in 1:200 dilution in 5% milk, C) Western blot of *P. falciparum* nuclear protein lysate probed with anti-PfHP1 antibody in 1:200 dilution in 5% milk.

2.4.2 Global profile and local profile of PfHP1 on var genes in comparison to other histone modifications

We performed ChIP sequencing on early troph stage parasites, targeting PfHP1, H3K9ac, H4K8ac, H3K79me2, H3K36me3 and PfSir2A to study epigenetic landscape of *var* genes and interplay of histone modifications in the regulation. It has been established in the field that *var* genes get expressed in the ring stage and in later stages the genes go into a transcriptionally poised state (Scherf et al. 1998). H3K9ac gave a large number of peaks in euchromatin and limited peaks in heterochromatin as expected. PfHP1 showed broad peaks restricted to the telomeric and sub-telomeric region, revealing antagonistic nature of H3K9ac and PfHP1. H4K8ac and H3K79me2 are relatively under-studied in the *P. falciparum*, both the modifications mimicked the global pattern of H3K9ac. Whereas H3K36me3 showed enrichment in the heterochromatic ends (Fig.2.3A).

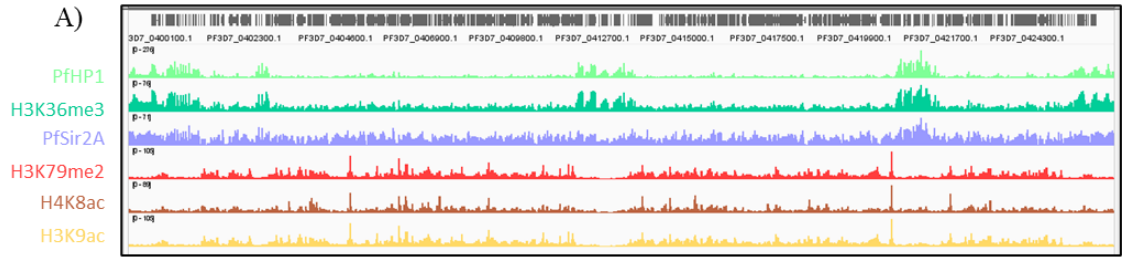


Figure 2.3: A) Global enrichment pattern of histone modifications H3K9ac, H4K8ac, H3K79me2, H3K36me3 and epigenetic factors PfHP1 and PfSir2A on chromosome 4 (continued in next page).

We further focused on the *var* gene regulation, to understand the histone modification dynamics at *var* loci. From previous bulk RNA sequencing and single cell RNA sequencing that multiple *var* genes are active in the 3D7 parasite population (Rawat et al. 2021b). We observed peaks of H3K9ac at the introns (mainly at the exon-intron junctions of both or one of the exons) of a subset of *var* genes. Interestingly it has been reported that introns of *var* genes harbor a promoter that can transcribe sense and antisense RNAs (Calderwood et al. 2003; Epp et al. 2009). Genes that had H3K9ac peak at *var* intron showed a co-localized peak for H3K79me2 and H4K8ac. PfHP1 showed a depletion on the intronic region of all *var* genes, but specifically genes that have H3K9ac at the introns showed further retraction in the heterochromatic boundaries in those *var* genes (Fig. 2.3B & C). These results are further alluding to the antagonistic nature of PfHP1 and the three-histone modifications, H3K9ac, H4K8ac and H3K79me2.

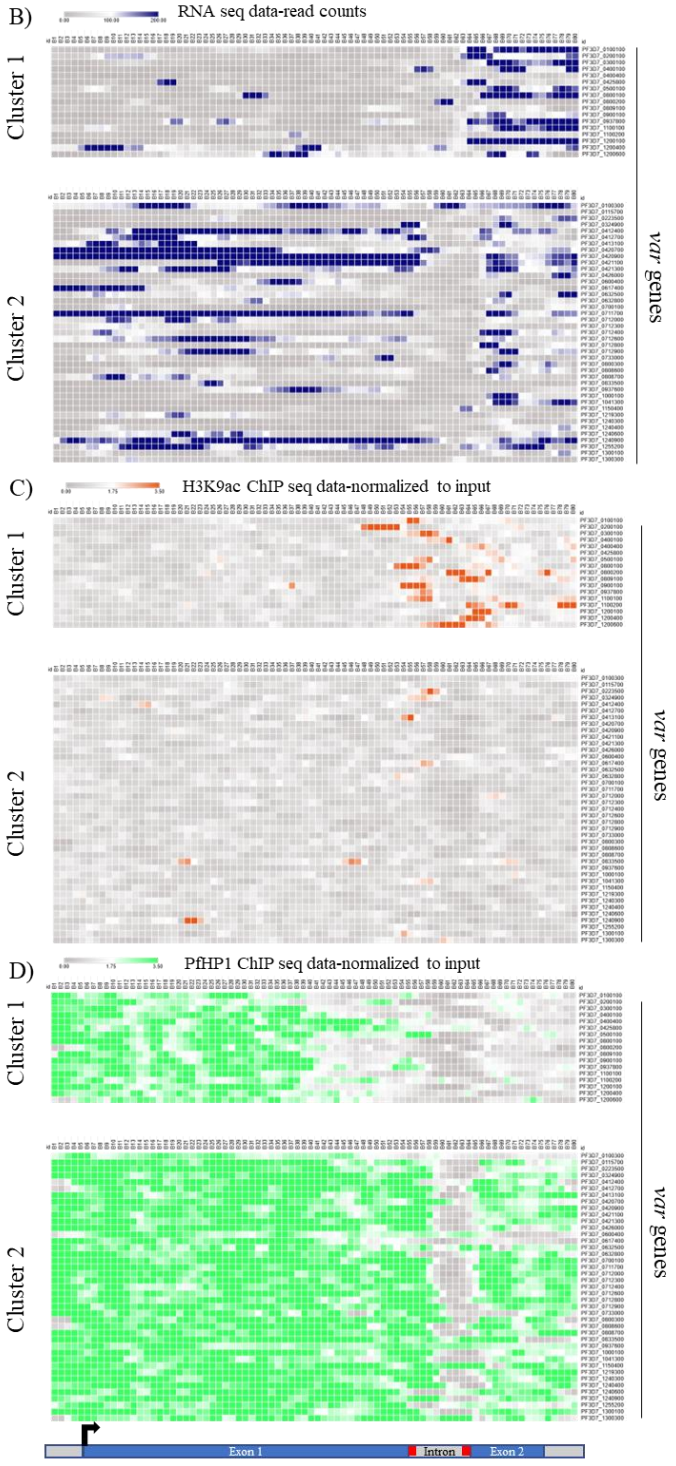
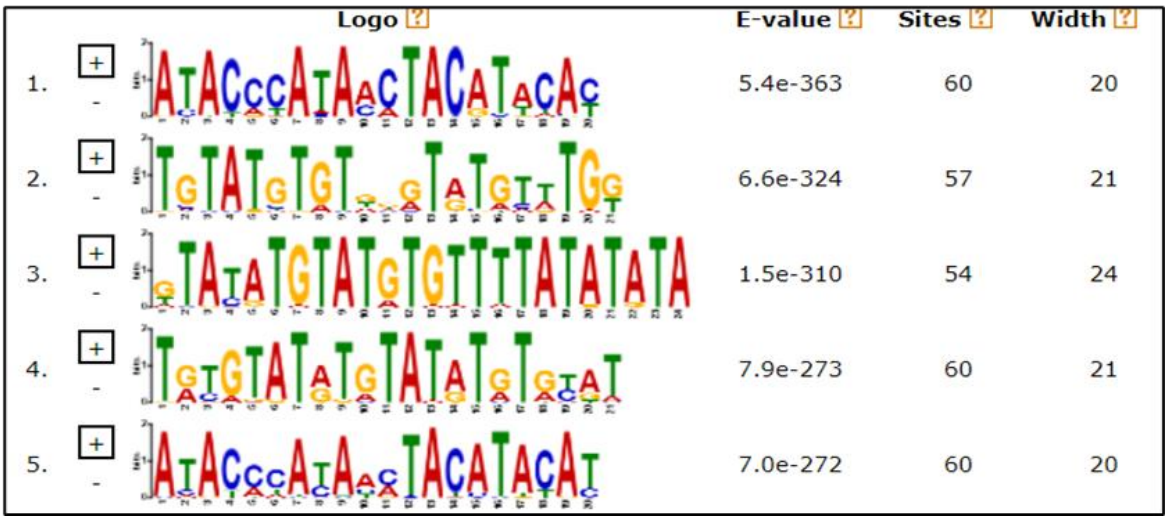


Figure 2.3: *Legend continued*, B) RNA sequencing data represented as read counts on the *var* gene body C) Normalised mean tag density of H3K9ac ChIP sequencing on *var* gene body D) Normalised mean tag density of PfHP1 ChIP sequencing on *var* gene body.

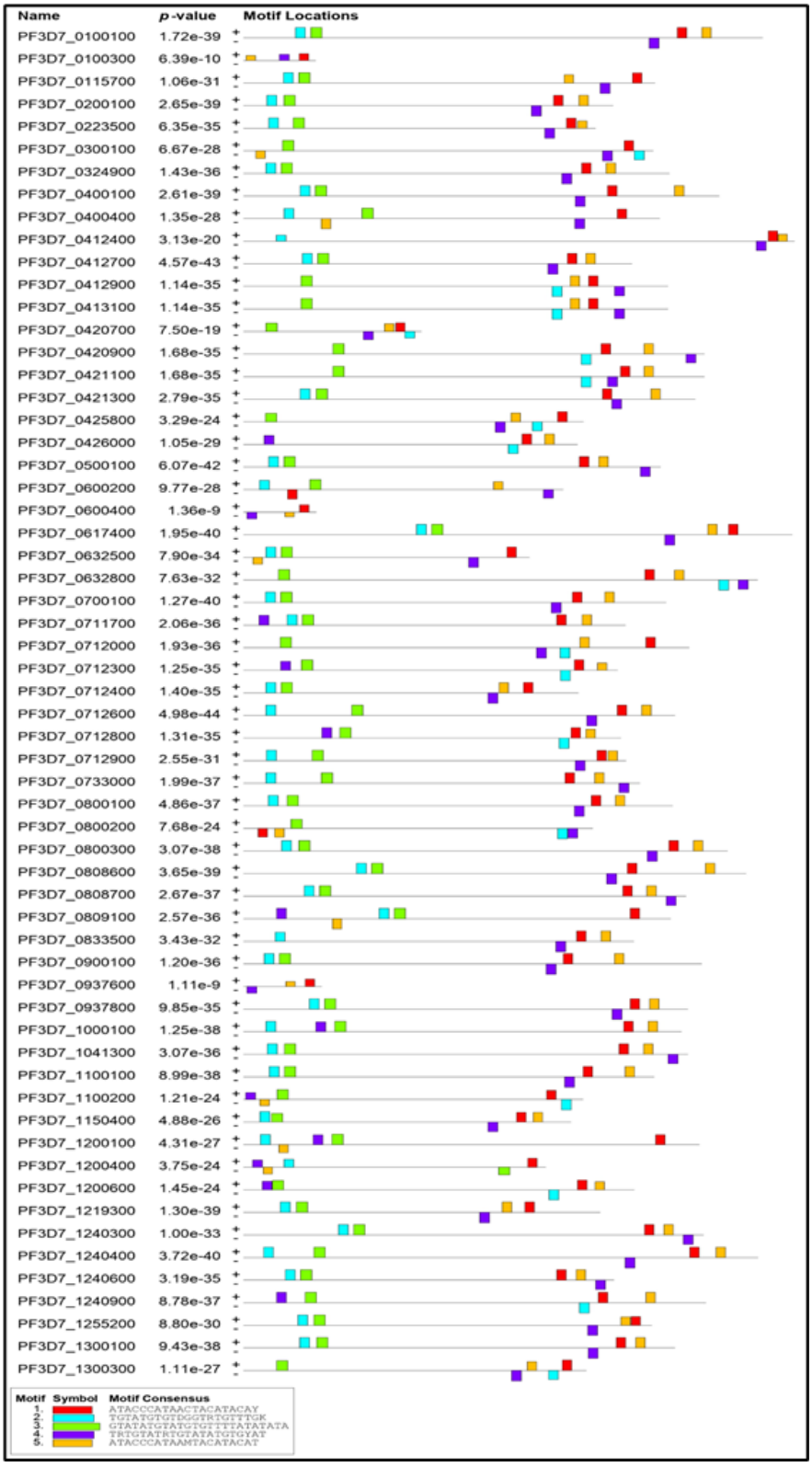
2.4.3 The insulator-like motifs on var introns show expansion and retraction of heterochromatin boundaries

Intronic regions of *var* genes show a limited heterochromatin spread. Also, there is an expansion and retraction of heterochromatin boundaries at the intron of different *var* genes (Fig. 2.4). *Var* introns can be a paradigm for studying heterochromatin spread and how it is controlled. Earlier it has been reported that *var* promoters and introns harbor a insulator like motif (Avraham et al. 2012). Motif discovery analysis on *var* genes have led us to identify five sequence motifs in the intronic region (Fig. 2.4A). Motif 1 and 5 are similar and present on the ends of the *var* introns. Motif 2, 3 and 4 are also very similar to each other and are present on start and end of the introns (Fig. 2.4B). These motifs have a repeat of TA-TG mainly and are present mostly on *var* genes and a few clonally variant genes. Interestingly, the activating histone modifications such as H3K9ac, H4K8ac, and H3K79me2 are centered at the intronic motif of a subset of *var* genes (Fig. 2.4C). However, in the absence of H3K9ac on the intronic motif, heterochromatin spreads into the motif (Fig. 2.4D). A previous study explores the effect of short-range interaction between the promoter and *var* introns via these insulator-like elements (Avraham et al. 2012). As we know in eukaryotes, DNA motif interactions can limit chromatin domains by marking their boundaries (Bell et al. 1999). The intronic motif and putative insulator proteins binding on to them might be regulating the heterochromatin dynamics.

A)



B)


Name
p-value
Motif Locations
PF3D7_0100100 1.72e-39
PF3D7_0100300 6.39e-10
PF3D7_0115700 1.06e-31
PF3D7_0200100 2.65e-39
PF3D7_0223500 6.35e-35
PF3D7_0300100 6.67e-28
PF3D7_0324900 1.43e-36
PF3D7_0400100 2.61e-39
PF3D7_0400400 1.35e-28
PF3D7_0412400 3.13e-20
PF3D7_0412700 4.57e-43
PF3D7_0412900 1.14e-35
PF3D7_0413100 1.14e-35
PF3D7_0420700 7.50e-19
PF3D7_0420900 1.68e-35
PF3D7_0421100 1.68e-35
PF3D7_0421300 2.79e-35
PF3D7_0425800 3.29e-24
PF3D7_0426000 1.05e-29
PF3D7_0500100 6.07e-42
PF3D7_0600200 9.77e-28
PF3D7_0600400 1.36e-9
PF3D7_0617400 1.95e-40
PF3D7_0632500 7.90e-34
PF3D7_0632800 7.63e-32
PF3D7_0700100 1.27e-40
PF3D7_0711700 2.06e-36
PF3D7_0712000 1.93e-36
PF3D7_0712300 1.25e-35
PF3D7_0712400 1.40e-35
PF3D7_0712600 4.98e-44
PF3D7_0712800 1.31e-35
PF3D7_0712900 2.55e-31
PF3D7_0733000 1.99e-37
PF3D7_0800100 4.86e-37
PF3D7_0800200 7.68e-24
PF3D7_0800300 3.07e-38
PF3D7_0808600 3.65e-39
PF3D7_0808700 2.67e-37
PF3D7_0809100 2.57e-36
PF3D7_0833500 3.43e-32
PF3D7_0900100 1.20e-36
PF3D7_0937600 1.11e-9
PF3D7_0937800 9.85e-35
PF3D7_1000100 1.25e-38
PF3D7_1041300 3.07e-36
PF3D7_1100100 8.99e-38
PF3D7_1100200 1.21e-24
PF3D7_1150400 4.88e-26
PF3D7_1200100 4.31e-27
PF3D7_1200400 3.75e-24
PF3D7_1200600 1.45e-24
PF3D7_1219300 1.30e-39
PF3D7_1240300 1.00e-33
PF3D7_1240400 3.72e-40
PF3D7_1240600 3.19e-35
PF3D7_1240900 8.78e-37
PF3D7_1255200 8.80e-30
PF3D7_1300100 9.43e-38
PF3D7_1300300 1.11e-27
Motif Symbol Motif Consensus
1. ATACCCATAACTACATACAY
2. TGTATGTGTDGGTRTGTTTGK
3. GTATATGTATGTGTTTTATATATA
4. TRTGTATRTGTATATGTGYAT
5. ATACCCATAAMTACATACAT

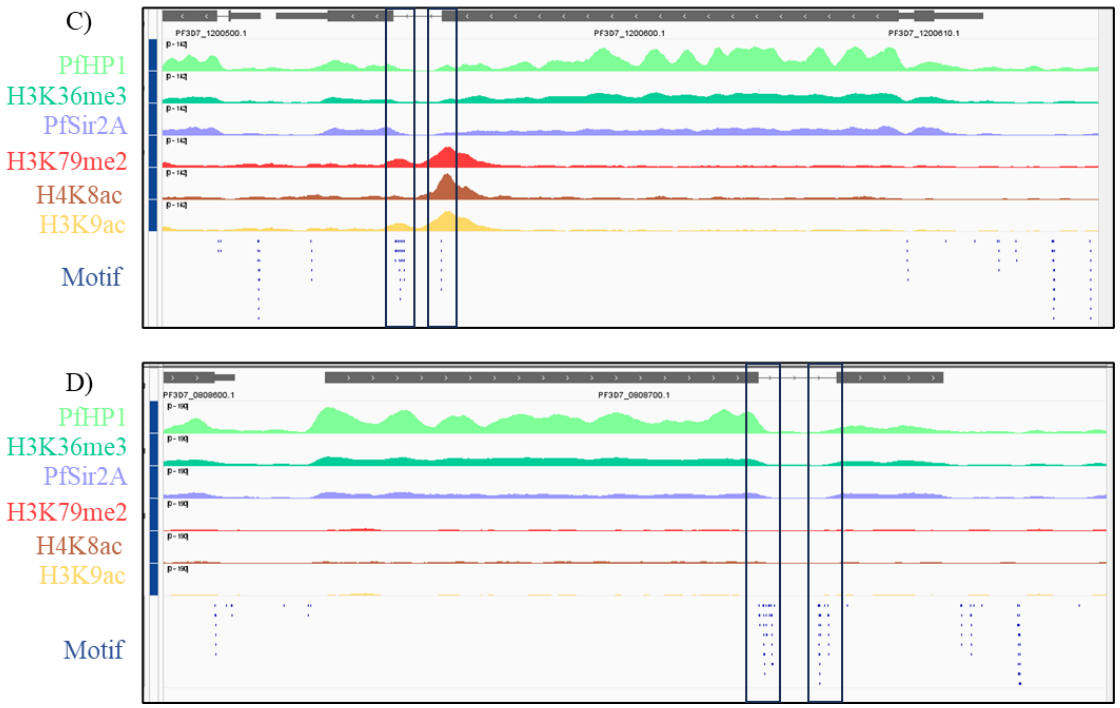


Figure 2.4: A) Consensus motifs identified in the MEME ChIP motif discovery analysis on *var* intronic sequences, B) Positioning of the identified motifs on the *var* introns, C) Examples of *var* gene with active modifications and, D) without active modifications in the intronic motif.

2.4.4 Interplay between PfHP1 and H3K9ac regulates the intronic transcription in the var genes

To further study the dynamicity of histone modifications on *var* intron loci, we performed RNA sequencing and ChIP sequencing of histone modifications and PfHP1 in control vs. temperature stress conditions. Temperature is known to cause upregulation of *var* genes and possibly switching as it mimics cyclic fevers that is the prominent symptom associated with falciparum malaria (Oakley et al. 2007). RNA sequencing showed an abundance of transcripts that map to the exon 1 and exon 2 (*Appendix* Fig. 1). But very few *var* genes showed reads that are panning the exon1 and 2 together which would indicate an event of splicing. Together these results indicate that in the trophozoite stage there is an abundance of intronic transcription whereas full-length transcription is at the basal level.

The RNA sequencing data of 3D7 shows that levels of H3K9ac in the intronic region are correlated with sense transcripts from there (Fig 2.3B&C). We observed a striking upregulation of transcription of upsB genes under temperature stress (Fig. 2.5A). On the contrary upsC type *var* genes showed a down regulation in the transcription (Fig. 2.5A). Interestingly most of the reads in upsB *var* genes mapped to exon 2 and in upsC *var* genes to exon 1, indicating that it is the intronic transcription that shows deregulation under temperature stress (Fig. 2.5B&C). When corresponding ChIP sequencing data was analyzed, intronic H3K9ac peaks showed an increase under temperature stress for upsB type (Fig. 2.6B&C). The level of H3K9ac in upsC *var* is greater in the control than in the temperature stress (Fig. 2.6A). However, the intronic region of upsC did not show much lower level of H3K9ac in both control and temperature conditions compared to upsB type (Fig. 2.6C). In contrast, PfHP1 levels on upsC introns were higher as compared to upsB. But contrary to the intuition, upsB introns showed slightly higher PfHP1 in the temperature stress in comparison to control (Fig. 2.6D).

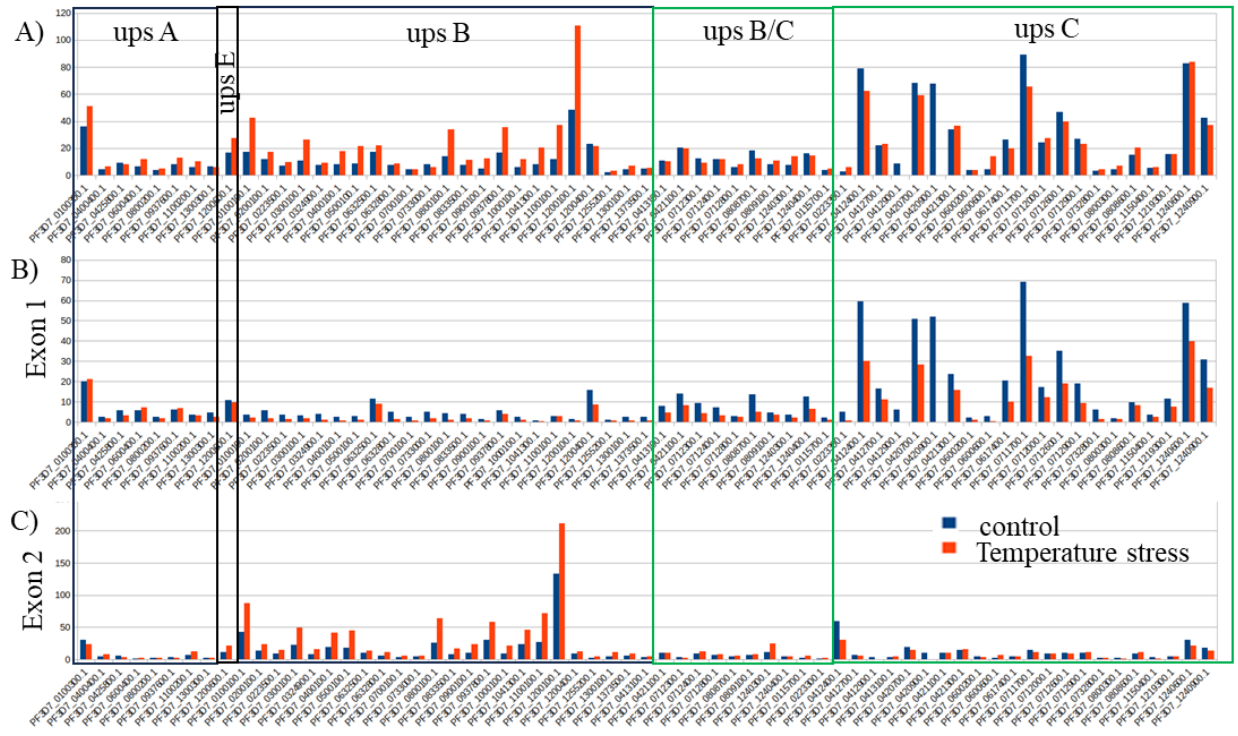


Figure 2.5: TPM of A) whole *var* gene body B) exon1 C) exon2 , arranged by the *ups* classes.

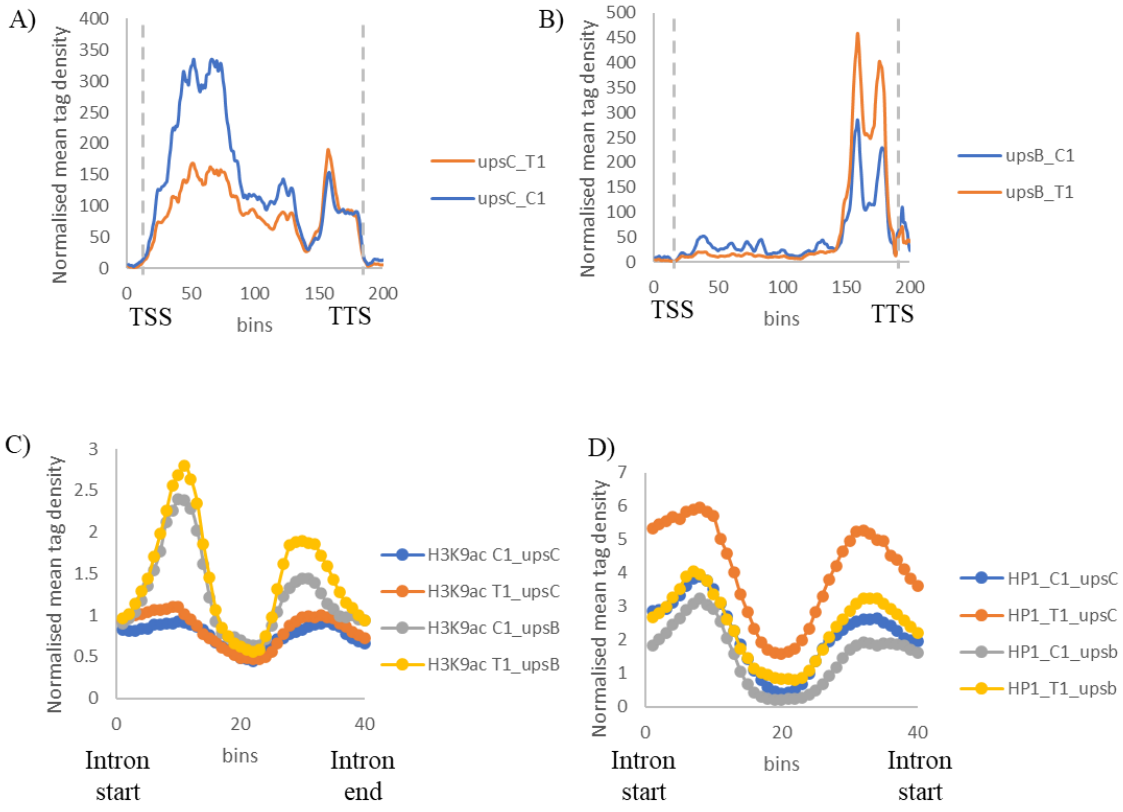


Figure 2.6: A) enrichment profile of H3K9ac on upsC *var* genes in comparison with control vs. temperature stress, B) enrichment profile of H3K9ac on upsB *var* genes in comparison with control vs. temperature stress, C) enrichment pattern of H3K9ac on *var* introns, D) enrichment profile of PfHP1 on *var* introns.

2.4.5 Intronic region of most of the var genes show an open chromatin/ euchromatic state

The striking absence of heterochromatin encouraged us to ponder the plausible mechanism of this restriction in heterochromatin spread. One argument could be that *var* introns might be nucleosome depleted regions, hence heterochromatin can't spread.

It has been earlier shown that H2A.Z marks the intergenic regions and activates transcription for different genes. H2A.Z is shown to be present in the intronic region of *var* genes (Petter et al. 2013), hence it is not an NDR. We re-analyzed the data from Bartfai et al. 2010 and found that H2A.Z is present on the intron of all *var* genes. Interestingly PfHP1 and H2A.Z have opposing

patterns of enrichment globally and locally at *var* loci (Fig. 2.7A&B). H2A.Z along with other binding proteins could be acting as an insulation for heterochromatin spread.

Also, we looked into the chromatin state of the *var* introns. The ATAC seq data from Toenhake et al. 2018, was re-visualized and we observe that the chromatin is open in the *var* introns. When this ATAC sequencing data is coupled with the strand specific RNA seq data from the same paper, it was observed that many *var* introns produce divergent transcripts (Fig. 2.7C). The association between the divergent transcription from the intron and the level of openness of chromatin is correlated.

Figure 2.7

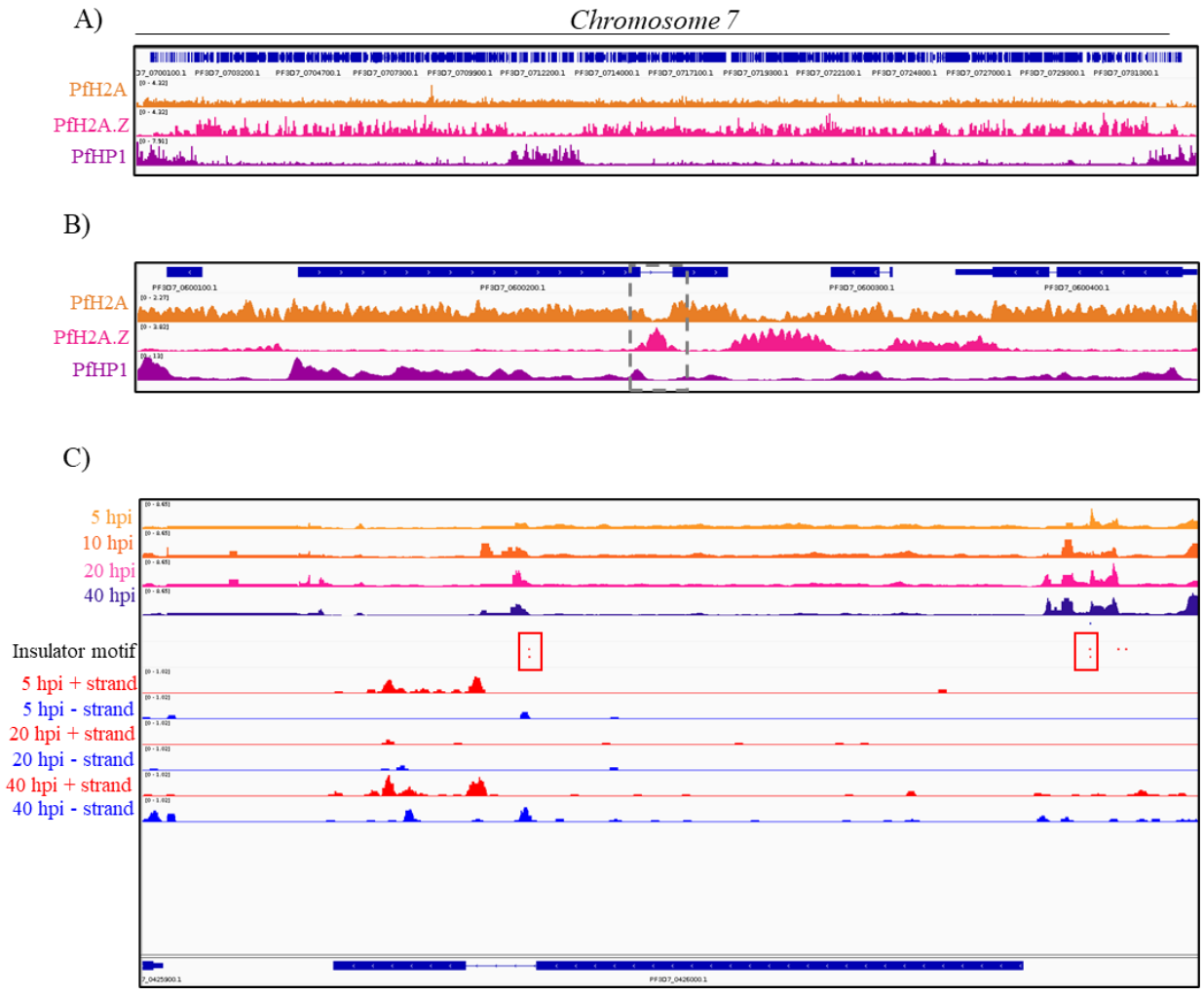


Figure 2.7: *Legend continued,* A) Global profile of PfH2A, PfH2A.Z and PfHP1, B) Profile of PfH2A, PfH2A.Z and PfHP1 on *var* loci, (intron is boxed), C) ATAC sequencing data showing chromatin state on *var* gene in different time points, corresponding strand specific RNA sequencing data is shown, the boxed region indicates the insulator like motif (PfH2A and PfH2A.Z ChIP sequencing data reanalyzed from Bartfai et al. 2010 , PfHP1 data is in-house and the ATAC sequencing and RNA sequencing data from Toenhake et al. 2018).

2.5 Discussion

Epigenetic mechanisms of *var* regulation have been studied for the past two decades. With the advantage of gene editing techniques like CRISPR-Cas9 and next generation sequencing many milestones could be achieved. Nevertheless, there were fewer attempts to comprehensively understand the regulatory network.

Here we have collectively presented various sequencing data from our own experiments and published data sets. Previously understudied H4K8ac and H3K79me2 showed an enrichment almost mimicking H3K9ac on *var* genes. We observed these three histone modifications on introns of 17 *var* genes that produce mostly sense transcript from there. However, we didn't observe the active modifications on the *var* promoters, owing to the transient nature of *var* full length transcription only during the early stage of IDC. While the *var* promoter is highly temporally regulated, the intronic promoter is almost always active during the entire IDC. There can be two role for the intronic sense transcripts;

1) The intronic transcripts might be helping in keeping *var* gene/s in a transcriptionally poised state.
2) The intronic transcription might be limiting the heterochromatin spread; it might have a role in heterochromatin maintenance rather than gene regulation directly.

The published observation that *var* intronic transcripts being incorporated to the chromatin, helps the second argument. Also, our own observation that PfHP1 boundaries are expanded and retracted in the introns correlated to the intronic transcription. Michel-Todó et al. identifies four types of pattern for heterochromatin transitions, 1) de novo formation 2) expansion or retraction from ends 3) expansion/retraction (internal) 4) localized closing/opening (Michel-Todó et al. 2023). The *var* introns can be classic examples of type 3.

There can be various factors that enable such a transition. The introns harbor insulator-like sequence motifs, these can be putative insulator proteins binding sites (Avraham et al. 2012). Insulator proteins can cause active exclusion of PfHP1 from the introns. The insulator proteins and what governs their binding needs to be seen.

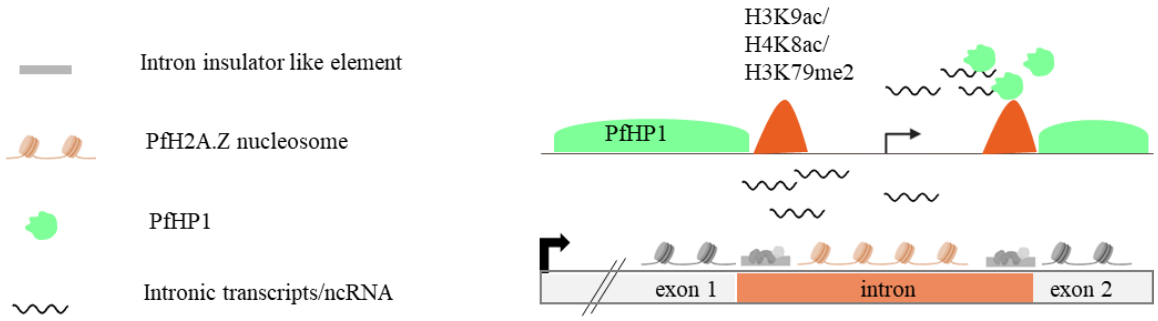


Figure 2.8: The proposed model from the results obtained in Chapter 2.

We also observed that PfHP1 and PfH2A.Z showed a striking antagonism in the enrichment saccpattern. The H2A.Z of *Saccharomyces cerevisiae* acts as an insulator from ectopic spread of the heterochromatin. PfH2A.Z could be a global regulator of heterochromatin spread. At a locus level also, we have seen that PfH2A.Z is present on the *var* introns where PfHP1 is excluded. PfH2A.Z along with other insulator proteins might be protecting heterochromatin boundaries (shown in our proposed model in Fig.2.8).

The mechanism behind heterochromatin formation and its spreading remains unexplored, uncovering it could provide valuable insights into *var* gene regulation.

Chapter 3: Heterochromatinization driven by liquid-liquid phase separation of PfHP1

3.1 Introduction

3.1.1 Heterochromatinization by HP1

Heterochromatin Protein 1 (HP1) is a multifaceted protein, which is essential for chromatin organization and genome regulation. Initially identified as key players in heterochromatin-mediated gene silencing, HP1 proteins are now recognized to participate in diverse nuclear processes, including transcriptional activation, chromosome segregation, telomere maintenance, DNA repair, and RNA splicing (Canzio et al. 2014). This functional versatility is particularly intriguing given the high sequence conservation among HP1 paralogs across various species. HP1a, a *Drosophila* HP1 paralog was the first identified HP1, showed position effect variegation (PEV), a phenomenon that is discussed in Chapter 1, section 1.2.3 (Muller 1930).

HP1 proteins play a central role in chromatin dynamics by interacting with specific histone marks, such as di- and trimethylation of lysine 9 on histone H3 (H3K9me2/3), which is critical for heterochromatin formation. Their structural architecture includes three main domains: the chromodomain (CD), which recognizes H3K9me2/3; the hinge region, which interacts with nucleic acids; and the chromoshadow domain (CSD), which facilitates dimerization and recruits ligands. Phosphorylation of serine residues in NTE (N-terminal extension) of mouse HP1α enhances its affinity for H3K9me3 (Canzio et al. 2011; Canzio et al. 2013; Canzio et al. 2014). These domains work in a modular fashion to mediate diverse interactions and functions (Fig.3.1A).

