UNIVERSITY OF CALIFORNIA
RIVERSIDE

Exploring Novel Drug Targets to Combat the Human Malaria Parasite, *Plasmodium falciparum*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Microbiology

by

Zeinab M. Chahine

September 2024

Dissertation Committee:
Dr. Karine Le Roch, Chairperson
Dr. Meera Nair
Dr. Adler Dillman

The Dissertation of Zeinab M. Chahine is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENT

I extend my sincere appreciation to the individuals who have played crucial roles throughout my graduate school experience, contributing significantly to my success. I owe a tremendous debt of gratitude to the remarkable individuals who have played pivotal roles in my life.

To my coworkers and friends; Jacques, Thomas, Desiree, Steven and Todd, your support and guidance have been indispensable, enriching my professional experience and providing a relaxing and engaging work environment. A special note of thanks to my Dr. Karine Le Roch. Your patience and understanding have been nothing short of outstanding. You have a unique ability to guide me without pressure, giving me the space I need to work independently while knowing exactly when to step in with encouragement or a gentle push. Your faith in me, often greater than my own, has been a constant source of inspiration. I am deeply grateful for your mentorship and for showing me what it means to lead with both strength and compassion.

To my closest friends, Karla, John, Marianne, Salam, Leila, Jacques, Thomas, Amanda, Todd, Ilse and all of my BioChem club friends, you're understanding and constant reminders to take a break and enjoy life have kept me balanced and focused. Thank you for being my pillars of strength and for believing in me even when I struggled to believe in myself. You made sure I never felt alone, pushing me forward when my mom was letting me have too much fun, and always being there to listen to my endless

complaints. Your unconditional love, patience, and belief in my abilities have been my greatest source of motivation. Thank you for standing by me through the late nights, the frustrations, and the triumphs. Your presence in my life has been invaluable, and I am thankful for every moment we've shared.

To my family, my father, Fatima, Addis, Hassan, Hussein, Ahmad and Ali you have been the foundation upon which all my achievements stand. Your unwavering support and encouragement have been my greatest motivators. Every step I took was fueled by the love and belief you have in me, and for that, I am eternally grateful.

A special appreciation goes to my mom. You were always there to lift me up when I needed it most, offering encouragement and the kind of love that only a mother can give. You had a unique way of reminding me to balance work with life, even when it meant convincing me to slack off from my responsibilities to enjoy a day out with you. You may have been a bad influence on my career, but you the greatest blessing in my life. I dedicate the rest of my life to showing you just how grateful I am for all you have sacrificed for us. Collectively, the guidance, knowledge, and moral support from my coworkers, family, and friends have been paved every step of my journey. I am profoundly thankful for the impact they have had on my success, and I look forward to paying this generosity forward in my future endeavors.

ABSTRACT OF THE DISSERTATION

Exploring Novel Drug Targets to Combat the Human Malaria Parasite, *Plasmodium falciparum*

by

Zeinab M. Chahine

Doctor of Philosophy, Graduate Program in Microbiology,
University of California, Riverside, September 2024
Dr. Karine Le Roch, Chairperson

The unicellular protozoan, *Plasmodium falciparum*, is the human malaria-causing parasite responsible for over half a million deaths each year. *P. falciparum* has a complex life cycle, alternating between human and invertebrate hosts, and is capable of not only evading the host immune responses but also developing resistance to all clinically administered treatments. Consequently, malaria remains a persistent global health concern. Recent advances in omics-based technologies, improved sequencing platforms, and enhanced CRISPR-based gene editing tools have opened new avenues for understanding the parasite biology, identifying new therapeutic strategies and combating the spread of malaria. This report delves into the systems-wide approaches employed to explore the drug mechanism of action of a new compound, MED6-189, an analogue of the kalihinol family of isocyanoterpene natural products, effective against drug-sensitive and -resistant *P. falciparum* strains. We also use advance multi-omics strategies to identify and functionally

characterize new potential drug targets involved in gene regulation. Those included targeting proteins coding for chromatin remodeling complexes or better understanding the role of long non-coding RNAs (lncRNA) in the parasite life cycle progression.

In essence, the goal of my PhD thesis was to identify promising and alternative targets to curb the spread of malaria, either through the development of new antimalarial agents or the discovery of novel, unexploited potential targets critical for the parasite's survival.

Table of Contents

List of Figures

List of Tables

Introduction

List of Schemes

Chapter 1

GLOSSARY

ACTs: Artemisinin-based combination therapies
APiAP2: APETALA2
As: Artemisinin
BSV: Blood stage vaccine
ChIP: chromatin immunoprecipitation
ChEP: chromatin enrichment for proteomics
CRISPR: Clustered regularly interspaced short palindromic repeats
COPI: Coat protein complex I
COPII: Coat protein complex II
CQ: Chloroquine
CSA: chondroitin sulfate A
DHAP: dihydroxyacetone phosphate
DHA: Dihydroartemisinin
DMAPP: dimethylallyl pyrophosphate
DMSO: dimethyl sulfoxide
DOXP: 1-deoxy-D-xylulose-5-phopshate pathway
ER: Endoplasmic reticulum
ERAD: ER-associated degradation system
FASII: Fatty acid synthesis type II
FIC: Fractional inhibitory concentration
FV: Food vacuole
GDV1: Gametocyte development protein1
GMEP: Global Malaria Eradication Program
HCQ: Hydroxychloroquine
Hi-C: chromosome conformation capture sequencing
HPI: Hours post invasion
HPS: Hours post synchronization
hRBC: human Red Blood Cell
IDC: intraerythrocytic developmental cycle
IEC: Intraerythrocytic cycle
IPP: isopentenyl pyrophosphate
ISC: Iron–Sulphur cluster formation
LncRNA: Long non-coding RNAs
LMF: lumefantrine
LPC: lysophosphatidylcholine
mAbs: monoclonal antibodies
MEP: mevalonate/methylerythritol phosphate pathway
MFQ: mefloquine
MoA: Mechanism of Action
MoR: Mechanism of Resistance
MORC: Microrchidia protein
NPCs: nuclear pore complexes

PC: phosphatidylcholine

PDH: Pyruvate dehydrogenase

PE: phosphatidylethanolamine

PEP: phosphoenolpyruvate

PEV: Pre-erythrocytic vaccine

***Pf*EMP1**: *P. falciparum* erythrocyte membrane protein 1

***pfmdr1*:** P-glycoprotein homologue

PMV: Placental malaria vaccine

PTM: Posttranslational modifications

PYR: pyrimethamine

QN: quinine

RUF6: RNAs of Unknown Function 6

SAM: S-adenosyl methionine

SAR: *Stramenopiles*, Alveolates and *Rhizaria*

SNP: single nucleotide mutations

SYN: syntaxin

TAG: triacylglyceride

TARE-lncRNAs: telomere-associated repetitive elements

TBV: Transmission-blocking vaccines

TF: Transcription factor

TG: triglyceride

TPP: thiamine pyrophosphate

TPP: Thermal Proteome Profiling

VSAs: variant surface antigens

WSV: Whole sporozoite vaccine

WHO: World health organization

WSV: Whole sporozoite vaccine

INTRODUCTION

Malaria is a life-threatening vector-borne disease endemic to roughly 84 countries, afflicting various higher eukaryotic hosts including humans and is caused by the unicellular protozoan parasite, *Plasmodium.* Mainly found in tropical and subtropical terrains, this disease held a devastating impact of over 619,000 human deaths and 247 million clinical infections globally in 2022 alone (*1*), with children under the age of five and the immunocompromised being the most vulnerable (*1*). The parasitic pathogen harbors a complex life cycle alternating between eukaryotic and invertebrate hosts and has adapted the means to effectively evade host immune responses as well as developing resistance to almost all clinically administered treatments (*2-6*).

The disease also harbors a significant economic impact, with roughly $4.3 billion spent globally in 2016 to treat and control its spread (*7*). Effective antimalarials and sensitive diagnostic kits are crucial in monitoring and controlling potential outbreaks and the spread of the disease. Currently, artemisinin-based combination therapies (ACTs), approved by WHO in 2006, has served as our most effective means of controlling the spread of malaria and in the treatment of the most infectious strains (*8*). Unfortunately, artemisinin resistance has been rising with cases of resistance detected since early 2009 (*8, 9*). Currently, these ACTs are believed to be among the last line of defense against the most virulent human infecting species of *Plasmodium, P. falciparum* (*6*).

Plasmodium belongs to the phylum *Apicomplexa*, obligate intracellular parasitic alveolates which also include *Toxoplasma gondii*, the causative agent toxoplasmosis; *Cryptosporidium parvum*, the leading cause of human cryptosporidiosis, and parasites that can cause acute, often fatal diseases in cattle, *Eimeria spp*. and *Theileria spp*. The *Plasmodium* species belong to the SAR supergroup of pathogens which includes *Stramenopiles*, Alveolates and *Rhizaria*. *Plasmodium* can be categorized by their major hosts, infecting humans (*P. vivax, P. malariae, P. ovale, and P. falciparum*), nonhuman primates (*P. reichenowi, P. cynomolgi* and *P. knowlesi* which infect chimpanzees and macaques, respectively), rodents (*P. berghei*, *P. chabaudi*, and *P. yoelii*), birds (*P. gallinaceum*) and even reptiles (*P. mexicanum*), each of which hold particular interest in the scientific community as they serve as important model systems for study (*10-13*) **(Fig. I-1).**

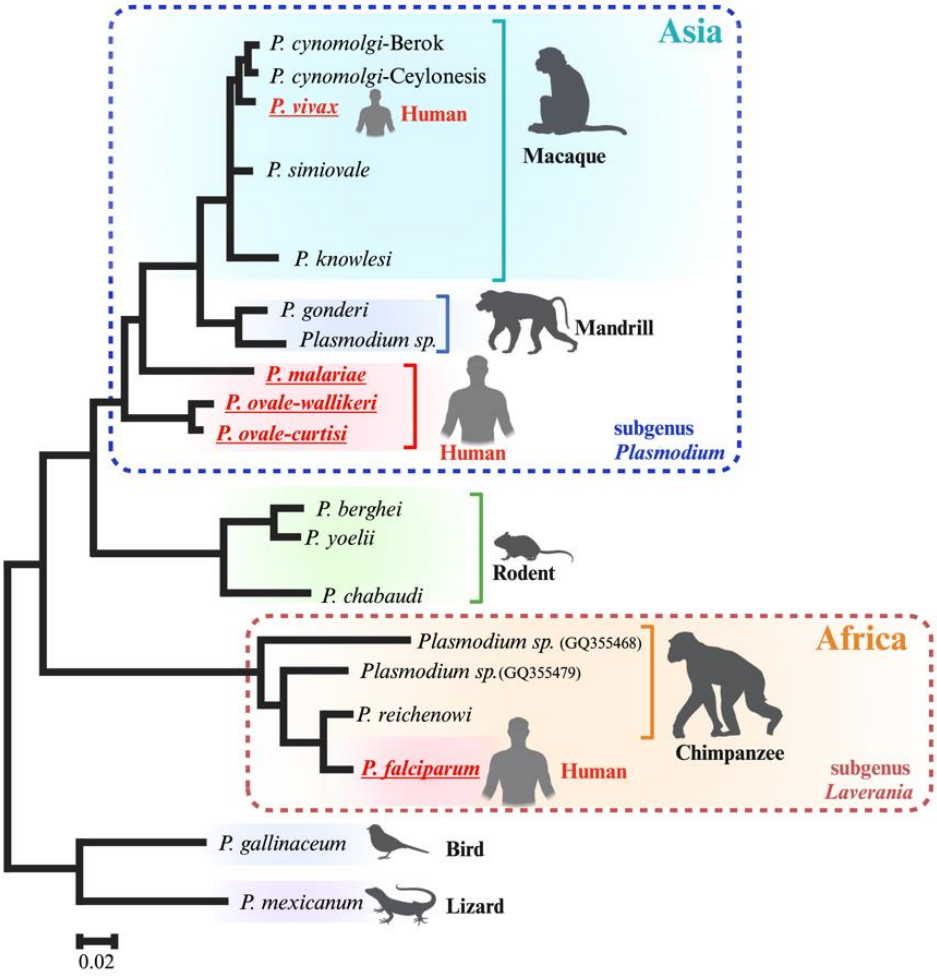


Figure I-1. *Plasmodium* Phylogeny Tree. *Plasmodium spp.* phylogenetic tree based on complete mitochondrial genomes. The phylogeny branches leading to human infecting malaria parasites are marked in red.

Epidemiology

Plasmodium species are widespread globally. In 2022 an estimated 94% of all malaria cases and 95% of deaths were reported in the African continent – holding a disproportionately higher share of the malaria burden. Four African countries in particular account for over half of all global malaria deaths: Nigeria (26.8%), the Democratic Republic of the Congo (12.3%), Uganda(5.1%) and Mozambique (4.2%). Additionally,

78% of all malaria deaths in the region occur in children under 5 years of age (*1*). A majority of malaria-related deaths that transpire in sub-Saharan Africa are caused by *P. falciparum* infections. *P. vivax,* in contrast, are more common in Asia and South America and often lead to mild anemia due to recurrent infections (*13*).

In 2022, roughly 7.4 million cases of malaria infection were reported in South America, South-East Asian, and the Western Pacific territories (*1*). The primary cause of these infections was due to *P. falciparum* and *P. vivax* species, with *P. vivax* estimated to have contributed to over 4.9 million cases of malaria. Despite wide geographical distributions, *P. ovale, P. malariae* and *P. knowlesi*, were only attributed to less than 0.5 million cases (*1*). *P. knowlesi* is usually found in macaque throughout Southeast Asia, but cases of fatal zoonotic infection have also been identified in humans. *P. malariae, P. ovale-curtisi* and *P. ovale-wallikeri* on the other hand, are commonly misdiagnosed, showing relatively low virulence and low parasitemia levels, making these species particularly difficult to diagnose and study.