Most eukaryotes have several paralogs of HP1; despite their shared structural framework, HP1 paralogs demonstrate remarkable functional diversity. For example, human HP1α and HP1β primarily localize to heterochromatic regions, i.e. mainly pericentromeric regions and mediate

gene silencing, whereas HP1γ is associated with euchromatic regions and contributes to transcriptional elongation. This divergence is attributed to subtle sequence differences among paralogs, which influence their binding affinities and interaction dynamics. Interestingly, even though the high sequence homology amongst these paralogs, small sequence variations can lead to significant differences in their biological roles, akin to the functional divergence observed in histone variants (Fig.3.1A) (Canzio et al. 2014).Another homolog of HP1 in *Saccharomyces pombe*, Swi6 possesses the ability to gene silencing. The interaction of Swi6 with anti-silencing proteins limits the heterochromatin spread, like an insulator (Braun et al. 2011; Zofall & Grewal 2006).

There are three main parts for heterochromatin formation; namely, nucleation, spreading and maintenance. HP1 seems to be 'bridging' all the three. The nucleation is largely determined by DNA sequence and/or RNA mediated mechanisms. The RNAi (RNA interference) machinery can recruit the factors for the heterochromatin nucleation (Yamada et al. 2005). The DNA sequence-based mechanism is based on various DNA binding proteins that can recognize and bind to the DNA. The heterochromatin spreading is largely guided by the oligomerization of HP1 protein (Canzio et al. 2011). Whereas the heterochromatin maintenance is a combined effort of histone modifying enzymes, chromatin binding factors, RNAi and HP1 (Haldar et al. 2011; Yamada et al. 2005; Sugiyama et al. 2007).

3.1.2 Biochemical properties of PfHP1

HP1 is a conserved protein that is present across different eukaryotic phyla; including evolutionarily primitive protozoans like *Plasmodium*. PfHP1 (discussed in section 2.1.3) is the only identified HP1 in *P. falciparum* unlike its metazoan counterparts with multiple paralogs. Also, PfHP1 is strikingly absent from the pericentromeric region and only present on the syntenic region that contains mostly genes that show antigenic variation such as *var* genes (Pérez-Toledo et al. 2009).

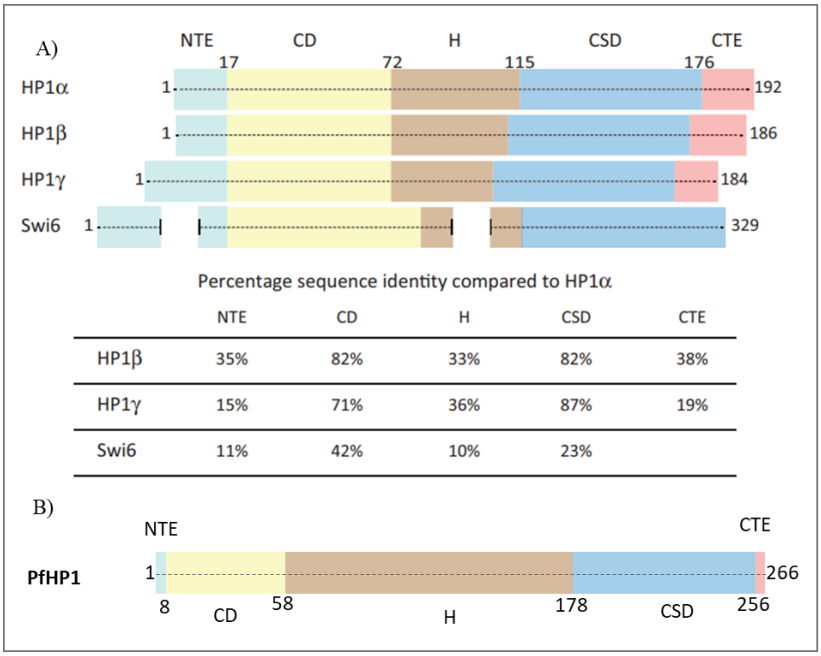


Figure 3.1: A) Domain organization and sequence similarity amongst HP1 paralogs (Image taken from Canzio et. al., 2014, Trends in Cell Biology) B) PfHP1 domain organization.

PfHP1 is a small protein with 266 amino acids that is coded from an intronless gene. Structure of PfHP1 is unavailable, but based on the sequence analysis, the CD is from amino acid 8 to 58, hinge from 58 to 178, and CSD from 178 to 256 (Fig. 3.1B) (Bui et al. 2021). The CD specifically binds to H3K9me3/2 and not any other histone modifications. *In vitro* studies show that CD is necessary and sufficient to this specific binding (Flueck et al. 2009; Pérez-Toledo et al. 2009). The PfHP1 can homodimerize, believed to be via pentapeptide PxVxL present in CSD (Pérez-Toledo et al. 2009).

The gene editing strategies on PfHP1 proved that it is essential for parasitic survival as well as transmission. Also, depletion of PfHP1 resulted in de-repression of *var* genes , indicating that PfHP1 is essential for heterochromatin formation and maintenance (Brancucci et al. 2014).

The localization of PfHP1 was found to be near the nuclear periphery as puncta or foci. The clustering of *var* genes are also shown to have a similar localization. And earlier it has been shown that H3K9me3 colocalize in *var* clusters, evidencing that PfHP1 foci could be the telomeric cluster. The domain deletion mutants of PfHP1, ΔCD, Δhinge and ΔCSD, show that only ΔCSD

shows a mislocalization, even though all three showed lethal phenotype. It is predicted that the RRKK motif in CSD could be the nuclear localization signal (NLS) for PfHP1 (Bui et al. 2021).

In silico analysis predicts potentially sumoylation, ubiquitinylation and phosphorylation could be present mainly on the hinge region of PfHP1 (Pérez-Toledo et al. 2009). So far only phosphorylation has been identified, that is catalyzed by kinase PfCK2 (*P. falciparum* Casein kinase 2). Even though several serine/threonine residues in various domains can be potentially phosphorylated, the hinge region cluster of three serines with phosphorylation in *in vitro* assay with PfCK2 and PfHP1. Interestingly, *in vivo* these phosphomutants neither affect PfHP1 localization nor the parasite survival (Bui et al. 2019). It is possible that PTMs including the phosphorylation might have a role in fine tuning PfHP1 gene regulatory function.

The current understanding of the biochemistry of the protein is limiting and future studies in this area potentially help in unraveling the mechanism of PfHP1 mediated gene silencing.

3.1.3 Liquid-liquid phase separation of HP1

3.1.3.1 Introduction to LLPS (liquid-liquid phase separation)

The major model for heterochromatinization was that the heterochromatinizing proteins like HP1 lead to chromatin compaction and bridging between the loops leading to steric exclusion of the activating factors and RNA pol II (Strom et al. 2017). However, heterochromatin loci are very dynamic and spatially confined. Liquid-liquid phase separation (LLPS) was one such testable hypothesis that accommodated the unusual behavior of heterochromatin. LLPS is a phenomenon where protein tends to demix from the cellular environment and form a separate compartment (Banani et al. 2017). One of the key hallmarks of the LLPS is the internal mixing, the physical properties of a liquid allow movement within as well as a selective permeability to external molecules. The internal pool of macromolecules can give rise to a distinct chemical environment inside and thereby this compartmentalization or partitioning can selectively regulate what's inside (Erdel et al. 2020).

Biomolecules in a solution interact with each other and solvent molecules in such a way that reduces the free energy of the system. By laws of thermodynamics, to achieve equilibrium any system will tend to move towards minimum free energy, by maximizing entropy in a given condition. Generally, this means molecules will be uniformly distributed in the solution in monomeric form or in small size complex to maximize entropy. If molecules can engage in a more energetically favorable condition with a condensed solution than in a dilute solution, the extra energy output can compensate for the entropic penalty due to clustering of biomolecules (Banani et al. 2017).

Multivalency is another factor that drives towards the condensate formation. The proteins remain soluble until they reach the solubility limit, where it starts to phase separate. Multivalent interactions can lead to large oligomer formation that decreases the solubility by reaching the threshold faster and hence phase separate. The entropic penalty for this clustering is lower than the cost for confining each macromolecule individually. The proteins with the multiple modular domains or disorder regions can form multiple weak ionic or Van-der-Waals interactions with other proteins, DNA or RNA , thus undergo phase separation (Banani et al. 2017).

Reversibility is another hallmark of phase separation, where the macromolecules can assemble and disassemble. Phase separation is a concentration-dependent assembly of the protein; hence the local concentration of the protein can act as a buffer for the formation and size of the droplets. The threshold needed for the phase separation can also be altered. For example, PTMs alter protein valency and intrinsic solubility. Also, RNA/DNA can also affect the protein solubility, hence modulating the assembly (Banani et al. 2017). Phase separation properties are also determined by environmental conditions, including temperature, pH, salt concentration and salt type.

3.1.3.2 History of LLPS

The spatiotemporal regulation of biochemical reactions is challenging in a densely packed cellular environment. Hence, there are specialized organelles that contain the reaction components in a compartment separated by lipid bilayer. However, these membranes are impermeable to most of the biomolecules.

In 1830s, the first membraneless organelle was identified, called the nucleolus (Pederson 2011). Since then, decades of FRAP (Fluorescent recovery after photobleaching), revealed the high dynamicity of these compartments. But there was no plausible explanation for this observation until in 2009 Tony Hyman and Cliff Brangwynne found that P granules in *Caenorhabditis elegans* exhibited liquid-like properties *in vivo* as well as *in vitro* (Brangwynne et al. 2009). In 2012, Michael Rosen and colleagues also identified polymeric transitions of multivalent proteins that lead to LLPS (Li et al. 2012). Since then many membraneless organelles were identified as phase separated including nucleolus, Cajal bodies, stress granules etc. (Banani et al. 2017; Wang et al. 2021a).

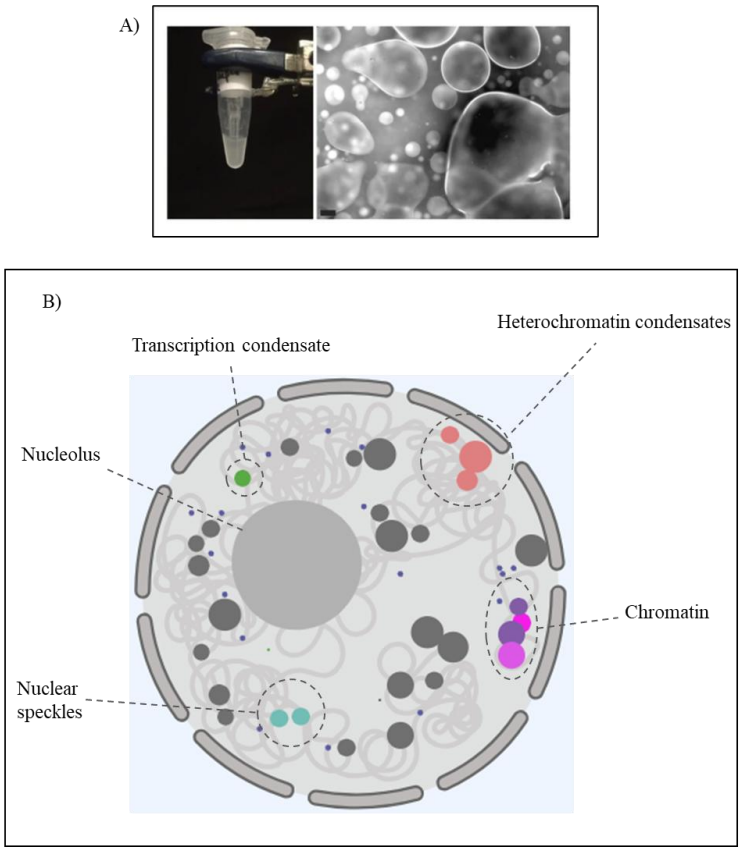


Figure 3.2: A) First observation of LLPS of HP1 *in vitro* (taken from Larson et al, 2017, Nature) , B) The schematic representing different membrane less organelles found inside a nucleus (made with biorender.com)

3.1.3.3 Homologs of HP1 show LLPS

HP1 exhibits liquid–liquid phase separation (LLPS) properties, which are critical for the formation of heterochromatin in human, fruit fly and yeast (Sanulli et al. 2019; Strom et al. 2017) (Fig.3.2A). The binding of human HP1α onto DNA facilitates the droplet formation *in vitro.* Phosphorylation of NTE also results in a similar effect (Larson et al. 2017). While HP1α-DNA interaction can drive the phase separation, binding of H3K9me3 is believed to increase the local concentration of the protein that facilitates the LLPS. The hinge region of HP1α, was necessary and sufficient for *in vitro* LLPS. Other paralogs of HP1 in humans, HP1β and HP1γ show little to no ability to phase separate (Keenen et al. 2021).

However, HP1β does show phase separation property, forming a complex with H3K9me3 nucleosomal array and H3K9 methyltransferase SUV39H1 or scaffolding protein TRIM28. This was also proposed as a mechanism for how H3K9me3 regulates heterochromatin formation (Wang et al. 2019b).

The phase separation model proposed for *Drosophila* HP1a suggests that the condensates act as a nucleation center for the heterochromatin formation, and as it matures, more droplets will fuse to make a heterochromatin compartment (Strom et al. 2017).

In contrast to the essential role of HP1-driven phase separation in other systems, mouse HP1α has shown only *in vitro* phase separation property; *in vivo,* the ability was much weaker. Also, HP1α was not necessary for the stability of the chromocenters (which are heterochromatin domains). Instead of a continuum of intermediate stages in the heterochromatin assembly, they observed a digital model for the heterochromatin compaction where it was either zero or one. This led to the conclusion that HP1-mediated LLPS was not important for heterochromatinization in mice (Erdel et al. 2020).

Swi6, the *S. pombe* HP1, has been shown to have phase separation properties and characterized for its biochemical nature (Sanulli et al. 2019). Recently Clr4 (H3K9 methyltransferase) along with Swi6 was also shown to have a phase separation property, revealing mechanism of

heterochromatin formation in *S. pombe*. The ncRNAs from centromeric regions modulate the phase separation properties (Kim et al. 2024).

In summary, despite the contradicting views on LLPS of HP1, HP1-mediated phase separation is a fundamental process that underpins the organization and function of heterochromatin. The ability of HP1 proteins to undergo LLPS facilitates the formation of stable nuclear compartments. Ongoing research continues to uncover the intricate biophysical mechanisms governing these processes and their implications for chromatin dynamics and gene regulation. Understanding these relationships will be crucial for elucidating the roles of heterochromatin in health and disease.

3.1.4 Phase separation in gene regulation

3.1.4.1 Phase separation of chromatin

The eukaryotic genome gets compacted in different size scales. The basic unit of chromatin is ~146 bp DNA, wrapped around the nucleosome. The histone octamer of nucleosomes may as well contain a diverse set of PTMs as discussed in the section 1.2.4. These PTMs often identify chromatin as active/inactive or heterochromatin/euchromatin. But the exact mechanism of this organization was unknown.

Gibson et al. 2019, demonstrated that nucleosomal arrays can phase separate and the linker DNA length, H1 histone or acetylation can give distinctive physical properties to the droplets and hence distinct identities. The acetylated histones caused dissolution of the droplets, whereas in presence of reader protein BRD4 (bromodomain 4), it formed droplets that are different from the other chromatin condensates. Hence it is speculated that different histone modification and histone variants may give rise to heterogeneity in liquid phases and consequent functions (Gibson et al. 2019).

3.1.4.2 Phase separation of other heterochromatin proteins

Mammalians have a constitutive as well as a facultative heterochromatin, usually by H3K27me3 mark. Polycomb repressive complexes (PRCs) are regulators of gene silencing and

heterochromatinization in the facultative heterochromatin. PcG (polycomb group) of proteins can form two complexes PRC1 and PRC2 (Obuse & Nakayama 2025). CBX2 (Chromobox 2), a protein that belongs to canonical PRC1, has shown to have LLPS properties. The proposed mechanism explained the compartmentalization of repressive loci and exclusion of active transcriptional machinery (Plys et al. 2019).

X chromosome inactivation (XCI) is another example of phase separation mediated compartmentalization that helps in selective regulation. Xist, a lncRNA that is central to the initiation of XCI, along with its partner binding proteins PTBP1, MATR3, TDP-43, and CELF1 self-assemble into a phase separated condensate. This process is essential for gene silencing in inactive X chromosomes (Pandya-Jones et al. 2020).

MeCP2 (Methyl CpG binding protein 2) is a component of constitutive heterochromatin that contains DNA methylation. This protein can form dynamic condensates and mutations in MeCP2 disrupts its LLPS property. The same mutation is associated with neurodevelopmental disorder Rett syndrome. The phase separation is imperative for the selective partitioning of heterochromatin and euchromatin (Li et al. 2020).

3.1.4.3 Phase separation in transcription

The transcription is an essential part of gene expression. The transcription factors, coactivators and enhancers are a major part of transcription machinery along with RNA polymerase. The transcription happens in clusters or foci (Cisse et al. 2013). These foci show properties of liquid condensates. The pluripotency factor OCT4, Yeast transcription factor GCN4 and disordered region in CTD (C-terminal domain) of RNA pol II can undergo phase separation (Boija et al. 2018).

There also exists super enhancers (SE), that are clusters of enhancers that harbor bulk of transcription machinery to produce robust transcriptional output in cell identity genes. These centers show a transition between assembly and disassembly. BRD4 and MED1 two transcription coactivators are shown to be driving the condensate formation of SEs and it has properties of LLPS (Sabari et al. 2018).

In addition, RNA itself can also regulate transcription condensate formation, using a feedback loop. During the transcription initiation the RNAs can induce condensate formation , but during late transcription/ elongation the burst of transcripts dissolve the condensates (Henninger et al. 2021).

The various examples of the condensate formation in the nucleus helps us to imagine an organelle with numerous membraneless organelles that undergo phase transitions in a dynamic manner (Fig. 3.2B).

3.2 Scope of this Chapter

PfHP1 binds to H3K9me3 and is exclusively present on the telomeres and CVGs. Even though it is known that PfHP1 represses *var* genes, the mechanism of gene silencing and heterochromatin formation is poorly understood. LLPS of PfHP1 is a possible mechanism of gene silencing.

In this Chapter we are testing the hypothesis whether PfHP1 can undergo LLPS *in vitro* or not. If yes, then what is the nature of this phase transition and is it a tunable process. The clues from this *in vitro* biochemical study will let us unravel the mechanism of *var* gene silencing.

We have standardized conditions for the PfHP1 purification and droplet formation assay. This study is also one of its kind, because phase separation of proteins is relatively under explored in malaria biology. Due to its rapidly changing dynamics and tunability the phase separation could be a general mechanism adapted by the parasites to help in disease progression.

3.3 Materials and Methods

3.3.1 Sequence analysis and alignment

The fasta sequence of HP1 homologs were obtained from NCBI database and the sequences were used to generate percent identity matrix, from sequence alignment algorithm of ClustalΩ (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) (Sievers et al. 2011). The percentage identity

matrix was further represented as a heat map using Morpheus (Broad institute) (https://software.broadinstitute.org/morpheus/).

3.2.2 Phase separation prediction

For the prediction of IDRs, IUPred (https://iupred2a.elte.hu/plot_new) and PONDR (https://www.pondr.com/) were used (Mészáros et al. 2018; Obradovic et al. 2003). CIDER was used for the plot charge distribution and hydropathy https://pappulab.wustl.edu/CIDER/) (Holehouse et al. 2017). The propensity to phase separation was predicted using PSPredictor (http://www.pkumdl.cn:8000/PSPredictor/) (Chu et al. 2022).

3.3.3 Site directed mutagenesis

The mutations generated were E6A/E7A/E8A, K17A/K18A/K19A/K20A, W29A, L254A, L254D in PfHP1-GFP (pET15b backbone) and PfHP1 (in pET28a+ backbone). In addition, Ser 33 and Ser206 were mutated to cysteine in PfHP1 (pET-Ht backbone) for the fluorescent labelling purpose. The backbone plasmid was amplified with the primers with the mutations (see the list of primers is provided *Appendix* Table 3). Pfu turbo (Agilent) enzyme was used in the PCR amplification. The PCR reaction was then subjected to DpnI (NEB) digestion (37^{o}C, overnight) to remove the parent vector. From the reaction 2µl was transformed into DH5α and plated on a luria agar plate with appropriate antibiotic selection. The clones were screened and sequence verified for the mutation.

3.3.4 Recombinant protein expression and purification

The clones generated for the recombinant protein expression were, PfHP1 with GFP and His-tag in pET15b, PfHP1, PfH2A and PfH2A.Z with his tag in pET28a+ backbone.

The protocol for recombinant protein expression and purification of PfHP1 was discussed in section 2.3.1. The same protocol was followed for PfHP1-GFP (and all the mutant PfHP1 made in the backbone) except we did not add β-mercaptoethanol in the buffers. After the purification, proteins were concentrated using amicon 10kDa and 30kDa membrane filters (millipore) for PfHP1 and PfHP1-GFP respectively. The concentrated proteins were desalted/buffer exchanged

(20mM HEPES pH 7.4, 300mM KCl, 5% glycerol, 2mM EDTA) using PD-10 column (Cytiva). The proteins were further concentrated if needed using amicon centrifugal filters.

The proteins were run on 10% polyacrylamide gel for SDS-PAGE and stained with Coomassie brilliant blue, to check the purity (Fig. 3.5A). The proteins were quantified using absorbance at 280 nm using a spectrophotometer. The freeze-thaw cycles were avoided and the experiments were performed with freshly desalted proteins.

The expression conditions used for PfH2A and PfH2A.Z were the same as that of PfHP1. The cell pellet was resuspended in sonication buffer (50mM Na_3PO_4 pH 8.0, 500mM NaCl, 20mM imidazole,10% glycerol, 3mM β-ME, 1mM PMSF, 0.5% TritonX100, 0.5mg/ml lysozyme; 40 ml sonication buffer for the pellet from 1L culture) and sonicated for 40 minutes at 60% amplitude, 2 sec ON and 5 sec OFF cycles in a probe sonicator (Vibra-cell, Sonics and Material Inc.). The lysate was spun at 12000 rpm, 4°C, 20 min. The supernatant was allowed to bind to Ni-NTA beads for 2 hours in rotation at 4°C. The unbound lysate was removed by gravity flow. The beads were washed with 5 CV (column volume) of wash buffer (40mM Imidazole, 50mM Na_3PO_4 pH 8.0, 500mM NaCl, 3mM β-ME, 0.1mM PMSF). The Proteins were eluted in 1 CV of elution buffer (300mM Imidazole, 50mM Na_3PO_4, 500mM NaCl, 3mM β-ME, 0.1mM PMSF) for a total of 5 elutions. After the purification, proteins were concentrated using amicon 10kDa membrane filters (millipore). The concentrated proteins were desalted/buffer exchanged (20mM HEPES pH 7.4, 300mM KCl, 5% glycerol, 2mM EDTA) using PD-10 column (Cytiva). The proteins were further concentrated if needed using amicon centrifugal filters. The purity of the protein was checked using SDS-PAGE. And quantified using absorbance at 280 nm in a spectrophotometer or a nanodrop. The protein was stored at 4°C for short term and -80°C for long term storage.

3.3.5 Fluorescence labelling of the protein

S33C and S206C mutants of PfHP1 were labeled with fluorophores under denaturing conditions at pH 7.4. The proteins were labelled by Alexa Fluor-488 dye (C5-maleimide), in a 2:1 molar ratio (dye:protein). The reaction mixtures were stirred for 2-3 hours in the dark at room temperature. After the labelling the excess dye was removed by PD-10 columns (20 mM HEPES

pH 7.4, 300mM KCl, 5% glycerol, 2mM EDTA), and concentrated using a 10 kDa MWCO Amicon membrane filter. The concentration of the labeled protein was estimated using absorbance at 280 nm and 495 nm (Alexa Fluor 488 C5-maleimide ε495 = 72,000 M-1cm-1). 1% of the labelled protein was added to the unlabeled PfHP1 for phase separation experiments.

3.3.6 YOYO-1 labelling of the DNA

The 601 Widom sequence was amplified from a pGEM 601 vector (Addgene) using specific primers. The PfDNA (*P. falciparum* DNA) is an 800 bp stretch amplified from the *P. falciparum* genome (primers are given in the *Appendix* Table 4). The 601 Widom DNA was labelled with YOYO-1 in a ratio of one molecule of the dye per 10bp of DNA.

Component	Amount
Template	1µg
5X T7 reaction buffer	10µl
10mM NTP mix	10µl
RNase OUT	1µL
0.1M DTT	2.5µL
T7 polymerase	1µl
Nuclease free water (NFW)	Make up to 50µl

Table 3.1: Protocol for the IVT reaction.

3.3.7 In vitro transcription of RNA

The template for the IVT is PCR amplified. The GC-rich RNA was coded from the RUF6 gene in the *P. falciparum*, the primer used contained overhangs with T7 promoter and terminator for the IVT reaction. The template for AU-rich RNA was difficult to amplify because of the extreme AT-richness. Overhanging oligos were used to assemble the template for the AU-rich RNA (oligos given in the *Appendix* Table 5). The PCR products were cleaned up using Phenol:chloroform:isoamyl alcohol method. The IVT reaction was set up according to the protocol given in Table 3.1. The reaction was kept overnight at 37ºC in a PCR machine. Next day

DNaseI (Promega) treatment at 65ºC for 10 minutes was performed. The reaction was stopped by adding DNaseI stop solution. The IVT RNA was cleaned up using precipitation by 2.5V 100% ethanol and 0.1V ammonium acetate. To ensure that the obtained product is RNA and not DNA, RNaseA (Himedia) digestion was performed. The purity of the RNA was inspected by running the samples with or without RNaseA on a 1.3% agarose gel. The quantification was done by a nanodrop instrument.

3.3.8 Droplet formation assay

Initial standardization and droplet assay experiments were done with Sandeep K Rai and Prof. Samrat Mukhopadhyay, IISER Mohali.

Before the droplet assay proteins were spun at a high speed to remove any possible debris or precipitate. In a clean PCR tube, mix 10 µl of buffer (20mM HEPES, pH 7.4) and 10 µl of PfHP1-GFP/mutants (30 µM stock) or PfHP1 (100 µM stock), this is the basic droplet assay reaction mix. Gently pipette and avoid bubbles. Observe under a microscope by putting 5µl of the reaction on a coverslip of 1.5 mm thickness. The microscopes used were ZEISS LSM 980 Elyra 7 or ZEISS LSM 710 confocal microscope. The proteins were imaged under 488 nm laser, 63x oil objective.

The other conditions tested were different temperatures and salt concentrations. Also effect of 601 Widom DNA, PfDNA, RNA (IVT AT-rich and GC-rich ncRNAs), PfH2A, PfH2A.Z, H3K9me3 peptide, H3K9ac peptide (Anaspec) and PEG8000 were tested. The test component was added to the pre-formed PfHP1 droplets. The amount of protein and other components used are mentioned along the figures.

For the droplet area quantification Fiji (Image J) software was used. The images were converted to gray scale (8 bit) first and changed the threshold setting to ‘intermodes’. The image was then analyzed using the ‘analyze particles’ option to obtain droplet area, number and mean intensity.

3.3.9 Fluorescence recovery after photobleaching (FRAP)

FRAP experiments were performed on ZEISS LSM 980 Elyra 7 super-resolution microscope (IISER Mohali). For all FRAP experiments, GFP tagged PfHP1 was used. FRAP measurements were performed for at least three independent samples. A region of interest (ROI) of 1µm diameter was defined within a selected droplet, and a same size ROI was defined outside as background. The ROI within the droplet was bleached using a 488 nm laser; five frames were taken before the bleaching and recovery was followed for 70 seconds. The recovery of the bleached spots was recorded using the in-built ZEN blue 3.2 (ZEISS) software. The five different data sets were background corrected using the outside ROI, normalized with standard deviation, and plotted using Origin 2020b.

3.3.10 Saturation concentration estimation

Prepare 500µl droplet assay reaction in a 1.5 ml tube with 15µM PfHP1 protein in 20mM HEPES pH 7.4, 150mM KCl, 2.5% glycerol and 1mM EDTA buffer. Incubate the reaction at room temperature for 5 minutes and spin at 16000 rpm, 25ºC for 35 minutes. Estimate the volume of supernatant after the spin and back calculate the volume of pellet or dense phase. Measure the absorbance at 280 nm of the light phase/supernatant, this shall give you the saturation concentration or $C_{sat.}$ Similarly measure the dense phase absorbance by resuspending the pellet in 80µl urea 8M (in 1X TBS). The concentration is calculated from the absorbance value. The experiment is performed in triplicates.

3.3.11 Turbidity assay

For the turbidity assay with DNA, 20µl reaction was prepared with varying concentration of PfDNA (1, 2, 5, 10, 20, 40, 80, 100 ng/µl) and 15µM PfHP1 in 20mM HEPES pH 7.4, 150mM KCl, 2.5% glycerol and 1mM EDTA buffer. The reaction was incubated at room temperature for 5 minutes and measured absorbance at 340 nm. All the experiments were performed with triplicates.

3.3.12 Pelleting assay with the nuclear extracts

The nuclear extracts were prepared as explained in section 2.2.3. The nuclear extract was dialyzed into a suitable buffer (50 mM Tris, 75 mM KCl, 10% glycerol, 1mM PMSF, 1 mM DTT, 1.5mM $MgCl_2$, 1mM EDTA) for the assays. The assay was performed by following the protocol from Lyon et al. 2023. Briefly, purified PfHP1-eGFP (30µM) was mixed with nuclear extract (100 ng/µl) in a reaction buffer (20mM HEPES pH 7.4, 150mM KCl, 2.5% glycerol and 1mM EDTA). Incubate the mix at room temperature for 15 minutes and spin at 14000 rpm, 10 minutes. The pellet and supernatant were separated and processed for mass spectrometry (protocol given in section 2.3.4).

3.4 Results

3.4.1 PfHP1 sequence analysis reveals that it has propensity to phase separate

The HP1 homologs of HP1 have the ability to phase separate. Multiple sequence alignment of PfHP1 with other HP1 homologs showed that there isn't much sequence identity (Fig.3.3A), even though the higher eukaryotic homologs are very similar to each other. The CD and CSD of PfHP1 had stretches of amino acids that are conserved across the species (Fig. 3.3B). Indicating that despite the low sequence identity amongst all the homologs they might have a common scheme for gene regulation via mechanisms such as phase separation.

Disorder region analysis particularly useful in identification of physicochemical basis for phase separation. IUPred and PONDR, two different algorithms have predicted the hinge region of PfHP1 to have disordered regions (Fig. 3.4A&B). In addition, PONDR prediction shows NTE, CSD and CTE to harbor stretches of disordered regions. CIDER, is another software that analyses the sequence of intrinsically disordered proteins. The hydropathy plot that shows hydrophobicity of the residues shows that PfHP1 has hydrophobic interaction throughout. It is important to note that hydrophobic interaction often drives phase separation (Fig. 3.4C). The NPCR (net charge per residue) shows alternating blocks of positive and negatively charged residues along the protein; these might help in phase separation by electrostatic interaction (Fig. 3.4D). Even though a neutral

NCPR is best for phase separation, the blocks of charged residues and interactions might result in an overall stable condensate.

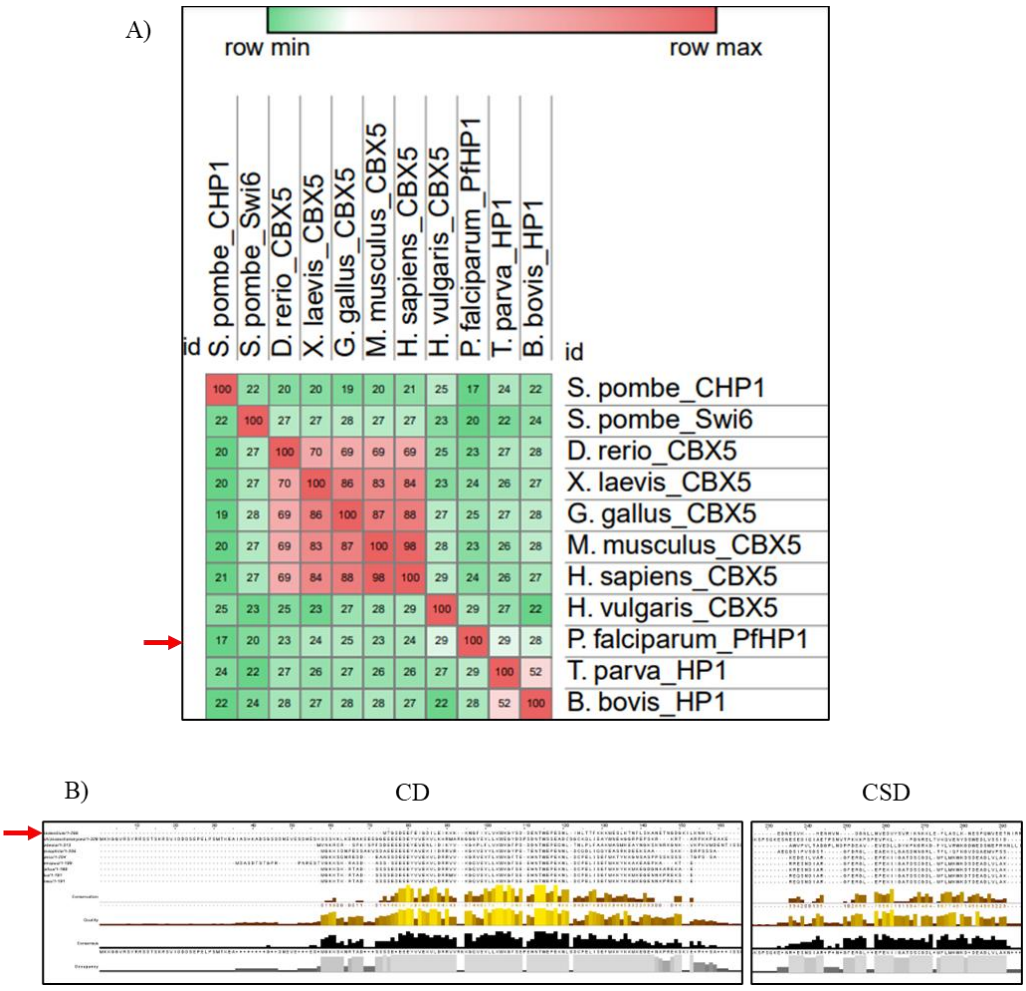


Figure 3.3: A) Percentage identity matrix of HP1 homologs, B) Multiple sequence alignment of HP1 homologs showing conserved residues in CD and CSD, PfHP1 is highlighted by a red arrow.

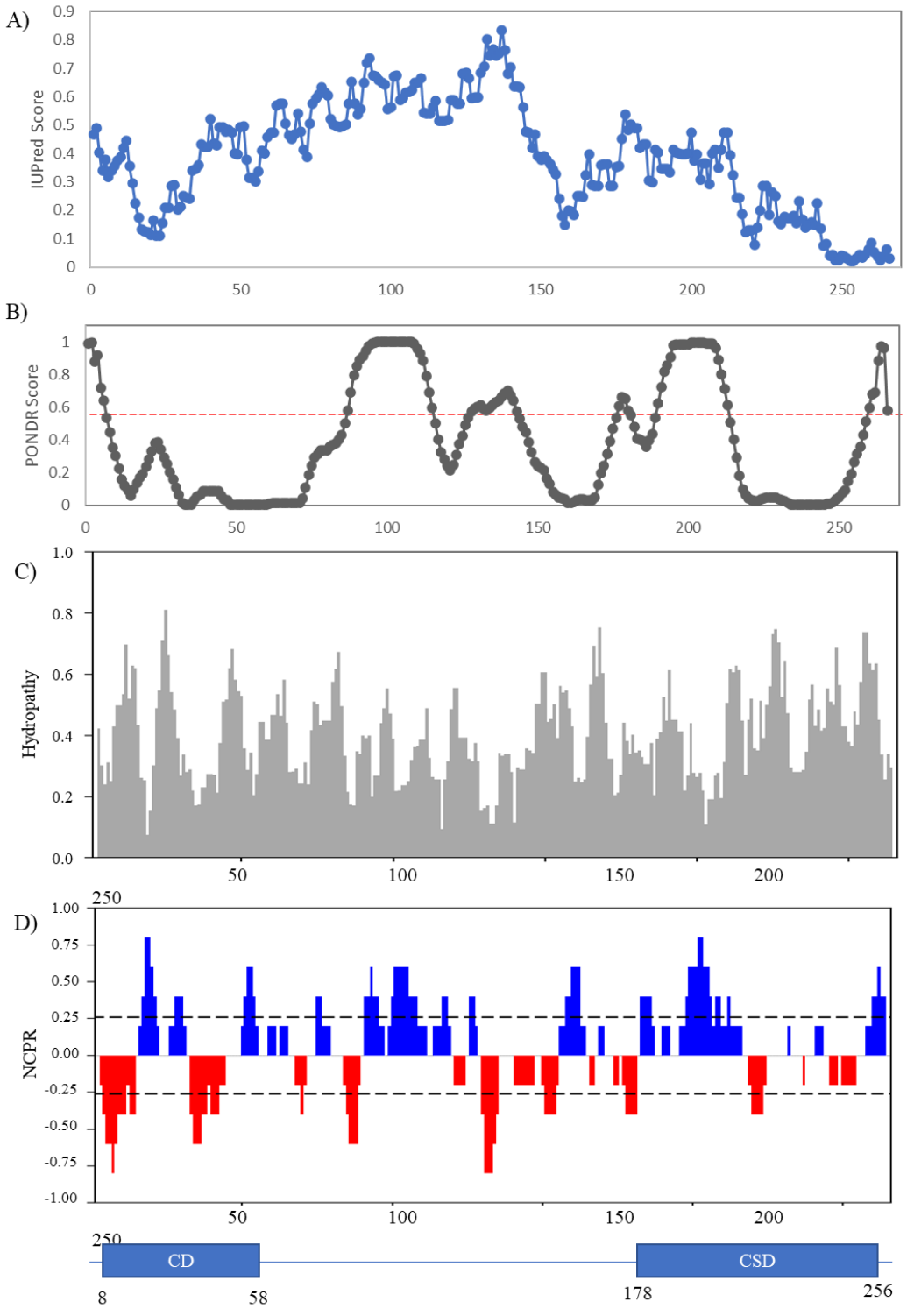


Figure 3.4: A) Disorder prediction of PfHP1 by IUPred, B) Disorder prediction of PfHP1 by PONDR, C) Hydropathy plot and D) NCPR of PfHP1 using CIDER online tool.

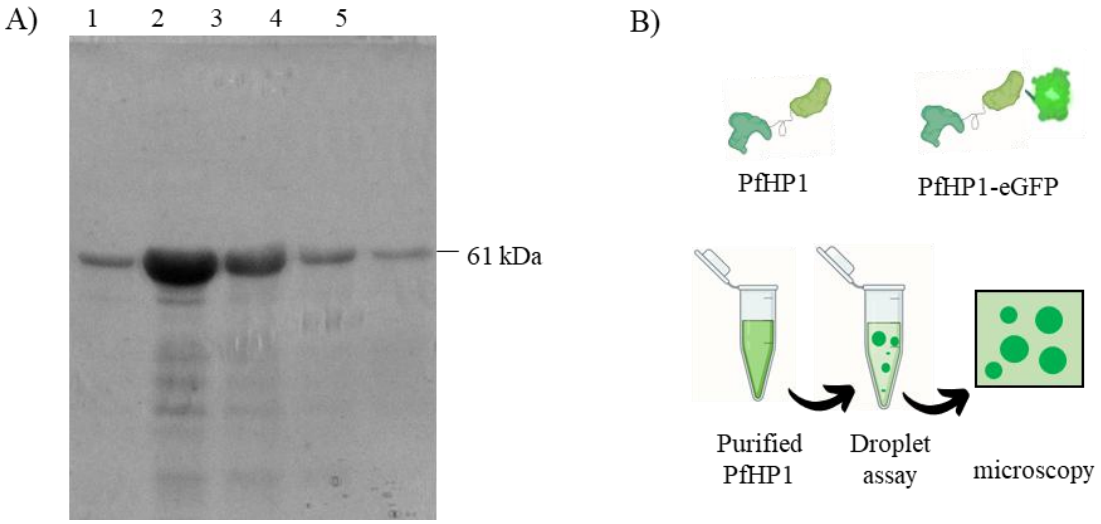


Figure 3.5: A) Protein purification profile of PfHP1-eGFP; 1- elution 1, 2- elution 2, 3- elution 3, 4- elution 4, 5- elution 5, B) Schematic showing a brief droplet assay protocol.

3.4.2 PfHP1 undergoes LLPS in vitro

Next, we test whether PfHP1 can undergo LLPS *in vitro*. We performed a classical phase separation assay where a purified PfHP1 protein was observed under varying concentrations and appropriate buffer conditions (physiologic pH and salt concentration) under microscope for droplet formation. PfHP1 with or without GFP tag (purification profile: Fig. 3.5A, experimental design: Fig. 3.5B) showed phase separated droplets under microscope (Fig. 3.6A&B). The assembly was concentration dependent and C_{sat} was estimated; for PfHP1-eGFP C_{sat} was 11.54µM and for PfHP1 C_{sat} was 37.4µM, however adding a crowding agent like PEG8000 improved the C_{sat} to 28.5µM (Fig. 3.6 C&D). The saturation concentrations indicate that GFP tagging has slightly improved the LLPS property. The interaction between hinge and CTE keeps human HP1α in an auto-inhibited state that can't form any oligomer. It is possible that PfHP1 might also have such states, the GFP tagging at the C-terminal might relieve such auto-inhibitory states and promote the oligomerization and hence LLPS.

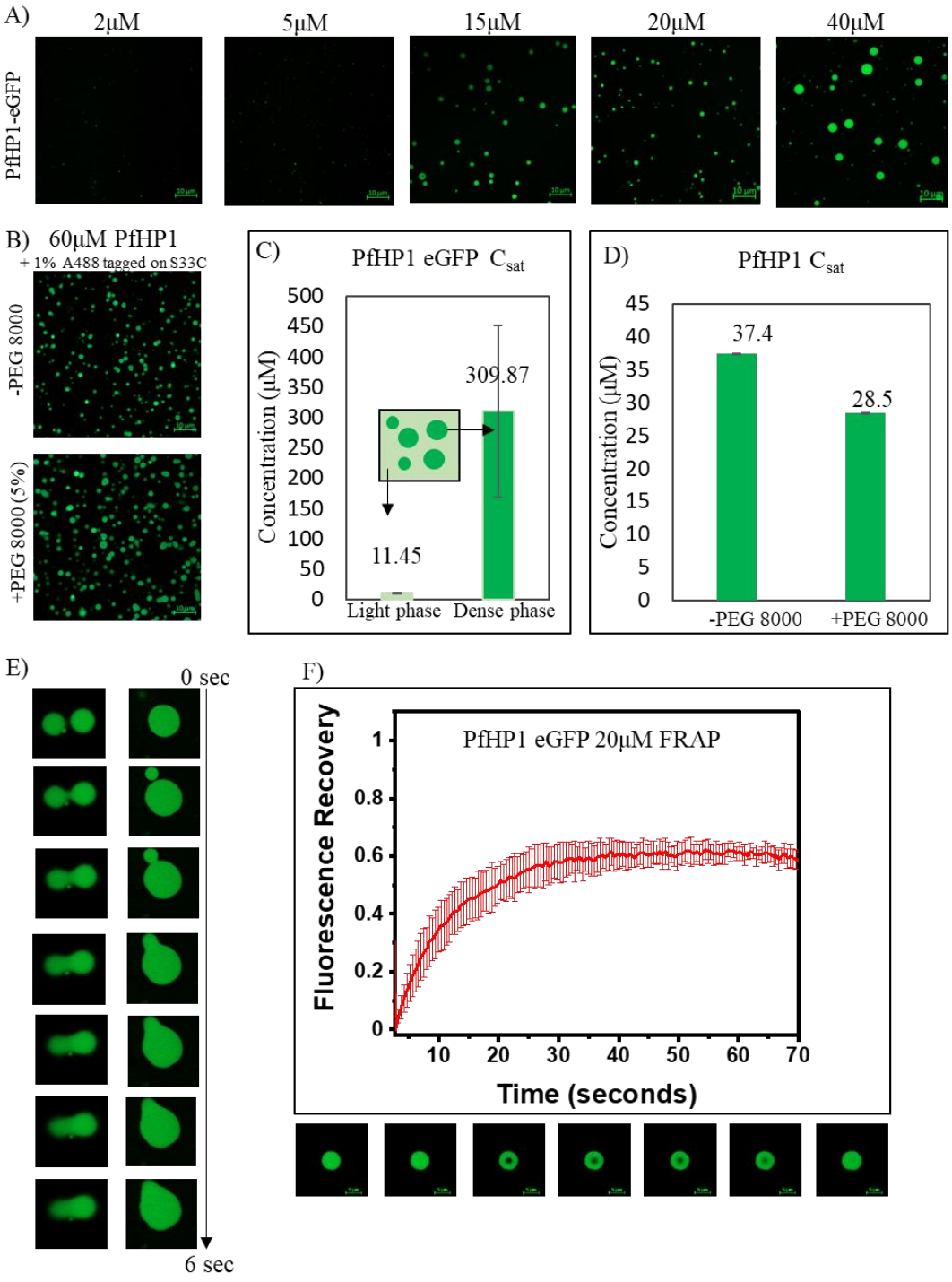


Figure 3.6: A) Droplet assay showing concentration dependent assembly of PfHP1-eGFP, B) Droplet assay of PfHP1, C) Saturation concentration estimation of PfHP1-eGFP, D) Saturation concentration of PfHP1 with and without PEG8000, E) Time lapse of PfHP1 eGFP droplets fusing F) FRAP assay with PfHP1-eGFP.

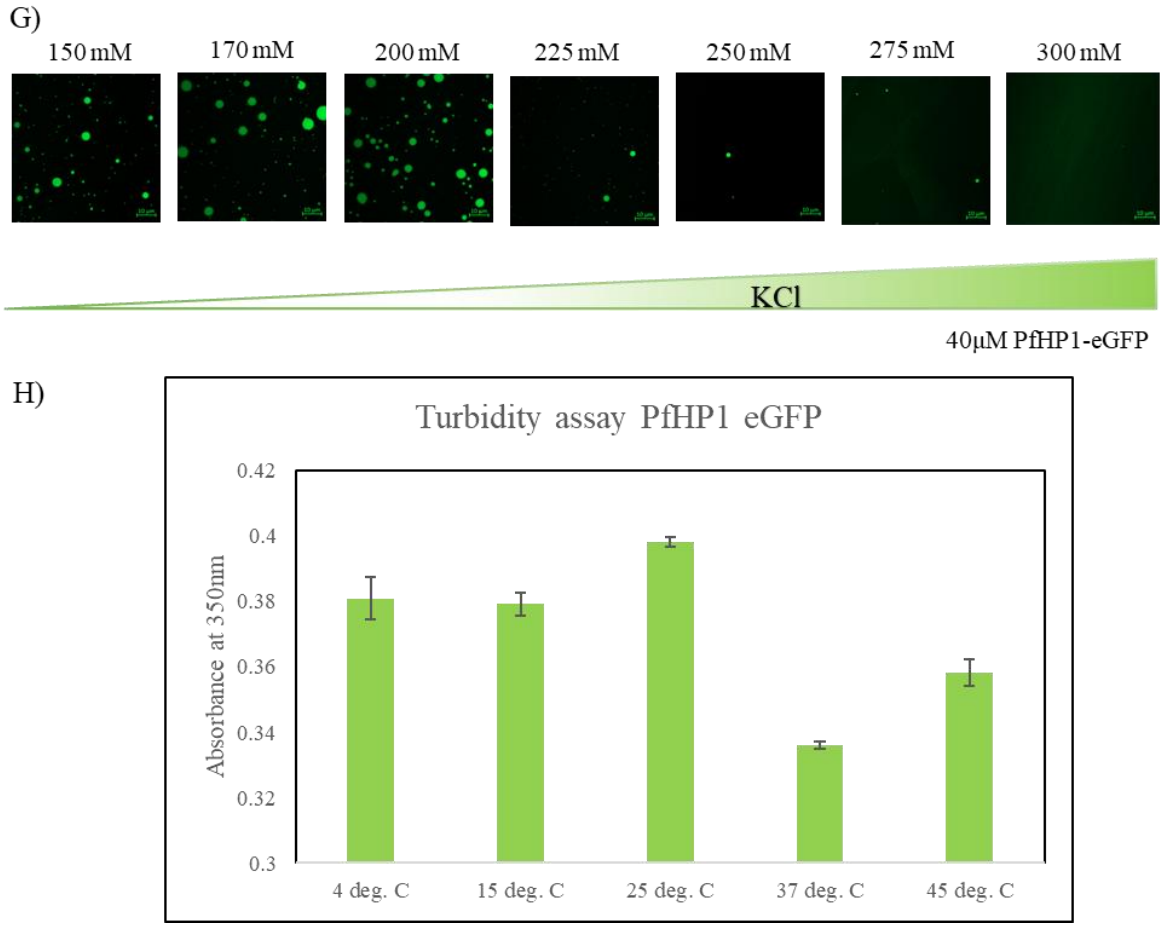


Figure 3.6: G) salt concentration dependence of phase separated droplets, H) Turbidity assay of PfHP1-eGFP droplets incubated in different temperatures.

The PfHP1 showed properties of a liquid droplet such as droplet fusion, surface wetting and dripping (Fig. 3.6E). Also, FRAP assay showed a fast recovery, indicating that within the droplet dynamics is fast like a liquid (Fig. 3.6F). Hence, we, for the first time, are showing the evidence for the PfHP1 phase separation.

We also checked the dependence of this phase separation on salt concentration and temperature. The salt concentration beyond 200mM affected the ability to phase separate, indicating the importance of electrostatic interactions in driving phase separation (Fig. 3.6 G). The room temperature was optimal for the phase separation, with a slight decrease in turbidity (indicative

of phase separation) at lower temperatures. The higher temperatures including physiological temperature, 37^{o}C decreased the droplet formation considerably (Fig. 3.6 H). Proteins can phase transition following a decreasing temperature or increasing temperature known as LCST (Lower critical solution temperature) and UCST (Upper critical solution temperature) respectively. PfHP1 might be following a phase transition that is UCST.

3.4.3 PfHP1 phase separation is modulated by nucleic acids

From the above results we assumed that PfHP1 will undergo LLPS in a highly crowded environment of nucleus, sensing local changes in local microenvironment. The nucleic acids are a major part of nuclear composition; hence DNA and RNA can regulate this phase transition. The HP1 homologs directly bind to DNA and compacts it in a cooperative manner. Also, DNA can drive phase separation of HP1 according to earlier studies. On the other hand, RNAs have a role in heterochromatin nucleation and formation. However, the role of RNAs in HP1 phase separation has not been studied.

We observe that increasing concentration of PfDNA (800 bp, amplified from *pfhp1* gene) promotes droplet formation of PfHP1-eGFP (Fig. 3.7 A,C & E). The droplets have peculiar morphology where the GFP fluorescence is unequally distributed (Fig. 3.7A). This could be indicative of sub domain formation by DNA inside the droplets. Also, PfDNA-PfHP1 have slightly faster and better FRAP recovery (Fig. 3.7B). A similar trend was observed with PfHP1 and 601 Widom sequence DNA (146 bp) (Fig. 3.7D). The DNA also reduces the critical concentration needed for the droplet formation (estimated by C_{sat}, Fig. 3.7G). Further we labelled the 601 DNA with YOYO-1 nucleic acid dye and demonstrated that DNA actually goes inside the droplets, and not just acting as a crowding agent. (Fig. 3.7F). Hence, it is safe to assume that the genomic DNA and clonally variant can be confined in a spatially separated droplet *in vivo*.

The DNA is the genetic material and is present in fixed copy numbers in a cell, hence, it is difficult to imagine a scenario where DNA regulates the phase separation in a dosage dependent manner. Whereas, RNA, especially ncRNAs are present in abundance and can affect the phase separation of various proteins in different systems.

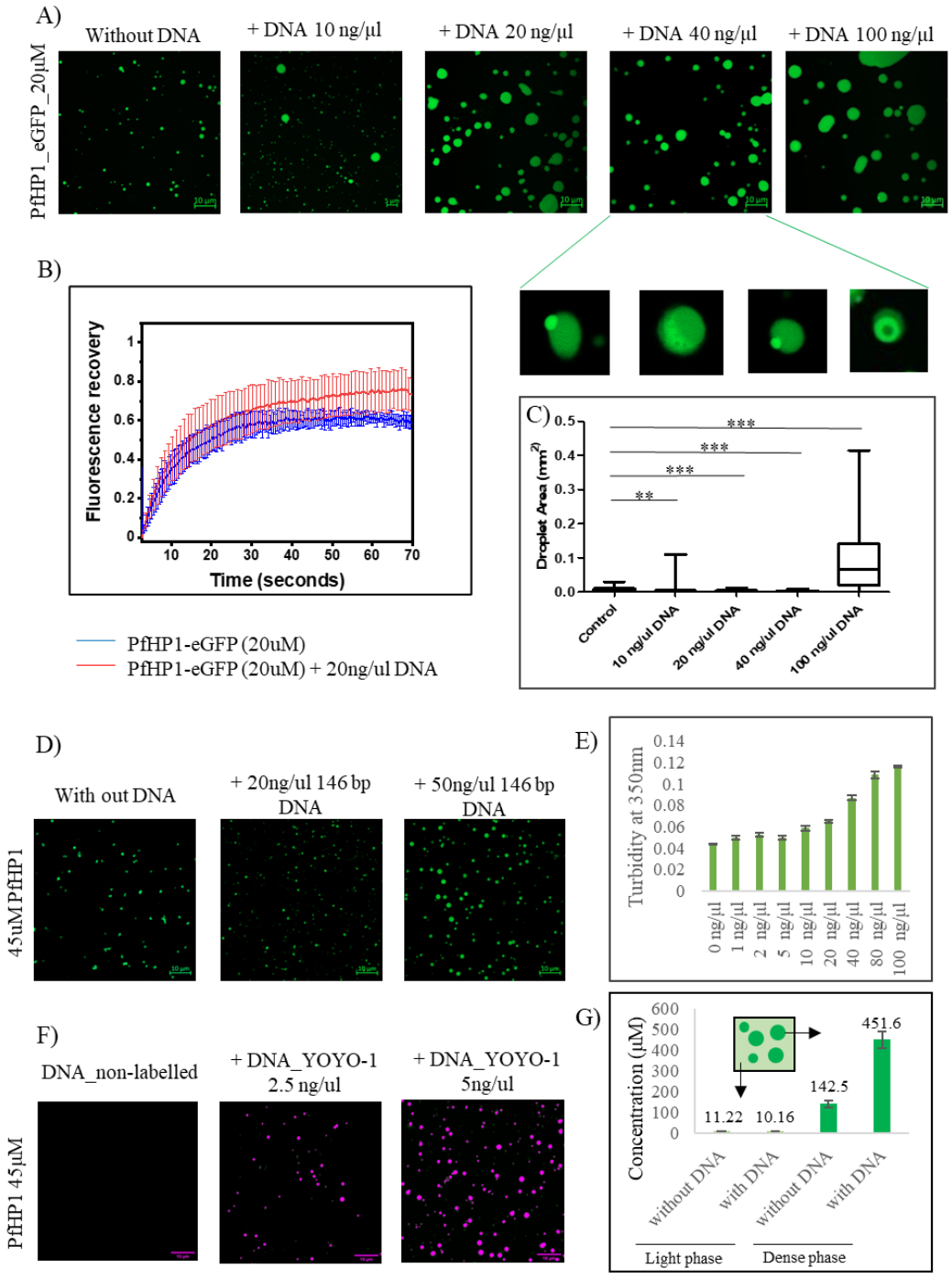


Figure 3.7: A) Droplet assay with PfHP1-eGFP with PfDNA B) FRAP of droplets with or without PfDNA, C) Quantification of droplet area when PfDNA was added, D) Droplet assay with PfHP1 (1% Alexa 488 S33C) and 601 Widom sequence, E) Turbidity assay with PfDNA and PfHP1-eGFP, F) Droplet assay with PfHP1 and YOYO-1 labelled 601 Widom DNA, G) Saturation concentration estimation of PfHP1-eGFP with or without DNA.

The clonally variant loci and proximal regions in *P. falciparum* harbors a lot of ncRNAs. Most of these RNAs are AU-rich because of the high AT-richness of the genome. But, strikingly, the RUF6 (also discussed in section 2.1.6) that are found proximal to upsC type *var* genes are GC-rich (~53%). We used IVT AU-rich RNA from the sub-telomeric regions and GC-rich ncRNA to test a dosage dependent effect on PfHP1 phase separation. The AU-rich RNA causes dissolution of PfHP1 droplets in a concentration dependent manner (Fig. 3.8A&B). Nevertheless, GC-rich RNA didn't affect the droplet formation in the concentration range that we tested (Fig. 3.7A). Also, at a saturating concentration poly U RNA caused complete dissolution of PfHP1-eGFP droplets (Fig. 3.8C). The AU-rich and GC-rich ncRNAs did not affect the PfHP1-eGFP droplets in the concentration range we tested (Fig. 3.8D). These results indicate that PfHP1 droplet formation might respond to the changing local concentration of ncRNAs in the nucleus and hence the number and size can be regulated.

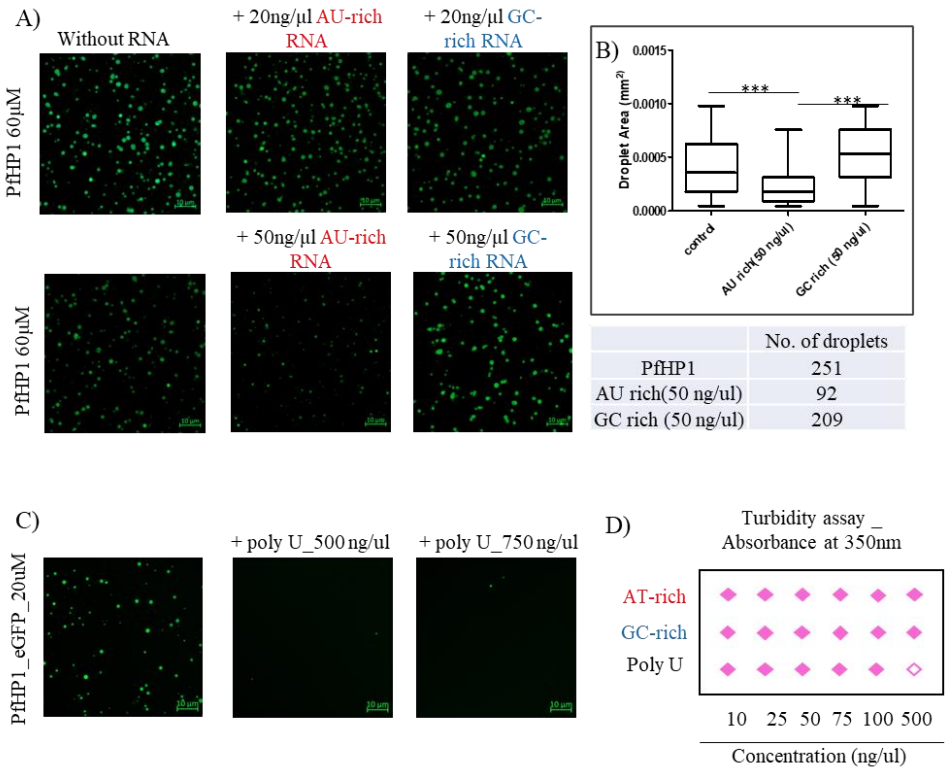


Figure 3.8: A) Droplet assay with PfHP1 (1% Alexa 488 S33C) and AU-rich and GC-rich ncRNA B) Quantification of droplet assay by droplet area and number of droplets C) Droplet assay with PfHP1-eGFP and poly U RNA, D) Turbidity assay with PfHP1-eGFP and RNAs.

3.4.4 PfH2A.Z acts as an antagonist of PfHP1 phase separation

PfHP1 and PfH2A.Z having an opposing pattern of enrichment on chromatin, makes us think about the antagonistic nature of this relationship. We hypothesized that H2A.Z might be hindering ectopic spread of the heterochromatin into the euchromatin. Earlier data from *S. cerevisiae* suggest a similar model (Meneghini et al. 2003).

The LLPS could also be inhibited by H2A.Z that may confine the droplets in a spatially restricted manner rather than engulfing the genome non-specifically. We used PfH2A.Z in the droplet assay while PfH2A serves as a control. The PfH2A.Z caused complete dissolution of the droplets while PfH2A caused slight reduction in the droplet formation (Fig. 3.9A). The buffer for PfH2A.Z and PfH2A had high salt concentrations (300mM) that might interfere with the LLPS. However, PfH2A.Z caused complete dissolution rather than a reduction in droplets indicating that PfHP1 and PfH2A.Z might indeed be the antagonist that controls the heterochromatin-euchromatin boundary. However, these experiments lack a context of nucleosomes, PfH2A.Z is always present with PfH2B.Z. And as known PfHP1 constantly interacts with the nucleosome. We recognize that this *in vitro* system lacks the sophistication but the biochemical nature of PfHP1 and PfH2A.Z can still be further explored in this context.

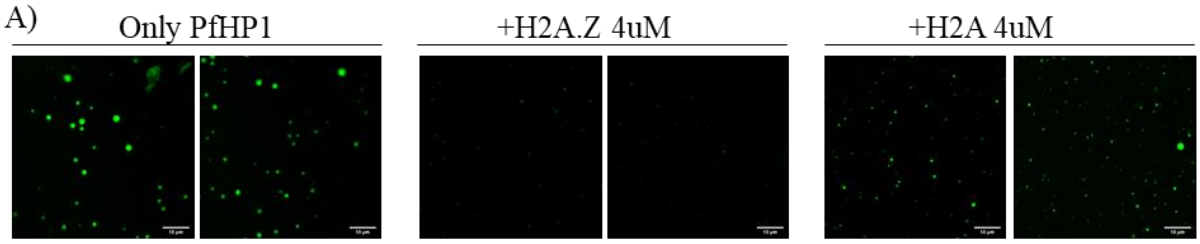


Figure 3.9: A) Droplet assay with PfHP1-eGFP (15 μM) and PfH2A.Z/PfH2A.

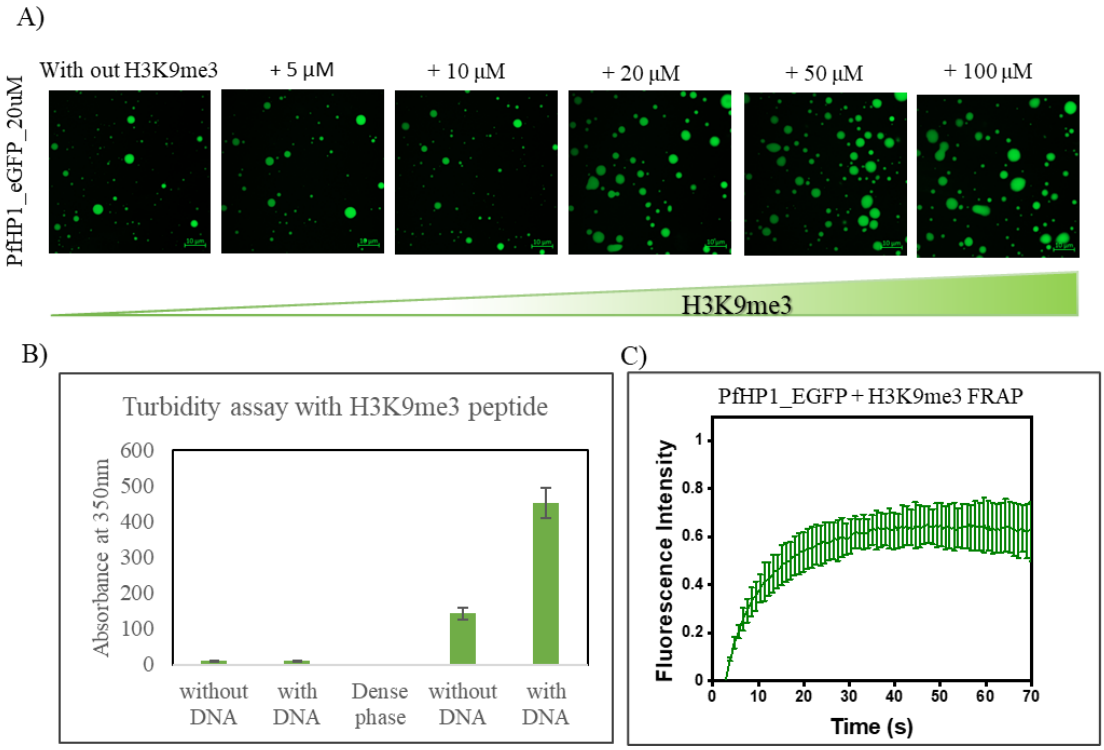


Figure 3.10: A) Droplet assay, B) Turbidity assay and C) FRAP with PfHP1-eGFP and H3K9me3 peptide.

3.4.5 PfHP1 phase separation selectively partitions nuclear proteins

The PfHP1 binds to only H3K9 methylated nucleosomes. There is a strong correlation between H3K9me3 and PfHP1, almost like a proxy for each other. In the nucleus PfHP1 forms puncta-like structures which could be phase separated droplets. But the composition of this punctum is not very well known.

It is intuitive that H3K9me3 nucleosomes will be a part of the droplet. To test this hypothesis, we use H3K9me3 peptide and observe that equimolar concentration of PfHP1 and H3K9me3 gives rise to more droplets, after that it hits a saturation with increasing peptide concentration (Fig. 3.10A&B).

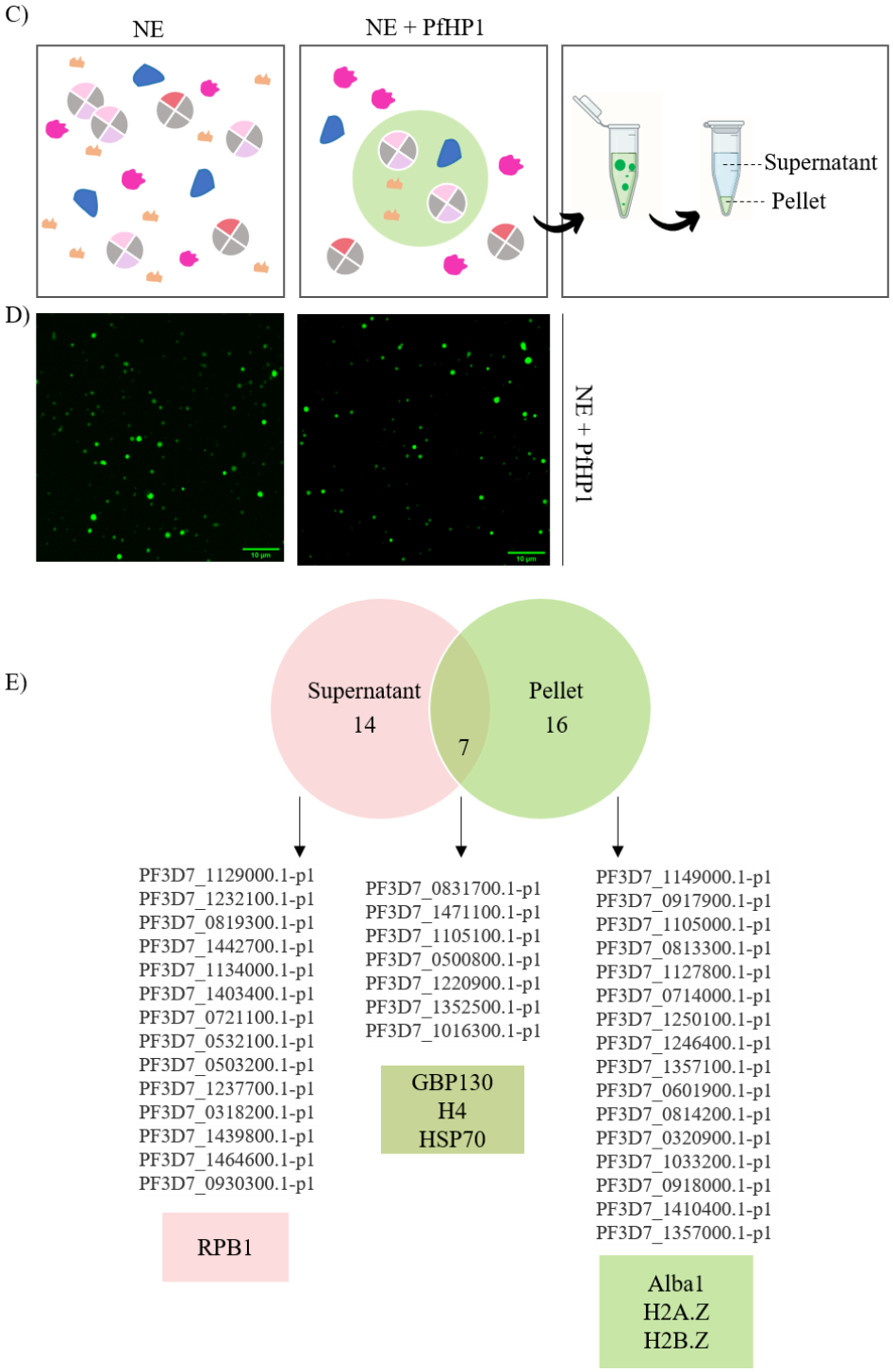


Figure 3.10: C) Schematic showing experimental design of the assay D) Droplet assay with *P. falciparum* nuclear extract E) Venn diagram showing proteins that are present in the supernatant and pellet.

To identify other proteins inside the droplet, we did a pelleting assay with nuclear protein extract. We observed that PfHP1 undergoes phase separation with nuclear extract (NE) (Fig. 3.10D). These condensates were separated from the soluble fraction by centrifugation at high speed (Fig. 3.10C). The proteins inside the droplets are selectively partitioned and the supernatant contains other proteins.

Some of the classical histones and histone variants PfH2A.Z and PfH2B.Z were present inside the droplet. Even though an activating function is attributed to the PfH2A.Z/PfH2B.Z nucleosome (discussed in section 2.1.4 and 3.3.4), they interact with PfHP1. Another interesting hit was Alba 1 (Fig. 3.10E and *Appendix* Table 6). Alba (Acetylation lowers binding affinity) proteins are an archaeal protein family consisting of RNA/DNA binding proteins. The PfAlba1-4 form a complex and bind to the sub-telomeric region. While PfAlba-1 is attributed to have functions in mRNA homeostasis, PfAlba2-4 have role in virulence gene regulation (Goyal et al. 2012). Selective partitioning of Albas with PfHP1 droplet could be an intriguing aspect of the *var* gene regulation.

The supernatant fraction had RNA Pol II subunit RPB1, which sits with the hypothesis that PfHP1 droplets can cause steric exclusion of transcription machinery (Fig. 3.10E and *Appendix* Table 7). More replicates will possibly give more insights into the understanding of the composition of PfHP1 mediated compartmentalization. The Sir2A, Sir2B, various methyl transferases, PfSIP2, ApiAP2-HC are all shown to be associated with heterochromatin or *var* repression (Discussed in Chapter 2 introduction), hence it is intuitive to assume them to be a part of one large PfHP1 droplet where anything positively regulate *var* gene expression is excluded.

3.4.6 Point mutations in NTE and CD disrupts the phase separation property of PfHP1

Despite the poor sequence homology between PfHP1 and HP1 homologs from other species, many residues in the chromodomain are conserved. Functional mutations in *S. pombe* Swi6, the closest well-characterized ortholog of PfHP1, have been reported and some of them cause disruption of LLPS. We hypothesize that mutations in Swi6, i.e. LoopX (mutation in the ARK loop that causes disruption of oligomerization of the protein), CageX (W104A), AcidX (E(74-80)A; both the mutations hampers binding to H3K9me3) and DimerX (L315D, mutation that disrupts Swi6 homodimerization), may have similar functional consequences in PfHP1 (Fig.

3.11C) (Canzio et al. 2011, 2013) . By sequence alignment we identified similar residues on PfHP1 and generated these mutations using site directed mutagenesis; E6A/E7A/E9A , K(17-20)A, W29A and L254A/D were the mutants (Fig. 3.11A). The sequence analysis of wild type and mutant PfHP1 predicted that E-A and K-A wouldn't phase separate, whereas W-A and L-D would (Fig. 3.11C). We did not observe an ARK loop in the CD of PfHP1 while analyzing the sequence. However, there was a stretch of four lysines that may have a role in multivalent interaction such as oligomerization or DNA binding because of the charge.

The purified mutant proteins were observed for the condensate formation *in vitro*. Among the four mutations we generated, E-A was in NTE, K-A and W-A was in CD and L-A/D was in the CSD of PfHP1(Fig. 3.11B). We observed that the AcidX mutant E-A did not form any droplets in the conditions where the wildtype does (Fig. 3.12A). However, by addition of PfDNA or PEG8000 there was a recovery in the droplet formation, indicating that C_{sat} might be higher for the mutant, hence lowering the critical concentration by adding DNA helps in the formation of the condensates. K-A mutation where the stretch of lysines were mutated showed the most dramatic effect in the droplet assay. The mutant protein did not form any droplets, even the DNA or PEG could not recover the phenotype (Fig. 3.12A). The CageX mutant, W-A showed no LLPS in the normal concentration range, however similar to E-A lowering the critical concentration by PEG8000 partially recovered the phenotype (Fig. 3.12A). For the dimerX mutant we generated two variations in the mutation; one is L-A where the leucine is mutated to alanine which is similar and non-polar; second is L-D, where non-polar is mutated to polar charged residue. We assumed that the second mutation might cause a much drastic phenotype. But contrary to our expectation both mutations did not show any effect on phase separation (Fig. 3.12B).

A)

Swi6
PfHP1

```
MKKGGVRSYRRSSTSKRSVIDDDSEPELPSMTKEAIASIKADSGSSDNEVESDIESKSSS      60
-----------------------------------------------------------      0

                  AcidX            LoopX      CageX
KKLKENAKEEEGGEEEEEDEYVVEKVLKHRMARKGGGYEYLLKWEGYDDPSDNTWSSEAD      120
------------MTGSDEEFEIGDIL--EIKKKKNGFIYLVKWKGYSDD-ENTWEPESN      44
             .::*: : .:*  .: :* .*: **:**:**.*  :***. *::

CSGCKQLIE-------AYW---NEHGGRPEPSKRKRTARPKKPEAKEPSPKSRK--T--D      166
LIHLTTFKKKMESLKTNFLSKANETNGDGKILKNH-ILAPT---QEDDSIKSKGRSSLAP      100
    . : :        :    ** .*  :  *.:    *.    :: * **:   :

EDK HDKDSNEKIEDVNEKTIKTADKSQEEFNENGPPSCQPNGHIESDNESKSPSQKESN      225
RRKMSRKSLTNKLF--NKKNLSLSDNSLKKSDFFDNFSVKHFNHVNDGNLLN------VF      152
. *   *. .:*:*  *:*.:.::*:* :: :*:.  * : :.*::..*  :        :

ESEDIQIAETPSN-VIPKKKPSPEVPK---LPDNRE--LTVKQVENYDSWEDLVSS-IDT      278
DVYSVRIKNKKLCFLASLKNESPQWVCETNIRRTGILNIKVNDFKRYVRRKKSSRGNRIV      212
:  .::* :.  : ::  *: **:  :   :  . .  :.*:.:.*   :.   .   .

                                    DimerX
IERKDDGTLEIYLTWK-----NGAISHHPSTITNKKCPQKMLQFYESHLTFRENE      328
IKNLHNVGDELYISVIHNINNKEIHSLYPSKVIEYIYPQELLNFLLSRLRYRTA-      266
*:. .:   *:*::       :   * :**.: :   **::*:*  *:* :*
```

B)

C)

Mutation	Swi6 residue	Possible function	PSP Score	Phase separation prediction (PSP)
WT	-	-	0.571	Yes
E(6,7,9)A	E(74-80)A (AcidX)	Binding to H3K9me3	0.3639	No
K(17-20)A	R93A-K93A (LoopX)	Oligomerization of PfHP1/ DNA binding	0.3663	No
W29A	W104A (CageX)	Binding to H3K9me3	0.5239	No
L254A/ L254D	L315D (DimerX)	Homodimerization of HP1	0.5735	No

Figure 3.11: A) Pairwise sequence alignment of PfHP1 and Swi6 B) Domain organization of PfHP1 with mutated residues C) Summary of mutations and there corresponding function along with phase separation prediction using PSPredictor.