The spread of malaria is contingent upon the presence of female *Anopheles* mosquitoes, which thrive in stagnant freshwater habitats in warm and humid climates. Progression of malaria is influenced by several factors, including the geographic distribution of mosquito species, their proximity to human populations, the effectiveness of vector control measures, and infection management. Consequently, malaria thrives in low-income tropical regions

and is often aggravated by social disturbances, pandemics and natural disasters which is often exasperated by the increasing global warming events **(Figure I-2).**

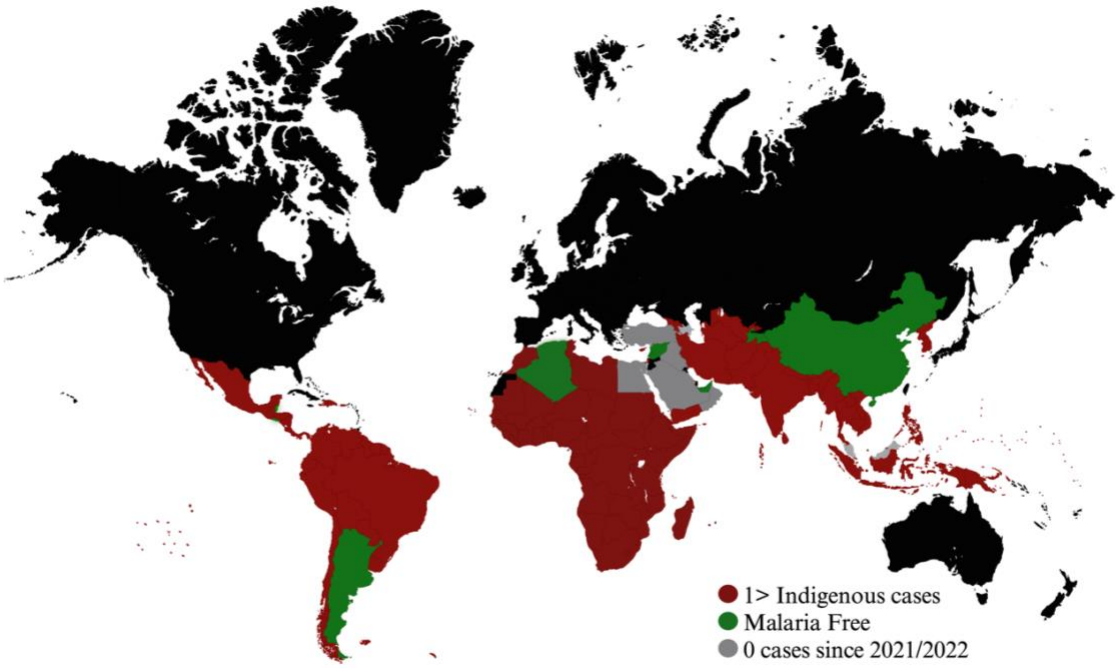


Figure I-2. Global malaria infections. Countries with indigenous malaria cases in 2000 with status as of 2022 (source: WHO database)

The *Plasmodium falciparum* Life cycle

Infection begins through the inoculation of sporozoites into human hosts from the bite of female *Anopheles* mosquitos. Here, mature sporozoites will enter the blood stream and circulate until they reach the liver within 30–60 minutes after first inoculation. Inside the liver, parasites initiate a series of differentiation and replication events to generate thousands of merozoites. It should be noted that the liver stages of the *Plasmodium* life

cycle can be unique to each species. For example, *P. vivax* and *P. ovale* form a dormant stage hypnozoite which can persist for months or even years in the liver before being reactivated and released into the bloodstream (*14-16*). After 6–10 days (sometimes longer depending on the species), the merozoites leave the liver to invade red blood cells where they initiate the asexual intraerythrocytic developmental cycle (IDC).

The asexual cycle consists of a remarkable series of morphological changes, from the relatively dormant ring stage to the highly dynamic and rapidly metabolizing trophozoite stage, to the dividing schizont stage. During this trophozoite stage the parasite decondenses its chromatin to allow intensive transcription and cellular activity leading to parasite replication during schizogony. The asexual cycle can last 24, 48 or 72 hours—depending on the species, and undergo asymmetric rounds of mitosis to produce 16 to 64 newly infective merozoites that are then released to infect new red blood cells (*17*). This phase of the parasite life cycle is symptomatic and can cause febrile waves, headache, vomiting, anemia, coma, and eventually death. The IDC is a recurrent cyclical event with favored environmental conditions. It is believed that signs of nutrient deprivation, changes in temperature or stress signaling, can signal parasites to shift from an asexual stage replicative cycle to sexual stage development. However, there is still dispute on whether this gametocytogenesis initiation event occurs in response to external tress stimuli or is an innate constitutive occurrence within a parasite subpopulation (*18-20*).

During sexual development, parasites differentiate into male and female gametocytes (*21*). In *P. falciparum*, this stage involves five morphologically unique steps that may last 2–3

weeks to maturation. Once mature, gametocytes enter the mammalian host circulatory systems awaiting uptake by a new *Anopheles* host. Inside mosquito gut lumen, increased Ca^{2+} levels, drop in pH and temperature levels along with other signaling cues such as xanthurenic acid, trigger the differentiation of the haploid gametocytes into gametes. A female gametocyte develops into a macro-gamete through a series of cytoplasmic reorganization events and remains immotile in the midgut. The male gametocyte on the other hand, transforms into a micro-gamete. The micro-gamete nucleus undergoes three rounds of replication to form 6–8 nuclei and undergoes a process of ex-flagellation. This cluster-like multinucleated cytoplasmic body subsequently breaks apart into individual male gametes that are directed to the macro-gamete. In the mosquito midgut male and female gametes fuse to form a diploid zygote that will subsequently develop into a motile ookinete. The newly formed ookinete then travels and penetrates the inner walls of the mosquito gut to undergo encystation. Once encystation occurs, the ookinete (now called an oocyst) continues to develop in the epithelium of the midgut forming sporozoites. Mature sporozoites will then erupt from the oocyst and migrate to the salivary glands to infect a new human host **(Figure I-3)** (*21, 22*). This elaborate life cycle is controlled by complex networks of proteins that are tightly regulated at the genomic, transcriptional and translational levels.

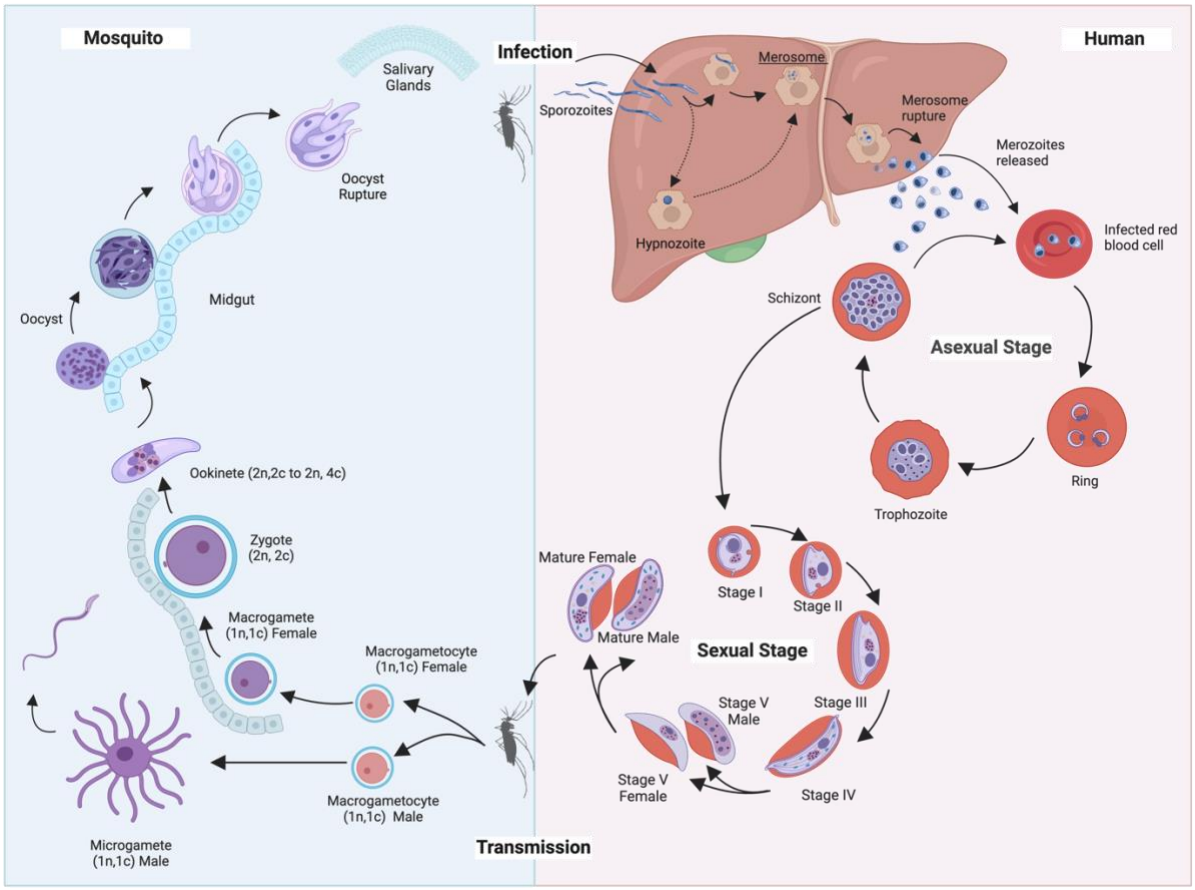


Figure I-3. The *Plasmodium* Life Cycle. Illustration depicting *Plasmodium* parasite life cycle alternating from an invertebrate Anopheles (Left) host to a mammalian human host (Right).

The Apicoplast Organelle

Apicomplexan parasites share a unique double membrane organelle called the apicoplast. This non-photosynthetic plastid originated through secondary endosymbiotic events derived from its red (*Chromera velia*) or green (*Chlorella*) ancestral algal cell (*23*). The apicoplast is composed of a 35 kb genome with only a handful of protein coding genes. In response to the reduced plastid proteome, a majority of the proteins are encoded within the nucleus and imported post-translationally from the cytoplasm(*24*). Synthesized through the ERAD-like pathway, the apicoplast is highly dependent on COPI and COPII vesicular

trafficking to acquire and uphold regular plastid function (*25*, *26*). The organelle is essential to apicomplexan survival, thus understanding the specific roles it plays in the parasite is of upmost significance in the field of malaria research. Moreover, since humans and animals lack the apicoplast organelle, it serves as a particularly attractive drug therapeutic target. Since its discovery in 1996 (*27*) the apicoplast has been shown to play a role in major metabolic pathways. One of the most important is attributed to the synthesis of Isoprenoids. Isoprenoids belong to a class of lipid molecules that serve as precursors for the production of essential compounds such as sterols, chlorophyll and quinones(*28*). Primarily, the apicoplast has been shown to participate in 4 major metabolic pathways:

1. *Isopentenyl pyrophosphate synthesis*

Isoprenoids are formed through the combination of 2 molecules: isopentenyl pyrophosphate (*29*) and dimethyl-allyl pyrophosphate (DMAPP). These components are synthesized by the mevalonate/methylerythritol phosphate (MEP) pathway, which is present in animal cells and plants, or the DOXP (1-deoxy-D-xylulose-5-phopshate) pathway present in bacteria and plastids. The apicoplast is a vital player in the MEP pathway through biosynthesis and export of such precursors. Evidence has also shown that IPP is important in an array of cellular developmental functions such as protein prenylation, vesicular trafficking, protein glycosylation, mitochondrial ubiquinone and heme-A biosynthesis (*30*-*33*).

2. *Iron–sulphur cluster formation (ISC)*

Iron is indispensable in the intraerythrocytic (IEC) stages of *Plasmodium* parasites. Iron is primarily used in the synthesis of iron-sulfur (FeS) clustering. These FeS clusters serve as cofactors, specifically the SufS and SufE cofactors. Studies have shown that FeS clusters participate in sulfur acquisition, have cysteine desulfurase activity, harbor roles in apicoplast organelle maintenance and isoprenoid biosynthesis.

3. *Type II fatty acid (FASII) pathway*

The FASII pathway consists of acquisition of phosphoenolpyruvate (PEP) and dihydroxyacetone phosphate (DHAP) into the apicoplast to synthesize FASII fatty acids, and acetyl-CoA. These components undergo an array of chemical steps to form lipoic acid, an enzyme cofactor (*34*) essential for membrane lipid synthesis in several cellular compartments. Another role of the FASII lies in the incorporation of fatty acids into precursors used for membrane lipid synthesis (*35*).

Antimalarials and Drug resistance

As previously mentioned, *Plasmodium* has adapted the means to effectively evade host immune responses that include a range of genetic changes including allelic variation. The parasite has also successfully developed resistance to all clinically administered treatments **(Table I-1)**. Chloroquine (CQ) was predominately used as the preferred method of malaria treatment to replace Quinine. Discovered in 1945, this antimalarial functions by disrupting heme deposit formation the food vacuole (FV). It was a welcoming drug

therapeutic as the only alternative (Quinine), used since the1600s, presented several adverse side effects such as: Blurred vision, confusion, hearing loss, stomach pain and vomiting. Unfortunately, the earliest signs of resistance to Chloroquine occurred in Southeast Asia in 1957 and spread globally just over a decade later (*36*). A similar outcome was presented upon administration of Pyrimethamines, where resistance outbreaks occurred within the same year of its use in 1957, or Mefloquines where resistance occurred only 5 years after its initial application in the 1990s. With emerging resistance to chloroquine and dwindling alternative drug therapies within reach, rates of infection underwent significant recidivism and malaria cases reached staggering morbidity and mortality numbers (*37*). In the 1950s several alternative compounds were developed but the potency was found to be short-lived, with the *P. falciparum* parasite developing resistance to all drug-based control efforts **(Table I-1)**.

In 2015, Professor Youyou Tu was awarded the Nobel Prize in Physiology and Medicine for her key contributions to the discovery of the common woodworm plant, Qinghao (*A. annua*) more commonly known as Artemisinin (*38*). Since its discovery, the spread and infection rates of *P. falciparum* saw its most significant drop in several years. Unfortunately, some evidence of erosion began to emerge in the only available defense against *P. falciparum*, with cases of resistance to Artemisinin slowly rising as early as 2009 in the Greater Mekong regions of Southeast Asia (western Cambodia, Thailand, Vietnam, Myanmar, and Laos) (*6, 37*). In the hopes of restraining resistance outbreaks, current control efforts recommend artemisinin-combination therapies (ACTs) (developed since

2001 and approved by the WHO in 2006) as a last resort to extend and preserve the potency and longevity of drugs at our arsenal. ACTs consist of the Artemisinin derivatives such as Dihydroartemisinin, artemether, and artesunate combined with piperaquine, lumefantrine, or mefloquine_ among others, resulting in a 95% efficacy (*39*). Partial resistance to artemisinin, was first confirmed in Cambodia in 2007 and has since spread across the Greater Mekong subregion. Alarmingly, clinical failure of even some combination treatments such as artemether-lumefantrine, dihydroartemisinin-piperaquine, and artesunate-mefloquine is now common in Thailand, Cambodia, Vietnam, and Laos. There are several hypotheses as to how the pathogen can overcome our most ardent efforts for controlling malaria infections **(Table I-1)**(*40-43*).

1. *Lack of structural diversity in drug therapie*s

Several of the available drug therapies are derived from structurally similar classes of chemicals. The significant structural similarities between CQ and its analogue Hydroxychloroquine (HCQ) for example, upended the widespread use of these compounds upon adaptation of single nucleotide mutations (SNP) mutations in the *Pfcrt* (K76T) or *Pfmdrg1* (N86Y, Y184F, S1034C, N1042D or D1246Y) genes(*44*) (*45*). These simple mutations allowed the pathogen to render both chemical compounds invaluable because their structural similarity had not altered their mechanism of action (MoA) significantly enough to overcome the resistance mechanism **(Table I-1).** Coincidently, several other classes of antimalarials have experienced similar outcomes, following identical pathways and thus being subject to the very same resistance outcomes, for example: lumefantrine

(LMF), amodiaquine (AQ), mefloquine (MFQ), quinine (QN) and artemisinin (AS) all have shown reduced efficacy in present of the *pfmdr1* gene mutation.