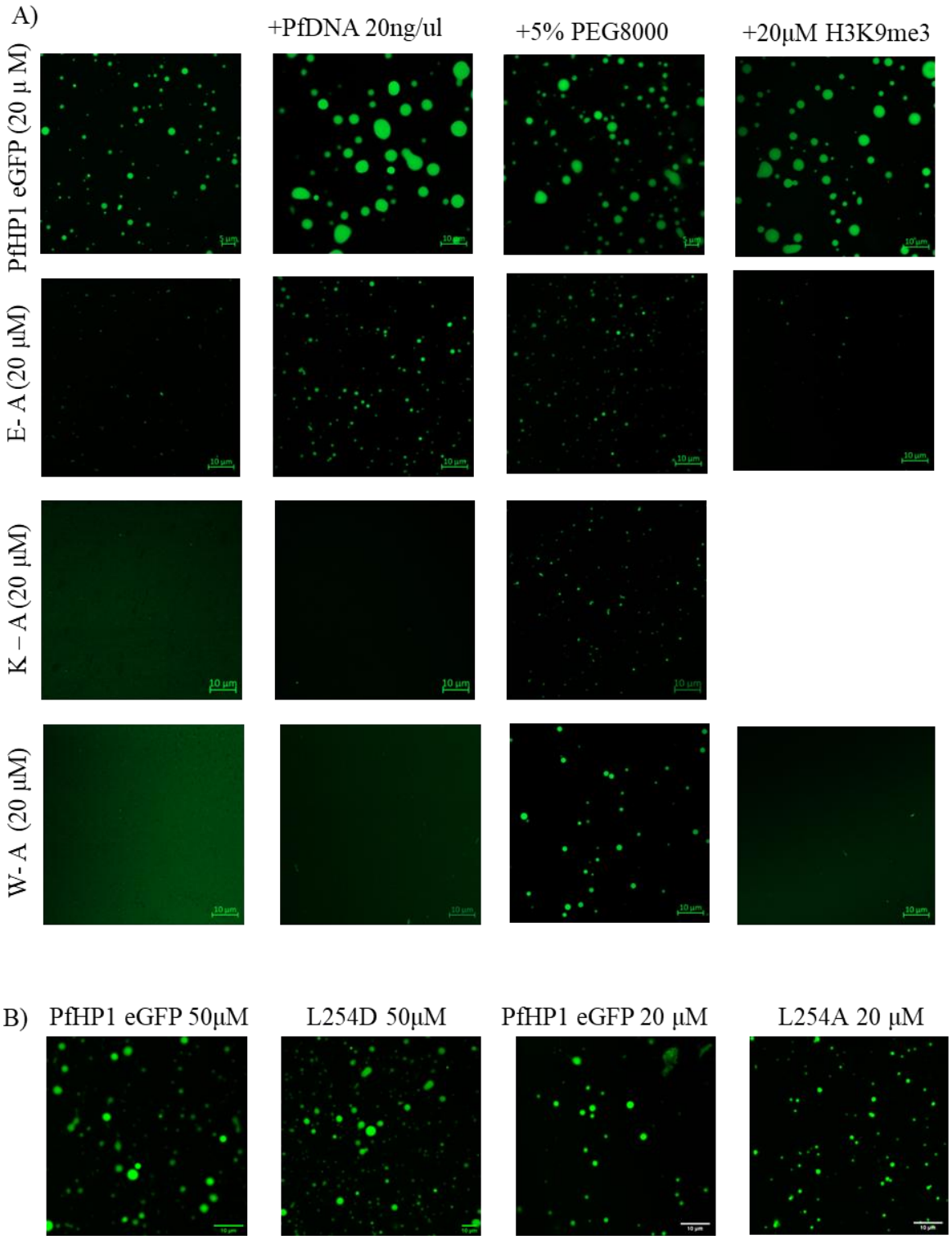


Figure 3.12: Droplet assay with PfHP1-eGFP mutants with mutations in A) NTE and CD, B) CSD; PfHP1 wildtype was used as a control in all assays.

3.5 Discussion

Plasmodium falciparum var genes form clusters near the nuclear periphery and are kept under suppression, except for the one which is expressed. However, the molecular mechanism underlying this geographically distinct loci coming together is poorly understood. Heterochromatin in *P. falciparum* is marked by H3K9me3 and bound by PfHP1 and shows a punctate localization in the nucleus. These puncta are positionally stable and don't show a complete disassembly throughout the IDC. However, the *var* genes do get expressed stage-specifically. Unlike the constitutive heterochromatin made by HP1 of higher eukaryotes, PfHP1 shows a rather dynamic and tunable heterochromatin. The LLPS of PfHP1 is a mechanism that can fit in all these aspects of gene regulation.

We have demonstrated that PfHP1 forms LLPS droplets *in vitro.* These droplets could be reminiscent of the PfHP1 mediated clustering and compartmentalization of *var* genes and other regulatory factors. The droplets are also tunable as they respond to ncRNAs and PfH2A.Z, in a dosage dependent manner. The PfHP1 droplets are highly dynamic and shows fast recovery within msec when 'FRAP'ed. The dynamicity of the droplets will let free diffusion of large macro molecules like DNA, RNA or other proteins.

The ncRNA are highly abundant in *P. falciparum*. As discussed in Chapter 2, the *var* introns itself produces the sense and antisense non-coding RNAs. The telomeric region also produces TARE (Telomere associated RNA elements) lncRNAs. Some of these RNAs are shown to integrate into chromatin; indicating the possible role of these ncRNAs in heterochromatin nucleation and formation. The concentration ranges that we tested showed poly U and AU-rich RNA dissolve the droplets whereas the GC-rich RNA did not show any significant difference. There can be two explanations to this observation; 1) all RNAs buffer the phase separation in a dosage dependent manner 2) GC-rich and AU-rich RNA might be differentially regulating the phase transition.

In the first scenario AU-rich and GC-rich RNA both will regulate the LLPS depending on their local concentrations inside the nucleus. At lower concentrations, RNA might help in heterochromatin assembly because the negatively charged RNA and the positive charge on PfHP1

might be favorable for the assembly. As the concentration of RNAs increases, the negative charge outcompetes and causes a repulsion that results in dissolution/disassembly of the droplets. A complete phase diagram with changing concentration of ncRNAs and PfHP1 might help to understand this phenomenon better.

The second scenario is that the AU-rich and GC-rich ncRNA might have different effects on heterochromatin assembly. The HP1 homologs have shown to interact with various RNAs via hinge region. However, there is no report of any HP1 homolog being able to recognize RNA sequence or AU-rich vs. GC-rich sequence. In the absence of a RNA recognition domain in PfHP1 it is safe to assume that even though it can bind to RNA, differentiating them based on sequence is unlikely.

The AU-rich RNAs are mostly produced from the telomeric or sub-telomeric region and GC-rich are produced from the RUF6 genes in the internal heterochromatin islands. Because of this spatial separation they might have differential concentrations that might locally dissolve the droplets. This theory supports the notion that *var* genes are regulated by its ups type, upsC might be regulated by GC-rich ncRNA and upsA/upsB might be regulated by AU-rich ncRNA.

We found that PfDNA promotes condensate formation. The *P. falciparum* DNA is mostly AT-rich (~80%) and the nucleosome wrapping is loose making it a more open chromatin. Hence the DNA might be accessible for the binding of PfHP1. These can be nucleation sites for the heterochromatin formation. Exploring the dynamics of PfDNA and PfHP1 could provide valuable insights into the steps involved in heterochromatin formation

Chapter 4: AT-rich DNA and PfHP1 driven DNA compaction

4.1 Introduction

4.1.1 DNA compaction by proteins

The extremely long DNA is compacted and condensed inside the nucleus in a microscale. Phase separation of chromatin binding proteins is emerging as a mechanism for this unusual condensation. The nature of this condensation involves heterotypic polymers and the polymer physics is applied to understand this process better. There are two polymer models proposed for the co-condensation of DNA and protein. The first one is BIPS (bridging-induced phase separation) where the DNA-protein interaction is driving the phase separation rather than protein-protein interaction. The DNA-protein interaction will act as a nucleation center for the bridging of two distinct parts of DNA and hence looping (Fig. 4.1). The second model is SIPS (self-association-induced phase separation), where protein-protein multivalent interactions will lead to the droplet formation. The protein polymer will bind to the DNA and form co-condensates (Park et al. 2024).

The classic example of BIPS is cohesin; it is a protein that helps in genome organization via forming DNA loops via loop extrusion mechanism (Ryu et al. 2021). The examples for SIPS include FUS condensates and mediator complexes that help in formation of transcription condensates (Cho et al. 2018; Renger et al. 2022).

HP1 shows multivalent interaction with itself and nucleosomes (Canzio et al. 2014). In addition, HP1 can also interact with bare DNA (Larson et al. 2017). It is suggested that HP1 phase separation can be both types BIPS or SIPS. The initial reports of phase separation suggest that HP1α (human) and HP1a (*Drosophila*) both follow SIPS (Larson et al. 2017; Strom et al. 2017). But a completely contrasting result was obtained in another structural study on HP1α , where it was shown that HP1 bound to H3K9me3 acts as a bridge between the nucleosomes and hence the

chromatin is compacted (Machida et al. 2018). It is suggested that HP1 droplets formed by SIPS can act as nucleation centers for the growth of the HP1-nucleosome condensates (Tortora et al. 2023).

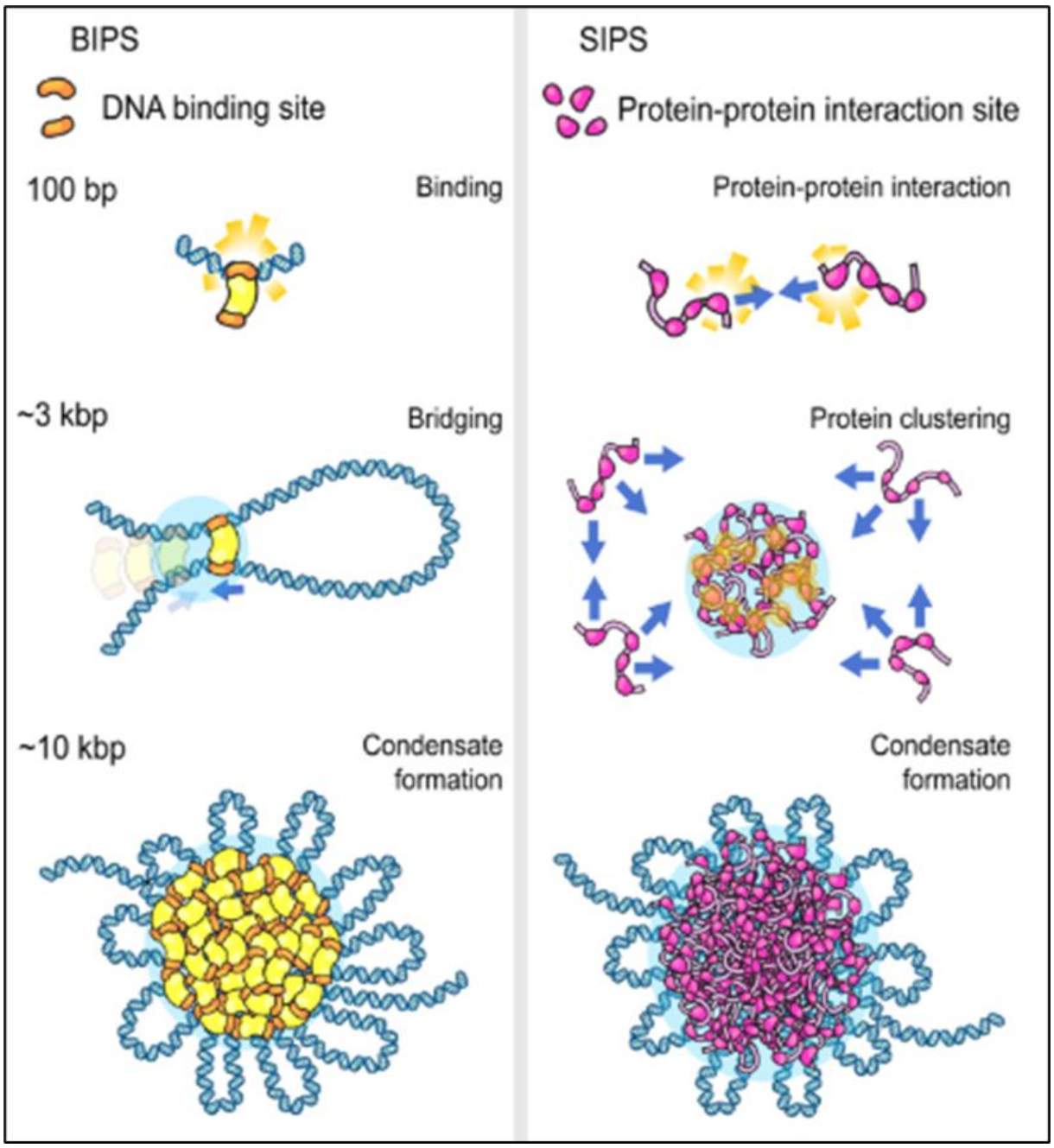


Figure 4.1: Schematic showing two models (BIPS and SIPS) of co-condensation of DNA and protein, image taken from Park et al. 2024, EMM.

Another recent study on DNA-HP1α dynamics showed that the DNA is compacted by a three step process, (1) HP1α assemble along the DNA prior to DNA condensation; (2) initiation of DNA compaction through capturing of nearby DNA fluctuations via HP1α-DNA and HP1α-HP1α interactions to form a proto-condensate, and (3) progression of DNA compaction through inclusion of un-compacted DNA into the growing condensate via HP1α-DNA and HP1α-HP1α interactions (Keenen et al. 2021).

We have shown in Chapter 3 that, PfHP1, like its homolog HP1α can also undergo phase separation. PfHP1 phase separates without any other additives. However, the droplets grow bigger as we add DNA. This leads to the speculation that PfHP1 might have a similar mechanism of DNA compaction like HP1α.

4.1.2 PfHP1 distribution on PfDNA

PfHP1 has a very peculiar distribution on the chromatin that is the sub-telomeric and telomeric region. The AT-richness of the heterochromatin genes is on an average is ~70%, however there is spike in AT-richness at the start and end of the gene (90-100%). The *var* genes in particular have highly AT-rich introns (~90%) and relatively low AT-rich exons (~70%). The euchromatic genes on an average have 80-90% AT-richness (Fig. 4.2).

PfHP1 mediated heterochromatin is restricted to canonical histone and relatively low AT regions (73% AT, much lower than the overall AT-richness of the genome). H2A.Z demarcates highly AT-rich intergenic regions (86% AT) including *var* introns (showed Chapter 2 results) (Hoeijmakers et al. 2012). However, the *var* introns show a dynamic pattern of heterochromatin spread. The AT-GC bias in the genome itself might be driving the distribution of the heterochromatin.

The PfDNA sequence determinants underlying the heterochromatin nucleation, spread, or maintenance was unknown until recently. The *de novo* heterochromatin nucleation was shown to be driven by *var* upstream (or promoter) sequence and/or ap2g coding sequence from the beginning of the gene (Pérez-Cantero et al. 2025). The *var* promoter sequences are highly AT-rich (Fig.4.2 B,C & D).

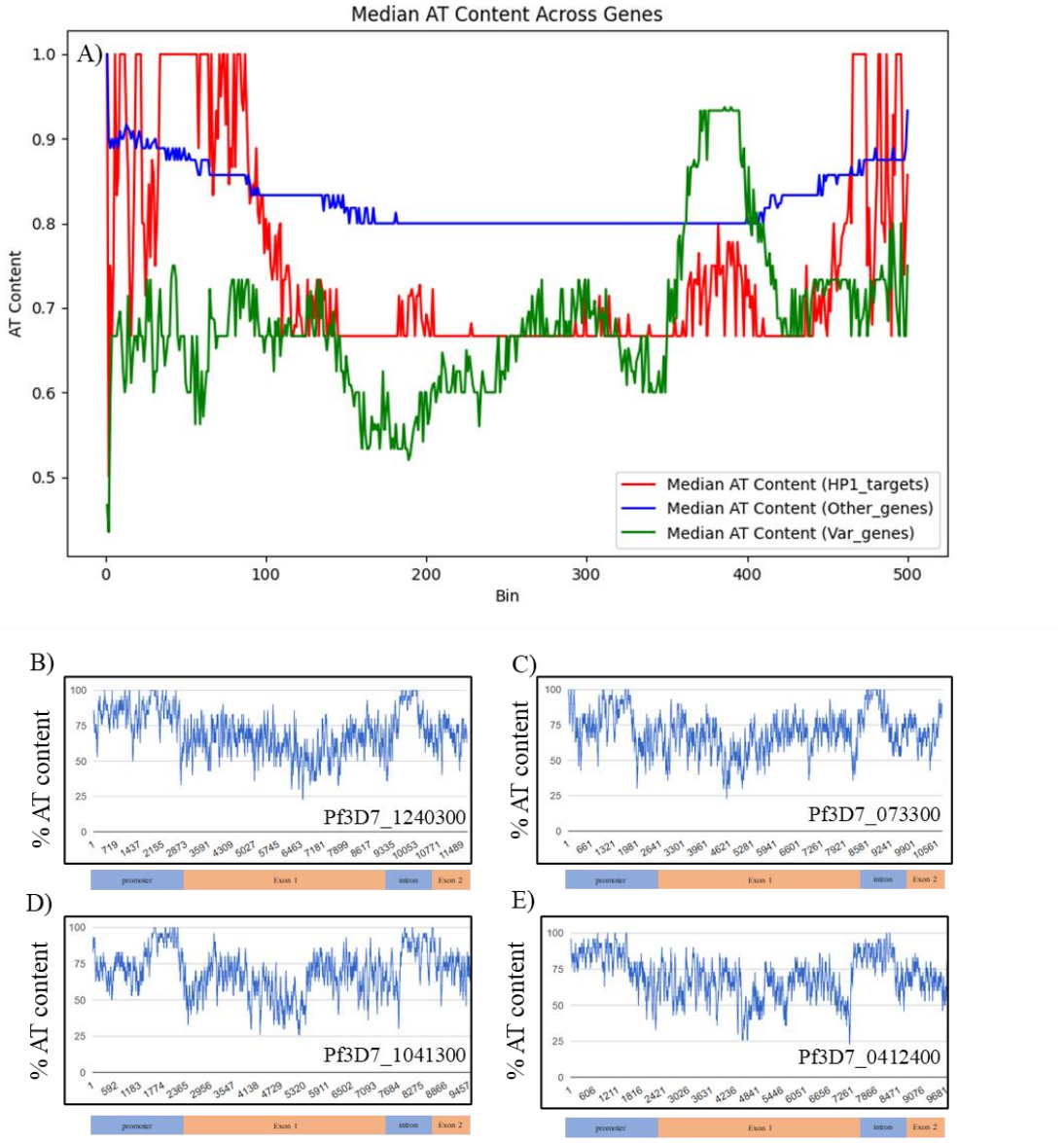


Figure 4.2: A) Median AT-richness calculated for euchromatic gene (blue), heterochromatic genes (red) and var genes (green); % AT-content of var genes B) Pf3D7_1240300, C) Pf3D7_073300, D) Pf3D7_1041300, E) Pf3D7_0412400.

4.2 Scope of this Chapter

P. falciparum has an unusually AT-rich (~81%) genome. The DNA compaction of such AT-rich genomes is an interesting problem, considering that TA nucleotides have more flexibility than

others. Despite that PfHP1 forms stable heterochromatin domains on the high AT rich genome. Also in a recent study, it is shown that highly AT-rich *var* promoters (88%) are the nucleation center for the heterochromatin formation, even though the mechanism is unknown (Pérez-Cantero et al. 2025).

Also, we observed that PfHP1 forms droplets with highly AT-rich PfDNA. This leads us to the speculation that PfHP1 might have evolved a way to condense the highly AT rich DNA of *P. falciparum*, unlike its other eukaryotic counterparts.

We are testing whether PfHP1 will form condensates with DNA in a sequence dependent manner; i.e., by differentiating AT-rich vs. GC-rich DNA sequences. Our observation from Chapter 2 suggests that introns are an interesting paradigm to study the heterochromatin dynamics. Moreover, *var* introns have similar AT-richness as that of *var* promoters. Hence it is possible that *var* introns can also act as nucleation centers similar to *var* promoters.

For this Chapter we adapted a single molecular DNA tethering strategy where we used a double tethered DNA with a *var* intron at the middle to test how PfHP1 behaves on the PfDNA. We are trying to uncover the molecular basis for the PfHP1 mediated DNA compaction.

4.3 Materials and Methods

4.3.1 AT-content calculation of the gene

For the AT content calculation of multiple genes together, a sliding window frequency of 50 bp was scanned across the gene length, and the AT content was plotted using the Python package Matplotlib.

For the calculation AT-content of individual genes, GC content calculator from biologics corp was used with a bin size of 30 bp (https://www.biologicscorp.com/tools/GCContent/). The gene sequences were retrieved from PlasmoDB (https://plasmodb.org/) and the promoter sequences were retrieved by selecting regions in IGV (integrative genomics viewer) against the *P. falciparum* genome.

4.3.2 PfHP1 expression and purification

The construct used was PfHP1-eGFP-pET15b and the mutants E-A and K-A. The same protocol was followed for the purification as given in section 3.3.4. For the expression and purification of PfH2A.Z the same protocol was used as given in the section 3.3.4.

4.3.3 Cloning of var intron into 18kbp plasmid

The *var* intron (Pf3D7_0412400) was amplified from gDNA of 3D7 line (Primers; F-TGTGGGGTGAAAGCTGTAAAAAATATATTGTGGCGTATTT, R-CGGCAGAGTCTCTAGCTACAAAAAAATGAACACATATGT). Due to the extreme AT-richness we had to use reduced extension temperature ($64^{O}C$) for the amplification (PCR protocol is given in the table xx). The PCR product was cloned into an 18kbp vector digested with the XbaI and HindIII restriction enzymes (NEB). The Takara snap assembly was used for the cloning; the vector and insert were added in 1:2 ratio. The clones were screened with restriction digestion and sequence confirmed for the insertion of *var* intron.

4.3.4 EMSA (Electrophoretic mobility shift assay)

The 1.3% agarose gel is made in 1X Sodium borate buffer (Stock: 20X SB buffer, pH 8.0, 8g NaOH and 48g Boric acid in 1L water). The PfDNA was made by amplifying the *var* intron (~1kbp) from a 19kbp vector using primers given in the section 4.3.3. The DNA was mixed in 20 ng/µl concentration with increasing concentration of PfHP1 (0nM, 125nM, 250nM, 500nM, 1µM, 2µM, 4µM and 10µM). The reaction was incubated for 30 minutes at room temperature and then run on the gel at 150V for 1 hour. Visualized the gel using EtBr.

4.3.5 Biotinylation of the 18kbp DNA and 19kbp DNA with var intron

All single molecular experiments and preparations are done with Prakshi Gaur and Dr. Mahipal Ganji, IISc, Bangalore.

The 18kbp vector (control), and 19kbp vector with *var* intron were digested with restriction enzymes Not1 and Xho1 (NEB) and cleaned up. The vectors were then biotinylated by ligating

with 500bp PCR amplified DNA with a biotin handle. The 500bp DNA was digested with Not1 or Xho1, and ligated with a vector using T4 ligase (NEB). The ligated product was separated from non-ligated by running on 0.8% agarose gel and gel extracted later.

For cy5 labelling of the vector, the same protocol was followed except that Not1 digested 500bp DNA was biotin-dUTP-cy5 (Jena Bioscience NU-803-CY5-L) labelled. During the ligation reaction the 18 kbp/19 kbp DNA will have cy5 labelling on one end, this will help in determining the directionality of the double tethered DNA.

4.3.6 Preparation of PEG passivated slides

For the microfluidics and single molecule experiments the slides were prepared as described in Satheesh et. al. (Mori & Nakashima 2023). Briefly the protocol is, two holes were drilled on the glass slide. The slides were cleaned and sonicated initially in 10% v/v dishwash detergent for 20 min and then in Milli-Q for 5 mins. Slides were then sonicated in acetone for 10 min. The coverslips and slides were then sonicated for 25 min in 1 M KOH solution for etching to generate free -OH groups. Followed by etching, the slides and coverslips were washed and sonicated with Milli-Q. To further clean the slides and coverslips, a Piranha etching process was carried out. This involved slowly adding 30% hydrogen peroxide (EMPLURA 107209) to sulfuric acid (Fisher Scientific 29997) in a 1:3 volume ratio. The mixture was stirred using Teflon holders and left for 30 minutes, or until the reaction ceased and bubbling stopped. After the etching, the Piranha solution was discarded, and the samples were rinsed with water. The slides and coverslips were then thoroughly washed with methanol and sonicated for 20 minutes.

Following this, aminosilanation was performed. A mixture was prepared using glacial acetic acid (EMPARTA ACS 1.93002.2521), APTES (3-aminopropyl triethoxysilane, TCI A0439), and methanol (EMPARTA ACS 1.07018.2521) in a 1:2:20 ratio. This freshly prepared solution was applied to the slides and coverslips and incubated for 25 minutes, allowing the surfaces to be functionalized with amine groups. Subsequently, the slides and coverslips were washed thoroughly with methanol at least five times, rinsed with Milli-Q water, and dried using compressed nitrogen gas.

Following aminosilanization, the surfaces were passivated with PEG using a mixture of biotin-PEG-SVA (Laysan Bio, Biotin-PEG-SVA-5000) and mPEG-SVA (Laysan Bio, mPEG-SVA-5000) in a 1:10 ratio. This blend was prepared in a freshly made 0.1 M potassium sulfate (K_2SO_4) buffer at pH 9–9.5, yielding a final concentration of 20 mM. A 60 µl volume of this solution was placed between a glass slide and a coverslip and incubated overnight in a dark, humid environment.

The next day, the slides and coverslips were separated, rinsed with Milli-Q water, and dried using compressed nitrogen gas. To enhance surface passivation, a second PEGylation step was performed prior to single-molecule imaging. For this step, $MS(PEG)_4$ (Thermo Scientific 22341) was used by dissolving 7 µl of a 250 mM $MS(PEG)_4$ solution in 63 µl of K_2SO_4 buffer.

4.3.7 Single molecule experiment

Microfluidic flow channels were created by placing double-sided transparent tape between a quartz slide and a coverslip. Inlet and outlet holes were introduced to allow for buffer exchange and sample loading. The channel had an approximate volume of 10 µl. To prevent non-specific binding and enable biotin-streptavidin interactions, the inner surfaces of the slide and coverslip were passivated with a 20% (w/v) PEG and biotin-PEG solution (mixed in a 50:1 ratio) prepared in fresh 100 mM potassium sulfate buffer.

For additional surface passivation, 10% BSA (Bovine Serum Albumin, GBiosciences RC1021) was incubated in the channel for 10 minutes, followed by an 800 µl T50 buffer wash (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA). This was followed by a 5-minute incubation with 0.1% Tween 20 (Sigma P9416), which was also washed out with 800 µl of T50 buffer.

To tether DNA to the surface, 20 µl of 0.1 mg/ml neutravidin (Sigma 31000) in T50 buffer was added and incubated for 10 minutes. Excess neutravidin was then washed off with 800 µl of T50 buffer. Next, 30 µl of biotinylated DNA (approximately 5–10 pM) was introduced into the channel at a controlled flow rate of 6–15 µl/min using a motorized syringe pump. The DNA molecules were immobilized on the PEG-passivated surface through biotin–neutravidin–biotin linkages. DNA with biotin tags on one or both ends remained bound, while untagged or unbound

DNA was flushed out with 100 µl of T50 buffer. This process resulted in the formation of single- or double-tethered DNA molecules of varying lengths, depending on the applied flow rate. The tethered DNA was visualized under the microscope using DNA intercalating dye 50nM Sytox orange ((Thermofisher Scientific S11368, excitation at 561 nm). Then the protein PfHP1-eGFP/ or mutants (excitation 488 nm) was applied in different concentrations (2µM, 1µM, 500nM). In the PfH2A.Z experiment, 500nM of the protein was used. The imaging was performed in a total internal reflection fluorescence microscope (TIRFM), Nikon Ti2 eclipse microscope. Used an oil immersion objective lens (Nikon Instruments Apo SR TIRF 100×numerical aperture 1.49, oil, IISc, Bangalore) under Total Internal Reflection conditions.

4.3.8 Data analysis of single molecule experiment

To analyze the 19 kb DNA molecules containing a central *var* intron AT-rich region, individual DNA strands were first captured using Fiji ImageJ. From these images, a time-averaged projection was created by averaging the fluorescence intensity across 200 frames. The end-to-end distances of the DNA molecules were then normalized on a scale from 0 to 1, allowing the precise positioning of the condensate along the DNA strand to be determined. The resulting data was further processed through interpolation and averaging to produce the plot shown in the results section.

The puncta number was calculated from the maximum intensity projections made using Fiji Image J. The data was plotted using graph pad Prism 10. The significance and P value were calculated by non-parametric t-test.

4.4 Results

4.4.1 PfHP1 forms co-condensates with AT-rich var intronic DNA

The biochemical characterization of PfHP1 showed it undergoes LLPS. Also, PfDNA promotes the phase separation and reduces critical concentration needed for the droplet formation (section 3.4.3). We observed that PfHP1 can directly bind to the AT-rich PfDNA in the gel shift assay

(Fig. 4.3). Earlier Chapter 2 we have also observed dynamic enrichment of PfHP1 on a *var* intronic region that is highly AT-rich (Fig. 4.2 B,C,D & E).

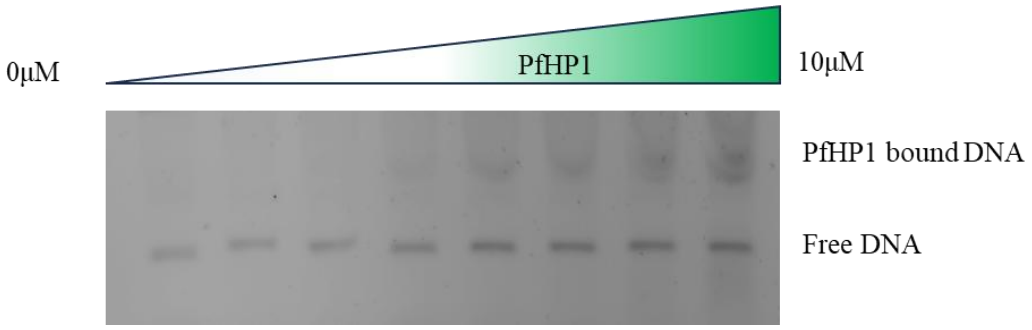


Figure 4.3: EMSA showing shift in PfDNA (20 ng/µl) upon addition of increasing concentration of PfHP1 (0nM, 125nM, 250nM, 500nM, 1µM, 2µM, 4µM and 10µM).

To study the real time dynamics of the PfHP1 on *var* intronic DNA we employed a single molecular DNA experiment with labelled PfDNA and PfHP1-GFP (see design given in Fig. 4.4A & B). We immobilized a control 18 kbp DNA and a 19 kbp DNA with 1 kbp *var* intron (Fig. 4.2 E) in the middle on a PEG passivated slide (Fig. 4.4B). The PfHP1-GFP was introduced in the flow cell and we observed puncta formation in the middle of the 19 kbp DNA and no puncta formation in 18 kbp control DNA (Fig. 4.4 C,D & E). We also noted that first PfHP1 coats the DNA and preferably forms puncta in the middle (Fig. 4.4 D & E). However, as we increase the concentration of the protein there is multiple puncta formation on a single DNA molecule (Fig. 4.7E).

We used Cy5 tagged DNA to determine the directionality of the DNA and relative position of puncta formation (Fig. 4.5 A & B). The intronic part of the 19kbp is ~90% AT-rich (Fig. 4.5 C). The estimated position of the puncta formation is around 10kbp that coincides with the highly AT-rich stretch. Also, we estimated that around 2000 bp length of DNA is compacted by PfHP1, centered around the AT-rich sequence.

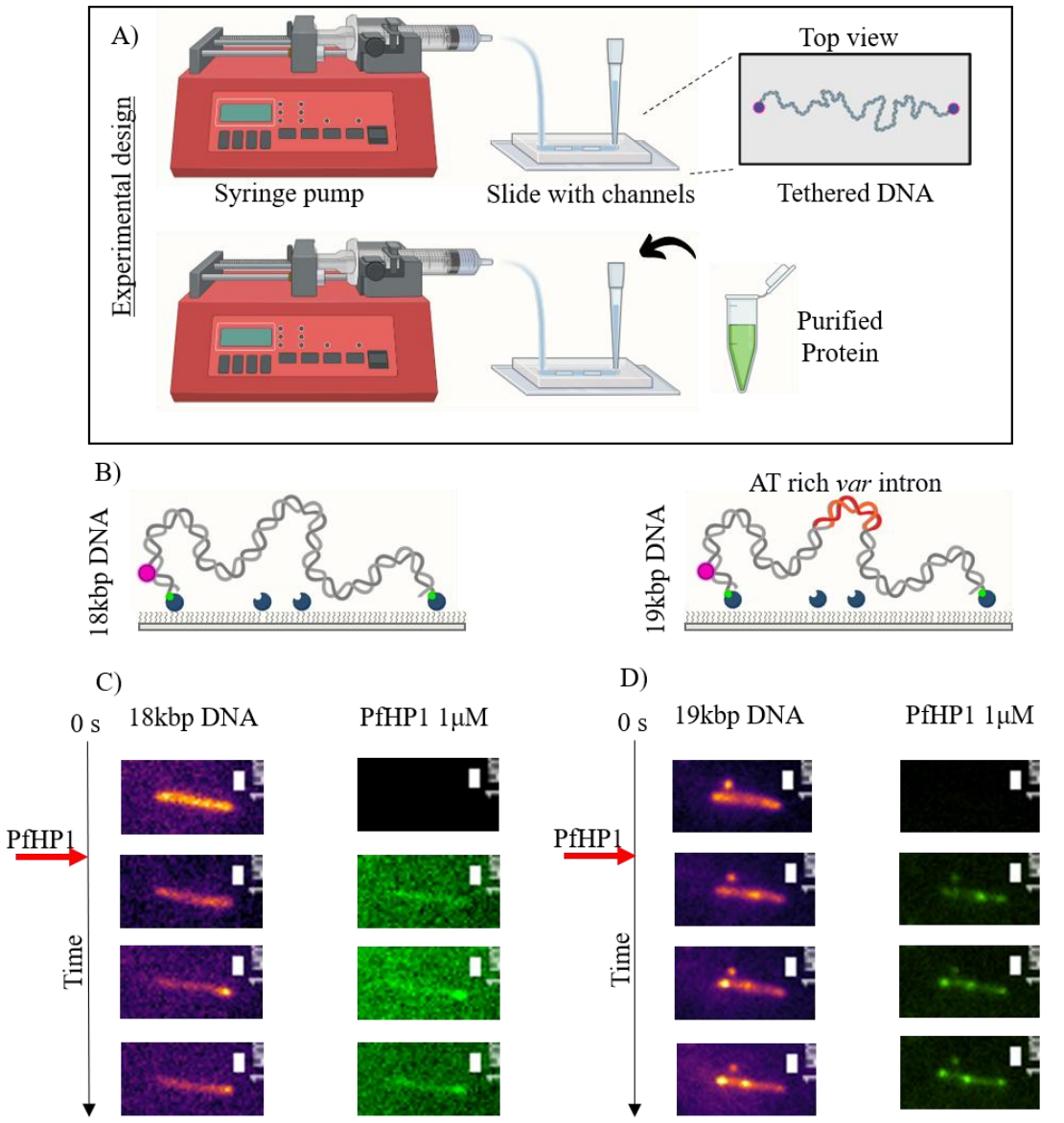


Figure 4.4: A) Schematic showing experimental design of the single molecule DNA experiment, B) Schematic of double tethered DNA with (right) or without (left) AT-rich stretch from the *var* intron; the DNA tethering was done by streptavidin (blue) and biotin (green) interaction; for determining directionality Cy5 (magenta) probe is attached on one end, C) Snaps from the time series of double labelling experiment done with control DNA and PfHP1, D) Snaps from the time series of double labelling experiment done with 19 kbp DNA and PfHP1 (schematics are made in biorender.com).