2. *Irrational Pharmacodynamic properties*

Conventional drugs are an outcome of coincidental discovery and oftentimes dispensed without complete understanding of their respective drug targets or the downstream implications of their use. For example: Mefloquine, Quinine, and Methylene Blue were once widely administered as first-line antimalarials, despite not fully understanding the precise biochemical and molecular implications of these naturally derived drug therapeutics. Some are now well understood to harbor toxic and harmful side effects on patients. Yet, at the time of their administration, the widespread use was unknowingly and irresponsibly causing unwanted side effects lasting months and even years after treatment. This was in part a result of lack of available and safe alternatives, coupled with the complexity of naturally derived compounds, making investments into their study a time consuming, costly and exhaustive endeavor. To date, details on the mechanism of action of some of our once most effective antimalarial therapies (Quinines, Methylene Blue, Artemisinins) remain a topic of debate.

3. *Improper drug administration*

Improper and overuse of therapeutics has resulted in tolerance and reduced efficacy. Irregular, or inconsistence drug intake, incomplete treatment courses, subtherapeutic dosage use and even mass drug administration, presents an opportunity for selective

pressure on parasite populations. These subpopulations can carry mutations conferring resistance with increased likelihood to survive and reproduce, rendering drugs less effective and complicating malaria control efforts.

4. *Disregarding Pharmacodynamic interactions*

Pharmacologists understand the importance of recognizing how a compound may impact an organism's biochemical pathways. Pharmacodynamics emphasizes dose-receptor relationships; these are the interactions between the drug's concentration , its effect on the target site as well as all other potential interactions in association. Knowing whether a specific combination of compounds may target divergent or conflicting pathways is imperative to designing effective and long-lasting treatment regimens. An example can be found with Atovaquone-Dihydroartemisinin(DHA) which function antagonistically, requiring higher dosages if administered in combination to offer proper treatment. These effects would result in increased unwanted side effects and reduced efficacy of the treatments.

5. *Characteristic Plasmodium traits*

The *P. falciparum* pathogen encompasses a 22.8-megabase nuclear genome. Holding one of the most A/T-rich genomes ever to be sequenced, with an estimated ~80% and ~90% A/T content for coding and intergenic regions, respectively (*46*). This unique parasite genome codes for ~5,500 annotated coding genes within its 14 chromosomes. Additionally, high replication events and well-developed immune evasion strategies have

allowed for significant recombination and random mutation events that can readily form opportunities for drug resistance and overwhelming all attempted control efforts. With such a unique arsenal conveyed by the pathogen, current drug regiments employ combinatory treatments of contrasting functionality in the hopes of countering the parasites resistance apparatuses.

Class/Family	Compound	Source	Discovery	Resistance	Time (yrs.)	Resistance Markers
4-Amino-quinolines	Chloroquine	Natural	1934	1957	12	***pfcrt*** (K76T, K76N, K76I, K72T) ***pfmdr1*** (N86Y, Y184F, S1034C, N1042D, D1246Y)
	Amodiaquine	Synthetic	1946	1971	19	
	Hydroxy chloroquine	Synthetic	1949	1982	33	
8-Amino-quinoline	Primaquine	Synthetic	1945	-	-	Causes acute hemolytic toxicity, especially in patients with G6PD deficiency
	Pamaquine	Synthetic	1920s	-	-	
	Pentaquine	Synthetic	1945	-	-	
	Isopentaquine	Synthetic	1940s	-	-	
4-quinolinemethanols	Quinine	Natural	1820	1910	90	***pfmdr1*** (N86Y, Y184F, S1034C, N1042D, D1246Y). ***pfmrp*** (Y191, A437S)
	Quinidine	Natural	1853	-	-	Depresses contractility, decreases systemic vascular resistance, causing arrythmia.
	Mefloquine	Synthetic	1974	1982	5	***pfcrt*** (K76T, A220S, Q271E, N326S, I356T, R371I). > ***pfmdr1*** copy number, ***pfmdr1*** (N86Y)
Artemisinin	Artemisinin	Natural	1972	2008	36	***pfk13*** (C580Y, R539T, I543T, F446L, N458Y, P547L, R561H, Y493H), ***pfatp6*** (A623E, S769N)
	Artemether	Synthetic	1981		<1	
	DHA	Synthetic	1997	2010	13	
	Artesunate	Synthetic	1977	2011	33	
	Arteether	Synthetic	1986		<1	

Table I-1. Approved antimalarial drugs with associated resistance markers (*40-43*).

Class/Family	Compound	Source	Discovered	Resistance	Time (yrs.)	Resistance Markers
Antimetabolites	Proguanil	Synthetic	1945	1949	1	*pfdhfr* (S108N, N51I, C59R), *pfcytb* (Y268S/C/N, M133I, L144S, G280D)
	Pyrimethamine	Synthetic	1952	1967	15	
	Atovaquone	Synthetic	1991	1996	<1	
Antibiotics	Tetracycline	Natural	1948	-	-	–
	Clindamycine	Semi-synthetic	1975	2010	35	
	Doxycycline	Synthetic	1967	-	-	
Diaminopyridines	Pyrimethamine	Synthetic	1952	1967	<1	*pfdhfr* (S108N, N51I, C59R, 164 I164L, A16V)
Phenanthrene methanol	Halofantrine	Synthetic	1975		<1	> *pfmdr1* copy number
Amino Alcohols	Lumefantrine	Synthetic	1970s	2005	35	*pfmdr1* (N86Y, Y184F, S1034C, N1042D, D1246Y) > *pfmdr1* copy number
Bisquinolones	Piperaquine	Synthetic	1960	2010	13	> *pfpm2*, *pfpm3* copy numbers
4-quinolinemethanols	Pyronaridine	Natural	1970	-	-	*pfmdr1* (N86Y, Y184F, S1034C, N1042D, D1246Y), *pfcrt* (K72T)
Sulfonamides	Sulfadoxine	Synthetic	1967	1967	<1	*pfdhps* (S436F/A, A437G, K540E, A581G, A613S/T)

Table I-1. Approved antimalarial drugs with associated resistance markers (*40-43*).

Control and Elimination Efforts

One of the earliest attempts to overcome the malaria burden began in 1955. The WHO's Global Malaria Eradication Program (GMEP) began with the optimistic ambition to create a '*malaria-free world*'. Initially, there were great headways achieved, seeing the successful elimination of malaria in many countries, yet global eradication was far-reaching. A significant reason for their shortcomings was due in part, to failure to fully implement adequate resources to sub-Saharan Africa, where the greatest malaria burden was found. In 1969 abandonment of GMEP led investigators in search of more practical and attainable means of controlling the spread and morbidity of the disease (*47*). Malaria elimination is now defined by successful disruption of local transmission of specific *Plasmodium* species within defined geographical margins. In 2017 the WHO adopted the E-2020 initiative, calling for reduction to zero incidence of indigenous malaria infection by 2020. The effort was comprised of 21 countries across 5 territories(*48*) (*49*). Results of these programs have been mixed, with promising reduction in infection rates seen in numerous countries, yet alarming increase in incidence rates (sometimes over to 50%) in others such as Botswana, South Africa, Saudi Arabia, Surinam and Ecuador. It is worth noting that declining control efforts were not a direct result of the global COVID-19 pandemic as progress against malaria control efforts had already been deteriorating. The implications of the pandemic, social upheaval, natural disasters, and civil unrest only contributed to the resurgence of malaria in certain regions. These hurdles, coupled with declining efficacy of antimalarials, insecticides as well as a simultaneous increase in resistance to said antimalarials, complete eradication seemed infeasible. Yet these

shortcomings are reflected in only a handful of countries. It is important to recognize global malaria deaths have been on a steady and gradual decline since 1980 from 1-1.2 million deaths to just over 500,000 deaths in 2022(*50, 51*). Following the COVID-19 pandemic, revival of control efforts, vector control strategies and increasing sensitivity of new diagnostic tools have resulted in a significant drop in malaria occurrence in many countries (WHO, 2023). These achievements are mainly attributed to increased political commitment and consequently, consistent funding being allocated to control efforts.

Effective malaria control efforts require adequate and rapid malaria surveillance tools to properly detect and respond to potential outbreaks and better control of transmission hotspots. We now have advanced approaches to vector-control strategies such as mosquito nets, safer insecticides and a greater global understanding of mosquito breeding grounds, which have collectively contributed to more successful control of transmission rates. Additionally, the gene-drive process, which strives to propagate genes conferring sterility to *Plasmodium* parasites through mosquito host populations remains a promising strategy. It is currently in its early phases of implementation yet has shown great success, with the complete wipe-out of captive mosquito populations in the city of Terni, Italy in as little as 8-12 generations (5-12 years). There are currently plans in motion to perform the first experimental releases in Burkina Faso, Mali, Ghana or Uganda (*52*).

Malaria Prophylaxis: Exploring Vaccines and Monoclonal Therapies

Over the past decade there have been significant advancements made in vaccine development strategies and monoclonal therapies against malaria. Malaria vaccines have

been under development since the 1960s. On October 2021, the WHO approved the RTS,S/AS01 (RTS,S) malaria vaccine for children in sub-Saharan Africa and regions with moderate to high *P. falciparum* malaria transmission. Shortly after, the R21/Matrix-M vaccine was also approved by the WHO for use in other malaria-endemic countries. Malaria vaccines are designed with one of four targeted mechanisms of action, each targeting distinct stages of the pathogen's life cycle in the hopes of disrupting replication and/or transmission:

1. Blood stage vaccine strategy (BSV)

These vaccines target the asexual stages of the parasite cell cycle, specifically merozoite replication. Targeting the blood stages of the pathogen is considered attractive as this stage is the fever inducing phase of the disease. However, several challenges are met in designing merozoite specific vaccines. Primarily, the window for antibody immune response is extremely brief as the time between merozoite egress to reinvasion can last only a few seconds. Additionally, the pathogen's well adapted mechanism of antigenic polymorphism and extreme titer load means these vaccines must elicit high enough titer antibodies against merozoite surface antigens to sufficiently impair parasite invasion. Between 2000 and 2015, over 30 BSV vaccine trials were implemented with a large majority targeting the MSP1 and AMA1 antigens. Other vaccines targeted antigens such as EBA175, GMZ2, MSP3, *Pf*RH5, AMA1-RON2 and *Pf*GARP antigens. Ultimately, for many of these trials, results showed scant evidence of protection, with some of the more promising trials having a meager 14% efficacy (GMZ2).

2. Pre-erythrocytic vaccine strategy (PEV)

Aimed to prevent blood-stage infections entirely. PEVs target the clinically silent stage of parasite infection. Antigens from *Plasmodium* sporozoite and liver stages that initiate human infection after mosquito inoculation are the target of PEVs. These vaccine strategies are intended to encourage antibody formation against surface antigens, clearing inoculated sporozoites traveling from the skin and bloodstream or block those invading hepatocytes. With PEVs, it is expected that high activity can completely clear pre-erythrocytic parasites before their release into the bloodstream. Current PEVs target the *P. falciparum* sporozoite antigen (*Pf*SPZ) or other major surface antigen of sporozoites such as the circum-sporozoite protein (CSP5). Preliminary results were performed in mice models targeting C-terminal-specific human mAb but were met with disappointing outcomes, failing to show functional activity *in vitro* or *in vivo* in mice. By far the most advanced of the PEVs is the RTS,S which incorporates a *P. falciparum* CSP fragment central repeat (**R**), a C-terminal region containing T-cell epitopes (**T**) and a hepatitis B surface antigen (**S**)(hence the "RTS").

3. Placental malaria vaccines (PMV)

PMVs are designed to specifically target parasites that sequester in the placenta, particularly those binding to chondroitin sulfate A (CSA), a placental proteoglycan present on the surface of the specific cells of the placenta. Unlike PEVs and BSVs, vaccines that aim to protect the general population, PMVs are formulated to provide more targeted protection. Natural antibodies to CSA-binding parasites, which protect against placental

malaria, are acquired over successive pregnancies, theoretically leading to increase resistance among women in endemic areas. VAR2CSA, a member of the *P. falciparum* erythrocyte membrane protein-1 (*Pf*EMP1) protein family mediates adhesion to placenta-specific CSA receptors. Over the past 5 years, the first trials of VAR2CSA-based PMVs have been underway. Results have shown the vaccine to be safe, immunogenic, and capable of inducing functional antibodies. However, cross-reactivity against VAR2CSA variants has been limited. Currently attempts aim to alternate immunization schedules, antigen dosages, and introduce VAR2CSA-combination based vaccines in the hopes of improving the cross-reactivity (*53*). In one study, researchers measure anti-adhesion Ab which limited parasite accumulation in the placenta in new mothers (Primigravidas) vs mothers who have been pregnant on previous occasions (multigravidas). Results showed primigravidas patients were more likely to have placental infections and significantly higher levels of parasitemia (*31*.6% infection with mean parasitemia: 7.3%) than or multigravidas (14.2% infections with mean parasitemia 1.9%).

4. Transmission-blocking vaccines (TBV)

With the notion of exterminating the sexual or mosquito stages of the parasite, TBVs integrate gametocyte specific surface antigens as a means to induce an immune response in mosquito vectors during blood meals. Thus far, there are 4 main candidates under study (a) surface proteins present in human blood stage gametocytes (such as *Pf*S230 and *Pf*S48/45) and (b) zygote surface proteins expressed in mosquito hosts post-fertilization (for example: *Pf*S25 and *Pf*S28). *Pf*S25 was the first TBV candidate to progress to clinical

trials. Most recently, trials are underway with *Pf*S230 based TBVs. Unfortunately, both *Pf*S25 and *Pf*S230 recombinant antigens have shown poor immunogenicity. Future candidates focus instead on the gametocyte surface antigen *Pf*S48/45 (*53*).

5. Whole sporozoite vaccines (WSV)

The concept of conferring sterilizing immunity against sporozoites posed a challenge since the 1970s. This was mainly because of the difficulties faced manufacturing irradiated sporozoites for a vaccine. In 2010, new technology made harvesting *Pf*SPZ from vector salivary glands of aseptic mosquitoes was made possible. After infecting mosquitos with lab-derived strains of the pathogen, followed by extraction and purification, *Pf*SPZ was followed by cryopreservation in liquid nitrogen vapor. *Pf*SPZ are then attenuated to create appropriate vaccine candidates. Studies have proven that the efficacy of WSV in humans shows some promise, with early studies showing 90% vaccine efficacy when compared to control infections in humans. Yet, efficacy was significantly reduced in heterologous strains, with efficacy dropping as low as 10% in some cases. Additionally, the level and duration of protection against presents another challenge, with long-term protection remaining lacking (*54*).

Among the listed vaccine candidates, PEVs have shown most promise and are currently under the most advanced stages of clinical study. Most recently, the RTS,S/AS01E, (a PEV) was recommended by the WHO in 2021. Currently in phase 3 clinical trials in Africa, three doses administered one month apart, followed by a booster dose 18 months later, has

shown reduced clinical malaria episodes over 3-4 years by 36% in young children vaccinated between 5 and 17 months of age and by 26% in infants vaccinated between 6 and 12 weeks of age. Additionally, long-acting monoclonal antibodies targeting CSP have proven effective in preventing *P. falciparum* malaria in human volunteers and endemic regions, offering a potential for a single-dose immunoprophylaxis. However, questions remain about the effectiveness of malaria vaccines and the best strategies for deploying them to benefit communities severely affected by malaria. Antigen discovery, structural studies, and refinement of platforms to enhance existing candidates remains a significant challenge. In the meantime, our current alternative remains to identify new therapeutic targets as a priority. With improved access to genomic, metabolomic, and sequencing technologies, we can uncover previously untapped information and vulnerabilities in the pathogen. Vulnerabilities that can only be identified with greater insight into parasite biology.

Epigenetic Regulation in Malaria Parasites: Chromosome landscape and the importance of chromatin remodeling

Growing interest has been placed in the field of epigenetics and the 3D architectural framework of the *Plasmodial* genome over the past two decades. Chromosomal configurations are now understood to play central roles in regulating the pathogen's transcriptomic rhythms, modulating major biological processes throughout the parasite's life cycle. (*55, 56*). *Plasmodium* was observed to adopt a dynamic chromatin structure, alternating between an extended, open chromatin to a collapsed and condensed, lowly

expressing genome depending on its transcriptional requirements. During its most active trophozoite state, the architectural landscape is open, allowing for increased chromosomal interactions and heightened transcriptional activity before adopting a more closed and relatively silent assembly of chromatin arrangements during its more dormant stage, i.e. ring stage (*57*). These patterns clearly correlate with the pathogen's transcription activity and chromatin condensation states (19, 92). Using chromatin immunoprecipitation (ChIP)-sequencing experiments against RNA polymerase II, authors in one study showed that RNA polymerase II bound to targeted promoters during early asexual stages paused in place until timely signaling thresholds were met for transcriptional activation at the trophozoite stage, resulting in a burst of transcriptional movement. This synchronous pattern of expression can only be managed by a well-orchestrated and coordinated assembly of transcription factors (TF) and other associated players including nucleosome positioning, histone modifications and their regulators, DNA methylation associated enzymes, ncRNAs, chromatin structural regulators and RNA-binding proteins. The combination of these features establishes a dynamic and highly regulated network that function in concert to police the cascade of gene expression patterns necessary for cell cycle progression(*58-60*).

1. Transcription factors

RNA polymerase II is responsible for the transcription of messenger RNAs along with several small nuclear RNAs **(Figure I-4)**. The polymerase requires general TFs to form a multiprotein complex that assembles onto targeted promoters. A unique feature of

Plasmodium biology is the relatively low number of transcription factors (TF) available to regulate and control gene expression within its genome. There are currently 73 known TFs in its arsenal(*61*). 20-27 of the most well-documented belong to the apicomplexan APETALA2 (ApiAP2) family of DNA-binding proteins present in all *Plasmodium spp*. The ApiAP2 protein family are known to modulate gene expression throughout its various developmental stages. Since their discovery, evidence has recognized their roles as master regulators of gene expression. The most well-documented TFs are the gametocyte-specific AP2-G and AP2-G2, which function as global repressors of genes involved in transmission and promote gametocyte maturation and sexual differentiation. It is worth noting that the process of gametocytogenesis encompasses a cascade of multiple ApiAP2 TFs with certain ApiAP2 TFs such as AP2-G5 (*62*), AP2-G, and AP2-G2 functioning as main contributors. These TFs work in congruence by regulating hundreds of genes and operating as a transcriptional switch from asexual to sexual programs. Also well characterized are the specific TFs which regulate the sporozoite stage (AP2-SP), and liver stage (AP2-L) and the AP2-O TF which activates gene expression in ookinetes.

Lesser studied TFs such as AP2-P are found to serve essential regulatory roles in trophozoite development, *var* gene regulation, merozoite development and parasite egress. Recent studies have focused on the regulatory functions of apicomplexan AP2 (ApiAP2) transcription factors associated with heterochromatin, specifically identifying and characterizing eight new *Pf*AP2-HFs. These factors were found to be preferentially enriched in heterochromatic regions and their knockout correlated with notable changes in *var* gene expression, affirming the regulatory role of the ApiAP2 family (*63*, *64*). Although

the majority of the ApiAP2 family TFs are associated with developmental-stage transitions, some of them seem to be required for supplementary mechanisms whose biological functions remain elusive. AP2-exp and AP2-exp2 mutants showed upregulation of antigenic variant genes (*rifin*, *stevor*, and *Pfmc-2TM*) and reduction of genes involved in cell remodeling, respectively (*62, 64, 65*). Different AP2 TFs are also associated with heterochromatin (*64*), such as AP2-HC (*66*), while AP2Tel and *Pf*SIP2 were found to be enriched in subtelomeric regions (*67, 68*). However, as the protist boasts just over 5,000 protein-coding genes, it is clear there exist other facets of regulatory control apart from the meager 73 TFs in its repertoire (1.46% of its protein coding genome in comparison to human (>7.5%) or yeast (3.1%) proteomes (*61, 69*) (*60*).

2. Nucleosomes

In *Plasmodium*, nucleosome occupancy is dynamic and varies considerably throughout the cell cycle. These drastic structural shifts are possible in part because of flexible nucleosome positioning regulated by TFs and heterochromatin markers. The nucleosomes compress roughly 147 DNA base pairs coiled tightly around a histone octamer. These histones are comprised of two of each of the following apparatuses H2A, H2B, H3, and H4. Comprehensive analysis has revealed that the *Plasmodium falciparum* genome encodes an established form of each core histone and four histone variants conferring distinct physical properties, dynamics, and alternative functions. Interestingly, these histone variants H2A.Z, H2B.Z, H2Bv, CenH3, and H3.3 were identified (*70-72*) and found to have specific and unique features, such as H2A.Z/H2B.Z, which is associated with intergenic regions of

P. falciparum (*73*, *74*), and *Pf*CENH3, shown to be linked with centromeres (*70*). H2A.Z also seems to be enriched at the transcription start site of the single active *var* gene, suggesting a contribution of this particular histone variant in regulation of this virulence gene family(*74, 75*). Although DNA compaction associated with nucleosomes is understood to hinder accessibility of the transcription machinery, the occupancy and positioning of these nucleosomes provide a complementary and dynamic regulatory mechanism crucial to controlling gene expression in malaria parasites **(Figure I-4)**(*60*).

3. Chromatin Organization and Histone Modifications

In addition, nucleosome organization, the histone core itself can be subject to an assortment of reversible epigenetic marks. These reversible posttranslational modifications (PTMs), which include acetylation, methylation, and phosphorylation, occur on histone tails and are distributed along the genome, mainly on promoter regions. In *Plasmodium*, several dozen PTMs have been identified through quantitative mass spectrometry for each histone, including variants, revealing some PTMs unique to the parasite (*71, 76-79*). Interestingly, acetylation of H4 seems to occur from the N terminus to C terminus, whereas deacetylation occurs in the opposite direction (*78, 79*). This model contrasts with other organisms and seems to be a feature unique to *Plasmodium*. Although the impact on transcriptional activity of these epigenetic modifications depends on global chromatin environments, some of them are considered to be hallmarks of euchromatin or heterochromatin (*60*).

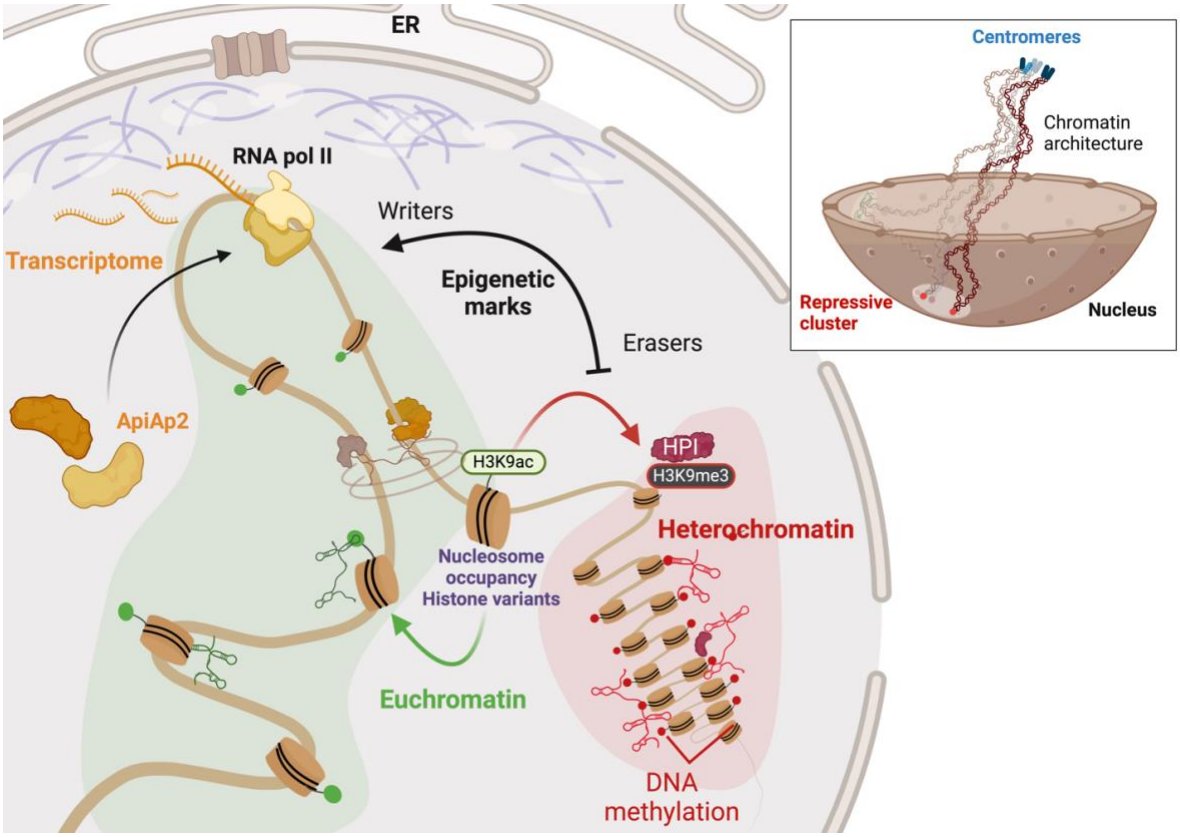


Figure I-4. Epigenetic mechanisms in *Plasmodium*. Gene expression patterns define cell function(s) and are under the control of master ApiAP2 transcription factors (yellow). Nucleosome occupancy, writers, erasers (black) and histone modifications are major factors regulating chromatin condensation. Other epigenetic mechanisms include lncRNAs and DNA methylation patterns (red). Finally, the overall chromatin landscape is well organized, forming a repressive cluster including telomeres and *var* genes positioned at opposing poles of centromeres within the nucleus (top right). Abbreviations: ApiAP2, apicomplexan Apetala2; HP1, heterochromatin protein 1; lncRNA, long noncoding RNA, RNA polymerase (*80*).

H3K9me3 and H3K36me3 are associated with inactive promoters and the heterochromatin and are found to be mutually exclusive of H3K9ac and H3K4me3, signatures of active transcription **(Figure I-4)**. H3K9me3 is recognized and bound by heterochromatin protein1 (HP1), an interaction conserved across eukaryotes, and their coupled role(s) in silencing *var* genes has been well studied in *Plasmodium*. A clear enrichment of H3K9me3 was

observed for inactive *var* genes located in the nuclear periphery of the parasite (*77, 81-83*). Conversely, the single active *var* gene was shown to be marked by H3K9ac and H3K4 di- and trimethylation (*84, 85*). A majority of these studies were performed on asexual stages, with some specific subsets of these modifications observed during distinct stages of the IDC. In ring stages, histones are enriched in H4K20me1 or H3K56me1, whereas trophozoites are associated with H3K9me2 and H3K18me1 (*76*). Although there is little information about other stages of *P. falciparum*, recent works have unveiled epigenetic modifications throughout different gametocyte stages (*86*). Various histone marks were identified as stage specific, such as H3K36me2/me3 for early gametocytes or H3K27me1 and H3K36me1 for late gametocytes (*76, 87*). Although additional studies will be required to decipher the complexity of the histone code, there is no doubt that these modifications are critical regulators of gene expression throughout the life cycle of *P. falciparum*(*60*).

4. Chromatin Organization and Histone-Modifying Enzymes

As discussed above, epigenetic modifications are essential for gene regulation in *Plasmodium*. The histone PTMs are reversible and themselves under the control of a wide range of enzymes. These proteins, identified as histone-modifying enzymes, are categorized as writers and erasers and respectively deposit and remove epigenetic marks **(Figure I-4).** Methylation of the histone lysines is controlled by the opposing activity of histone lysine methyltransferases (HKMTs) and demethylases (HKDMs)(*88*) . Similarly, histone acetyltransferases, including MYST and GNAT families, along with histone deacetylases (HDACs), composed of HDAC and sirtuin proteins, have antagonistic effects

on the acetylation of histone tails (*89*). The activities of these enzymes were extensively studied in the regulation of *var* genes in *P. falciparum*. Complementary enzymatic activities orchestrate the mutually exclusive expression of *var* genes. Apicidin and curcumin, respective inhibitors of HDACs and histone acetyltransferase *Pf*GCN5, have been shown to drastically alter the epigenome and transcriptome of *P. falciparum* during the IDC (*90, 91*). Thus, inhibitors of histone-modifying enzymes are considered potent antimalarial compounds retaining efficacy against multiple stages of *P. falciparum* (*92*). Taken together, histone-modifying enzymes have proven to be crucial in regulation of histone modifications and warrant further investigation and characterization, as they are promising therapeutic targets(*60*).

5. Beyond Histones: Regulation and Function of DNA Methylation

DNA methylation is another crucial regulatory player and requires direct chemical modification occurring of DNA. In mammalian genomes, this reaction is catalyzed by DNA methyltransferases, which transfer a methyl group to the fifth carbon of a cytosine residue to form 5-methylcytosine (5mC) (*93*). An additional oxidized intermediate, 5-hydroxymethylcytosine (5hmC), can also be generated during the demethylation process. Recruitment of readers to these epigenetic marks affects gene repression or hinders the binding of TFs **(Figure I-4)**. In *Plasmodium*, the contribution of this mechanism remains poorly understood. Although a unique DNA methyltransferase was identified in the parasite, a pioneering study provided a genome-wide map of 5mC distribution(*94*). Using bisulfite sequencing (BS-seq) technology, the authors detected that 0.58% of the total

genomic cytosines were methylated. However, this method does not distinguish between 5mC and 5hmC. More recently, 5mC levels were detected at 0.01–0.02% of the *P. falciparum* genome, while a new modification, 5hmC-like, was identified as predominant, with 0.19–0.38% of total genomic cytosines (*60, 95*). While additional data will be needed to confirm the presence of these modifications in the genome, they were found in gene bodies during the IDC.