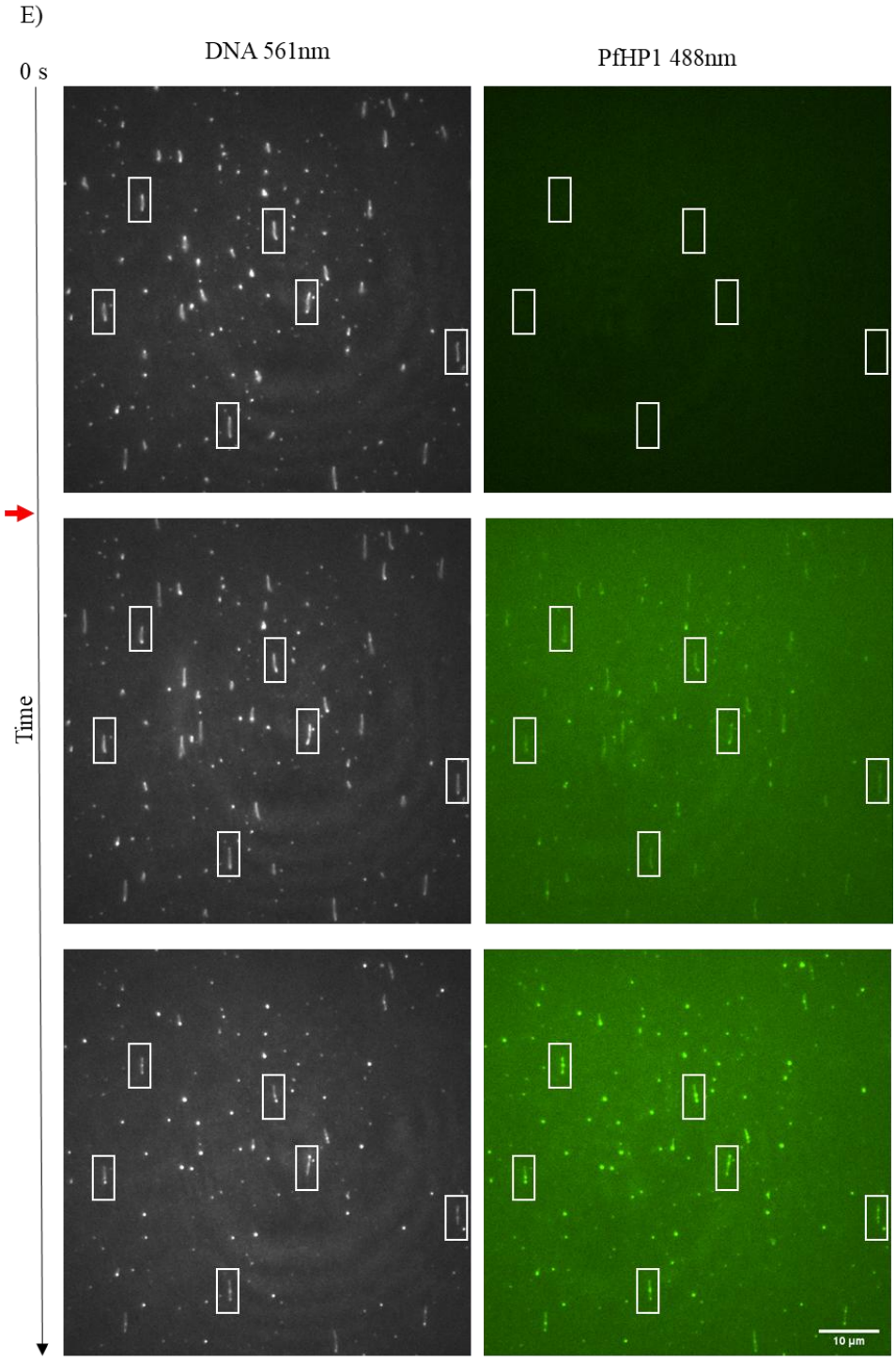


Figure 4.4: E) Snaps indicating multiple puncta formation from the time series of double labelling experiment done with control DNA and PfHP1 (the red arrow indicates the point at which the proteins were added during the time series).

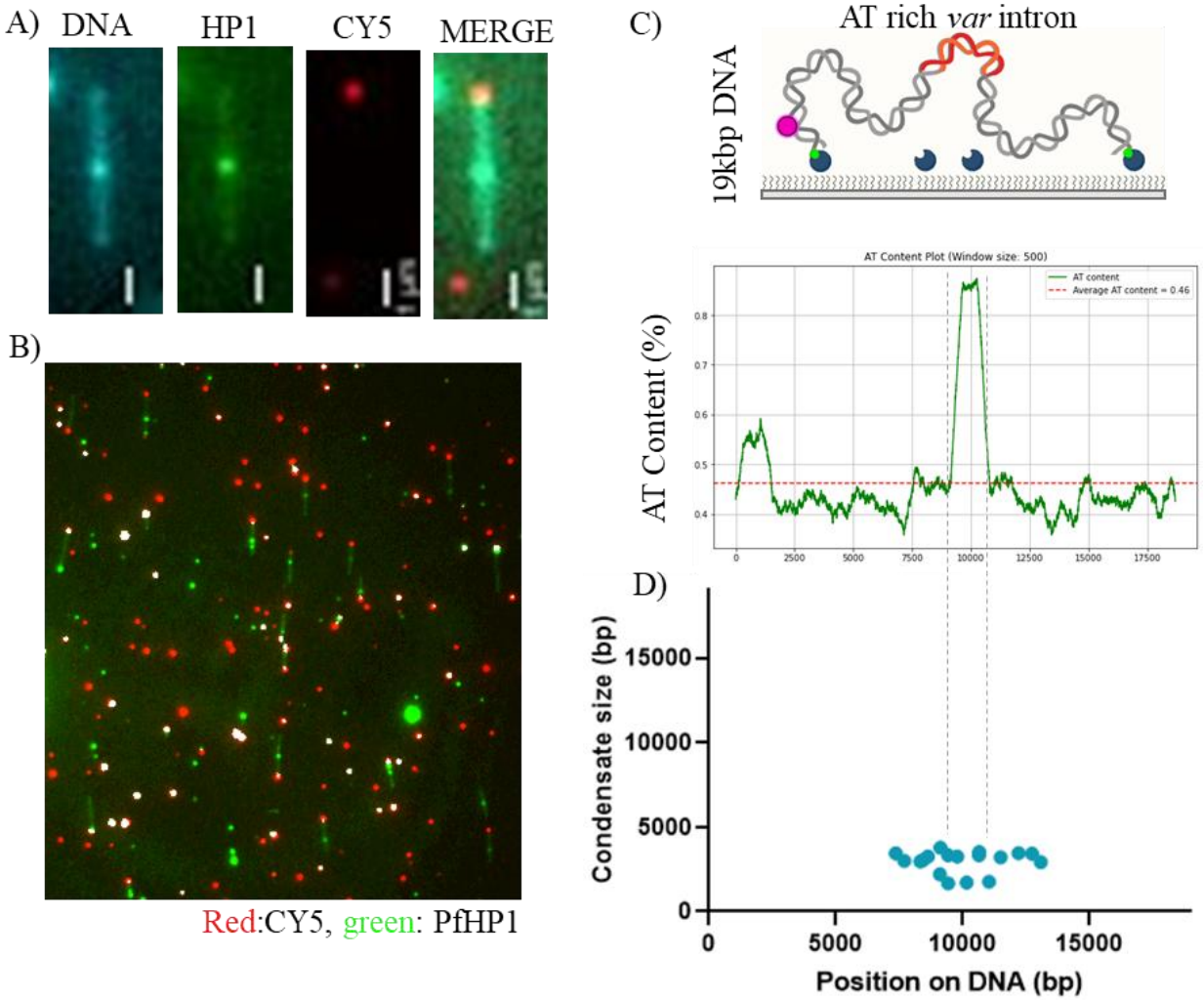


Figure 4.5: A) & B) Co-localization of Cy5 tag with DNA molecule for determining the directionality, C) Percentage AT-richness of the 19 kbp DNA, D) Graph indicating position of the puncta formation along the 19 kbp DNA, also shows the DNA length that is compacted inside the droplet on Y axis (schematics are made in biorender.com).

4.4.2 PfHP1 puncta formation on the PfDNA is hampered by PfH2A.Z

In Chapter 2 we discussed the possibility of PfH2A.Z limiting the heterochromatin spread. To test this hypothesis, we added PfH2A.Z in the single molecule flow cell. The fluorescence intensity of the DNA was reduced upon the addition of PfH2A.Z, indicating the binding of the protein and hence the DNA becomes inaccessible to the intercalating dye. We observed that

PfH2A.Z stiffened the 19 kbp immediately upon the addition; however, we didn't observe any sequence dependent co-condensation of the DNA with PfH2A.Z (Fig. 4.6 A). This was counter intuitive given that PfH2A.Z binds to highly AT-rich regions of the *P. falciparum* genome.

Also, presence of PfH2A.Z hampered the co-condensation of PfDNA and PfHP1 (Fig. 4.6 B&C). The PfHP1 coated on the DNA in presence of the PfH2A.Z however the puncta formation was not stable.

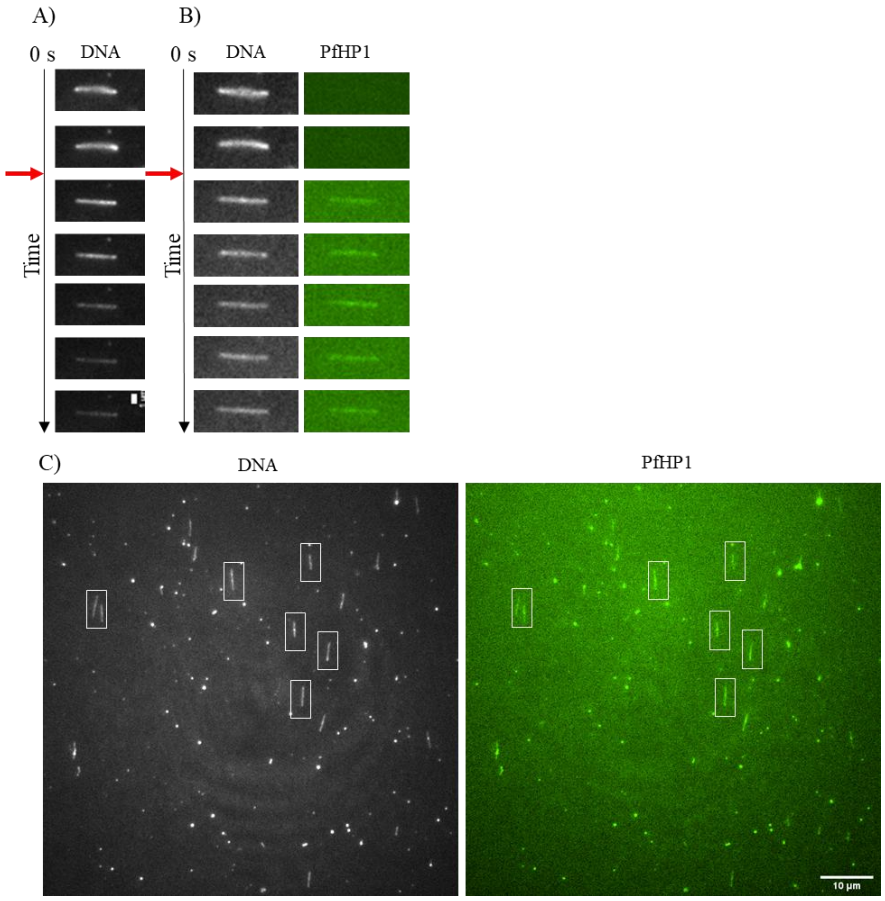


Figure 4.6: A) Snaps from single molecule experiment with sytox orange labelled DNA and unlabeled PfH2A.Z, B) Snaps from single molecule experiment with sytox orange labelled DNA, PfH2A.Z followed by PfHP1-eGFP (the red arrow indicates the point at which the proteins were added during the time series), C) Snap from single molecule experiment with sytox orange labelled DNA, PfH2A.Z followed by PfHP1-eGFP.

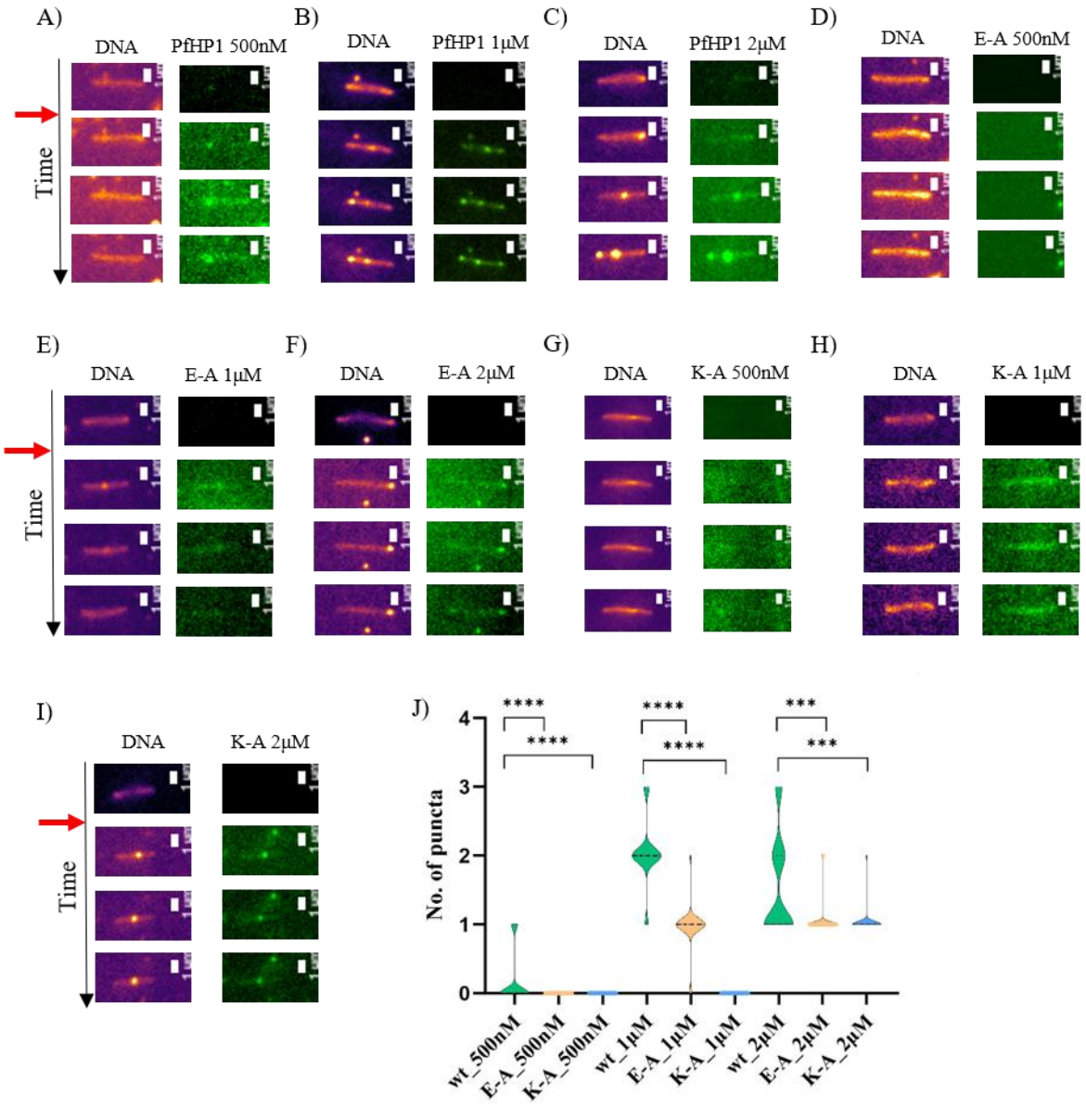


Figure 4.7: Snaps from the single molecule experiment with labelled DNA and PfHP1 WT vs. mutants A) PfHP1 500nM, B) PfHP1 1µM, C) PfHP1 2µM, D) E-A 500nM, E) E-A 1µM, F) E-A 2µM G) K-A 500nM, H) K-A 1µM and I) K-A 2µM, J) Violin plot of no. of puncta in WT vs. mutants (the red arrow indicates the point at which the proteins were added during the time series).

4.4.3 PfHP1 phase separation mutants show subpar puncta formation

Next, we decided to look at the co-condensation properties of the PfHP1 phase separation mutants E-A and K-A (Section 3.4.6). The E-A mutant is predicted to be H3K9me3 binding mutant and

K-A is predicted to be the oligomerization mutant. Notably, the mutated stretch of four lysines in the K-A mutant might also hamper with the DNA binding. As we observed, addition of DNA also didn't rescue the phenotype for this mutant.

The E-A mutant didn't form any condensate at 500nM concentration, at higher concentrations it did form co-condensates, however the puncta were highly unstable at 1μM (Fig. 4.7 A-E). Also, at 2μM it didn't form multiple puncta as observed in the wildtype (Fig. C,F). Interestingly the puncta formed, were at the one end of the DNA instead of AT-rich sequence at the middle (Fig. 4.7 E,F). Possibly due to loss of function to do a sequence dependent co-condensation like wildtype.

The K-A mutant didn't form any condensate at 500nM and 1μM concentrations, nonetheless the mutant coated the DNA at 1μM (Fig. 4.7 G,H). At a higher concentration of 2μM, K-A showed puncta formation, but only one punctum instead of multiple puncta observed in the case wildtype (Fig. 4.7 C,I,J). These results suggest that weak intermolecular interaction between stretch of lysines to DNA might be important for the condensate formation.

Condition	% of DNA with puncta
WT 500nM	17.03
WT 1μM	100
WT 2μM	100
E-A 500nM	0
E-A 1μM	92.85
E-A 2μM	100
K-A 500nM	0
K-A 1μM	0
K-A 2μM	100

Table 4.1: Percentage of double tethered DNA showing at least one punctum for PfHP1 wildtype and mutants.

We also calculated the percentage of the double tethered DNA that showed puncta formation. We observed that mutants mostly show a binary phenotype; where it is either forming puncta or not forming at all in a given condition (Table 4.1).

4.5 Discussion

We show the co-condensation of PfHP1 and PfDNA in an AT-rich sequence dependent manner. Earlier it has been shown that human HP1α can form co-condensates with DNA and start compaction. However, there is no report of a sequence dependent binding or condensation by HP1 homologs. Also, HP1 lacks a DNA sequence recognition domain. That brings us to the major question that how does PfHP1 compacts AT-rich DNA preferentially.

The TA dinucleotide by nature has more flexibility than other kinds. The HP1 homologs have an intrinsic property to compact DNA. Because of the AT rich genome of *P. falciparum* (80%), PfHP1 might have acquired an ability to compact the AT-rich DNA. Even though HP1 protein doesn't have an ability to recognize sequence it might be able to recognize the minor groove AT DNA by its flexibility and curvature.

Another interesting observation is that PfHP1 mediated heterochromatin is restricted to canonical histone and relatively low AT regions (73% AT). PfH2A.Z demarcates highly AT-rich intergenic regions (86% AT) including *var* introns. It is highly probable that PfHP1's spread to these regions is insulated by PfH2A.Z. So far it is believed that the *P. falciparum* genome is indexed by nucleosome positioning of histone variants such as PfH2A.Z/PfH2AB.Z.

PfHP1 shows preferential compaction of highly AT rich DNA (~90%) *in vitro* (while in the absence of context of nucleosomes). We also observed that PfHP1 forms puncta over one end of the tethered DNA in addition to the *var* intron in the middle. Notably one end of the 19 kbp DNA has ~60% AT-rich stretch. It is safe to assume that PfHP1 has preference for AT-rich sequence, while there might be fine tuning based on the percentage of AT-richness. Also, our experiments with PfH2A.Z reveal that this histone variant interferes with the condensate formation with

PfHP1. But a major drawback is that PfH2A.Z is used here as a single protein whereas nucleosomes exist in octamer. If we do the DNA tethering assay in the presence of PfH2A.Z/PfH2B.Z double variant histone octamers, results might be much clearer and it will put things in much more context. Nonetheless, the ability of PfHP1 to compact the DNA in a base-composition dependent manner indicates that it can also bookmark the genome into domains aside from the histone variants.

The knowledge of physical and chemical properties of PfHP1 can be used in understanding the mechanism of heterochromatin formation. We propose a model where co-condensation of PfHP1 and *var* intron/promoter might be acting as a nucleation center (Fig. 4.8). This aligns with an earlier study that has identified highly AT-rich *var* promoters to be nucleation sites for the heterochromatin formation (Pérez-Cantero et al. 2025). The heterochromatin formation is a multi-step process where;

1) Intermolecular interaction with PfDNA-PfHP1 helps in nucleation, and these condensates capture more DNA and start the compaction. While there could be additional DNA sequence determinants and DNA binding proteins involved in the heterochromatin nucleation, it remains largely unstudied in the field.

2) The multivalent interaction of PfHP1 with itself and H3K9me3 helps in the heterochromatin spread and LLPS droplet formation.

3) Further growth and maturation of multiple droplets by fusion leads to heterochromatin domain formation (Fig. 4.8). Additionally, the boundary of these droplets might be maintained by PfH2A.Z. Further, this model of heterochromatin formation needs to be discussed and tested in the context of virulence gene regulation.

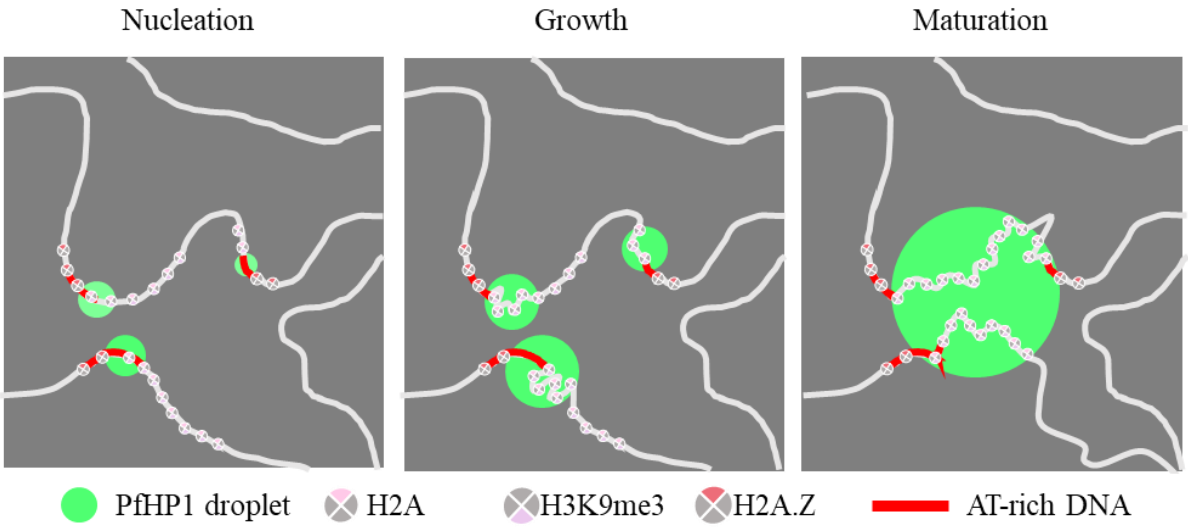


Figure 4.8: Proposed model for DNA condensation by PfHP1.

Chapter 5: The role of PfHP1 mediated phase separation in *var* gene regulation by conditionally expressing PfHP1 mutants**

5.1 Introduction

5.1.1 Gene editing techniques in P. falciparum

The functional studies of any protein heavily depend on genetic manipulations in the genome. The discovery of CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats-CRISPR associated protein 9) has revolutionized the gene editing approaches (Jinek et al. 2012).

In *P. falciparum in vitro* transfection have been employed since 1995 (Wu et al. 1995). Emergence of homologous recombination-based gene disruption has largely helped in advancement of malaria research (Nunes et al. 1999; Pace et al. 2000). Initially researchers came up with gene disruption based on stochastic homologous recombination by integration of the homology region provided by an episomally maintained homology region. But this is a time taking process which requires continuous culturing of *P. falciparum* lines for several months. By the establishment of CRISPR-Cas9 based gene editing techniques in *P. falciparum,* the targeting of various genes in the genome has been simplified (Wagner et al. 2014) (Ghorbal et al. 2014) (Fig. 5.1). CRISPR-Cas9 relies on the ability of Cas9 enzymes to cut the genome to make double stranded breaks (DSB) in a target guide RNA dependent manner. In higher eukaryotes the DSB repair can occur through two pathways: non-homologous end joining (NHEJ) and homology-based repair. However, *Plasmodium spp.* lacks NHEJ, hence CRISPR gene editing solely depends on the homology-based repair.

CRISPR-Cas9 requires mainly three components, a nuclease that can cut (Cas9), a guide RNA (gRNA) and a homology arm containing donor template. There are several considerations and challenges in getting a successful CRISPR line.

While designing the gRNA, identify a 20-nucleotide long sequence that is upstream of an NGG/ PAM (Protospacer Adjacent Motif) sequence. There are tools available online to make this easier, like CRISPOR, benchling, etc. This software often gives scores based on off target effects and on target efficiency in activity. For an AT-rich genome like *P. falciparum* on target score might not be very accurate. Because these scores are determined based on the modelling done on the gRNAs from mammalian systems. The most efficient gRNAs are ones that bind closest to the site of the modification and with the least off targets. For the point mutations, it is highly recommended to use gRNAs within the 100-200bp of the mutation. Choice of multiple gRNAs for a single editing will also increase the probability of success (Lee et al. 2019).

Another important challenge is donor plasmid design. Out of different lengths tested for the homology arms 250 bp to 1000 bp is shown to be highly efficient (Ribeiro et al. 2018). Also, the gRNA binding site and site of mutation are often 100s of base pairs far, hence, there is a chance that this 100bp might act as a homology region and not insert the mutation at all. To avoid this re-codonised region between the gRNA site and the desired mutation is used (Lee et al. 2019).

The replicating donor plasmids and CRISPR plasmids can result in multiple copies of the same insertions, making a concatemer. Linearizing the plasmid will not let it replicate for another cycle. Also, there are negative selection markers available; to eliminate episomes, once the editing is done (Lee et al. 2019).

But even by using CRISPR-Cas9, generating transgenics is still challenging because of the low efficiency of transfection protocols and lethality associated with the disruption of essential genes. Hence, inducible CRISPR-Cas9 systems have been of great interest. Researchers have developed drug inducible Cas9 and expression of sgRNA under inducible promoters, but mostly ineffective due to low editing efficiency and leakiness (de Solis et al. 2016; Zhao et al. 2018) (Nuñez et al. 2016). Cre-LoxP is a tool based on site specific recombination that enables stimulus dependent deletion of genes. By using the Cre-loxP system rapid and highly efficient inducible systems can be achieved. Recently, by combining CRISPR-Cas9 and Cre-loxP technologies parasites were engineered to conditionally express mutants (Collins et al. 2013; Knuepfer et al. 2017) (Jones et al. 2016). In this strategy they used a 3D7 line that has a rapamycin inducible DiCre (Dimerisable

Cre recombinase) and using CRISPR two loxP introns were introduced flanking the wildtype. In a sequential CRISPR editing mutant gene was introduced downstream of the second loxP, hence upon induction of DiCre the loxP sites will recombine to delete the wildtype, and only the mutant will be expressed. This is a powerful tool to study essential genes such as PfHP1.

5.1.2 Role of PfHP1 in P. falciparum survival and transmission

Previous studies have revealed that PfHP1 is essential for parasite survival. One of the first studies was Brancucci et. al, 2014, that used a conditional depletion/ knock down of PfHP1 (by DD :destabilizing domain) (Brancucci et al. 2014). The Shield-1 ligand stabilizes the DD and hence PfHP1.They showed that parasites undergo a growth arrest in schizogony. They observed using live imaging upon addition of Shield-1 the PfHP1 was depleted within 12 hours. However, in western blotting and immunofluorescence assay, PfHP1 was undetectable in the next generation. Interestingly, in the generation 2 along with the growth arrest, sexual conversion was also abundant (Fig. 5.2). This led to the conclusion that PfHP1 is at the core of the gametocytogenesis switch. Later GDV1 was shown to be evicting PfHP1 from the ap2-g loci for its activation and subsequent gametocytogenesis. These results suggest that PfHP1 is essential for asexual proliferation and the regulation of sexual conversion. The exact mechanism of PHP1 mediated cell cycle progression is unknown.

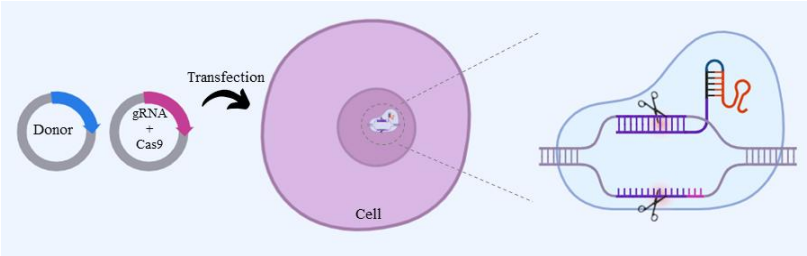


Figure 5.1: A schematic representing a simplified design of CRISPR-Cas9 gene editing in *P. falciparum* (made with bioRender.com).

Another important observation in their study was, PfHP1 depletion leads to de-repression of the *var* genes. This was in support of the hypothesis for *var* gene activation that; rather than the enrichment of active transcription machinery at a *var* loci, it is the histone modifying and

remodeling that disassemble the heterochromatin for singular *var* activation. Additionally, it was stated that PfHP1 is essential for heterochromatin inheritance to the daughter cells, indicating that PfHP1 is essential for heterochromatin formation and re-establishment.

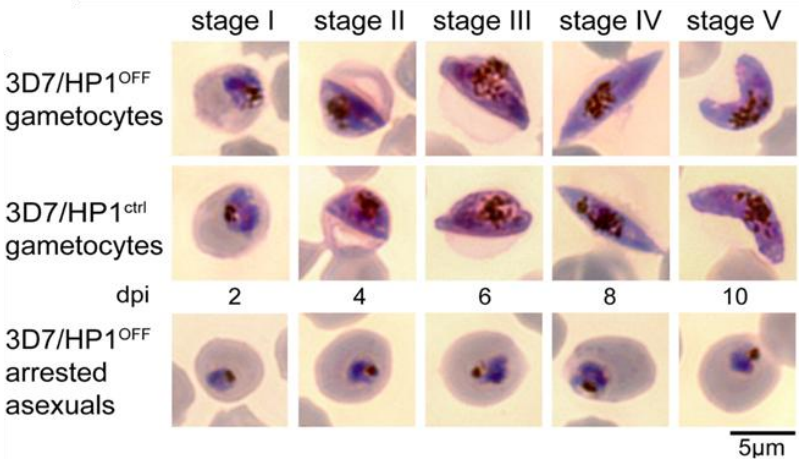


Figure 5.2: Phenotypic changes in PfHP1 depleted parasites (taken from Brancucci et. al., 2014, Cell Host and Microbe).

Another study on Serine phosphorylation of PfHP1 also used a conditional mutant expression approach (Bui et al. 2019). As discussed in the section 5.1.1, they utilized Cre/lox and CRISPR-cas9 to achieve this (Fig. 5.3). Multiple serine in the hinge region was mutated to alanine to get a non-phosphorylated form of the protein. However, the phosphomutants showed no growth defect with no obvious role in gene regulation.

Domain deletion mutants of PfHP1 were also successfully generated by using a conditional system (Bui et al. 2021). While ΔCD and ΔHinge showed nuclear localization, ΔCSD showed mis localization in cytoplasm along with some in nucleus. The PRRK stretch in CSD was predicted to be the nuclear localization signal (NLS). All three domain deletions were lethal and led to higher sexual conversion, indicating that all of them are essential for PfHP1 function.

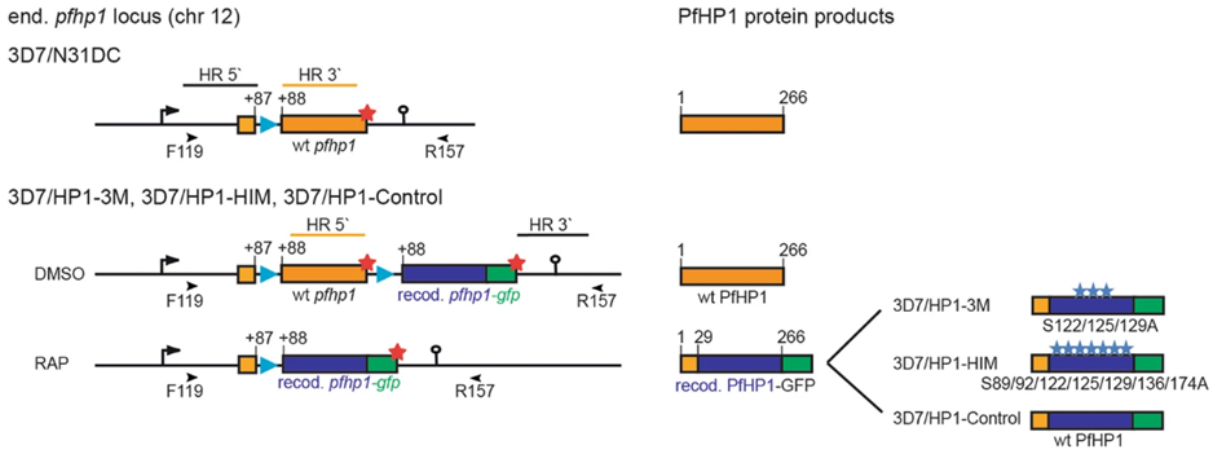


Figure 5.3: Schematic representation of conditional expression of PfHP1 mutants, by Cre-loxP and CRISPR-Cas9 (taken from Bui et. al., 2019, Sci Rep).

5.2 Scope of this Chapter

We have seen PfHP1 undergoes LLPS and forms co-condensates with DNA in Chapter 3 and 4 respectively. In this Chapter we are trying to see functional consequences of such unique biochemical and biophysical properties of PfHP1. The point mutations that were identified in Chapter 3 are used here to generate PfHP1 mutants in cell lines. Subsequently, we are trying to study the effect of these mutants on chromatin binding and *var* gene expression.

We have utilized three strategies to generate the mutant transgenics;

1) Controlled over-expression lines of PfHP1 mutants: The cell lines will over-express mutant PfHP1 but in the background of wild type PfHP1 from endogenous loci. Helpful to study the dominant negative effect of PfHP1 mutation on *var* gene expression. This approach has a caveat; it will produce the mutant protein in large amounts and may suffer from leaky expression.

2) Gene mutation at the endogenous gene loci of PfHP1: The cells will only express mutant PfHP1 and can study the effect on *var* gene expression in endogenous expression levels of PfHP1. The mutant protein could be lethal to the parasites as PfHP1 is an essential gene. An inducible PfHP1

mutant line will enable us to study the effect of the mutations by expressing PfHP1 for a small window of time.

3) Conditional expression of GFP tagged mutant PfHP1 from the endogenous loci: Conditional expression of mutant PfHP1 will enable growth of cell lines by only expressing the mutants upon induction, in case the mutation is toxic to the cells. The GFP tagging will help us to observe the changes in localization of PfHP1 upon mutations.

5.3 Materials and Methods

5.3.1 Cell lines maintenance

The transgenic lines were generated in 3D7/IG5/DiCre background (all lines were generated with Dr. Igor Niederwieser and Prof. Till Voss at Swiss TPH, Switzerland). All the lines were maintained in 5% hematocrit in RPMI 1640 medium supplemented with 25 mM HEPES, 450 µM hypoxanthine, 24 mM sodium bicarbonate, and 0.5% AlbuMAX II. Also, to reduce background conversion into sexual stages (Brancucci et al. 2017), the complete medium was complimented with 2mM Choline Chloride. The parasites were incubated at 37 °C in an atmosphere of 3% oxygen, 4% carbon dioxide, and 93% nitrogen.

The transgenic lines were generated were maintained (at IISER Pune) in 5% hematocrit in RPMI 1640 medium supplemented with 25 mM HEPES, 100 µM hypoxanthine, 12.5 µg/ml gentamycin, 41.6 mM sodium bicarbonate, and 0.5% AlbuMAX II, 2mM Choline Chloride. The parasites were incubated at 37 °C in an atmosphere of 5% oxygen, 5% carbon dioxide, and 93% nitrogen.

5.3.2 Transfection constructs: Controlled over-expression lines expressing PfHP1 mutants

Inducible expression lines of wildtype PfHP1 and mutants were generated by targeting the cg6 locus (PF3D7_0709200, a dispensable locus) of 3D7. The guide RNA against the cg6 locus was cloned in the pBF_gC plasmid. The transgenics are made by co-transfecting 50 ug each of pBF_gC_cg6 and donor plasmid pD_gC_cg6_hp1_gfp_dd_glmS. Original pBF_gC and pD

donor plasmids were obtained from the work published by Filarsky and colleagues (Filarsky et al. 2018). The mutant PfHP1 proteins are under Calmodulin (CAM) promoter and conditionally regulated by glmS ribozyme and DD (destabilizing domain). Upon addition of Shield-1, the DD domain will be stabilized and will start expressing the mutant PfHP1 protein (Fig. 5.4A).

5.3.3 Transfection constructs: Conditional expression of gene mutation at the endogenous gene loci of PfHP1

The donor plasmid for the CRISPR is cloned in the pUC backbone using Gibson assembly cloning technique. The design of the construct is given in Fig. 5.7A. The guide RNA targeting the PfHP1 loci was cloned in a pHF_gC backbone. The CRISPR strategy was based on a self-cleaving ribozyme that will simultaneously give rise to two gRNAs once transcribed inside the cell. The two gRNAs (gRNA 1 and gRNA 2) will target the cas9 to cut and homogeneously recombine at the beginning and end of the gene. The gRNAs were designed using CRISPOR (https://crispor.gi.ucsc.edu/) online tool (Concordet & Haeussler 2018).

The stable knock-in lines will code for the wild type PfHP1 with the mScarlette-I tag. Upon rapamycin treatment, the 3D7/IG5/DiCre lines will start dimerizing Cre recombinase, which will recombine loxP sites and produce GFP-tagged mutant PfHP1 protein.

50µg of the plasmids were precipitated with 0.1V sodium acetate and 2.5V absolute ethanol at room temperature for 10 minutes. Wash with 70% ethanol. Take the pellet inside the hood (in sterile conditions), remove the ethanol, and let it dry. Resuspend it in 1X TE (25 µl).

5.3.4 Transfection

The transgenic lines were generated in 3D7/IG5/DiCre background. Before transfection, sorbitol synchronizes the 3D7/DiCre line. Early rings are most amenable for this transfection; while this protocol does not demand very tight synchrony, it is good to start with a culture with at least 5% early ring stages.

Take 250 µl of culture pellet with ~5% parasitemia and wash (600g, 4 min) the pellet with 2 ml cytomix (120mM KCl, 0.15mM $CaCl_2$, 2mM EGTA, 5mM $MgCl_2$, 10mM K_2HPO_4/KH_2PO_4

25mM HEPES, pH 7.6). Remove the supernatant and add 400µl cytomix and 25 µl each of the plasmids (50µg) that need to be co-transfected. Resuspend with pipetting and incubate at room temperature for 5-10 minutes. While waiting prepare cuvettes, Pasteur pipettes, dishes etc. Transfer the mix to a cuvette with 1ml pipette without any bubbles. Immediately give electric pulse in a Biorad electroporator with the following program 310V/960µF. Immediately after the pulse, the cells are transferred to a 100 mm dish with 10 ml media and 400 µl of fresh blood. Rinse the cuvette once with 1 ml of media and add that to the culture dish. Keep it in an incubator in 37^{O}C temperature and 92% N_2/4% CO_2/ 3% O_2.

Change media in the next 24 hrs and start drug selection with the appropriate drug (WR 20 uM stock: use 5nM, Bsd 10 mg/ml: use 5 µg/ml). Add WR drug every day until 7-8 days, make sure the parasites are cleared. Add Bsd drug for every 9-10 days. Add fresh blood every week to ensure healthy growth of parasites. The transgenic will start appearing in 20 days. Closely observe the culture for 30 days; if parasites don't appear, discard. Even if parasites appear after 30 days, they are most likely non-transfectants.

5.3.5 Freezing down the parasites

It's essential to freeze down the culture and make enough stocks for the future. 500 µl of the parasite pellet in an equal volume of freezing solution (0.0065% of NaCl, 0.0302% of sorbitol and 0.28% of glycerine) and either snap freeze in liquid nitrogen or store at -80^{O}C. -80^{O}C is suitable for fast revival later, and liquid nitrogen is good for long-term storage.

5.3.6 Thawing/reviving of the parasite lines

Prewarm the thawing solution (3.5% NaCl) and media. Take out the stocks and thaw them at 37^{O}C. Transfer it to a 15ml falcon and very slowly add 1 ml of thawing solution. Mix well and spin at 300g for 5 min. Wash the cells with 1-2 ml of thawing solution and give a spin. Transfer the pellet to a 100mm dish/ t25 flask and add 200-300 ml fresh blood in the media.

5.3.7 Sorbitol synchronization

Prepare 5% sorbitol in Milli-Q and pre-warm at 37^{O}C. Spin down the culture that contains at least 5% early ring stages. Add 5-10 volumes of sorbitol to the cell pellet and mix well. Let it incubate at 37^{O}C for 5 minutes. Spin at 300g for 5 min. Wash the cell pellet with media. Transfer the cell pellet to a new culture dish and add the appropriate number of media and fresh blood if needed.

5.3.8 Parasite harvesting

Spin down the culture at 650 g, 5min. To the culture pellet add 0.015% saponin and spin at 2000 rpm, 5 min. Wash with 0.015% saponin for 1-2 times Wash with 1x PBS 1-2 times. If harvesting for RNA isolation after this step, add trizol and thoroughly resuspend the pellet, and store at -80°C. If harvesting for protein, lyse the cell pellet with appropriate lysis buffer and store the cell lysate at -80°C.

For ChIP: Spin down the culture at 650 g, 5min. Use at least 500ul culture pellet with 10% parasitemia per ChIP sample for ring stage ChIP. For troph and schizont half of this will be sufficient. After spinning down, remove half of the media and add 1% formaldehyde (16% stock Formaldehyde methanol free, ThermoFisher scientific) dropwise. Keep on a rocker at room temperature for 10 minutes, 15 rpm. For quenching, add 150mM glycine (2M stock). Keep on a rocker at room temperature for 10 minutes, 15 rpm. Spin at 6000 rpm, 4 °C, 5 minutes. Wash with chilled 1X PBS and spin the pellet at the same speed for 5 minutes (just reverse the pellet direction so that the pellet travels through the 1X PBS. Repeat the previous step for a total of 3 washes. Store the pellet at -80°C.

5.3.9 PCR validation

The genomic DNA was isolated from the wildtype and mutant lines using QIAGEN blood easy kit. PCR was set up using the primers upstream and downstream the insertions. NEB Q5 DNA polymerase was used for the amplification. The amplification products were run on 1% agarose gel and stained with EtBr for the visualization.

For the sequencing confirmation the PCR amplified product was sent for Sanger sequencing after the cleanup.

5.3.10 Western blot validation

The saponin harvested parasites were lysed in 8M Urea/ 10% SDS buffer. The supernatant after lysis was taken and quantified using Qubit Protein assay kit (High sensitivity or broad range for quantification). And ~25µg was used for Western blot gel run. PVDF membrane (Bio Rad) was used for western transfer and 5% skimmed milk was used for blocking. Primary antibodies used were α-PfHP1 (in-house, 1:200), α-CD (Prof. Till Voss, 1:1000), α-GFP (abcam, ab290), and α-RFP (Rockaland).

5.3.11 Live imaging

The parasites were labelled with Hoechst/DAPI (Invitrogen) for 30 min at 37°C. Cells were washed with 1X PBS and mounted on a slide with vectashield/slow fade antifade (Invitrogen). The images were taken at Leica DM6 at Swiss TPH/ IISER Pune microscopy facility. The brightfield images are taken in Leica DM750 microscope

5.3.12 NGS and data analysis

All the RNA sequencing and ChIP sequencing were performed at the NGS facility at IISER Pune. The sequencing and data analysis were done following the protocols given in the sections 2.3.6-2.3.11.

5.3.13 Immunofluorescence assay (IFA)

We used 100µl of RBC pellet with parasites and washed it with 1X PBS thrice. Resuspend it in 4% PFA (Paraformaldehyde) made in 1X PBS / 0.0075% glutaraldehyde and incubate at 37^{O}C water bath for 30 min. Removed the solution and washed with 1X PBS thrice. For the permeabilization, added 300µl of 0.1% Triton-X100 in 1X PBS and incubated for 10 min at room temperature. Washed with 1X PBS, thrice. Blocked with 3% BSA in 1X PBS for 1 hour at room temperature. Washed with 1X PBS and incubated overnight in the primary antibody (serine-2-

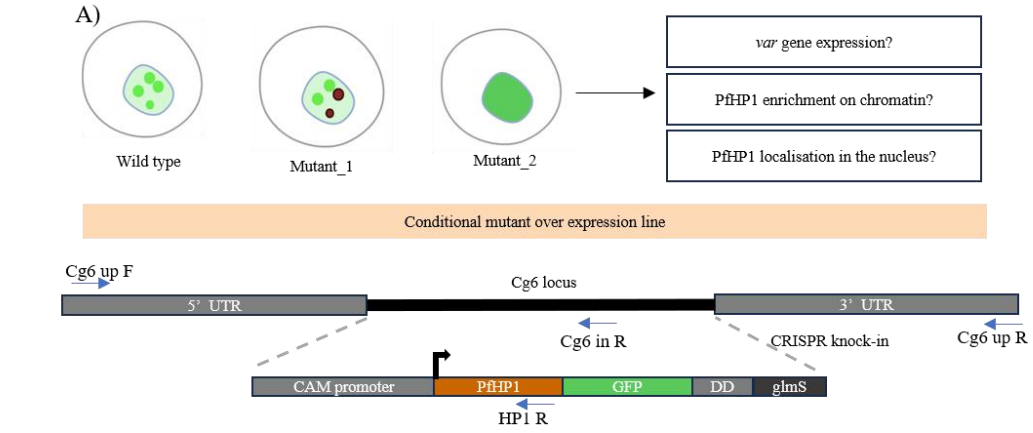
phosphorylation RNA pol II, abcam) made in 3% BSA. 1:200 antibody dilution was used. Washed with 1X PBS thrice and incubated in the secondary antibody (Alexa 568, secondary antibody, Invitrogen) for 1 hour. Washed with 1X PBS thrice. Made a smear, air dry, methanol fix (45 s), and mounted with Prolong gold antifade with DAPI (Invitrogen). The images are taken in Leica DM6 at the IISER Pune microscopy facility.

5.4 Results

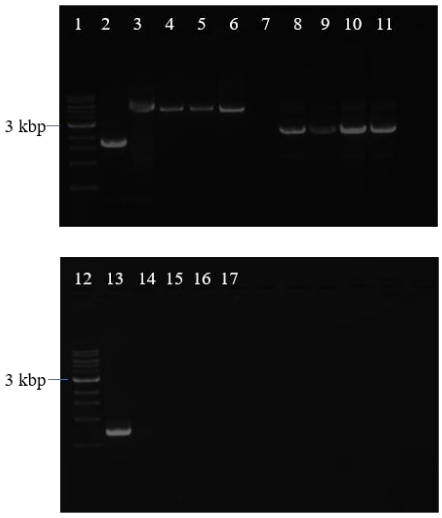
5.4.1 Controlled over expression of PfHP1 mutants

We first decided to go ahead with mutant overexpression strategy, because of the fact that PfHP1 is an essential gene and directly mutating the endogenous loci might be lethal.

The PfHP1 mutant over-expressing lines were generated by inserting it in a dispensable locus called cg6 (Fig. 5.4A). The mutants were expressed with a strong promoter, calmodulin along with a conditional switch of glmS (riboswitch) and DD (Destabilizing domain) (Fig. 5.4A). The parasites start appearing in 20-25 days of transfection for PfHP1 wild-type, K-A, W-A and L-D lines. However, we could not generate a line for E-A mutant. The genomic DNA was isolated and PCR amplified using outside primers to the cg6 locus. We observed the desired bands for all the four lines (Fig. 5.4B). Also, as a negative control we used cg6 outside primer and cg6 inside primer (the part that's deleted upon insertion) (Fig. 5.4B). Next, we isolated protein lysate from the parasites with or without addition of Shield-1 to see inducible expression of PfHP1 and its mutant. In the western blot analysis with the anti-CD antibody, we see that there are two bands in the overexpression lines corresponding to wild type PfHP1 (wildtype, ~31 kDa) and PfHP1-GFP (mutant, ~75 kDa) (Fig. 5.4C). The addition of Shield-1 showed a slight over expression for wildtype PfHP1 and a robust overexpression for the K-A mutant. L-D showed



B)



Also confirmed by sanger sequencing.

no.	sample	Primers	Expected size(approx.)
2	Control 3D7	Cg6 up F- cg6 down R	1.7 kbp
3	HP1 (clone 1)	Cg6 up F- cg6 down R	5.7 kbp
4	HP1 (clone 2)	Cg6 up F- cg6 down R	5.7 kbp
5	W29A	Cg6 up F- cg6 down R	5.7 kbp
6	L254D (clone 1)	Cg6 up F- cg6 down R	5.7 kbp
7	Control 3D7	Cg6 up F- HP1 R	-
8	HP1 (clone 1)	Cg6 up F- HP1 R	1.8 kbp
9	HP1 (clone 2)	Cg6 up F- HP1 R	1.8 kbp
10	W29A	Cg6 up F- HP1 R	1.8 kbp
11	L254D (clone 1)	Cg6 up F- HP1 R	1.8 kbp
13	Control 3D7	Cg6 up F- cg6 in R	950 bp
14	HP1 (clone 1)	Cg6 up F- cg6 in R	-
15	HP1 (clone 2)	Cg6 up F- cg6 in R	-
16	W29A	Cg6 up F- cg6 in R	-
17	L254D (clone 1)	Cg6 up F- cg6 in R	-

C)

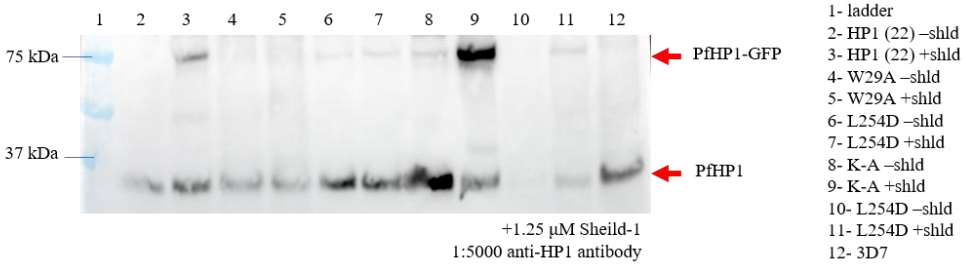


Figure 5.4: A) Experimental design of controlled over-expression lines, B) PCR validation of controlled over-expression lines, the primers used and the expected sizes are given in the table, C) Western blot validation of the lines using anti-PfHP1 (against CD) antibody, the dilution and other details are given in the figure.

lower expression than the wildtype in both the clones. However, W-A didn't show a band for the GFP tagged mutant (Fig. 5.4C). There can be two reasons for this observation, 1) W-A expression is overall very weak, 2) the stages we used for protein lysate preparation might not have robust expression of PfHP1. We know from the transcriptome data from the previous studies that PfHP1 expresses highest during the late stages of development.

The lines were GFP tagged, hence the live microscopy images also confirmed the presence of GFP-tagged mutant proteins (Fig. 5.5). The wildtype HP1 and K-A showed puncta-like localization. Whereas W-A and L-D mutation showed a more dispersed localization all over the nucleus, also the GFP signal was much weaker for these two mutants as compared to the wild type. This observation could be reminiscent of disturbed LLPS of HP1 observed *in vitro* upon point mutations in the protein. Notably, even though W-A didn't show expression in the western blot we observed GFP signal in the microscopy.

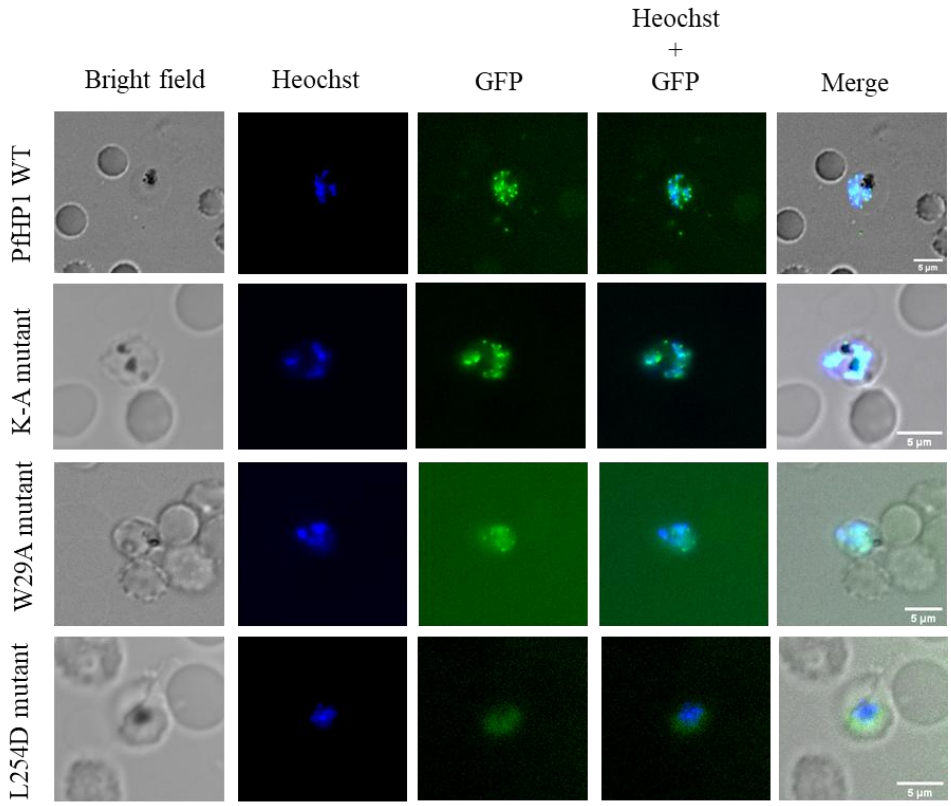


Figure 5.5: live imaging of the controlled over-expression line.

5.4.2 Controlled over expression of PfHP1 mutants shows altered chromatin binding

Next, we isolated chromatin from mutant and wildtype lines and performed ChIP followed by NGS. We used anti-GFP antibody and anti-PfHP1 antibody for the pull down. At a global level, we observed that PfHP1 distribution was similar for the wildtype, K-A and two clones L-D. However, W-A mutant showed distribution in the euchromatin regions as well as the heterochromatic regions (Fig. 5.6A). Then we calculated overall PfHP1 enrichment on its target, *var* genes. All mutants had reduced chromatin binding with a strikingly low tag density for W-A (Fig. 5.6B&C).

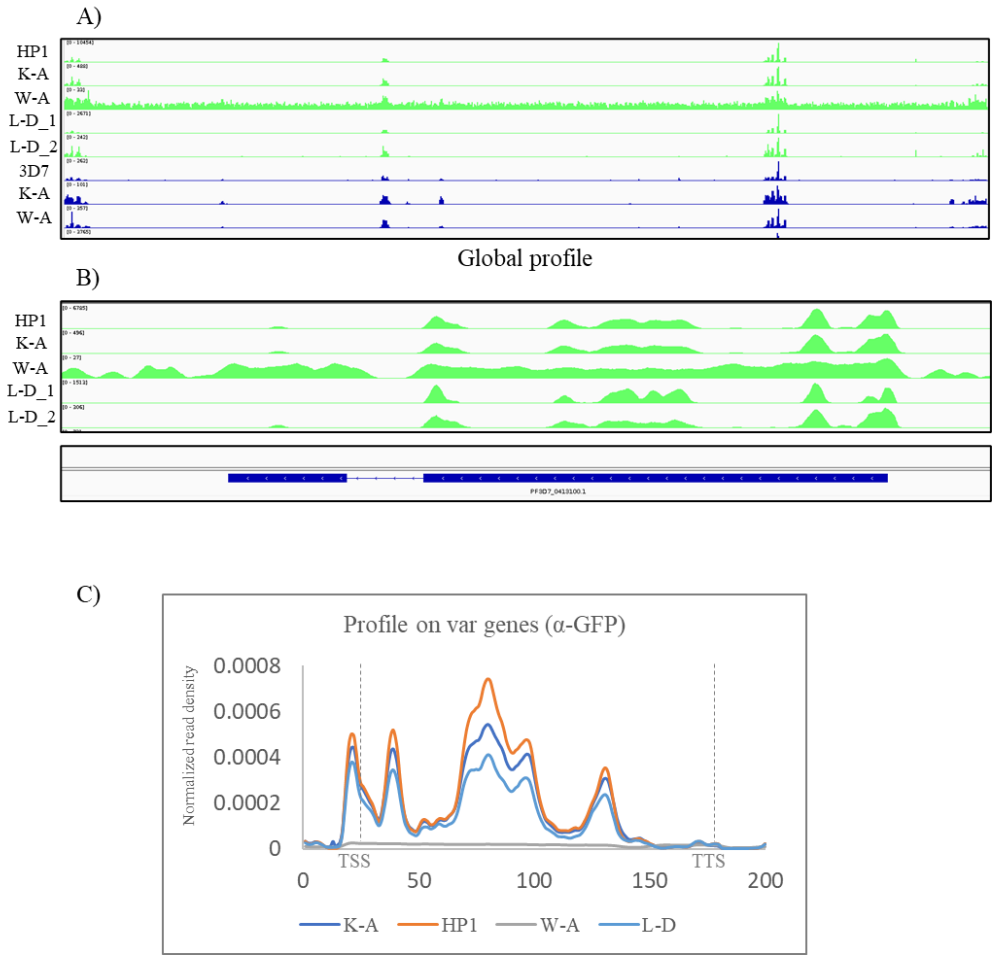


Figure 5.6: A) global profile of PfHP1 wildtype and mutants (green color represent ChIP pull down with α-GFP antibody and blue represents the pull down with α-PfHP1 antibody), B) Wildtype and mutant profile on *var* loci, C) Mean tag density plotted against gene length of *var* genes, normalized to number of reads.

Even though we observed reduced binding affinity for PfHP1 mutants, we could not find any growth defect or phenotypic manifestation of this overexpression. Probably, because the mutants are expressed in a background of the wildtype PfHP1. To test whether wildtype PfHP1 has any change in the mutant lines we did another ChIP sequencing in the W-A and K-A lines with anti-PfHP1 antibody (Fig. 5.6A). We observed proper PfHP1 localization and distribution in the genome. The contribution from the mutants might be very little owing to its low expression level (for W-A and L-D) at the protein level. Hence, we decided to generate mutations in the PfHP1 loci completely eliminating the effect of the wildtype protein.

5.4.3 Generation and validation of conditional mutants of PfHP1

Initially, we started with a CRISPR-Cas9 strategy that deletes the wildtype PfHP1 and inserts the point mutation. However, we could not generate the mutants even after several attempts. We speculated that the point mutations could be detrimental. Next, we decided to go for a conditional expression of mutant strategy as shown in the schematic in the Figure 5.7A. For the gene editing we used CRISPR-Cas9 with self-cleaving gRNAs that can target the start and the end of the gene in a single transfection (Fig. 5.7B) (Gao and Zhao 2014). In this strategy, we will have dual tagging of the PfHP1, where wildtype is tagged with mScarlet-I and mutants are tagged with the GFP. Upon addition of rapamycin drug, the DiCre gets activated and recombines the loxP flanking the wildtype PfHP1-mScarlet (Fig. 5.7C). This results in a conditionally expressing a PfHP1-GFP.

We could successfully generate PfHP1 mutant lines for all four mutations: E-A, K-A, W-A, L-D and wildtype PfHP1 (where with or without rapamycin treatment wildtype will express, Fig. 5.7C). The lines showed expression of mSca-I tagged PfHP1 (Fig. 5.7C). Upon rapamycin treatment, it showed GFP-tagged-PfHP1 expression (Fig. 5.7C). Further, PCR and western blotting was done to successfully validate the recombinant lines. We used 5’ UTR and 3’ UTR for amplification of the loci, the expected size of PfHP1 in 3D7 wildtype, the transgenics with or without rapamycin treatment vary (Fig. 5.8A). We confirmed that all lines have insertion of the correct size. Further, we confirmed the rapamycin treatment is working in the 100nM for 4 hrs, in PfHP1, W-A and K-A lines by PCR (Fig. 5.8A). Also, western blot of the PfHP1 line showed

the PfHP1-GFP band, the PfHP1-mScarlet-I band at the correct size (Fig.5.8B). We also show that K-A mutant with rapamycin treatment shows a band at the size of PfHP1-GFP, whereas no band without rapamycin treatment when probed with anti-GFP antibody (Fig. 5.8C).

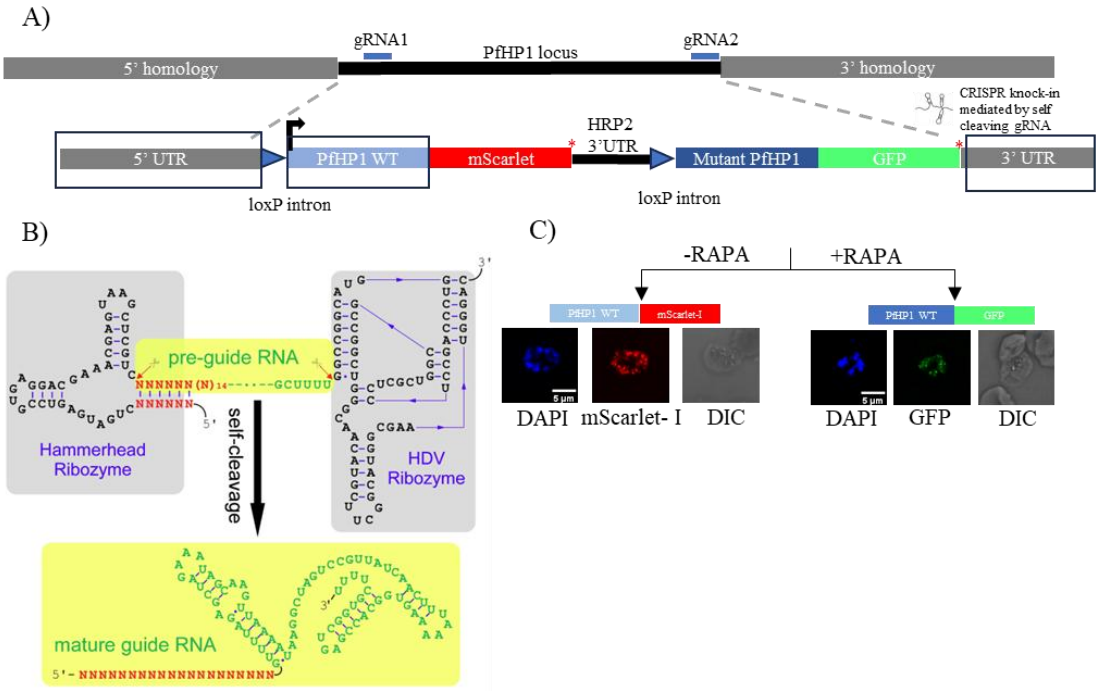


Figure 5.7: A) Schematic representing design of conditional expression of PfHP1 mutant lines (the black boxes indicate the homology region and red asterisk indicate stop codon), B) self-cleaving RNA for the delivery of two gRNA (image taken from Gao and Zhao 2014, Integr. Plant Biol), C) PfHP1 dual color tagging with or without rapamycin treatment.

5.4.4 Point mutations in PfHP1 disrupt its puncta formation in the nucleus

Further we did live cell imaging on these lines, and observed mScarlet-I signals and puncta, like localization of PfHP1 in the nucleus (Fig. 5.9). We also did an IFA on these lines using anti-Ser-2-phosphorylation antibody for RNA Pol II to see if PfHP1and RNA Pol II have exclusive localization (Fig. 5.11). We observe that our speculation was indeed true. This further supports our hypothesis that there is an exclusion of transcription machinery from the PfHP1 droplets and active *var* genes are also part of these 'transcription factories'.

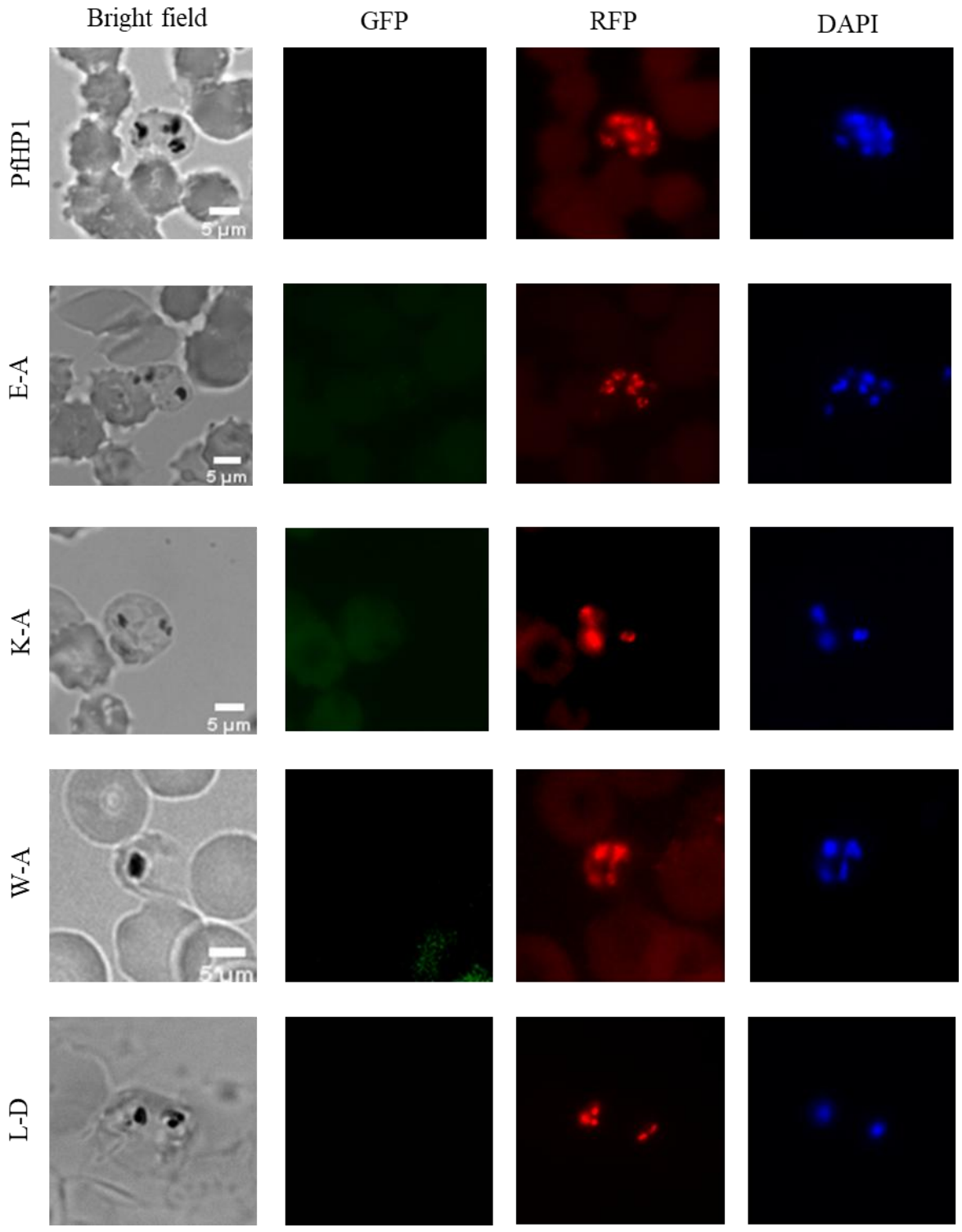


Figure 5.9: Live image showing mScarlet signal of PfHP1 wildtype when there is no rapa treatment.

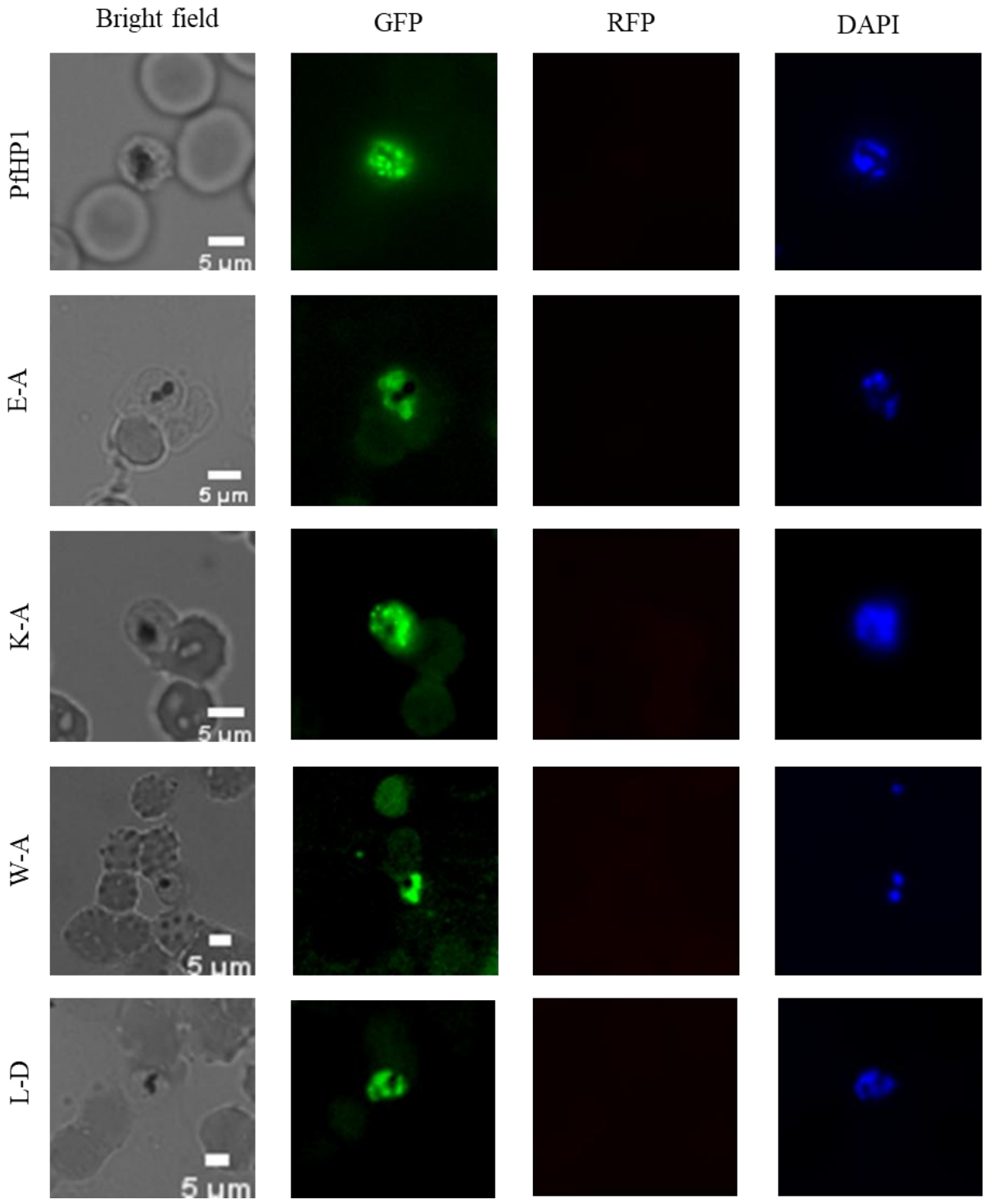


Figure 5.10: A) Live image showing GFP signal when the lines were given rapa treatment.

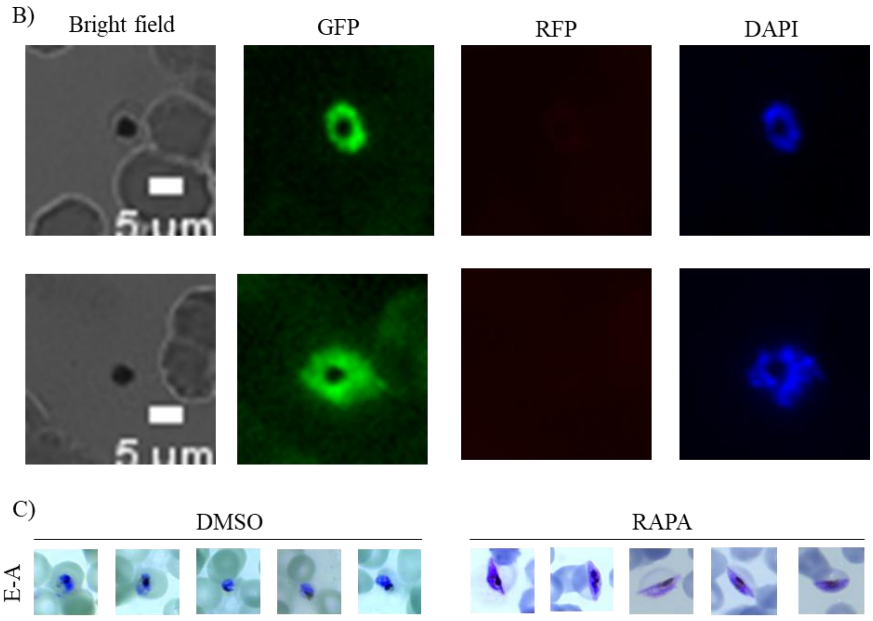


Figure 5.10: B) live imaging of K-A mutant line, 4 days after the rapa treatment, C) Bright field images of E-A mutant showing sexual conversion upon rapa treatment, images were taken 8 days after the rapa treatment.

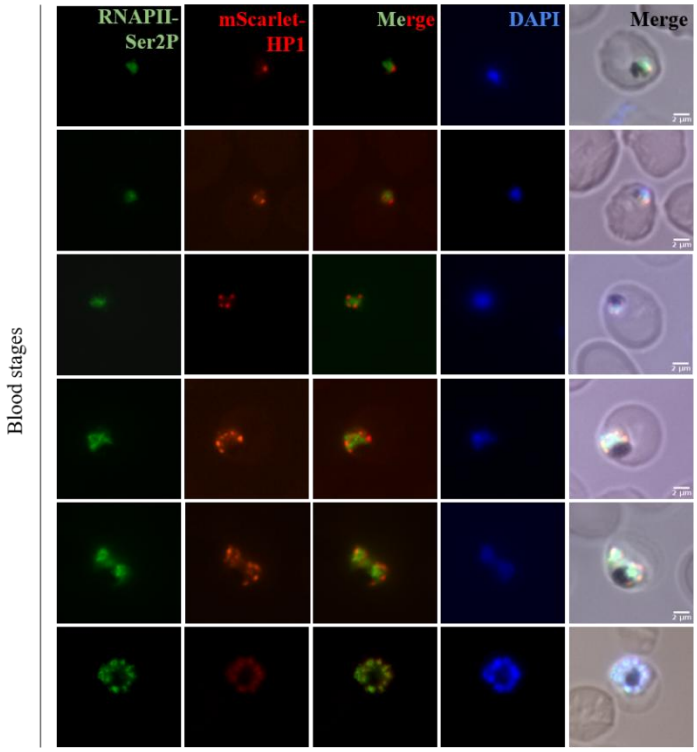


Figure 5.11: IFA with Serine-2-phosphorylation antibody on PfHP1 mScarlet-I line. The punctate localization of PfHP1 was observed.

To study the localization of the mutant PfHP1s we did live imaging of rapamycin treated cells in their late blood stages in the same generation (gen1) as of rapamycin induction. The mutants as well as wildtype gave GFP with no or very little background from the mScarlet-I, indicating that our rapa treatment is working efficiently. We observed very compact puncta for PfHP1 and K-A in gen1. E-A and W-A mutations showed a more dispersed localization in the nucleus and no puncta. Whereas, L-D showed puncta formation however slightly bigger and dispersed than the wildtyp e (Fig. 5.10A).

The K-A mutation showed dramatic changes when performed *in vitro* droplet experiments, where we observed a complete lack of LLPS even after adding PEG8000 or DNA. But no such phenotype was observed in our live cell imaging experiment. Hence, we imaged K-A in the gen2 (next cycle after rapa treatment) to see if there is any deviation in the localization in comparison to the wildtype (Fig. 5.10B). We observed that the puncta formation was disrupted (Fig. 5.10B) and the mutation was lethal to the parasites (data not shown).

We also observed that the E-A mutant line doesn't grow beyond gen2 and showed a very high sexual conversion rate (Fig. 5.10C). Another mutant W-A shows growth defects by very slow growth upon rapamycin induction; this observation needs to be studied in more detail to understand whether the merozoite formation or invasion rates are affected.

5.5 Discussion

In this Chapter we have tried to delineate the role of PfHP1 phase separation in the parasite biology. We have combined the information gained from the *in vitro* assays to design transgenics and draw important conclusions about the PfHP1 function.