6. The Role of Noncoding RNAs in Gene Regulation

Over the last decades evidence has been mounting on the importance of ncRNAs in cell regulation and gene expression. Although they were initially overlooked and regarded as transcriptional noise, the advent of deep RNA-seq platforms provided an unprecedented advantage to uncover the roles of these previously unacknowledged non-protein-coding molecules. They are now a major focus of study and promising targets for drug therapies, vaccine targets, or diagnostic biomarkers. In eukaryotes, small ncRNAs such as small interfering RNA (siRNA), microRNA (miRNA), and Piwi-interacting RNA (piRNA) average 18 to 200 nucleotides and function primarily in posttranscriptional modification or silencing and degradation processes. Long ncRNAs (lncRNAs), in contrast, are more than 200 nucleotides and vary in structure, forming complex hairpins and loops that dictate their niche epigenetic and regulatory functions. In malaria parasites, there is no evidence for the presence of RNA interference mechanisms (miRNAs, siRNAs, piRNAs)(*96*). Instead, these protozoans have adapted a higher-than-average repertoire of RNA-binding proteins

and lncRNAs that have been implicated in a wide array of regulatory functions (*97*) **(Figure I-4).**

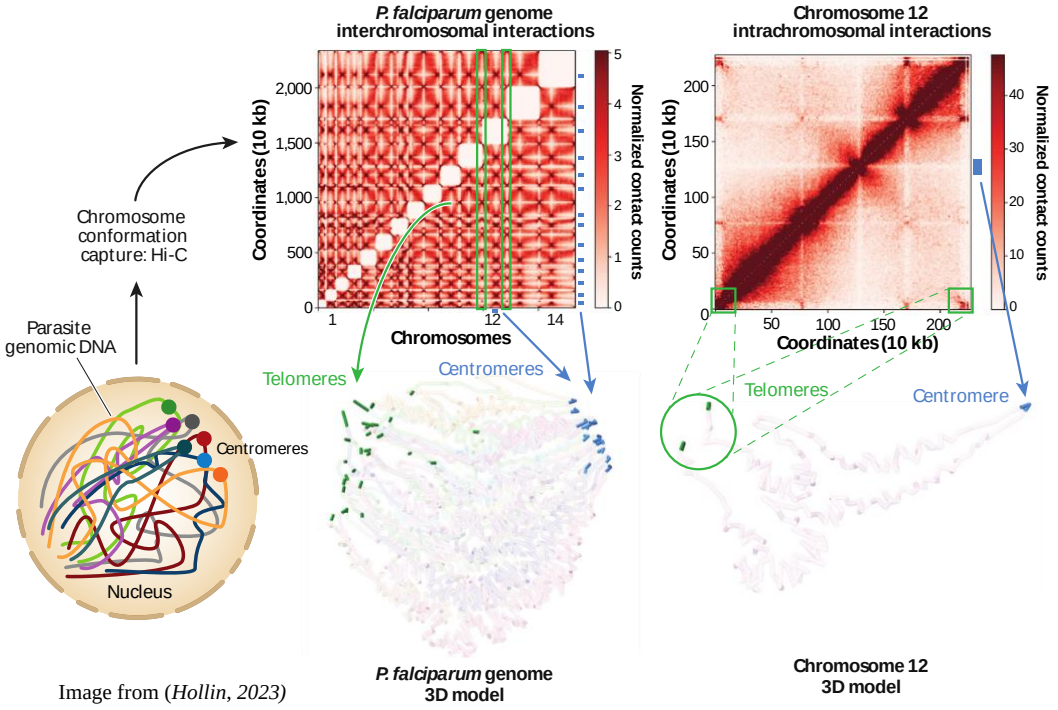


Image from (*Hollin, 2023)*

Figure I-5. Hi-C data and 3D modeling of the *Plasmodium falciparum* genome. The chromatin conformation of the parasite genome is captured by Hi-C. Analysis of Hi-C data allows the generation of chromosomal contact maps depicting inter- and intrachromosomal interactions. This technique also enables reliable computer modeling of genome 3D structure. Interchromosomal interactions (left) between telomeres (green) and centromeres (blue) of the parasite genome are highlighted. The intrachromosomal interactions and the 3D model of chromosome 12 have been emphasized (right) *(80).*

Usually transcribed by RNA polymerase II, lncRNAs are capped, spliced, and polyadenylated as are their mRNA counterparts, yet do not code for proteins. They can be classified as: sense, antisense, bidirectional, intronic, or intergenic. lncRNAs can bind to DNAs, RNAs, and proteins, playing various functional roles. By attaching to genomic DNA, lncRNAs can alter 3D chromatin structures, facilitate long-range interactions

between regulatory elements and regulate gene expression. Through transcriptional interference or by signaling downstream gene expression, lncRNAs can play fundamental roles in many essential cellular processes. For example, the intergenic lncRNA transcript (TR2) is involved in activation of the *clag3* gene switching and/or expression which plays a role in resistance to antimalarials such as Blasticidin S. Additionally, lncRNAs can influence gene promoter accessibility by recruiting chromatin remodeling enzymes such as histone acetylases and deacetylases. Regulatory ncRNAs, have also been implicated in governing *var* gene expression and other subtelomeric variant surface antigens (VSAs). lncRNAs are implicated in the modulation of antigenic switching of the *Pf*EMP1 variants through transcription of both a sense lncRNA and an antisense lncRNA. These lncRNAs originate from a bidirectional promoter in the *var* intron and are extended to exon I and exon II for antisense and sense lncRNAs, respectively (*60, 81*).

Expression of antisense lncRNAs from a particular silent *var* increases the activation of this specific gene and triggers, in trans, this *var* gene switching (*98, 99*). Several studies have also been performed on RUF6, categorized as a family of ncRNAs of Unknown Function (*100*). There are roughly 15 RUF6 ncRNAs in *Plasmodium* located at or near internal *var* genes and are expressed in a clonally variant manner. They colocalize with telomeric and internal *var* genes, and their overexpression can upregulate active *var* gene expression. Recent studies also suggest that several ApiAP2 TF2, termed heterochromatin-associated factors, may mediate the influence of RUF6-ncRNA on *var* gene expression. LncRNAs transcribed from the *P. falciparum* telomere-associated repetitive elements

(TARE-lncRNAs) may also play a role in forming and maintaining local heterochromatin near key immune evasion genes. These TARE-DNA sequences span from the telomeres to the first subtelomeric *var* genes. TARE-lncRNAs are found to be transcribed throughout the IDC and may regulate telomeric heterochromatin by recruiting chromatin modifiers and repressive hallmarks, controlling telomere maintenance, heterochromatin clusters, and *var* gene expression(*60*).

LncRNAs are also serve as key constituents of early sexual commitment in *Plasmodium.* Gametocytogenesis is regulated by the master transcriptional regulator *Pfap2-g*, triggering sexual commitment. The AP2-Gprotein in turn, is regulated by the GDV1 protein. The GDV1protein displaces HP1 which normally mediates *Pfap2-g* gene silencing during asexual conditions. With displacement of HP1, *Pfap2-g* expression is permitted, and sexual commitment is accessible. During asexual stages a five-exon antisense lncRNA is transcribed downstream from the *Pfgdv1* locus repressing *Pfgdv1*.Additionally, knockout of this lncRNA increased expression of several genes associated with sexual differentiation (*ap2-g*, *dblmsp2*, *Pfg14_748*, and *Pf3d7_1477400*)(*101*). Recently, an inducible gametocyte producer line was engineered to overexpress GDV1, leading to a sexual conversion rate of 75%, a striking difference to the 8% found in the control population (*102*), validating the essentiality of this gene-lncRNA in this mechanism. At the post-transcriptional level, lncRNAs have also been shown to affect mRNA export, turnover, decay, and translation. In the cytoplasm, they serve as scaffolds for protein complexes involved in gene expression, interacting directly with translation factors, thus influencing

the frequency and efficiency of translation. Despite the perplexity and debate maintained in the study of ncRNAs, the relative importance of these molecules across parasite development and transmission is undeniable. lnRNAs have been proven to serve an array of functional roles in the pathogen, regulating all aspects of regular cell development. Yet substantial efforts are still needed to better understand the role these ncRNAs have on the greater mechanistic and functional implications they have in parasite biology(*60*).

7. Nuclear And Chromatin Architecture

As discussed previously, the transcriptional status of malaria parasites is dynamic and controlled by various and interconnected layers of regulatory mechanisms associated with phenotypic and morphological changes. In the ring stage, for example, the genome is compacted and enriched in nucleosomes, with each nucleus encompassing 3 to 7 tightly clustered nuclear pores (*103-105*). This condensation hampers the binding of the transcriptional machinery and is consistent with the low transcription activity detected during ring stages. Conversely, the nuclear volume and the number of nuclear pore complexes (NPCs) are significantly increased in trophozoite stages, validating the open chromatin structure and active transcriptional activity (*103, 105*). These NPCs are roughly 60 in number and seem not to colocalize with HP1 and H3K9me3, two main heterochromatin markers, suggesting they are largely distributed in euchromatin and active regions (*106*). The number of pores gradually decreases throughout schizogony, with 6–16 NPCs for mid-schizonts and 2–6 for late schizonts, clustering once more to a single location near the euchromatin region (*105*). At this stage, nucleosomes are repackaged, and the

chromatin is compacted to facilitate the processes of cytokinesis and reinvasion, leading to a reduced nuclear volume akin to the ring stage. Female gametocytes have been shown to develop a smaller and more compacted nucleus, with transcripts stored and translationally repressed, while male gametocytes, needing to prepare for mitosis and gamete formation, show larger nucleus volumes. In *P. berghei*, nuclear pores are distant from DNA, but *P .falciparum* has morphologically atypical gametocytes, which has hindered the ability to draw reliable conclusions(*107*). The remarkable plasticity and patterns of nuclear landscapes between *Plasmodium spp*. and within transitional stages warrant further analysis to help describe these complex parasites(*60*).

8. Chromatin Architecture

The pathogen is also reliant on the overall chromosomal architecture to mechanistically control the spatial proximity between essential genome components such as promotors and enhances cross talk between regions otherwise located distantly apart on the same chromosome or on different chromosomes. These contact points were first recorded during early studies employing FISH, revealing important loci belonging to *var* gene families in the subtelomeric regions of heterologous chromosomes at the nuclear periphery(*82*, *108*). Clustering of the chromosomal telomeres as a repressive center enriched in H3K9me3 and HP1 protein enables interaction and communication between the *var* gene families, crucial for mutual exclusive expression. These perinuclear repressive clusters congregate additional subtelomeric multigene families associated with reinvasion such as stevor and *Pf*mc-2TM. Like their *var* gene neighbors, these clonally variant gene families are subject

to mutually exclusive expression during the parasite cell cycle(*109*). The intra- and interchromosomal organizational patterns were further divulged following the development of 3D chromosome capture tools. The unique patterns of chromosomal condensation were found to reflect the transcriptomic activity of parasite cell cycle progression. This genomic landscape adopts tight condensation formation throughout the pathogen's early ring stage and during late schizogony, consistent with the low transcriptional activity associated with these parasite stages (*60*).

Conversely, contact maps and 3D models uncovered a more open chromatin throughout the trophozoite stages, when transcription peaks. Hi-C studies revealed the spatiotemporal nature of the parasite's genome whereby centromeres are directed to one end of the nuclear periphery along with discrete telomeric heterochromatin clusters allocated to the opposing end(*103, 110, 111*) **(Figure I-5)**. Although no well-defined topologically associating domains were identified in *Plasmodium*, these Hi-C analyses have helped highlight the formation of domain-like structures surrounding the genes involved in antigenic variation bringing the subtelomeric and internal *var* genes loci closer together. Distinct changes were also observed for the transmission stages, such as rearrangement of the *ap2-g* locus moving it from the repressive heterochromatin cluster during the asexual stage to the active euchromatin at the early gametocyte stage, which led to its expression (*112*). Large chromosomal rearrangements were also observed for genes encoding proteins involved in invasion and exported proteins associated with gametocytogenesis and erythrocyte remodeling, indicating their relocation with the repressed virulence genes at the

gametocyte stage (*112*). With the advancements in chromatin capture technologies, crucial features have been revealed about the importance of the dynamic nature of the *Plasmodium* chromosomal landscape driving developmental stages of the parasite **(Figure I-5)**(*60*).

CONCLUDING STATEMENTS

This thesis report details the pathogenicity of the *Plasmodium falciparum* parasite and outlines the global impact of this debilitating disease. A significant focus is placed on the unique attributes of the pathogen that enables it to survive, evading all control efforts and host immune responses. Here I bring to attention the control mechanisms and eradication efforts over the decades, highlighting the pathogen's resilience and adaptability. The report also reviews the apicoplast organelle, common to all *Apicomplexans* including *Plasmodium*, emphasizing its significance for parasite survival and in response, the importance of this plastid in malaria research. Additionally, it describes the unique characteristics of the *Plasmodium* chromatin landscape and the epigenetic mechanisms regulating this enigmatic parasite.

The first chapter brings to attention the early steps involved in drug discovery research. In collaboration with the University of California, Irvine's Department of Chemistry, I tested various chemical compounds produced in attempts to refine a potential antimalarial agent. This work involved a continuous exchange of information with the Vanderwal lab to help them better understand the impact of their compound manipulation on the potency of the drug in *Plasmodium* parasites. Though our lab’s partnership, the Vanderwal lab was capable of simplifying and upscaling the compound to levels required for future drug

mechanism of action (MoA) studies. These efforts helped expose the active site of the compound, conserving its potency, whilst increasing its stability and solubility in solution. Our combined efforts resulted in the published work described below. The second chapter of the thesis delves into the subsequent stages involved in drug discovery and drug refinement. In this second chapter we explore systems-wide approaches applied for identifying new therapeutic targets. I investigate the mechanism of action of a novel antimalarial agent, Kalihinol, and its potential resistance mechanisms in the pathogen. Using multi-omics platforms, I gained a comprehensive understanding of the compound's effects on both parasite and host biology. I conducted all *in vitro* experiments on *P. falciparum*, providing significant evidence of the compound's mechanism of action. My work included drug-based assays along with precisely designed experiments tailored to validate our scientific hypothesis on the compound localization and effect on the apicoplast organelle. While our hypothesis was confirmed, our transcriptomic analyses, proteomic, metabolomic assays, and CRISPR-Cas9 genome editing provided an unbiased and comprehensive understanding of the complexity of the compound within the parasite biology. I applied an array of molecular approaches, not only to identify the molecular pathways impacted by the Kalihinol analogue, but I also confirm the role of *Pfsec13* in its resistance. This research laid the foundation for a larger collaborative project, which including several national and international researchers to provide greater *in vivo* testing on humanized mouse models as well as several other *Plasmodium spp.* Collectively this ambitious and collaborative project brought forth a comprehensive and detailed analysis

on our compound. Our efforts were ultimately acknowledged and was accepted for publication.