Two previous studies on PfHP1 have used Cre-loxP based conditional expression of the mutants. We used a similar approach, but we have made small improvements by dual fluorescent tagging of the mutant vs. wildtype. Also, we have used an innovative approach for generating the transgenics by using a self-cleaving RNA that can give rise to two gRNAs that can simultaneously target the start and the end of the PfHP1 gene.

In the *in vitro* studies we learned that E-A, K-A and W-A mutations affect the phase separation property of PfHP1 and the former two mutants affect the co-condensation of the PfHP1 with DNA. We see that these mutations indeed affect the chromatin binding and specific localization of the protein. We deduce that the phase separation property of PfHP1 is important for its function *in vivo*.

We observe that, E-A mutant shows drastic changes in localization and overall survival of the parasites. This phenotype almost mimics the PfHP1 knock-out / depletion (Brancucci et al. 2014). The K-A mutant also affects the survival of the parasites but only after a few cycles after the rapa treatment. Interestingly, the localization gets disrupted only after two cycles. The transcriptomic change that might be bringing about such a temporal difference will be interesting to look at.

The W-A mutant also showed a defective nuclear localization and growth. The overexpression lines also showed a noticeable change in the chromatin binding of this mutant. We propose that W29 amino acid could be instrumental for the H3K9me3 binding via hydrophobic interactions; hence the observations. This is also supported by the alignment with Swi6 that shows the W29 residue aligns with the hydrophobic cage in the Swi6 that helps in H3K9me3 binding (section 3.4.6).

While we observe defective chromatin binding for the mutants in the overexpression line, this phenotype has been rescued by the wildtype expressing in the background. In the second strategy we overcome this limitation by knocking out the wildtype PfHP1 completely. It will be interesting to look at the *var* gene expression and chromatin level changes in such a system. We plan to perform transcriptomics and ChIP studies on these mutant lines (**These set of experiments are ongoing, hence we have not included it in the thesis).

The unequivocal evidence for the *in vivo* phase separation of PfHP1 and physical nature of it would come from fluorescence microscopy based studies (Alberti et al. 2019). However, our progress is hindered by the technology limitations and the nanoscale size of the PfHP1 puncta. While we acknowledge the limitation, we have devised inventive strategies, and developed powerful tools to study the mechanism of heterochromatin formation and *var* regulation.

Chapter 6: Conclusion and Future perspectives

Humanity has long been suffering from *P. falciparum* infections and the lethality associated with it. The core problem lies in the shortcomings in the current strategies to combat malaria. The emergence of drug-resistant parasites has been a major setback (Pandit et al. 2023). Understanding the parasite biology better is the only way to design a better cure. One major missing part of the puzzle is virulence gene switching and the mechanism of regulation. Despite the vast literature available on *var* genes, we are in shortage of a unifying mechanism that accommodates the unusual observations and contrasting results from various studies. We have employed a multidisciplinary approach where we utilized NGS, gene editing techniques, biochemistry, proteomics and single molecular experiments to understand *var* gene regulation holistically.

We have tried to understand the *var* regulation from the perspective of heterochromatin. It is an oversimplification to believe that heterochromatin 'just' represses *var* genes. PfHP1 lies at the heart of the *var* gene regulation.

Firstly, we looked at the epigenetic landscape of the *var* gene and this laid the foundation for the study. PfHP1 is a proxy for heterochromatin, showing dynamic changes in the distribution across the *var* genes. These might be marking the different 'states' of the *var* genes. We identified the *var* introns to be an important paradigm to study the heterochromatin dynamics.

Next, we looked at the biochemical properties of PfHP1 and showed that it can undergo LLPS like its homologs. This process is highly tunable, and can be affected by local concentrations of DNA, RNA and proteins (like H2A.Z). PfHP1 can also form co-condensates with the PfDNA. This process largely depends on the AT-rich DNA stretches as in the *var* introns. This is a unique observation that connects the atypically AT-rich genome and its compaction by the PfHP1.

In Chapter 5 we have applied the knowledge gained from our previous findings to understand the gene regulation in heterochromatin domains. We observe that PfHP1 point mutations that disrupt the phase separation also disrupts its chromatin binding function.

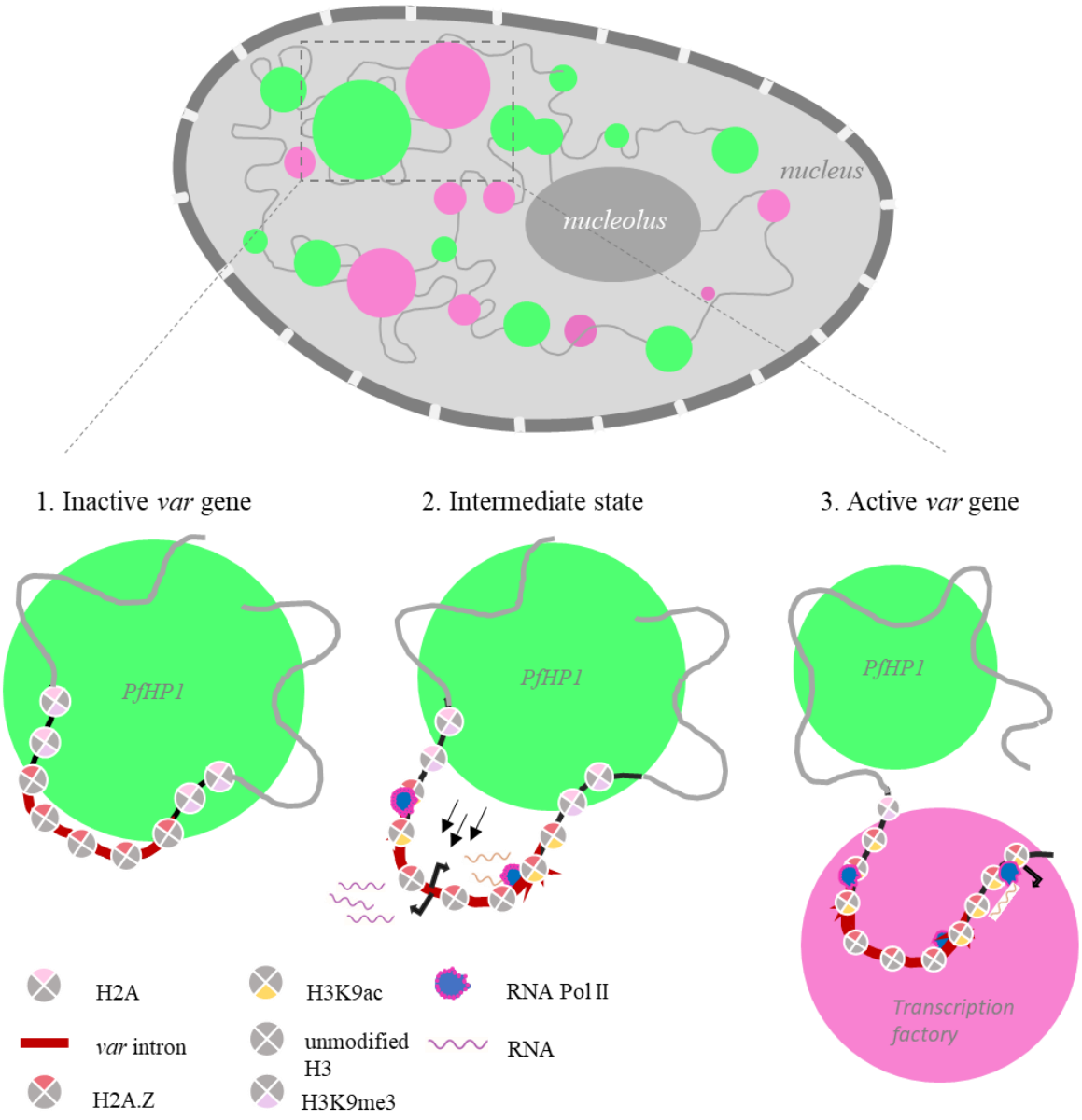


Figure 6.1: Proposed model of *var* gene regulation by LLPS of PfHP1.

With these new findings we propose a speculative model (Fig. 6.1). *P. falciparum* nuclei have LLPS droplets of PfHP1 that condenses DNA into a compact structure. The *var* genes are clustered in such droplets and maintained in a repressive environment by sequestration of other epigenetic regulators. These droplets also exclude active transcription machinery.

The introns of *var* genes might be acting as a 'door handle', which is always out of the PfHP1 droplet (even in the 'inactive state'). When the door knob is 'pulled' out of the droplets, it leads to 'opening' of *var* genes which is the active state. The *var* intronic transcription might be the tunable mechanism that causes local dissolution of the droplets (a continuum of 'intermediate states') and pulls the 'door' open. And this lets *var* genes loop out of the droplets.

The boundary of the droplets could be maintained by either PfH2A.Z alone or in combination with other insulator proteins. The promoter and intron of *var* genes contain insulator-like motifs that can be binding sites for many insulators. In *P. falciparum* there are very few insulator proteins identified so far. This is an interesting open question in the field.

Further, this study opens new gates for the expansion of our molecular understanding of the epigenetic regulation in *P. falciparum*. We have listed a few future directions that can be pursued;

1) The olfactory receptors (ORs) are an interesting paradigm of monoallelic expression much like *var* genes. The ORs are under suppression of H3K9me3 (Lyons et al. 2014). Recently, it has been identified that active ORs are associated with phase separated droplets that contain enhancer elements that enable active transcription (Florini et al. 2022). This is a concept that can be applied to *var* genes as well. The perinuclear expression site of *var* genes can also be a phase separated droplet clusters active *var* gene/s and transcription machinery (Fig. 6.1).
2) This study opens a new arena for *Plasmodium* epigenetics where phase separation of various proteins can be an important mechanism of gene regulation. One can imagine that phase separation could be a dynamic mechanism that parasites would adapt given the short time scale of each IDC and tight temporal regulation of transcription. The main hurdle is to overcome the technical limitation in understanding the physical nature of these droplets inside the cell.
3) Another aspect of PfHP1 mediated regulation, which we haven't explored in this study, is gametocytogenesis. We know that GDV1 causes expulsion of PfHP1 from the ap2-g loci (Filarsky et al. 2018). However, the biochemical nature of this process is still not known.

It is a proposition that GDV1 might be causing local dissolution of PfHP1 droplets and hence activates the gene.

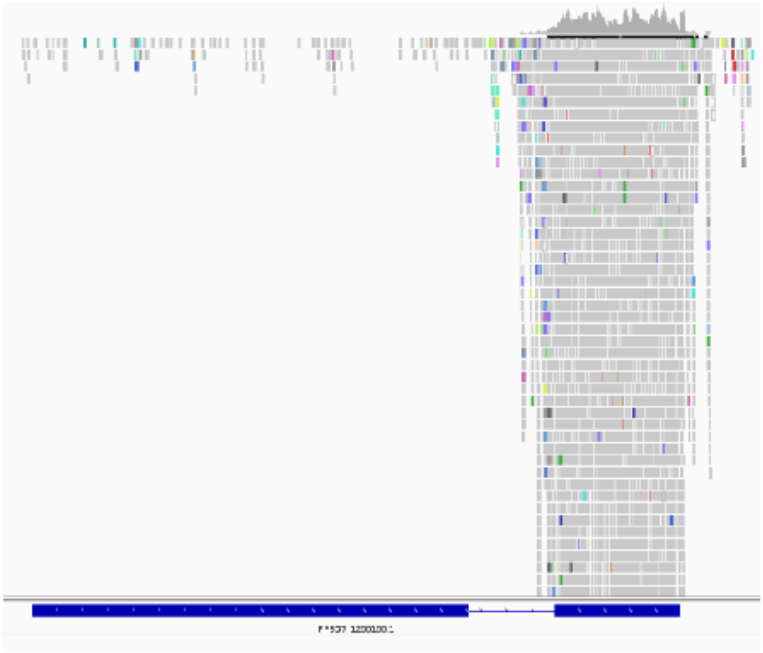
4) The *var* genes have very similar intronic and exon 2 sequences, giving rise to similar sequences for ncRNAs originated from there. The intronic transcription has always been studied in a cis-acting mechanism of regulation. It is time that we rethink this in the light of new information we have, that is LLPS of PfHP1. The ncRNAs can be cis and trans-acting by nature, that fine tunes the phase separation. This can be the reason why it has been hard to establish a one-on-one relation between the intronic transcription and full-length transcription of *var* genes. Knocking out all/ multiple introns at once can be a way to study the trans-acting nature of the intronic transcripts.
5) The intergenic regions and *var* introns are highly AT-rich, while the exons are moderately AT-rich. We identify the intronic regions to be nucleation sites of heterochromatin formation. It will be interesting to look at a range of AT-rich and GC-rich sequences for its co-condensation with PfHP1. We have developed an important paradigm to study the dynamics of PfHP1 compaction on *P. falciparum* specific sequences on single molecules of DNA.
6) The LLPS of proteins are also emerging as a new avenue for therapeutic targeting. This has already been extensively studied in diseases associated with protein phase separation such as cancer and Alzheimer's (Tong et al. 2022; Wang et al. 2021a; Wheeler 2020). The PTMs associated with the phase separation are often targeted in such cases. The point mutations we identified on PfHP1 can also be used as targets for drugs. This is underexplored in malaria drug research due to the lack of basic understanding of biochemical/physical properties of the parasite proteins. While phase separation is a promising target, the human body also contains numerous phase-separated proteins. This poses an enormous challenge for therapeutic development in this direction.

LIST OF PUBLICATIONS

1) Mamatharani DV, Purkayastha D*, Niederwieser I*, Rai SK, Gaur P, Ganji M, Mukhopadhyay S, Voss T, Karmodiya K. Phase separation of heterochromatin protein 1 (HP1) controls virulence gene silencing in *Plasmodium falciparum.* 2025 (*_equal contribution, manuscript submitted_).
2) Dave B, Kanyal A, Mamatharani DV, Karmodiya K. Pervasive sequence-level variation in the transcriptome of Plasmodium falciparum. NAR Genomics and Bioinformatics. 2022 Jun 1;4(2): lqac036.
3) Rawat M, Kanyal A, Choubey D, Deshmukh B, Malhotra R, Mamatharani DV, Rao AG, Karmodiya K. Identification of co-existing mutations and gene expression trends associated with K13-mediated artemisinin resistance in Plasmodium falciparum. Frontiers in Genetics. 2022 Apr 6; 13:824483.
4) Kanyal A, Deshmukh B, Davies H, Mamatharani DV, Farheen D, Treeck M, Karmodiya K. PfHDAC1 is an essential regulator of P. falciparum asexual proliferation and host cell invasion genes with a dynamic genomic occupancy responsive to artemisinin stress. MBio. 2024 Jun 12;15(6): e02377-23.

APPENDIX

A)



B)

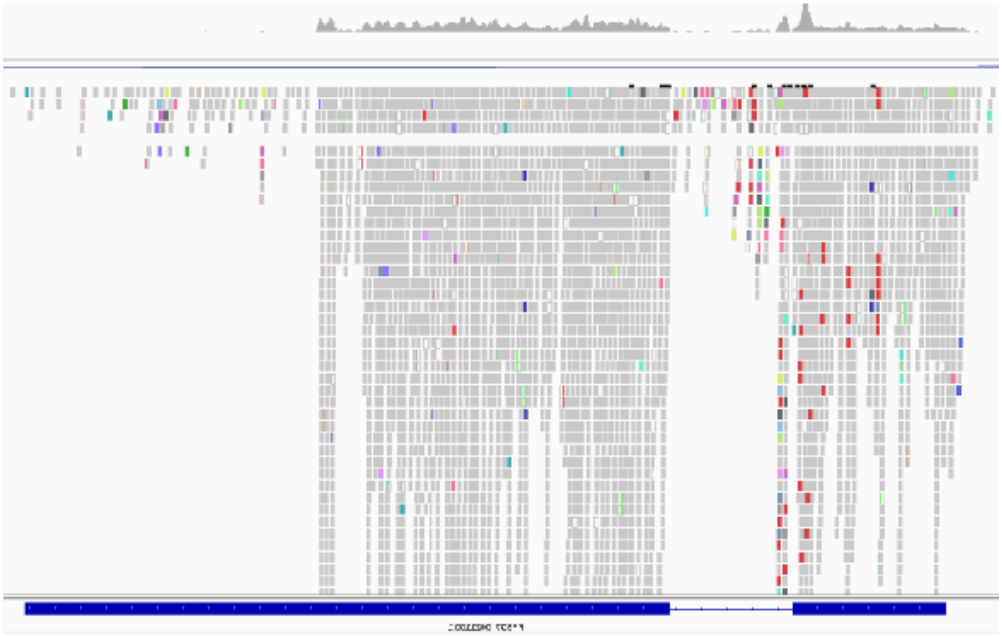


Figure 1: A) RNA sequencing reads mapping to only exon 2 of *var* gene Pf3D7_1200100, B) RNA sequencing reads mapping to exon1 and exon2 , but no splicing events in *var* gene Pf3D7_04211001.

Protein ID	Protein Name	Peptide count
PF3D7_1220900	Heterochromatin protein 1	9
PF3D7_1246200	actin I	5
PF3D7_0424600	Plasmodium exported protein (PHISTb)	2
PF3D7_1446600	centrin-2	2
PF3D7_0817900	High mobility group protein B2 (HMGB2)	2
PF3D7_0516200	40S ribosomal protein S11	2
PF3D7_0501000	Plasmodium exported protein, unknown function	2

Table 1: Results of IP-mass spectrometry showing that anti-PfHP1 antibody has pulled down PfHP1 along with other interacting proteins.

Protein ID	Protein name	Peptide count
PF3D7_1105000	H4	9
PF3D7_1105100	H2B	5
PF3D7_0320900	H2A.Z	4
PF3D7_1471100	exported protein 2	3
PF3D7_1006200	Alba3	2
PF3D7_0714000	H2B variant	4
PF3D7_0424600	PHISTb	1

Table 2: Top hits for pull down with pre-immune sera in IP-mass spectrometry.

Primer name	Sequence
SDM primers	
PfHP1_E(6-9)A_FP	GACAGGGTCAGATGCTGCTTTTGCTATTGGTGATATAC
PfHP1_E(6-9)A_RP	GTATATCACCAATAGCAAAAGCAGCATCTGACCCTGTC
PfHP1_S33C_FP	CTTGGGAACCCGAATGTAATTTAATAC
PfHP1_S33C_RP	GTATTAAATTACATTCGGGTTCCCAAG
PfHP1_S206C_FP	GAAGAAAAAAAAAGTTGTAGGGGTAATAG
PfHP1_S206C_RP	CTATTACCCCTACAACTTTTTTTTCTTC
PfHP1_K(17-20)A_FP	GATATACTTGAAATAGCTGCTGCTGCTAATGGTTTTATTTATTTAG
PfHP1_K(17-20)A_FP	CTAAATAAATAAAACCATTAGCAGCAGCAGCTATTTCAAGTATATC
L254A_FP	ATCCACAGGAACTCGCAAATTTTTTATTATC
L254A_RP	GATAATAAAAAATTTGCGAGTTCCTGTGGAT
W29A_FP	GGTTTTATTTATTTAGTAAAAGCGAAAGGATATTCAGATGATGAG
W29A_RP	CTCATCATCTGAATATCCTTTCGCTTTTACTAAATAAATAAAACC

Table 3 : List of primers for the SDM in Chapter 3.

Primer name	Sequence
601 Widom amplification primers	
pGEM-601_FP	GATGTATATATCTGACACGTGC
pGEM-601_RP	GCCCTGGAGAATCCCGGTGC
PfDNA amplification primers	
HP1_FP	CGCGCGAATTCTATGACAGGGTCAGATGAAGAA
HP1_RP	ATGCGGCCGCTTAAGCTGTACGGTATCTT

Table 4: List of primers for PCR amplification of DNA needed for the droplet assay.

Primer name	Sequence
GC-rich RNA_IVT PCR Primers	
PF3D7_0712100 F	TAATACGACTCACTATAGAAGCTGCCTCAGTAGCCCAGTGGTTAG
PF3D7_0712100 R	CTGCGCCACCCCCCTCCGAGT
AU-rich RNA_IVT PCR Primers	
T7 promoter FP	TAATACGACTCACTATAG
T7 terminator RP	CAAAAAACCCCTCAAGAC
Pf3D7_07:340595-340821 FP	TAATACGACTCACTATAGTATATATATATGTAGTTAAAAAAATGTAGTTTTTCTTTTTACATTTAGATGGTGTCATATATTAACTGCTAAGATAAACCTAATAAAAATTATAGTTTCATT
Pf3D7_07:340595-340821 RP	CAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAGGCTTTATAATATTGTGTTTTATATCATTTTTGAGGAATTGTTAAAGGAATAATGAAACTATAATTTTTATTAGG

Table 5: List of primers for the PCR amplification of IVT template

Accession	Name	Peptides(95%)
PF3D7_1220900.1-p1	| transcript=PF3D7_1220900.1 | gene=PF3D7_1220900 | organism=Plasmodium_falciparum_3D7 | gene_product=heterochromatin protein 1 | transcript_product=heterochromatin protein 1 | location=Pf3D7_12_v3:831252-832052(-) | protein_length=266 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	216
PF3D7_1105100.1-p1	| transcript=PF3D7_1105100.1 | gene=PF3D7_1105100 | organism=Plasmodium_falciparum_3D7 | gene_product=histone H2B | transcript_product=histone H2B | location=Pf3D7_11_v3:226065-226418(+) | protein_length=117 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	4
PF3D7_0500800.1-p1	| transcript=PF3D7_0500800.1 | gene=PF3D7_0500800 | organism=Plasmodium_falciparum_3D7 | gene_product=mature parasite-infected erythrocyte surface antigen | transcript_product=mature parasite-infected erythrocyte surface antigen | location=Pf3D7_05_v3:43327-47761(-) | protein_length=1434 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	4
PF3D7_1105000.1-p1	| transcript=PF3D7_1105000.1 | gene=PF3D7_1105000 | organism=Plasmodium_falciparum_3D7 | gene_product=histone H4 | transcript_product=histone H4 | location=Pf3D7_11_v3:221818-222129(-) | protein_length=103 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	6
PF3D7_0320900.1-p1	| transcript=PF3D7_0320900.1 | gene=PF3D7_0320900 | organism=Plasmodium_falciparum_3D7 | gene_product=histone H2A.Z | transcript_product=histone H2A.Z | location=Pf3D7_03_v3:875213-876295(+) | protein_length=158 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_0814200.1-p1	| transcript=PF3D7_0814200.1 | gene=PF3D7_0814200 | organism=Plasmodium_falciparum_3D7 | gene_product=DNA/RNA-binding protein Alba 1 | transcript_product=DNA/RNA-binding protein Alba 1 | location=Pf3D7_08_v3:687340-688086(+) | protein_length=248 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_1352500.1-p1	| transcript=PF3D7_1352500.1 | gene=PF3D7_1352500 | organism=Plasmodium_falciparum_3D7 | gene_product=thioredoxin-related protein, putative | transcript_product=thioredoxin-related protein, putative | location=Pf3D7_13_v3:2093365-2094337(-) | protein_length=208 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_1357100.1-p1	| transcript=PF3D7_1357100.1 | gene=PF3D7_1357100 | organism=Plasmodium_falciparum_3D7 | gene_product=elongation factor 1-alpha | transcript_product=elongation factor 1-alpha | location=Pf3D7_13_v3:2268901-2270232(+) | protein_length=443 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2

PF3D7_1357000.1-p1	| transcript=PF3D7_1357000.1 | gene=PF3D7_1357000 | organism=Plasmodium_falciparum_3D7 | gene_product=elongation factor 1-alpha | transcript_product=elongation factor 1-alpha | location=Pf3D7_13_v3:2265817-2267148(-) | protein_length=443 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_0831700.1-p1	| transcript=PF3D7_0831700.1 | gene=PF3D7_0831700 | organism=Plasmodium_falciparum_3D7 | gene_product=heat shock protein 70 | transcript_product=heat shock protein 70 | location=Pf3D7_08_v3:1365467-1367506(-) | protein_length=679 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_0917900.1-p1	| transcript=PF3D7_0917900.1 | gene=PF3D7_0917900 | organism=Plasmodium_falciparum_3D7 | gene_product=heat shock protein 70 | transcript_product=heat shock protein 70 | location=Pf3D7_09_v3:737985-740276(+) | protein_length=652 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1016300.1-p1	| transcript=PF3D7_1016300.1 | gene=PF3D7_1016300 | organism=Plasmodium_falciparum_3D7 | gene_product=GBP130 protein | transcript_product=GBP130 protein | location=Pf3D7_10_v3:650898-653567(-) | protein_length=824 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_1410400.1-p1	| transcript=PF3D7_1410400.1 | gene=PF3D7_1410400 | organism=Plasmodium_falciparum_3D7 | gene_product=rhoptry-associated protein 1 | transcript_product=rhoptry-associated protein 1 | location=Pf3D7_14_v3:420442-422790(+) | protein_length=782 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_0813300.1-p1	| transcript=PF3D7_0813300.1 | gene=PF3D7_0813300 | organism=Plasmodium_falciparum_3D7 | gene_product=NPL domain-containing protein, putative | transcript_product=NPL domain-containing protein, putative | location=Pf3D7_08_v3:656078-657630(+) | protein_length=315 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1149000.1-p1	| transcript=PF3D7_1149000.1 | gene=PF3D7_1149000 | organism=Plasmodium_falciparum_3D7 | gene_product=antigen 332, DBL-like protein | transcript_product=antigen 332, DBL-like protein | location=Pf3D7_11_v3:1950210-1968726(+) | protein_length=6093 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1471100.1-p1	| transcript=PF3D7_1471100.1 | gene=PF3D7_1471100 | organism=Plasmodium_falciparum_3D7 | gene_product=exported protein 2 | transcript_product=exported protein 2 | location=Pf3D7_14_v3:2905163-2906565(-) | protein_length=287 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1246400.1-p1	| transcript=PF3D7_1246400.1 | gene=PF3D7_1246400 | organism=Plasmodium_falciparum_3D7 | gene_product=myosin A-tail interacting protein | transcript_product=myosin A-tail interacting protein | location=Pf3D7_12_v3:1931970-1932584(+) | protein_length=204 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1

PF3D7_1033200.1-p1	| transcript=PF3D7_1033200.1 | gene=PF3D7_1033200 | organism=Plasmodium_falciparum_3D7 | gene_product=early transcribed membrane protein 10.2 | transcript_product=early transcribed membrane protein 10.2 | location=Pf3D7_10_v3:1335222-1336289(+) | protein_length=355 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_0714000.1-p1	| transcript=PF3D7_0714000.1 | gene=PF3D7_0714000 | organism=Plasmodium_falciparum_3D7 | gene_product=histone H2B variant | transcript_product=histone H2B variant | location=Pf3D7_07_v3:643101-643472(-) | protein_length=123 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_0601900.1-p1	| transcript=PF3D7_0601900.1 | gene=PF3D7_0601900 | organism=Plasmodium_falciparum_3D7 | gene_product=conserved Plasmodium protein, unknown function | transcript_product=conserved Plasmodium protein, unknown function | location=Pf3D7_06_v3:78190-78784(+) | protein_length=128 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1250100.1-p1	| transcript=PF3D7_1250100.1 | gene=PF3D7_1250100 | organism=Plasmodium_falciparum_3D7 | gene_product=osmiophilic body protein G377 | transcript_product=osmiophilic body protein G377 | location=Pf3D7_12_v3:2045070-2054429(-) | protein_length=3119 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_0918000.1-p1	| transcript=PF3D7_0918000.1 | gene=PF3D7_0918000 | organism=Plasmodium_falciparum_3D7 | gene_product=glideosome-associated protein 50 | transcript_product=glideosome-associated protein 50 | location=Pf3D7_09_v3:742494-743684(-) | protein_length=396 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1127800.1-p1	| transcript=PF3D7_1127800.1 | gene=PF3D7_1127800 | organism=Plasmodium_falciparum_3D7 | gene_product=TFIIS central domain-containing protein, putative | transcript_product=TFIIS central domain-containing protein, putative | location=Pf3D7_11_v3:1077288-1081550(-) | protein_length=1420 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1

Table 6: Mass spectrometry result of protein hits and peptide counts for pellet fraction in droplet assay with *P. falciparum* nuclear extract.

Accession	Name	Peptides(95%)
PF3D7_1220900.1-p1	| transcript=PF3D7_1220900.1 | gene=PF3D7_1220900 | organism=Plasmodium_falciparum_3D7 | gene_product=heterochromatin protein 1 | transcript_product=heterochromatin protein 1 | location=Pf3D7_12_v3:831252-832052(-) | protein_length=266 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	161
PF3D7_0930300.1-p1	| transcript=PF3D7_0930300.1 | gene=PF3D7_0930300 | organism=Plasmodium_falciparum_3D7 | gene_product=merozoite surface protein 1 | transcript_product=merozoite surface protein 1 | location=Pf3D7_09_v3:1201812-1206974(+) | protein_length=1720 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	4
PF3D7_0500800.1-p1	| transcript=PF3D7_0500800.1 | gene=PF3D7_0500800 | organism=Plasmodium_falciparum_3D7 | gene_product=mature parasite-infected erythrocyte surface antigen | transcript_product=mature parasite-infected erythrocyte surface antigen | location=Pf3D7_05_v3:43327-47761(-) | protein_length=1434 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	4
PF3D7_1232100.1-p1	| transcript=PF3D7_1232100.1 | gene=PF3D7_1232100 | organism=Plasmodium_falciparum_3D7 | gene_product=60 kDa chaperonin | transcript_product=60 kDa chaperonin | location=Pf3D7_12_v3:1328311-1331371(-) | protein_length=718 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_0721100.1-p1	| transcript=PF3D7_0721100.1 | gene=PF3D7_0721100 | organism=Plasmodium_falciparum_3D7 | gene_product=conserved protein, unknown function | transcript_product=conserved protein, unknown function | location=Pf3D7_07_v3:916434-917168(+) | protein_length=244 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_1134000.1-p1	| transcript=PF3D7_1134000.1 | gene=PF3D7_1134000 | organism=Plasmodium_falciparum_3D7 | gene_product=heat shock protein 70 | transcript_product=heat shock protein 70 | location=Pf3D7_11_v3:1320114-1322105(+) | protein_length=663 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	2
PF3D7_1352500.1-p1	| transcript=PF3D7_1352500.1 | gene=PF3D7_1352500 | organism=Plasmodium_falciparum_3D7 | gene_product=thioredoxin-related protein, putative | transcript_product=thioredoxin-related protein, putative | location=Pf3D7_13_v3:2093365-2094337(-) | protein_length=208 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_0318200.1-p1	| transcript=PF3D7_0318200.1 | gene=PF3D7_0318200 | organism=Plasmodium_falciparum_3D7 | gene_product=DNA-directed RNA polymerase II subunit RPB1 | transcript_product=DNA-directed RNA polymerase II subunit RPB1 | location=Pf3D7_03_v3:748465-755838(+) | protein_length=2457 |	1

	sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	
RRRRRPF3D7_0503200.1-p1	REVERSED | transcript=PF3D7_0503200.1 | gene=PF3D7_0503200 | organism=Plasmodium_falciparum_3D7 | gene_product=cell division control protein 6, putative | transcript_product=cell division control protein 6, putative | location=Pf3D7_05_v3:130453-133392(+) | protein_length=979 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1471100.1-p1	| transcript=PF3D7_1471100.1 | gene=PF3D7_1471100 | organism=Plasmodium_falciparum_3D7 | gene_product=exported protein 2 | transcript_product=exported protein 2 | location=Pf3D7_14_v3:2905163-2906565(-) | protein_length=287 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1464600.1-p1	| transcript=PF3D7_1464600.1 | gene=PF3D7_1464600 | organism=Plasmodium_falciparum_3D7 | gene_product=serine/threonine protein phosphatase UIS2, putative | transcript_product=serine/threonine protein phosphatase UIS2, putative | location=Pf3D7_14_v3:2619679-2624187(+) | protein_length=1442 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1442700.1-p1	| transcript=PF3D7_1442700.1 | gene=PF3D7_1442700 | organism=Plasmodium_falciparum_3D7 | gene_product=conserved Plasmodium protein, unknown function | transcript_product=conserved Plasmodium protein, unknown function | location=Pf3D7_14_v3:1742113-1749790(+) | protein_length=2511 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1439800.1-p1	| transcript=PF3D7_1439800.1 | gene=PF3D7_1439800 | organism=Plasmodium_falciparum_3D7 | gene_product=vesicle-associated membrane protein, putative | transcript_product=vesicle-associated membrane protein, putative | location=Pf3D7_14_v3:1618819-1619794(+) | protein_length=241 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1237700.1-p1	| transcript=PF3D7_1237700.1 | gene=PF3D7_1237700 | organism=Plasmodium_falciparum_3D7 | gene_product=conserved protein, unknown function | transcript_product=conserved protein, unknown function | location=Pf3D7_12_v3:1572371-1573557(+) | protein_length=210 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1016300.1-p1	| transcript=PF3D7_1016300.1 | gene=PF3D7_1016300 | organism=Plasmodium_falciparum_3D7 | gene_product=GBP130 protein | transcript_product=GBP130 protein | location=Pf3D7_10_v3:650898-653567(-) | protein_length=824 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_0819300.1-p1	| transcript=PF3D7_0819300.1 | gene=PF3D7_0819300 | organism=Plasmodium_falciparum_3D7 | gene_product=conserved Plasmodium protein, unknown function | transcript_product=conserved Plasmodium protein, unknown function | location=Pf3D7_08_v3:874367-877498(+) | protein_length=1043 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1

PF3D7_0532100.1-p1	| transcript=PF3D7_0532100.1 | gene=PF3D7_0532100 | organism=Plasmodium_falciparum_3D7 | gene_product=early transcribed membrane protein 5 | transcript_product=early transcribed membrane protein 5 | location=Pf3D7_05_v3:1301224-1301769(+) | protein_length=181 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1129000.1-p1	| transcript=PF3D7_1129000.1 | gene=PF3D7_1129000 | organism=Plasmodium_falciparum_3D7 | gene_product=spermidine synthase | transcript_product=spermidine synthase | location=Pf3D7_11_v3:1121930-1123377(+) | protein_length=321 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1403400.1-p1	| transcript=PF3D7_1403400.1 | gene=PF3D7_1403400 | organism=Plasmodium_falciparum_3D7 | gene_product=conserved Plasmodium protein, unknown function | transcript_product=conserved Plasmodium protein, unknown function | location=Pf3D7_14_v3:125986-129877(+) | protein_length=647 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_1105100.1-p1	| transcript=PF3D7_1105100.1 | gene=PF3D7_1105100 | organism=Plasmodium_falciparum_3D7 | gene_product=histone H2B | transcript_product=histone H2B | location=Pf3D7_11_v3:226065-226418(+) | protein_length=117 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	1
PF3D7_0831700.1-p1	| transcript=PF3D7_0831700.1 | gene=PF3D7_0831700 | organism=Plasmodium_falciparum_3D7 | gene_product=heat shock protein 70 | transcript_product=heat shock protein 70 | location=Pf3D7_08_v3:1365467-1367506(-) | protein_length=679 | sequence_SO=chromosome | SO=protein_coding_gene | is_pseudo=false	4

Table 7: Mass spectrometry result of protein hits and peptide counts for supernatant fraction in droplet assay with *P. falciparum* nuclear extract.