The final two chapters examine the chromatin landscape of *Plasmodium*, identifying the previously unknown *Pf*MORC protein's role in chromosomal rearrangement and gene silencing, making it a promising new therapeutic target. My work involved proteomics, phenotypic assays, genomic verification, and conducting ChIP-sequencing to validate the role of this protein in heterochromatin landscape and antigenic gene silencing. I played a significant role in both the molecular, benchwork experimentation, and in writing of the finalized manuscript. Additionally, I explored the functional significance of ncRNAs in gene regulation and sexual stage induction, using various molecular and genomic approaches. My role included CRISPR-Cas9-based knockout of the previously uncharacterized ncRNA, lncRNA-14, followed by genomic validation, phenotypic and transcriptomic analysis to assess its impact on gametocyte development and sexual differentiation. This research was also published collaboratively. Overall, this project aims to identify new targets to combat malaria, whether through developing new antimalarial agents or discovering novel targets critical for the parasite's survival.

REFERENCES

1. WHO, World malaria report 2023. (2023).

2. X. Wang *et al.*, Molecular Epidemiology of Drug Resistance Genes in Plasmodium falciparum Isolates Imported from Nigeria between 2016 and 2020: Continued Emergence of Fully Resistant Pfdhfr-Pfdhps Alleles. *Microbiol Spectr*, e0052822 (2022).

3. W. H. Organization, Global report on antimalarial drug efficacy and drug resistance, 2000–2010. *WHO*, (2010).

4. A. Uwimana *et al.*, Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda. *Nat Med* **26**, 1602-1608 (2020).

5. A. Uwimana *et al.*, Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study. *Lancet Infect Dis* **21**, 1120-1128 (2021).

6. N. Noreen *et al.*, New insights into the spread of resistance to artemisinin and its analogues. *J Glob Antimicrob Resist* **27**, 142-149 (2021).

7. M. V. Andrade *et al.*, The economic burden of malaria: a systematic review. *Malar J* **21**, 283 (2022).

8. A. M. Dondorp *et al.*, Artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med* **361**, 455-467 (2009).

9. A. Haakenstad *et al.*, Tracking spending on malaria by source in 106 countries, 2000-16: an economic modelling study. *Lancet Infect Dis* **19**, 703-716 (2019).

10. U. Bohme *et al.*, Complete avian malaria parasite genomes reveal features associated with lineage-specific evolution in birds and mammals. *Genome Res* **28**, 547-560 (2018).

11. R. Thomson-Luque *et al.*, Plasmodium falciparum transcription in different clinical presentations of malaria associates with circulation time of infected erythrocytes. *Nat Commun* **12**, 4711 (2021).

12. W. Liu *et al.*, Multigenomic Delineation of Plasmodium Species of the Laverania Subgenus Infecting Wild-Living Chimpanzees and Gorillas. *Genome Biol Evol* **8**, 1929-1939 (2016).

13. J. R. Poespoprodjo, N. M. Douglas, D. Ansong, S. Kho, N. M. Anstey, Malaria. *Lancet* **402**, 2328-2345 (2023).

14. I. Mueller, P. A. Zimmerman, J. C. Reeder, Plasmodium malariae and Plasmodium ovale--the "bashful" malaria parasites. *Trends Parasitol* **23**, 278-283 (2007).

15. C. J. Merrick, Hypnozoites in Plasmodium: Do Parasites Parallel Plants? *Trends Parasitol* **37**, 273-282 (2021).

16. S. Mehra, E. Stadler, D. Khoury, J. M. McCaw, J. A. Flegg, Hypnozoite dynamics for Plasmodium vivax malaria: The epidemiological effects of radical cure. *J Theor Biol* **537**, 111014 (2022).

17. T. B. Vieira, T. P. Astro, R. R. de Moraes Barros, Genetic Manipulation of Non-Falciparum Human Malaria Parasites. *Front Cell Infect Microbiol* **11**, 680460 (2021).

18. L. Tilley, M. McConville, Malaria: Sensing when it's time for sex. *Nature* **499**, 38-40 (2013).

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

20. S. Eksi *et al.*, Plasmodium falciparum gametocyte development 1 (Pfgdv1) and gametocytogenesis early gene identification and commitment to sexual development. *PLoS Pathog* **8**, e1002964 (2012).

21. J. Chawla, J. Oberstaller, J. H. Adams, Targeting Gametocytes of the Malaria Parasite Plasmodium falciparum in a Functional Genomics Era: Next Steps. *Pathogens* **10**, (2021).

22. J. D. Dvorin, D. E. Goldberg, Plasmodium Egress Across the Parasite Life Cycle. *Annu Rev Microbiol* **76**, 67-90 (2022).

23. G. I. McFadden, E. Yeh, The apicoplast: now you see it, now you don't. *Int J Parasitol* **47**, 137-144 (2017).

24. K. Keegstra, K. Cline, Protein import and routing systems of chloroplasts. *Plant Cell* **11**, 557-570 (1999).

25. D. W. Chung, N. Ponts, J. Prudhomme, E. M. Rodrigues, K. G. Le Roch, Characterization of the ubiquitylating components of the human malaria parasite's protein degradation pathway. *PLoS One* **7**, e43477 (2012).

26. J. D. Fellows, M. J. Cipriano, S. Agrawal, B. Striepen, A Plastid Protein That Evolved from Ubiquitin and Is Required for Apicoplast Protein Import in Toxoplasma gondii. *mBio* **8**, (2017).

27. G. I. McFadden, M. E. Reith, J. Munholland, N. Lang-Unnasch, Plastid in human parasites. *Nature* **381**, 482 (1996).

28. J. I. MacRae, E. Marechal, C. Biot, C. Y. Botte, The apicoplast: a key target to cure malaria. *Curr Pharm Des* **18**, 3490-3504 (2012).

29. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation. *Science* **295**, 1306-1311 (2002).

30. A. M. Guggisberg, R. E. Amthor, A. R. Odom, Isoprenoid biosynthesis in Plasmodium falciparum. *Eukaryot Cell* **13**, 1348-1359 (2014).

31. H. Jomaa *et al.*, Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285**, 1573-1576 (1999).

32. G. G. van Dooren, L. M. Stimmler, G. I. McFadden, Metabolic maps and functions of the Plasmodium mitochondrion. *FEMS Microbiol Rev* **30**, 596-630 (2006).

33. R. M. Simao-Gurge *et al.*, Biosynthesis of heme O in intraerythrocytic stages of Plasmodium falciparum and potential inhibitors of this pathway. *Sci Rep* **9**, 19261 (2019).

34. N. Thomsen-Zieger, J. Schachtner, F. Seeber, Apicomplexan parasites contain a single lipoic acid synthase located in the plastid. *FEBS Lett* **547**, 80-86 (2003).

35. C. Bisanz *et al.*, Toxoplasma gondii acyl-lipid metabolism: de novo synthesis from apicoplast-generated fatty acids versus scavenging of host cell precursors. *Biochem J* **394**, 197-205 (2006).

36. R. M. Packard, The origins of antimalarial-drug resistance. *N Engl J Med* **371**, 397-399 (2014).

37. N. Ma, Z. Zhang, F. Liao, T. Jiang, Y. Tu, The birth of artemisinin. *Pharmacol Ther* **216**, 107658 (2020).

38. X. Z. Su, L. H. Miller, The discovery of artemisinin and the Nobel Prize in Physiology or Medicine. *Sci China Life Sci* **58**, 1175-1179 (2015).

39. S. Oguche *et al.*, Efficacy of artemisinin-based combination treatments of uncomplicated falciparum malaria in under-five-year-old Nigerian children. *Am J Trop Med Hyg* **91**, 925-935 (2014).

40. A. Djimde *et al.*, A molecular marker for chloroquine-resistant falciparum malaria. *N Engl J Med* **344**, 257-263 (2001).

41. K. J. Wicht, S. Mok, D. A. Fidock, Molecular Mechanisms of Drug Resistance in Plasmodium falciparum Malaria. *Annu Rev Microbiol* **74**, 431-454 (2020).

42. T. Hollin, Z. Chahine, K. G. Le Roch, Epigenetic Regulation and Chromatin Remodeling in Malaria Parasites. *Annu Rev Microbiol* **77**, 255-276 (2023).

43. Malaria eradication: benefits, future scenarios and feasibility: a report of the Strategic Advisory Group on Malaria Eradication. *World Health Organization*, (2021).

44. WHO, Global Malaria Programme. (2024).

45. L. S. Garcia, Malaria. *Clin Lab Med* **30**, 93-129 (2010).

46. WHO. (2021).

47. M. Prof Christopher JL Murray *et al.*, Global malaria mortality between 1980 and 2010: a systematic analysis. *The Lancet Journal* **379**, 413-431 (2012).

48. A. Hammond *et al.*, Gene-drive suppression of mosquito populations in large cages as a bridge between lab and field. *Nat Commun* **12**, 4589 (2021).

49. P. E. Duffy, J. Patrick Gorres, Malaria vaccines since 2000: progress, priorities, products. *NPJ Vaccines* **5**, 48 (2020).

50. B. Mordmuller *et al.*, A PfSPZ vaccine immunization regimen equally protective against homologous and heterologous controlled human malaria infection. *NPJ Vaccines* **7**, 100 (2022).

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

52. S. Z. Davis, T. Hollin, T. Lenz, K. G. Le Roch, Three-dimensional chromatin in infectious disease-A role for gene regulation and pathogenicity? *PLoS Pathog* **17**, e1009207 (2021).

53. G. Batugedara, K. G. Le Roch, Unraveling the 3D genome of human malaria parasites. *Semin Cell Dev Biol* **90**, 144-153 (2019).

54. T. Hollin, K. G. Le Roch, From Genes to Transcripts, a Tightly Regulated Journey in Plasmodium. *Front Cell Infect Microbiol* **10**, 618454 (2020).

55. T. Hollin, M. Gupta, T. Lenz, K. G. Le Roch, Dynamic Chromatin Structure and Epigenetics Control the Fate of Malaria Parasites. *Trends Genet* **37**, 73-85 (2021).

56. T. Hollin, Z. Chahine, K. G. Le Roch, Epigenetic Regulation and Chromatin Remodeling in Malaria Parasites. *Annu Rev Microbiol*, (2023).

57. M. T. Watzlowik, S. Das, M. Meissner, G. Langst, Peculiarities of Plasmodium falciparum Gene Regulation and Chromatin Structure. *Int J Mol Sci* **22**, (2021).

58. X. Shang *et al.*, PfAP2-EXP2, an Essential Transcription Factor for the Intraerythrocytic Development of Plasmodium falciparum. *Front Cell Dev Biol* **9**, 782293 (2021).

59. T. L. Campbell, E. K. De Silva, K. L. Olszewski, O. Elemento, M. Llinas, Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. *PLoS Pathog* **6**, e1001165 (2010).

60. X. Shang *et al.*, Genome-wide landscape of ApiAP2 transcription factors reveals a heterochromatin-associated regulatory network during Plasmodium falciparum blood-stage development. *Nucleic Acids Res* **50**, 3413-3431 (2022).

61. R. M. Martins *et al.*, An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite Plasmodium falciparum. *Sci Rep* **7**, 14042 (2017).

62. E. Carrington *et al.*, The ApiAP2 factor PfAP2-HC is an integral component of heterochromatin in the malaria parasite Plasmodium falciparum. *iScience* **24**, 102444 (2021).

63. C. Flueck *et al.*, A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in chromosome end biology. *PLoS Pathog* **6**, e1000784 (2010).

64. M. Sierra-Miranda *et al.*, PfAP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds to Plasmodium falciparum telomeres. *Cell Microbiol* **19**, (2017).

65. Z. Chahine, K. G. Le Roch, Decrypting the complexity of the human malaria parasite biology through systems biology approaches. *Front Syst Biol* **2**, (2022).

66. W. A. Hoeijmakers *et al.*, Plasmodium falciparum centromeres display a unique epigenetic makeup and cluster prior to and during schizogony. *Cell Microbiol* **14**, 1391-1401 (2012).

67. J. Miao *et al.*, The malaria parasite Plasmodium falciparum histones: organization, expression, and acetylation. *Gene* **369**, 53-65 (2006).

68. W. J. Sullivan, Jr., Histone H3 and H3.3 variants in the protozoan pathogens Plasmodium falciparum and Toxoplasma gondii. *DNA Seq* **14**, 227-231 (2003).

69. W. A. Hoeijmakers *et al.*, H2A.Z/H2B.Z double-variant nucleosomes inhabit the AT-rich promoter regions of the Plasmodium falciparum genome. *Mol Microbiol* **87**, 1061-1073 (2013).

70. M. Petter *et al.*, H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy var gene promoters in the malaria parasite Plasmodium falciparum. *Mol Microbiol* **87**, 1167-1182 (2013).

71. M. Petter *et al.*, Expression of P. falciparum var genes involves exchange of the histone variant H2A.Z at the promoter. *PLoS Pathog* **7**, e1001292 (2011).

72. N. Coetzee *et al.*, Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual Plasmodium falciparum parasites. *Sci Rep* **7**, 607 (2017).

73. A. M. Salcedo-Amaya *et al.*, Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. *Proc Natl Acad Sci U S A* **106**, 9655-9660 (2009).

74. A. Saraf *et al.*, Dynamic and Combinatorial Landscape of Histone Modifications during the Intraerythrocytic Developmental Cycle of the Malaria Parasite. *J Proteome Res* **15**, 2787-2801 (2016).

75. M. B. Trelle, A. M. Salcedo-Amaya, A. M. Cohen, H. G. Stunnenberg, O. N. Jensen, Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum. *J Proteome Res* **8**, 3439-3450 (2009).

76. C. Epp, F. Li, C. A. Howitt, T. Chookajorn, K. W. Deitsch, 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**, 116-127 (2009).

77. J. J. Lopez-Rubio, L. Mancio-Silva, A. Scherf, Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* **5**, 179-190 (2009).

78. S. A. Ralph *et al.*, Tropical infectious diseases: metabolic maps and functions of the Plasmodium falciparum apicoplast. *Nat Rev Microbiol* **2**, 203-216 (2004).

79. J. J. Lopez-Rubio *et al.*, 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol Microbiol* **66**, 1296-1305 (2007).

80. L. H. Freitas-Junior *et al.*, Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121**, 25-36 (2005).

81. S. Shrestha *et al.*, Distinct Histone Post-translational Modifications during Plasmodium falciparum Gametocyte Development. *J Proteome Res* **21**, 1857-1867 (2022).

82. J. Connacher *et al.*, H3K36 methylation reprograms gene expression to drive early gametocyte development in Plasmodium falciparum. *Epigenetics Chromatin* **14**, 19 (2021).

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

84. A. Kanyal *et al.*, Genome-wide survey and phylogenetic analysis of histone acetyltransferases and histone deacetylases of Plasmodium falciparum. *FEBS J* **285**, 1767-1782 (2018).

85. B. K. Chaal, A. P. Gupta, B. D. Wastuwidyaningtyas, Y. H. Luah, Z. Bozdech, Histone deacetylases play a major role in the transcriptional regulation of the Plasmodium falciparum life cycle. *PLoS Pathog* **6**, e1000737 (2010).

86. L. Cui *et al.*, PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in Plasmodium falciparum. *Eukaryot Cell* **6**, 1219-1227 (2007).

87. N. Coetzee *et al.*, Epigenetic inhibitors target multiple stages of Plasmodium falciparum parasites. *Sci Rep* **10**, 2355 (2020).

88. M. V. C. Greenberg, D. Bourc'his, The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* **20**, 590-607 (2019).

89. N. Ponts *et al.*, Genome-wide mapping of DNA methylation in the human malaria parasite Plasmodium falciparum. *Cell Host Microbe* **14**, 696-706 (2013).

90. E. Hammam *et al.*, Discovery of a new predominant cytosine DNA modification that is linked to gene expression in malaria parasites. *Nucleic Acids Res* **48**, 184-199 (2020).

91. J. Baum *et al.*, Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Res* **37**, 3788-3798 (2009).

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

93. Q. Jing *et al.*, Plasmodium falciparum var Gene Is Activated by Its Antisense Long Noncoding RNA. *Front Microbiol* **9**, 3117 (2018).

94. I. Amit-Avraham *et al.*, Antisense long noncoding RNAs regulate var gene activation in the malaria parasite Plasmodium falciparum. *Proc Natl Acad Sci U S A* **112**, E982-991 (2015).

95. E. L. Schymanski *et al.*, Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* **48**, 2097-2098 (2014).

96. M. Filarsky *et al.*, GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* **359**, 1259-1263 (2018).

97. S. D. Boltryk *et al.*, CRISPR/Cas9-engineered inducible gametocyte producer lines as a valuable tool for Plasmodium falciparum malaria transmission research. *Nat Commun* **12**, 4806 (2021).

98. F. Ay *et al.*, Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. *Genome Res* **24**, 974-988 (2014).

99. E. M. Bunnik *et al.*, DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite Plasmodium falciparum. *BMC Genomics* **15**, 347 (2014).

100. A. Weiner *et al.*, 3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite Plasmodium falciparum. *Cell Microbiol* **13**, 967-977 (2011).

101. N. Dahan-Pasternak *et al.*, PfSec13 is an unusual chromatin-associated nucleoporin of Plasmodium falciparum that is essential for parasite proliferation in human erythrocytes. *J Cell Sci* **126**, 3055-3069 (2013).

102. J. Kehrer *et al.*, Nuclear Pore Complex Components in the Malaria Parasite Plasmodium berghei. *Sci Rep* **8**, 11249 (2018).

103. L. H. Freitas-Junior *et al.*, Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature* **407**, 1018-1022 (2000).

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

105. J. E. Lemieux *et al.*, Genome-wide profiling of chromosome interactions in Plasmodium falciparum characterizes nuclear architecture and reconfigurations associated with antigenic variation. *Mol Microbiol* **90**, 519-537 (2013).

106. E. M. Bunnik *et al.*, Comparative 3D genome organization in apicomplexan parasites. *Proc Natl Acad Sci U S A* **116**, 3183-3192 (2019).

107. E. M. Bunnik *et al.*, Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. *Nat Commun* **9**, 1910 (2018).

108. A. M. Samuels *et al.*, Efficacy of RTS,S/AS01E malaria vaccine administered according to different full, fractional, and delayed third or early fourth dose regimens in

children aged 5-17 months in Ghana and Kenya: an open-label, phase 2b, randomised controlled trial. *Lancet Infect Dis* **22**, 1329-1342 (2022).

109. K. E. Ward, D. A. Fidock, J. L. Bridgford, Plasmodium falciparum resistance to artemisinin-based combination therapies. *Curr Opin Microbiol* **69**, 102193 (2022).

110. C. C. W. J. P. A. R. D. M. S. P. J. M. G. K. C. J., Kalihinol-A, a Highly Functionalized Diisocyano Diterpenoid Antibiotic from a Sponge. *J. Am. Chem. Soc.* **106**, 4644–4646 (1984).

111. L. H. Yang, O. O. Lee, T. Jin, X. C. Li, P. Y. Qian, Antifouling properties of 10beta-formamidokalihinol-A and kalihinol A isolated from the marine sponge Acanthella cavernosa. *Biofouling* **22**, 23-32 (2006).

112. H. Miyaoka, Y. Abe, E. Kawashima, Synthesis of marine diterpene isocyanide (-)-kalihinol Y and diterpene isothiocyanate (-)-10-epi-kalihinol I. *Chem Pharm Bull (Tokyo)* **60**, 1224-1226 (2012).

113. D. Wolf, F. J. Schmitz, New diterpene isonitriles from the sponge Phakellia pulcherrima. *J Nat Prod* **61**, 1524-1527 (1998).

114. M. E. Daub, J. Prudhomme, K. Le Roch, C. D. Vanderwal, Synthesis and potent antimalarial activity of kalihinol B. *J Am Chem Soc* **137**, 4912-4915 (2015).

115. M. E. Daub, J. Prudhomme, C. Ben Mamoun, K. G. Le Roch, C. D. Vanderwal, Antimalarial Properties of Simplified Kalihinol Analogues. *ACS Med Chem Lett* **8**, 355-360 (2017).

116. A. D. Wright *et al.*, Antimalarial activity: the search for marine-derived natural products with selective antimalarial activity. *J Nat Prod* **59**, 710-716 (1996).

117. M. J. Schnermann, R. A. Shenvi, Syntheses and biological studies of marine terpenoids derived from inorganic cyanide. *Nat Prod Rep* **32**, 543-577 (2015).

118. A. D. Wright *et al.*, Inhibition of heme detoxification processes underlies the antimalarial activity of terpene isonitrile compounds from marine sponges. *J Med Chem* **44**, 873-885 (2001).

119. H. H. Lu *et al.*, Synthesis of (+)-7,20-Diisocyanoadociane and Liver-Stage Antiplasmodial Activity of the Isocyanoterpene Class. *J Am Chem Soc* **138**, 7268-7271 (2016).

120. D. F. Taber, B. P. Gunn, Control elements in the intramolecular Diels-Alder reaction. Synthesis of (.+-.)-torreyol. *Journal of the American Chemical Society* **101**, 3992-3993 (1979).

121. S. V. Pronin, C. A. Reiher, R. A. Shenvi, Stereoinversion of tertiary alcohols to tertiary-alkyl isonitriles and amines. *Nature* **501**, 195-199 (2013).

122. E. L. Dahl, P. J. Rosenthal, Multiple antibiotics exert delayed effects against the Plasmodium falciparum apicoplast. *Antimicrob Agents Chemother* **51**, 3485-3490 (2007).

123. K. Kennedy *et al.*, Delayed death in the malaria parasite Plasmodium falciparum is caused by disruption of prenylation-dependent intracellular trafficking. *PLoS Biol* **17**, e3000376 (2019).

124. S. C. Nair *et al.*, Apicoplast isoprenoid precursor synthesis and the molecular basis of fosmidomycin resistance in Toxoplasma gondii. *J Exp Med* **208**, 1547-1559 (2011).

125. C. J. Tonkin, N. S. Struck, K. A. Mullin, L. M. Stimmler, G. I. McFadden, Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites. *Mol Microbiol* **61**, 614-630 (2006).

126. R. F. Waller, M. B. Reed, A. F. Cowman, G. I. McFadden, Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. *EMBO J* **19**, 1794-1802 (2000).

127. J. Meletiadis, S. Pournaras, E. Roilides, T. J. Walsh, Defining fractional inhibitory concentration index cutoffs for additive interactions based on self-drug additive combinations, Monte Carlo simulation analysis, and in vitro-in vivo correlation data for antifungal drug combinations against Aspergillus fumigatus. *Antimicrob Agents Chemother* **54**, 602-609 (2010).

128. T. C. Chou, Derivation and properties of Michaelis-Menten type and Hill type equations for reference ligands. *J Theor Biol* **59**, 253-276 (1976).

129. S. Steinbacher *et al.*, Structural basis of fosmidomycin action revealed by the complex with 2-C-methyl-D-erythritol 4-phosphate synthase (IspC). Implications for the catalytic mechanism and anti-malaria drug development. *J Biol Chem* **278**, 18401-18407 (2003).

130. E. Yeh, J. L. DeRisi, Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage Plasmodium falciparum. *PLoS Biol* **9**, e1001138 (2011).

131. T. Uddin, G. I. McFadden, C. D. Goodman, Validation of Putative Apicoplast-Targeting Drugs Using a Chemical Supplementation Assay in Cultured Human Malaria Parasites. *Antimicrob Agents Chemother* **62**, (2018).

132. S. E. Lindner *et al.*, Enzymes involved in plastid-targeted phosphatidic acid synthesis are essential for Plasmodium yoelii liver-stage development. *Mol Microbiol* **91**, 679-693 (2014).

133. G. Pessi, G. Kociubinski, C. B. Mamoun, A pathway for phosphatidylcholine biosynthesis in Plasmodium falciparum involving phosphoethanolamine methylation. *Proc Natl Acad Sci U S A* **101**, 6206-6211 (2004).

134. N. Kilian, J. Y. Choi, D. R. Voelker, C. Ben Mamoun, Role of phospholipid synthesis in the development and differentiation of malaria parasites in the blood. *J Biol Chem* **293**, 17308-17316 (2018).

135. X. W. Chan *et al.*, Chemical and genetic validation of thiamine utilization as an antimalarial drug target. *Nat Commun* **4**, 2060 (2013).

136. J. Knockel *et al.*, Filling the gap of intracellular dephosphorylation in the Plasmodium falciparum vitamin B1 biosynthesis. *Mol Biochem Parasitol* **157**, 241-243 (2008).

137. E. E. Reid, P. Thompson, C. R. Lyttle, D. T. Dennis, Pyruvate dehydrogenase complex from higher plant mitochondria and proplastids. *Plant Physiol* **59**, 842-848 (1977).

138. J. P. Musabyimana *et al.*, Plasmodium falciparum S-Adenosylmethionine Synthetase Is Essential for Parasite Survival through a Complex Interaction Network with Cytoplasmic and Nuclear Proteins. *Microorganisms* **10**, (2022).

139. Y. Ouyang, Q. Wu, J. Li, S. Sun, S. Sun, S-adenosylmethionine: A metabolite critical to the regulation of autophagy. *Cell Prolif* **53**, e12891 (2020).

140. H. Choi, D. Fermin, A. I. Nesvizhskii, Significance analysis of spectral count data in label-free shotgun proteomics. *Mol Cell Proteomics* **7**, 2373-2385 (2008).

141. Y. Zhang, Z. Wen, M. P. Washburn, L. Florens, Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem* **82**, 2272-2281 (2010).

142. C. D. Broeckling, F. A. Afsar, S. Neumann, A. Ben-Hur, J. E. Prenni, RAMClust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data. *Anal Chem* **86**, 6812-6817 (2014).

143. H. Choi, S. Kim, D. Fermin, C. C. Tsou, A. I. Nesvizhskii, QPROT: Statistical method for testing differential expression using protein-level intensity data in label-free quantitative proteomics. *J Proteomics* **129**, 121-126 (2015).

144. M. Deponte *et al.*, Wherever I may roam: protein and membrane trafficking in P. falciparum-infected red blood cells. *Mol Biochem Parasitol* **186**, 95-116 (2012).

145. R. Tuteja, Unraveling the components of protein translocation pathway in human malaria parasite Plasmodium falciparum. *Arch Biochem Biophys* **467**, 249-260 (2007).

146. A. S. Mamidi, A. Ray, N. Surolia, Structural Analysis of PfSec62-Autophagy Interacting Motifs (AIM) and PfAtg8 Interactions for Its Implications in RecovER-phagy in Plasmodium falciparum. *Front Bioeng Biotechnol* **7**, 240 (2019).

147. H. Fang, N. Green, Nonlethal sec71-1 and sec72-1 mutations eliminate proteins associated with the Sec63p-BiP complex from S. cerevisiae. *Mol Biol Cell* **5**, 933-942 (1994).

148. A. Tripathi, E. C. Mandon, R. Gilmore, T. A. Rapoport, Two alternative binding mechanisms connect the protein translocation Sec71-Sec72 complex with heat shock proteins. *J Biol Chem* **292**, 8007-8018 (2017).

149. D. Martinez Molina *et al.*, Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84-87 (2013).

150. J. M. Dziekan *et al.*, Cellular thermal shift assay for the identification of drug-target interactions in the Plasmodium falciparum proteome. *Nat Protoc* **15**, 1881-1921 (2020).

151. A. Florentin, D. W. Cobb, H. M. Kudyba, V. Muralidharan, Directing traffic: Chaperone-mediated protein transport in malaria parasites. *Cell Microbiol* **22**, e13215 (2020).

152. M. M. Savitski *et al.*, Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784 (2014).

153. C. Nagaraj, N. K. Chakrabarati, M. V. Narasimham, A longitudinal study of P. falciparum response to chloroquine in relation to the epidemiology. *J Commun Dis* **21**, 229-240 (1989).

154. T. Masserey *et al.*, The influence of biological, epidemiological, and treatment factors on the establishment and spread of drug-resistant Plasmodium falciparum. *Elife* **11**, (2022).

155. A. L. Malisa *et al.*, Drug coverage in treatment of malaria and the consequences for resistance evolution--evidence from the use of sulphadoxine/pyrimethamine. *Malar J* **9**, 190 (2010).

156. T. J. Green *et al.*, Cloning and characterization of Plasmodium falciparum FCR-3/FMG strain. *Am J Trop Med Hyg* **34**, 24-30 (1985).

157. S. Siniossoglou *et al.*, A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. *Cell* **84**, 265-275 (1996).

158. M. V. Katti, R. Sami-Subbu, P. K. Ranjekar, V. S. Gupta, Amino acid repeat patterns in protein sequences: their diversity and structural-functional implications. *Protein Sci* **9**, 1203-1209 (2000).

159. M. B. Jimenez-Diaz *et al.*, Improved murine model of malaria using Plasmodium falciparum competent strains and non-myelodepleted NOD-scid IL2Rgammanull mice engrafted with human erythrocytes. *Antimicrob Agents Chemother* **53**, 4533-4536 (2009).

160. S. Tachibana *et al.*, Plasmodium cynomolgi genome sequences provide insight into Plasmodium vivax and the monkey malaria clade. *Nat Genet* **44**, 1051-1055 (2012).

161. N. M. B. Brancucci *et al.*, Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite Plasmodium falciparum. *Cell* **171**, 1532-1544 e1515 (2017).