BIBLIOGRAPHY

Alberti S, Gladfelter A, Mittag T. 2019. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell*. 176(3):419–34

Alghamdi JM, Al-Qahtani AA, Alhamlan FS, Al-Qahtani AA. 2024. Recent advances in the treatment of malaria. *Pharmaceutics*. 16(11)

Avraham I, Schreier J, Dzikowski R. 2012. Insulator-like pairing elements regulate silencing and mutually exclusive expression in the malaria parasite *Plasmodium* falciparum. *Proc Natl Acad Sci USA*. 109(52):E3678-86

Ay F, Bunnik EM, Varoquaux N, Bol SM, Prudhomme J, et al. 2014. Three-dimensional modeling of the *P. falciparum* genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. *Genome Res.* 24(6):974–88

Balaji S, Babu MM, Iyer LM, Aravind L. 2005. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* 33(13):3994–4006

Banani SF, Lee HO, Hyman AA, Rosen MK. 2017. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18(5):285–98

Bártfai R, Hoeijmakers WAM, Salcedo-Amaya AM, Smits AH, Janssen-Megens E, et al. 2010. H2A.Z demarcates intergenic regions of the *Plasmodium falciparum* epigenome that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog.* 6(12):e1001223

Baruch DI, Pasloske BL, Singh HB, Bi X, Ma XC, et al. 1995. Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of

parasitized human erythrocytes. *Cell*. 82(1):77–87

Beet EA. 1946. Sickle cell disease in the Balovale District of Northern Rhodesia. *East Afr. Med. J.* 23:75–86

Bell AC, West AG, Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*. 98(3):387–96

Bhowmick K, Tehlan A, Sunita, Sudhakar R, Kaur I, et al. 2020. *Plasmodium falciparum* GCN5 acetyltransferase follows a novel proteolytic processing pathway that is essential for its function. *J. Cell Sci.* 133(1):

Bian G, Joshi D, Dong Y, Lu P, Zhou G, et al. 2013. Wolbachia invades Anopheles stephensi populations and induces refractoriness to *Plasmodium* infection. *Science*. 340(6133):748–51

Blasco B, Leroy D, Fidock DA. 2017. Antimalarial drug resistance: linking *Plasmodium falciparum* parasite biology to the clinic. *Nat. Med.* 23(8):917–28

Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, et al. 2018. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*. 175(7):1842-1855.e16

Bozdech Z, Llinás M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. 2003. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* 1(1):E5

Brancucci NMB, Bertschi NL, Zhu L, Niederwieser I, Chin WH, et al. 2014. Heterochromatin protein 1 secures survival and transmission of malaria parasites. *Cell Host Microbe*. 16(2):165–76

Brancucci NMB, Gerdt JP, Wang C, De Niz M, Philip N, et al. 2017. Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite *Plasmodium*

falciparum. *Cell*. 171(7):1532-1544.e15

Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, et al. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*. 324(5935):1729–32

Braun S, Garcia JF, Rowley M, Rougemaille M, Shankar S, Madhani HD. 2011. The Cul4-Ddb1(Cdt)^2 ubiquitin ligase inhibits invasion of a boundary-associated antisilencing factor into heterochromatin. *Cell*. 144(1):41–54

Brown IN, Brown KN, Hills LA. 1968. Immunity to malaria: the antibody response to antigenic variation by *Plasmodium knowlesi. Immunology*. 14(1):127–38

Brown KN, Brown IN. 1965. Immunity to malaria: antigenic variation in chronic infections of *Plasmodium knowlesi*. *Nature*. 208(5017):1286–88

Brumpt E. Les parasites du paludisme des chimpanzés

Bryant JM, Regnault C, Scheidig-Benatar C, Baumgarten S, Guizetti J, Scherf A. 2017. CRISPR/Cas9 Genome Editing Reveals That the Intron Is Not Essential for var2csa Gene Activation or Silencing in *Plasmodium falciparum*. *MBio*. 8(4):

Bui HTN, Niederwieser I, Bird MJ, Dai W, Brancucci NMB, et al. 2019. Mapping and functional analysis of heterochromatin protein 1 phosphorylation in the malaria parasite *Plasmodium falciparum*. *Sci. Rep.* 9(1):16720

Bui HTN, Passecker A, Brancucci NMB, Voss TS. 2021. Investigation of Heterochromatin Protein 1 Function in the Malaria Parasite *Plasmodium falciparum* Using a Conditional Domain Deletion and Swapping Approach. *mSphere*. 6(1):

Bunnik EM, Cook KB, Varoquaux N, Batugedara G, Prudhomme J, et al. 2018. Changes in genome organization of parasite-specific gene families during the *Plasmodium*

transmission stages. *Nat. Commun.* 9(1):1910

Calderwood MS, Gannoun-Zaki L, Wellems TE, Deitsch KW. 2003. *Plasmodium falciparum var* genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J. Biol. Chem.* 278(36):34125–32

Canzio D, Chang EY, Shankar S, Kuchenbecker KM, Simon MD, et al. 2011. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol. Cell*. 41(1):67–81

Canzio D, Larson A, Narlikar GJ. 2014. Mechanisms of functional promiscuity by HP1 proteins. *Trends Cell Biol.* 24(6):377–86

Canzio D, Liao M, Naber N, Pate E, Larson A, et al. 2013. A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature*. 496(7445):377–81

Carrington E, Cooijmans RHM, Keller D, Toenhake CG, Bártfai R, Voss TS. 2021. The ApiAP2 factor PfAP2-HC is an integral component of heterochromatin in the malaria parasite *Plasmodium falciparum*. *iScience*. 24(5):102444

Carter R, Mendis KN. 2002. Evolutionary and historical aspects of the burden of malaria. *Clin. Microbiol. Rev.* 15(4):564–94

Chahine Z, Le Roch KG. 2022. Decrypting the complexity of the human malaria parasite biology through systems biology approaches. *Front. Syst. Biol.* 2:

Chen PB, Ding S, Zanghì G, Soulard V, DiMaggio PA, et al. 2016. *Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites. *Sci. Rep.* 6:21802

Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, et al. 2007. Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci USA*. 104(3):899–902

Chowdhury DR, Angov E, Kariuki T, Kumar N. 2009. A potent malaria transmission blocking vaccine based on codon harmonized full length Pfs48/45 expressed in Escherichia coli. *PLoS ONE*. 4(7):e6352

Cho W-K, Spille J-H, Hecht M, Lee C, Li C, et al. 2018. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science*. 361(6400):412–15

Chu X, Sun T, Li Q, Xu Y, Zhang Z, et al. 2022. Prediction of liquid-liquid phase separating proteins using machine learning. *BMC Bioinformatics*. 23(1):72

Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, et al. 2013. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science*. 341(6146):664–67

Ciuca M, Ballif L, Chelarescu-Vieru M. 1934. Immunity in malaria. *Trans. R. Soc. Trop. Med. Hyg.* 27(6):619–22

Coleman BI, Skillman KM, Jiang RHY, Childs LM, Altenhofen LM, et al. 2014. A *Plasmodium falciparum* histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host Microbe*. 16(2):177–86

Collins CR, Das S, Wong EH, Andenmatten N, Stallmach R, et al. 2013. Robust inducible Cre recombinase activity in the human malaria parasite *Plasmodium falciparum* enables efficient gene deletion within a single asexual erythrocytic growth cycle. *Mol. Microbiol.* 88(4):687–701

Concordet J-P, Haeussler M. 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* 46(W1):W242–45

Connacher J, von Grüning H, Birkholtz L. 2022. Histone modification landscapes as a roadmap for malaria parasite development. *Front. Cell Dev. Biol.* 10:848797

Cox FE. 2010. History of the discovery of the malaria parasites and their vectors. *Parasit.*

Vectors. 3(1):5

Cox HW. 1959. A study of relapse *Plasmodium berghei* infections isolated from white mice. *J. Immunol.* 82(3):209–14

Crawford JE, Clarke DW, Criswell V, Desnoyer M, Cornel D, et al. 2020. Efficient production of male Wolbachia-infected Aedes aegypti mosquitoes enables large-scale suppression of wild populations. *Nat. Biotechnol.* 38(4):482–92

Crompton PD, Pierce SK, Miller LH. 2010. Advances and challenges in malaria vaccine development. *J. Clin. Invest.* 120(12):4168–78

Cui L, Fan Q, Cui L, Miao J. 2008. Histone lysine methyltransferases and demethylases in *Plasmodium falciparum*. *Int. J. Parasitol.* 38(10):1083–97

Deitsch KW, Dzikowski R. 2017. Variant gene expression and antigenic variation by malaria parasites. *Annu. Rev. Microbiol.* 71:625–41

de Solis CA, Ho A, Holehonnur R, Ploski JE. 2016. The Development of a Viral Mediated CRISPR/Cas9 System with Doxycycline Dependent gRNA Expression for Inducible *In vitro* and *In vivo* Genome Editing. *Front. Mol. Neurosci.* 9:70

Diffendall GM, Barcons-Simon A, Baumgarten S, Dingli F, Loew D, Scherf A. 2023. Discovery of RUF6 ncRNA-interacting proteins involved in *P. falciparum* immune evasion. *Life Sci. Alliance*. 6(1):

Diffendall G, Claës A, Barcons-Simon A, Nyarko P, Dingli F, et al. 2024. RNA polymerase III is involved in regulating *Plasmodium falciparum* virulence

Diffendall G, Scherf A. 2024. Deciphering the *Plasmodium falciparum* perinuclear *var* gene expression site. *Trends Parasitol.* 40(8):707–16

Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, et al. 2009. Artemisinin resistance in

Plasmodium falciparum malaria. *N. Engl. J. Med.* 361(5):455–67

Doolan DL, Dobaño C, Baird JK. 2009. Acquired immunity to malaria. *Clin. Microbiol. Rev.* 22(1):13–36, Table of Contents

Duffy PE, Gorres JP, Healy SA, Fried M. 2024. Malaria vaccines: a new era of prevention and control. *Nat. Rev. Microbiol.* 22(12):756–72

Duffy PE, Patrick Gorres J. 2020. Malaria vaccines since 2000: progress, priorities, products. *npj Vaccines*. 5:48

Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, et al. 2005. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum. Cell*. 121(1):13–24

Eaton MD. 1938. The agglutination of *Plasmodium knowlesi* by immune serum. *J. Exp. Med.* 67(6):857–70

Epp C, Li F, Howitt CA, Chookajorn T, Deitsch KW. 2009. Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite *Plasmodium falciparum*. *RNA*. 15(1):116–27

Erdel F, Rademacher A, Vlijm R, Tünnermann J, Frank L, et al. 2020. Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation. *Mol. Cell*. 78(2):236-249.e7

Escalante AA, Ayala FJ. 1994. Phylogeny of the malarial genus *Plasmodium,* derived from rRNA gene sequences. *Proc Natl Acad Sci USA*. 91(24):11373–77

Esen M, Kremsner PG, Schleucher R, Gässler M, Imoukhuede EB, et al. 2009. Safety and immunogenicity of GMZ2 - a MSP3-GLURP fusion protein malaria vaccine candidate. *Vaccine*. 27(49):6862–68

Fan Q, An L, Cui L. 2004. *Plasmodium falciparum* histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. *Eukaryotic Cell*. 3(2):264–76

Filarsky M, Fraschka SA, Niederwieser I, Brancucci NMB, Carrington E, et al. 2018. GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science*. 359(6381):1259–63

Florini F, Visone JE, Deitsch KW. 2022. Shared mechanisms for mutually exclusive expression and antigenic variation by protozoan parasites. *Front. Cell Dev. Biol.* 10:852239

Flueck C, Bartfai R, Niederwieser I, Witmer K, Alako BTF, et al. 2010. A major role for the *Plasmodium falciparum* ApiAP2 protein PfSIP2 in chromosome end biology. *PLoS Pathog.* 6(2):e1000784

Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, et al. 2009. *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog.* 5(9):e1000569

Frank M, Dzikowski R, Costantini D, Amulic B, Berdougo E, Deitsch K. 2006. Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* 281(15):9942–52

Fraschka SA-K, Henderson RWM, Bártfai R. 2016. H3.3 demarcates GC-rich coding and subtelomeric regions and serves as a potential memory mark for virulence gene expression in *Plasmodium falciparum*. *Sci. Rep.* 6:31965

Freitas-Junior LH, Hernandez-Rivas R, Ralph SA, Montiel-Condado D, Ruvalcaba-Salazar OK, et al. 2005. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell*. 121(1):25–36

Gatton ML, Peters JM, Fowler EV, Cheng Q. 2003. Switching rates of *Plasmodium falciparum* var genes: faster than we thought? *Trends Parasitol.* 19(5):202–8

Genton B, Betuela I, Felger I, Al-Yaman F, Anders RF, et al. 2002. A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea. *J. Infect. Dis.* 185(6):820–27

Ghorbal M, Gorman M, Macpherson CR, Martins RM, Scherf A, Lopez-Rubio J-J. 2014. Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat. Biotechnol.* 32(8):819–21

Gibson BA, Doolittle LK, Schneider MWG, Jensen LE, Gamarra N, et al. 2019. Organization of chromatin by intrinsic and regulated phase separation. *Cell*. 179(2):470-484.e21

Goyal M, Alam A, Iqbal MS, Dey S, Bindu S, et al. 2012. Identification and molecular characterization of an Alba-family protein from human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Res.* 40(3):1174–90

Gupta AP, Zhu L, Tripathi J, Kucharski M, Patra A, Bozdech Z. 2017. Histone 4 lysine 8 acetylation regulates proliferation and host-pathogen interaction in *Plasmodium falciparum. Epigenetics Chromatin*. 10(1):40

Haldar S, Saini A, Nanda JS, Saini S, Singh J. 2011. Role of Swi6/HP1 self-association-mediated recruitment of Clr4/Suv39 in establishment and maintenance of heterochromatin in fission yeast. *J. Biol. Chem.* 286(11):9308–20

Half a million children die of malaria every year. Finally we can change that. 2023. *Nature*. 622(7982):218–218

Heitz E. 1928. “Das” Heterochromatin der Moose

Hempelmann E, Krafts K. 2013. Bad air, amulets and mosquitoes: 2,000 years of changing perspectives on malaria. *Malar. J.* 12:232

Henninger JE, Oksuz O, Shrinivas K, Sagi I, LeRoy G, et al. 2021. RNA-Mediated Feedback Control of Transcriptional Condensates. *Cell*. 184(1):207-225.e24

Hermsen CC, Verhage DF, Telgt DSC, Teelen K, Bousema JT, et al. 2007. Glutamate-rich protein (GLURP) induces antibodies that inhibit *in vitro* growth of *Plasmodium falciparum* in a phase 1 malaria vaccine trial. *Vaccine*. 25(15):2930–40

Hoeijmakers WAM., Salcedo-Amaya AM, Smits AH, Françoijs KJ, Treeck M, Gilberger TW, et al. (2013). H2A.Z/H2B.Z Double-variant Nucleosomes Inhabit the AT-rich Promoter Regions of the *Plasmodium falciparum* Genome. Mol. Microbiol. 87, 1061–1073

Hoeijmakers WAM, Stunnenberg HG, Bártfai R. 2012. Placing the *Plasmodium falciparum* epigenome on the map. *Trends Parasitol.* 28(11):486–95

Holehouse AS, Das RK, Ahad JN, Richardson MOG, Pappu RV. 2017. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys. J.* 112(1):16–21

Hollin T, Le Roch KG. 2020. From genes to transcripts, a tightly regulated journey in *Plasmodium*. *Front. Cell. Infect. Microbiol.* 10:618454

Hommel M, David PH, Oligino LD. 1983. Surface alterations of erythrocytes in *Plasmodium falciparum* malaria. Antigenic variation, antigenic diversity, and the role of the spleen. *J. Exp. Med.* 157(4):1137–48

Hommel M, David PH. 1981. *Plasmodium knowlesi* variant antigens are found on schizont-infected erythrocytes but not on merozoites. *Infect. Immun.* 33(1):275–84

Horii T, Shirai H, Jie L, Ishii KJ, Palacpac NQ, et al. 2010. Evidences of protection against

blood-stage infection of *Plasmodium falciparum* by the novel protein vaccine SE36. *Parasitol. Int.* 59(3):380–86

Jabeena CA, Govindaraju G, Rawat M, Gopi S, Sethumadhavan DV, et al. 2021. Dynamic association of the H3K64 trimethylation mark with genes encoding exported proteins in *Plasmodium falciparum*. *J. Biol. Chem.* 296:100614

Jiang L, Mu J, Zhang Q, Ni T, Srinivasan P, et al. 2013. PfSETvs methylation of histone H3K36 represses virulence genes in *Plasmodium falciparum. Nature*. 499(7457):223–27

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 337(6096):816–21

Jones ML, Das S, Belda H, Collins CR, Blackman MJ, Treeck M. 2016. A versatile strategy for rapid conditional genome engineering using loxP sites in a small synthetic intron in *Plasmodium falciparum. Sci. Rep.* 6:21800

Kafsack BFC, Rovira-Graells N, Clark TG, Bancells C, Crowley VM, et al. 2014. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature*. 507(7491):248–52

Kanyal A, Deshmukh B, Davies H, Mamatharani DV, Farheen D, et al. 2024. PfHDAC1 is an essential regulator of *P. falciparum* asexual proliferation and host cell invasion genes with a dynamic genomic occupancy responsive to artemisinin stress. *MBio*. 15(6):e0237723

Kanyal A, Rawat M, Gurung P, Choubey D, Anamika K, Karmodiya K. 2018. Genome-wide survey and phylogenetic analysis of histone acetyltransferases and histone deacetylases of *Plasmodium falciparum*. *FEBS J.* 285(10):1767–82

Karmodiya K, Pradhan SJ, Joshi B, Jangid R, Reddy PC, Galande S. 2015. A comprehensive

epigenome map of *Plasmodium falciparum* reveals unique mechanisms of transcriptional regulation and identifies H3K36me2 as a global mark of gene suppression. *Epigenetics Chromatin*. 8:32

Keenen MM, Brown D, Brennan LD, Renger R, Khoo H, et al. 2021. HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. *eLife*. 10:

Kengne-Ouafo JA, Bah SY, Kemp A, Stewart L, Amenga-Etego L, et al. 2023. The global transcriptome of *Plasmodium falciparum* mid-stage gametocytes (stages II-IV) appears largely conserved and gametocyte-specific gene expression patterns vary in clinical isolates. *Microbiol. Spectr.* 11(5):e0382022

Kensche PR, Hoeijmakers WAM, Toenhake CG, Bras M, Chappell L, et al. 2016. The nucleosome landscape of *Plasmodium falciparum* reveals chromatin architecture and dynamics of regulatory sequences. *Nucleic Acids Res.* 44(5):2110–24

Kim H-S, Roche B, Bhattacharjee S, Todeschini L, Chang A-Y, et al. 2024. Clr4SUV39H1 ubiquitination and non-coding RNA mediate transcriptional silencing of heterochromatin via Swi6 phase separation. *Nat. Commun.* 15(1):9384

Knuepfer E, Napiorkowska M, van Ooij C, Holder AA. 2017. Generating conditional gene knockouts in *Plasmodium* - a toolkit to produce stable DiCre recombinase-expressing parasite lines using CRISPR/Cas9. *Sci. Rep.* 7(1):3881

Kraemer SM, Kyes SA, Aggarwal G, Springer AL, Nelson SO, et al. 2007. Patterns of gene recombination shape var gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. *BMC Genomics*. 8:45

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, et al. 2017. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*.

547(7662):236–40

Leech JH, Barnwell JW, Aikawa M, Miller LH, Howard RJ. 1984. *Plasmodium falciparum* malaria: association of knobs on the surface of infected erythrocytes with a histidine-rich protein and the erythrocyte skeleton. *J. Cell Biol.* 98(4):1256–64

Lee JT. 2011. Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nat. Rev. Mol. Cell Biol.* 12(12):815–26

Lee MCS, Lindner SE, Lopez-Rubio J-J, Llinás M. 2019. Cutting back malaria: CRISPR/Cas9 genome editing of *Plasmodium. Brief. Funct. Genomics*. 18(5):281–89

Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, et al. 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*. 301(5639):1503–8

Li CH, Coffey EL, Dall'Agnese A, Hannett NM, Tang X, et al. 2020. MeCP2 links heterochromatin condensates and neurodevelopmental disease. *Nature*. 586(7829):440–44

Li P, Banjade S, Cheng H-C, Kim S, Chen B, et al. 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature*. 483(7389):336–40

Lopez-Rubio J-J, Mancio-Silva L, Scherf A. 2009. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe*. 5(2):179–90

Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R, Scherf A. 2007. 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol. Microbiol.* 66(6):1296–1305

Loy DE, Liu W, Li Y, Learn GH, Plenderleith LJ, et al. 2017. Out of Africa: origins and evolution of the human malaria parasites *Plasmodium falciparum* and *Plasmodium vivax. Int. J. Parasitol.* 47(2–3):87–97

Lyons DB, Magklara A, Goh T, Sampath SC, Schaefer A, et al. 2014. Heterochromatin-mediated gene silencing facilitates the diversification of olfactory neurons. *Cell Rep.* 9(3):884–92

Machida S, Takizawa Y, Ishimaru M, Sugita Y, Sekine S, et al. 2018. Structural basis of heterochromatin formation by human HP1. *Mol. Cell*. 69(3):385-397.e8

Malaria. www.who.int

Mancio-Silva L, Slavic K, Grilo Ruivo MT, Grosso AR, Modrzynska KK, et al. 2017. Nutrient sensing modulates malaria parasite virulence. *Nature*. 547(7662):213–16

Martins RM, Macpherson CR, Claes A, Scheidig-Benatar C, Sakamoto H, et al. 2017. An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite *Plasmodium falciparum*. *Sci. Rep.* 7(1):14042

McLean SA, Pearson CD, Phillips RS. 1982. *Plasmodium chabaudi:* Antigenic variation during recrudescent parasitaemias in mice. *Exp. Parasitol.* 54(3):296–302

Meneghini MD, Wu M, Madhani HD. 2003. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell*. 112(5):725–36

Mészáros B, Erdos G, Dosztányi Z. 2018. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* 46(W1):W329–37

Miao J, Fan Q, Cui L, Li X, Wang H, et al. 2010. The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite *Plasmodium falciparum*. *Mol. Microbiol.* 78(4):883–902

Michel-Todó L, Bancells C, Casas-Vila N, Rovira-Graells N, Hernández-Ferrer C, et al. 2023. Patterns of Heterochromatin Transitions Linked to Changes in the Expression of

Plasmodium falciparum Clonally Variant Genes. *Microbiol. Spectr.* 11(1):e0304922

Monahan K, Lomvardas S. 2015. Monoallelic expression of olfactory receptors. *Annu. Rev. Cell Dev. Biol.* 31:721–40

Mori T, Nakashima M. 2023. Sequence-dependent heterochromatin formation in the human malaria parasite *Plasmodium falciparum*. *Heliyon*. 9(9):e19164

Morrison O, Thakur J. 2021. Molecular complexes at euchromatin, heterochromatin and centromeric chromatin. *Int. J. Mol. Sci.* 22(13):

Muller HJ. 1930. Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* 22(3):299–334

Musabyimana J-P, Musa S, Manti J, Distler U, Tenzer S, et al. 2024. The *Plasmodium falciparum* histone methyltransferase SET10 participates in a chromatin modulation network crucial for intraerythrocytic development. *mSphere*. 9(11):e0049524

Ngwa CJ, Gross MR, Musabyimana J-P, Pradel G, Deitsch KW. 2021. The role of the histone methyltransferase pfset10 in antigenic variation by malaria parasites: a cautionary tale. *mSphere*. 6(1):

Niang M, Bei AK, Madnani KG, Pelly S, Dankwa S, et al. 2014. STEVOR is a *Plasmodium falciparum* erythrocyte binding protein that mediates merozoite invasion and rosetting. *Cell Host Microbe*. 16(1):81–93

Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, et al. 2008. Evidence of artemisinin-resistant malaria in western Cambodia. *N. Engl. J. Med.* 359(24):2619–20

Nunes A, Thathy V, Bruderer T, Sultan AA, Nussenzweig RS, Ménard R. 1999. Subtle mutagenesis by ends-in recombination in malaria parasites. *Mol. Cell. Biol.* 19(4):2895–2902

Nuñez JK, Harrington LB, Doudna JA. 2016. Chemical and biophysical modulation of cas9 for tunable genome engineering. *ACS Chem. Biol.* 11(3):681–88

O'Neill SL, Ryan PA, Turley AP, Wilson G, Retzki K, et al. 2018. Scaled deployment of Wolbachia to protect the community from dengue and other Aedes transmitted arboviruses. [version 3; peer review: 2 approved]. *Gates Open Res.* 2:36

Oakley MSM, Kumar S, Anantharaman V, Zheng H, Mahajan B, et al. 2007. Molecular factors and biochemical pathways induced by febrile temperature in intraerythrocytic *Plasmodium falciparum* parasites. *Infect. Immun.* 75(4):2012–25

Obeng-Adjei N, Larremore DB, Turner L, Ongoiba A, Li S, et al. 2020. Longitudinal analysis of naturally acquired PfEMP1 CIDR domain variant antibodies identifies associations with malaria protection. *JCI Insight*

Obradovic Z, Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK. 2003. Predicting intrinsic disorder from amino acid sequence. *Proteins*. 53 Suppl 6:566–72

Obuse C, Nakayama J-I. 2025. Functional involvement of RNAs and intrinsically disordered proteins in the assembly of heterochromatin. *Biochim. Biophys. Acta Gen. Subj.* 1869(6):130790

Ollomo B, Durand P, Prugnolle F, Douzery E, Arnathau C, et al. 2009. A new malaria agent in African hominids. *PLoS Pathog.* 5(5):e1000446

Pace T, Scotti R, Janse CJ, Waters AP, Birago C, Ponzi M. 2000. Targeted terminal deletions as a tool for functional genomics studies in *Plasmodium*. *Genome Res.* 10(9):1414–20

Pandit K, Surolia N, Bhattacharjee S, Karmodiya K. 2023. The many paths to artemisinin resistance in *Plasmodium falciparum. Trends Parasitol.* 39(12):1060–73

Pandya-Jones A, Markaki Y, Serizay J, Chitiashvili T, Mancia Leon WR, et al. 2020. A protein

assembly mediates Xist localization and gene silencing. *Nature*. 587(7832):145–51

Park J, Kim J-J, Ryu J-K. 2024. Mechanism of phase condensation for chromosome architecture and function. *Exp. Mol. Med.* 56(4):809–19

Pederson T. 2011. The nucleolus. *Cold Spring Harb. Perspect. Biol.* 3(3):

Pérez-Cantero A, Llorà-Batlle O, Pelaez-Conde I, Martínez-Guardiola C, Cortés A. 2025. Heterochromatin *de novo* formation and maintenance in *Plasmodium falciparum*. *BioRxiv*

Pérez-Toledo K, Rojas-Meza AP, Mancio-Silva L, Hernández-Cuevas NA, Delgadillo DM, et al. 2009. *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(8):2596–2606

Petter M, Lee CC, Byrne TJ, Boysen KE, Volz J, et al. 2011. Expression of *P. falciparum* var genes involves exchange of the histone variant H2A.Z at the promoter. *PLoS Pathog.* 7(2):e1001292

Petter M, Selvarajah SA, Lee CC, Chin WH, Gupta AP, et al. 2013. H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy var gene promoters in the malaria parasite *Plasmodium falciparum*. *Mol. Microbiol.* 87(6):1167–82

Pickford AK, Michel-Todó L, Dupuy F, Mayor A, Alonso PL, et al. 2021. Expression Patterns of *Plasmodium falciparum* Clonally Variant Genes at the Onset of a Blood Infection in Malaria-Naive Humans. *MBio*. 12(4):e0163621

Plys AJ, Davis CP, Kim J, Rizki G, Keenen MM, et al. 2019. Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev.* 33(13–14):799–813

Price M. 2017. Dramatic evolution within human genome may have been caused by malaria

parasite. *Science*

Quakyi IA, Carter R, Rener J, Kumar N, Good MF, Miller LH. 1987. The 230-kDa gamete surface protein of *Plasmodium falciparum* is also a target for transmission-blocking antibodies. *J. Immunol.* 139(12):4213–17

Ralph SA, Bischoff E, Mattei D, Sismeiro O, Dillies M-A, et al. 2005. Transcriptome analysis of antigenic variation in *Plasmodium falciparum--var* silencing is not dependent on antisense RNA. *Genome Biol.* 6(11):R93

Rambhatla JS, Turner L, Manning L, Laman M, Davis TME, et al. 2019. Acquisition of Antibodies Against Endothelial Protein C Receptor-Binding Domains of *Plasmodium falciparum* Erythrocyte Membrane Protein 1 in Children with Severe Malaria. *J. Infect. Dis.* 219(5):808–18

Rawat M, Kanyal A, Sahasrabudhe A, Vembar SS, Lopez-Rubio J-J, Karmodiya K. 2021a. Histone acetyltransferase PfGCN5 regulates stress responsive and artemisinin resistance related genes in *Plasmodium falciparum*. *Sci. Rep.* 11(1):852

Rawat M, Srivastava A, Johri S, Gupta I, Karmodiya K. 2021b. Single-Cell RNA Sequencing Reveals Cellular Heterogeneity and Stage Transition under Temperature Stress in Synchronized *Plasmodium falciparum* Cells. *Microbiol. Spectr.* 9(1):e0000821

Reichenow E. Ueber das Vorkommen der Malaria Parasiten des Menschen bei den afrikanischen Menschenaffen

Renger R, Morin JA, Lemaitre R, Ruer-Gruss M, Jülicher F, et al. 2022. Co-condensation of proteins with single- and double-stranded DNA. *Proc Natl Acad Sci USA*. 119(10):e2107871119

Reyes RA, Raghavan SSR, Hurlburt NK, Introini V, Bol S, et al. 2024. Broadly inhibitory

antibodies to severe malaria virulence proteins. *Nature*. 636(8041):182–89

Reyser T, Paloque L, Augereau J-M, Di Stefano L, Benoit-Vical F. 2024. Epigenetic regulation as a therapeutic target in the malaria parasite *Plasmodium falciparum. Malar. J.* 23(1):44

Ribeiro JM, Garriga M, Potchen N, Crater AK, Gupta A, et al. 2018. Guide RNA selection for CRISPR-Cas9 transfections in *Plasmodium falciparum*. *Int. J. Parasitol.* 48(11):825–32

Rich SM, Ayala FJ. 2006. Evolutionary origins of human malaria parasites. In *Malaria: Genetic and Evolutionary Aspects*, pp. 125–46. Boston: Kluwer Academic Publishers

Roestenberg M, McCall M, Hopman J, Wiersma J, Luty AJF, et al. 2009. Protection against a malaria challenge by sporozoite inoculation. *N. Engl. J. Med.* 361(5):468–77

Rowe JA, Claessens A, Corrigan RA, Arman M. 2009. Adhesion of *Plasmodium falciparum*-infected erythrocytes to human cells: molecular mechanisms and therapeutic implications. *Expert Rev. Mol. Med.* 11:e16

RTS,S Clinical Trials Partnership. 2015. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. *Lancet*. 386(9988):31–45

Ryu J-K, Bouchoux C, Liu HW, Kim E, Minamino M, et al. 2021. Bridging-induced phase separation induced by cohesin SMC protein complexes. *Sci. Adv.* 7(7):

Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, et al. 2018. Coactivator condensation at super-enhancers links phase separation and gene control. *Science*. 361(6400):eaar3958

Sagara I, Dicko A, Ellis RD, Fay MP, Diawara SI, et al. 2009. A randomized controlled phase 2 trial of the blood stage AMA1-C1/Alhydrogel malaria vaccine in children in Mali. *Vaccine*. 27(23):3090–98

Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AMG, et al. 2009.

Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc Natl Acad Sci USA*. 106(24):9655–60

Santos JM, Josling G, Ross P, Joshi P, Orchard L, et al. 2017. Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor. *Cell Host Microbe*. 21(6):731-741.e10

Sanulli S, Trnka MJ, Dharmarajan V, Tibble RW, Pascal BD, et al. 2019. HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature*. 575(7782):390–94

Saraf A, Cervantes S, Bunnik EM, Ponts N, Sardiu ME, et al. 2016. Dynamic and Combinatorial Landscape of Histone Modifications during the Intraerythrocytic Developmental Cycle of the Malaria Parasite. *J. Proteome Res.* 15(8):2787–2801

Saul A. 1999. Circumsporozoite polymorphisms, silent mutations and the evolution of *Plasmodium falciparum. Parasitol Today (Regul Ed)*. 15(1):38–40

Scherf A, Hernandez-Rivas R, Buffet P, Bottius E, Benatar C, et al. 1998. Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in *Plasmodium falciparum. EMBO J.* 17(18):5418–26

Scherf A, Lopez-Rubio JJ, Riviere L. 2008. Antigenic variation in *Plasmodium falciparum. Annu. Rev. Microbiol.* 62:445–70

Schneider VM, Visone JE, Harris CT, Florini F, Hadjimichael E, et al. 2023. The human malaria parasite *Plasmodium falciparum* can sense environmental changes and respond by antigenic switching. *Proc Natl Acad Sci USA*. 120(17):e2302152120

Shang X, Wang C, Fan Y, Guo G, Wang F, et al. 2022. Genome-wide landscape of ApiAP2 transcription factors reveals a heterochromatin-associated regulatory network during

Plasmodium falciparum blood-stage development. *Nucleic Acids Res.* 50(6):3413–31

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539

Silberhorn E, Schwartz U, Löffler P, Schmitz S, Symelka A, et al. 2016. *Plasmodium falciparum* Nucleosomes Exhibit Reduced Stability and Lost Sequence Dependent Nucleosome Positioning. *PLoS Pathog.* 12(12):e1006080

Simantov K, Goyal M, Dzikowski R. 2022. Emerging biology of noncoding RNAs in malaria parasites. *PLoS Pathog.* 18(7):e1010600

Sinha A, Hughes KR, Modrzynska KK, Otto TD, Pfander C, et al. 2014. A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature*. 507(7491):253–57

Sluiter CP. 1912. De dierlijke parasieten van den mensch en van onze huisdieren

Smith JD, Chitnis CE, Craig AG, Roberts DJ, Hudson-Taylor DE, et al. 1995a. Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell*. 82(1):101–10

Smith JD, Chitnis CE, Craig AG, Roberts DJ, Hudson-Taylor DE, et al. 1995b. Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell*. 82(1):101–10

Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. 2017. Phase separation drives heterochromatin domain formation. *Nature*. 547(7662):241–45

Stubbs J, Simpson KM, Triglia T, Plouffe D, Tonkin CJ, et al. 2005. Molecular mechanism for switching of P. falciparum invasion pathways into human erythrocytes. *Science*.

309(5739):1384–87

Subudhi AK, Green JL, Satyam R, Salunke RP, Lenz T, et al. 2023. DNA-binding protein PfAP2-P regulates parasite pathogenesis during malaria parasite blood stages. *Nat. Microbiol.* 8(11):2154–69

Sugiyama T, Cam HP, Sugiyama R, Noma K, Zofall M, et al. 2007. SHREC, an effector complex for heterochromatic transcriptional silencing. *Cell*. 128(3):491–504

Su XZ, Heatwole VM, Wertheimer SP, Guinet F, Herrfeldt JA, et al. 1995. The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell*. 82(1):89–100

Tabassum W, Bhattacharyya S, Varunan SM, Bhattacharyya MK. 2021. Febrile temperature causes transcriptional downregulation of *Plasmodium falciparum* Sirtuins through Hsp90-dependent epigenetic modification. *Mol. Microbiol.* 115(5):1025–38

Tang J, Chisholm SA, Yeoh LM, Gilson PR, Papenfuss AT, et al. 2020. Histone modifications associated with gene expression and genome accessibility are dynamically enriched at *Plasmodium falciparum* regulatory sequences. *Epigenetics Chromatin*. 13(1):50

Tong X, Tang R, Xu J, Wang W, Zhao Y, et al. 2022. Liquid-liquid phase separation in tumor biology. *Signal Transduct. Target. Ther.* 7(1):221

Tonkin CJ, Carret CK, Duraisingh MT, Voss TS, Ralph SA, et al. 2009. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in *Plasmodium falciparum. PLoS Biol.* 7(4):e84

Tortora MMC, Brennan LD, Karpen G, Jost D. 2023. HP1-driven phase separation recapitulates the thermodynamics and kinetics of heterochromatin condensate formation. *Proc Natl Acad Sci USA*. 120(33):e2211855120

Tuteja R. 2007. Malaria - an overview. *FEBS J.* 274(18):4670–79

van der Pluijm RW, Tripura R, Hoglund RM, Pyae Phyo A, Lek D, et al. 2020. Triple artemisinin-based combination therapies versus artemisinin-based combination therapies for uncomplicated *Plasmodium falciparum* malaria: a multicentre, open-label, randomised clinical trial. *Lancet*. 395(10233):1345–60

Venkatesan P. 2025. WHO world malaria report 2024. *Lancet Microbe*. 101073

Vettermann C, Schlissel MS. 2010. Allelic exclusion of immunoglobulin genes: models and mechanisms. *Immunol. Rev.* 237(1):22–42

Volz JC, Bártfai R, Petter M, Langer C, Josling GA, et al. 2012. PfSET10, a *Plasmodium falciparum* methyltransferase, maintains the active var gene in a poised state during parasite division. *Cell Host Microbe*. 11(1):7–18

von Grüning H, Coradin M, Mendoza MR, Reader J, Sidoli S, et al. 2022. A dynamic and combinatorial histone code drives malaria parasite asexual and sexual development. *Mol. Cell. Proteomics*. 21(3):100199

Voss TS, Healer J, Marty AJ, Duffy MF, Thompson JK, et al. 2006. A var gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature*. 439(7079):1004–8

Wagner JC, Platt RJ, Goldfless SJ, Zhang F, Niles JC. 2014. Efficient CRISPR-Cas9-mediated genome editing in *Plasmodium falciparum. Nat. Methods*. 11(9):915–18

Wang B, Zhang L, Dai T, Qin Z, Lu H, et al. 2021a. Liquid-liquid phase separation in human health and diseases. *Signal Transduct. Target. Ther.* 6(1):290

Wang G-H, Gamez S, Raban RR, Marshall JM, Alphey L, et al. 2021b. Combating mosquito-borne diseases using genetic control technologies. *Nat. Commun.* 12(1):4388

Wang J, Xu C, Wong YK, Li Y, Liao F, et al. 2019a. Artemisinin, the Magic Drug Discovered from Traditional Chinese Medicine. *Engineering*. 5(1):32–39

Wang L, Gao Y, Zheng X, Liu C, Dong S, et al. 2019b. Histone modifications regulate chromatin compartmentalization by contributing to a phase separation mechanism. *Mol. Cell*. 76(4):646-659.e6

Weiner A, Dahan-Pasternak N, Shimoni E, Shinder V, von Huth P, et al. 2011. 3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite *Plasmodium falciparum. Cell. Microbiol.* 13(7):967–77

Wheeler RJ. 2020. Therapeutics-how to treat phase separation-associated diseases. *Emerg. Top. Life Sci.* 4(3):307–18

Wong ML, Liew JWK, Wong WK, Pramasivan S, Mohamed Hassan N, et al. 2020. Natural Wolbachia infection in field-collected Anopheles and other mosquito species from Malaysia. *Parasit. Vectors*. 13(1):414

Wu Y, Sifri CD, Lei HH, Su XZ, Wellems TE. 1995. Transfection of *Plasmodium falciparum* within human red blood cells. *Proc Natl Acad Sci USA*. 92(4):973–77

Wyss M, Kanyal A, Niederwieser I, Bartfai R, Voss TS. 2024. The *Plasmodium falciparum* histone methyltransferase PfSET10 is dispensable for the regulation of antigenic variation and gene expression in blood-stage parasites. *mSphere*. 9(11):e0054624

Yamada T, Fischle W, Sugiyama T, Allis CD, Grewal SIS. 2005. The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Mol. Cell*. 20(2):173–85

Yeoh LM, Goodman CD, Mollard V, McFadden GI, Ralph SA. 2017. Comparative transcriptomics of female and male gametocytes in *Plasmodium berghei* and the evolution

of sex in alveolates. *BMC Genomics*. 18(1):734

Zhao C, Zhao Y, Zhang J, Lu J, Chen L, et al. 2018. HIT-Cas9: A CRISPR/Cas9 Genome-Editing Device under Tight and Effective Drug Control. *Mol. Ther. Nucleic Acids*. 13:208–19

Zimmerman PA, Woolley I, Masinde GL, Miller SM, McNamara DT, et al. 1999. Emergence of FY*A(null) in a *Plasmodium vivax*-endemic region of Papua New Guinea. *Proc Natl Acad Sci USA*. 96(24):13973–77

Zofall M, Grewal SIS. 2006. Swi6/HP1 recruits a JmjC domain protein to facilitate transcription of heterochromatic repeats. *Mol. Cell*. 22(5):681–92