162. P. K. Sheokand *et al.*, A Plasmodium falciparum lysophospholipase regulates host fatty acid flux via parasite lipid storage to enable controlled asexual schizogony. *Cell Rep* **42**, 112251 (2023).

163. G. G. Holz, Jr., Lipids and the malarial parasite. *Bull World Health Organ* **55**, 237-248 (1977).

164. B. J. Foth *et al.*, The malaria parasite Plasmodium falciparum has only one pyruvate dehydrogenase complex, which is located in the apicoplast. *Mol Microbiol* **55**, 39-53 (2005).

165. A. Mateus, T. A. Maatta, M. M. Savitski, Thermal proteome profiling: unbiased assessment of protein state through heat-induced stability changes. *Proteome Sci* **15**, 13 (2016).

166. Y. Harada, H. Li, J. S. Wall, H. Li, W. J. Lennarz, Structural studies and the assembly of the heptameric post-translational translocon complex. *J Biol Chem* **286**, 2956-2965 (2011).

167. D. Prusty *et al.*, Single-stranded DNA binding protein from human malarial parasite Plasmodium falciparum is encoded in the nucleus and targeted to the apicoplast. *Nucleic Acids Res* **38**, 7037-7053 (2010).

168. R. Prekeris, J. Klumperman, Y. A. Chen, R. H. Scheller, Syntaxin 13 mediates cycling of plasma membrane proteins via tubulovesicular recycling endosomes. *J Cell Biol* **143**, 957-971 (1998).

169. L. Ayong, G. Pagnotti, A. B. Tobon, D. Chakrabarti, Identification of Plasmodium falciparum family of SNAREs. *Mol Biochem Parasitol* **152**, 113-122 (2007).

170. L. A. Parish, J. C. Rayner, Plasmodium falciparum secretory pathway: characterization of PfStx1, a plasma membrane Qa-SNARE. *Mol Biochem Parasitol* **164**, 153-156 (2009).

171. J. Y. van der Meer, A. K. Hirsch, The isoprenoid-precursor dependence of Plasmodium spp. *Nat Prod Rep* **29**, 721-728 (2012).

172. C. Ungermann, D. Kummel, Structure of membrane tethers and their role in fusion. *Traffic* **20**, 479-490 (2019).

173. T. Wang, L. Li, W. Hong, SNARE proteins in membrane trafficking. *Traffic* **18**, 767-775 (2017).

174. J. Fu *et al.*, Apicoplast biogenesis mediated by ATG8 requires the ATG12-ATG5-ATG16L and SNAP29 complexes in Toxoplasma gondii. *Autophagy*, 1-19 (2022).

175. F. Rohdich *et al.*, Biosynthesis of terpenoids. 2C-Methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspF) from Plasmodium falciparum. *Eur J Biochem* **268**, 3190-3197 (2001).

176. E. L. Flannery, D. A. Fidock, E. A. Winzeler, Using genetic methods to define the targets of compounds with antimalarial activity. *J Med Chem* **56**, 7761-7771 (2013).

177. G. LaMonte *et al.*, Mutations in the Plasmodium falciparum Cyclic Amine Resistance Locus (PfCARL) Confer Multidrug Resistance. *mBio* **7**, (2016).

178. S. Menard *et al.*, Induction of Multidrug Tolerance in Plasmodium falciparum by Extended Artemisinin Pressure. *Emerg Infect Dis* **21**, 1733-1741 (2015).

179. F. Ariey *et al.*, A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* **505**, 50-55 (2014).

CHAPTER 1

Concise Synthesis of the Antiplasmodial Isocyanoterpene 7,20-Diisocyanoadociane

Karns AS, Ellis BD, Roosen PC, **Chahine Z**, Le Roch KG, Vanderwal CD.Angew Chem Int Ed Engl. 2019 Sep 23;58(39):13749-13752. doi: 10.1002/anie.201906834. Epub 2019 Aug 27.

[1] 102 Natural Sciences II, Department of Chemistry, University of California, Irvine, CA, 92697-2025, USA.
[2] Institute for Integrative Genome Biology, Center for Infectious Disease and Vector Research, Department of Molecular, Cell, and Systems Biology, University of California, Riverside, 900 University Avenue, Riverside, CA, 92521, USA.

PREFACE

Identification and synthesis of new and effective antimalarial drugs include empirical screening, examining plant extracts from traditional medicines, creating analogs of compounds with known antimalarial (or antimicrobial or anti-tumor) activity. Many compounds are modified to target specific parasite enzymes through rational drug design. Antimalarial therapeutics begin at a chemist's bench, progressing through efficacy testing, toxicity testing, and clinical trials. The development of novel antimalarial drugs is influenced by the structural diversity, number of available compounds for screening and the strategies used to identify them. While discovering new compounds targeting known parasite mechanisms can be beneficial, identifying new targets is equally crucial. This may involve understanding the functional roles of essential parasite molecular biology and pathogen development. In this first chapter, we collaborate with Dr. Christopher Vanderwal at the Department of Chemistry in the University of California, Irvine to test

the efficacy of a potential antimalarial isocyanoterpene compound termed 7,20-diisocyanoadociane (DICA).

My role in this project consisted of assessing the solubility, stability, and potency of the compounds synthesized by the Vanderwal lab. Compounds received by the lab were tested on numerous wild-type and drug-resistant lines to compare efficacy. Results were communicated back to the University of California, Irvine, where the lab was able to reassess and refine their compound reactions based upon the findings we provided. This chapter sheds light on the preliminary steps involved in drug development and the screening processes that are needed to produce simple, stable and effective drug therapies.

ABSTRACT

The flagship member of the antiplasmodial isocyanoterpenes, 7,20-diisocyanoadociane (DICA), was synthesized from dehydrocryptone in 10 steps, and 13 from commercially available material. Our previous formal synthesis was reengineered, leveraging only productive transformations to deliver DICA in fewer than half the number of steps of our original effort. Important contributions, in addition to the particularly concise strategy, include a solution to the problem of axial nucleophilic methylation of a late-stage cyclohexanone, and the first selective synthesis and antiplasmodial evaluation of the DICA stereoisomer with both isonitriles equatorial.

INTRODUCTION

Sponge-derived isocyanoterpene (ICT) secondary metabolites are striking for their isonitrile functional groups—rarely found in natural products—and their potent antiplasmodial activity, which appear correlated (1). This family includes a range of complex polycyclic architectures **(Figure 1-1.a)** that has inspired our group to develop a unified strategy toward seemingly topologically disparate ICT diterpenes. Among them, 7,20-diisocyanoadociane (DICA, **1**) stands out as arguably the most stereochemically complex (1). DICA was described in 1976, (2)but it wasn't until studies by König and colleagues two decades later (3)that DICA's potent antiplasmodial activity was revealed. The compelling combination of its structure and activity originated our interest in DICA and, by extension, the rest of the ICT family (4-7).

RESULTS

Upon initiation of our efforts in this area, we identified 2-isocyanodecalin fragment **6** **(Figure 1-1.b)** as a common motif in many potent ICT diterpenes and developed a general strategy around retron **5** (8). An intermolecular conjugate addition/enolate trapping strategy was devised with the goal of maximizing convergency and overall efficiency; this approach appeared to offer the flexibility to be generalizable to many ICTs (9). Indeed, this strategy was central to our synthesis of kalihinol B (**2**) (4),

many simplified yet active kalihinol analogues (5), and our previous formal synthesis of DICA (**1**) (6a).

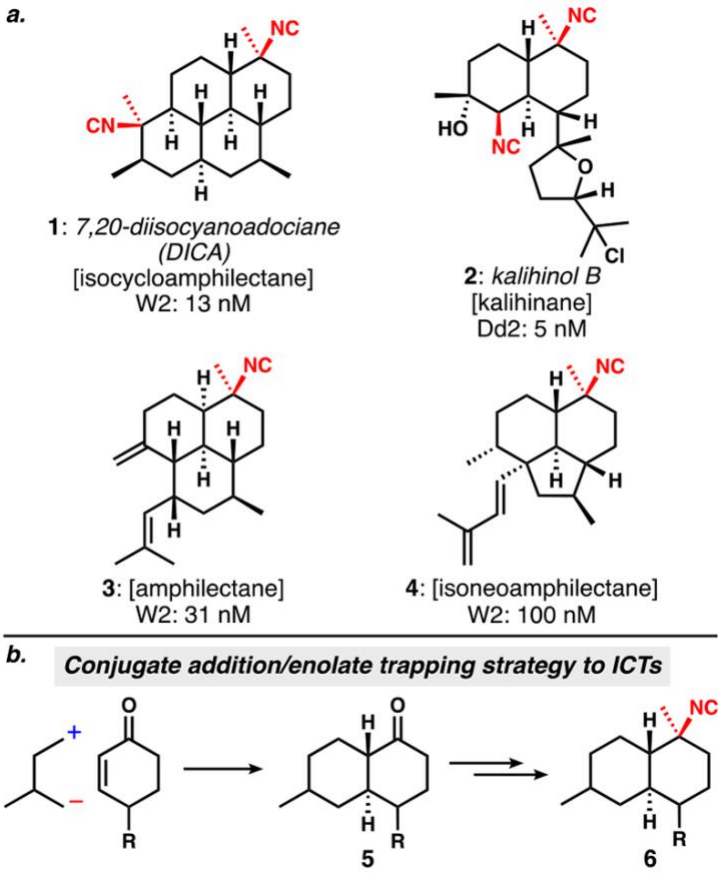
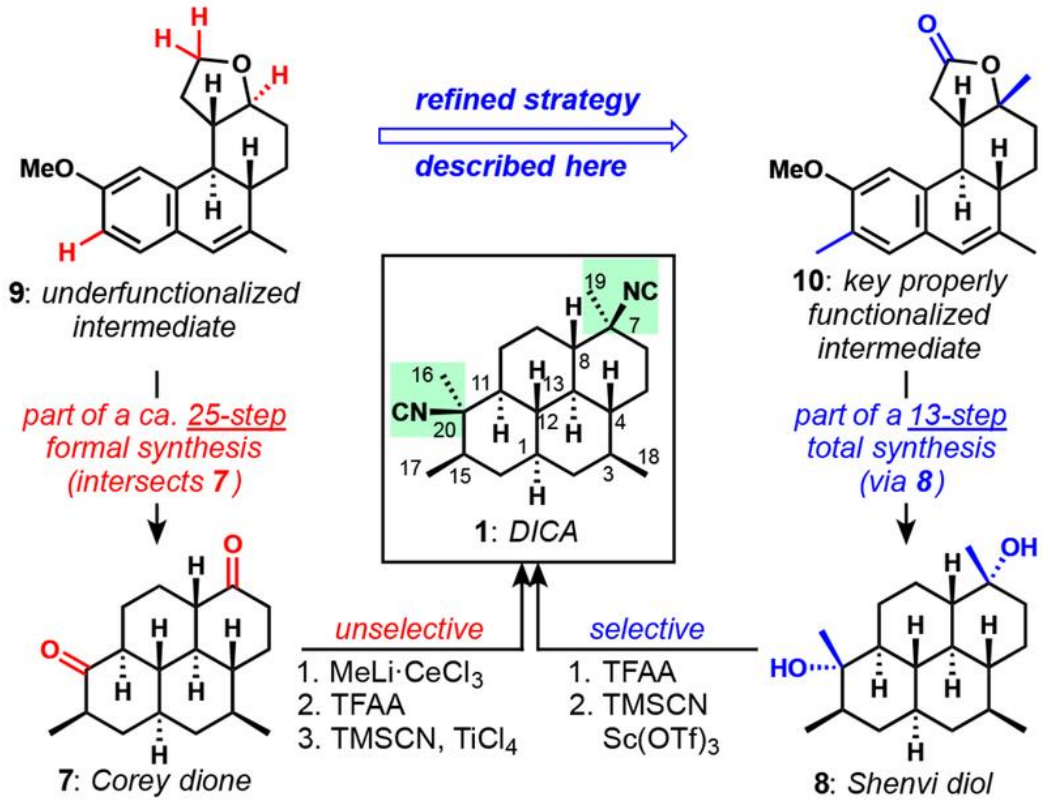


Figure 1-1. a) Select ICTs including 7,20-diisocyanoadociane and their antiplasmodial activities. W2 and Dd2 are drug-resistant strains of the malaria-causing parasite *Plasmodium falciparum*; the concentrations given are IC_{50} values. These data are taken from refs and . **(b)** A general strategy to access various ICT architectures.

In 1987, Corey and Magriotis described the first synthesis of DICA via a logical, Diels–Alder based strategy that generated tetracyclic diketone (Corey dione) **7** in about two dozen steps **(Scheme 1-1)** (10). Double equatorial nucleophilic methylation, activation by trifluoroacetylation and treatment with TMSCN and $TiCl_4$ led to stereochemically

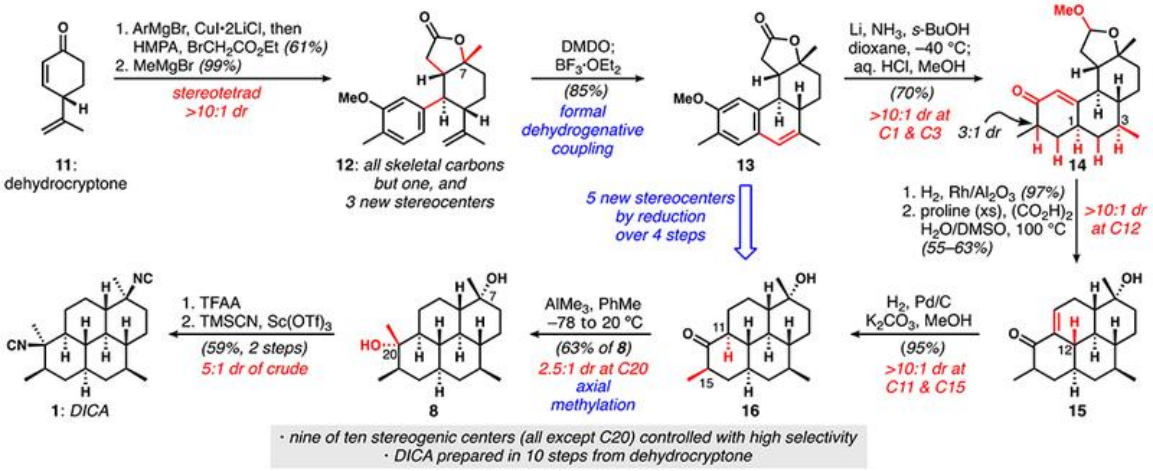
uncontrolled formation of **1** and its three stereoisomers. Several subsequent formal syntheses were based on the formation of the Corey dione, and thus inherited its troublesome endgame. These achievements include a Diels–Alder approach strategically similar to Corey's from the Miyaoka group (11), an oxidative enolate coupling/ring-closing metathesis approach from Robinson and Thomson(12) and our own conjugate addition/enolate trapping strategy(6a). Two total syntheses sought to solve the stereocontrol conundrum at the C7 and C20 isonitrile-bearing carbons: Fairweather and Mander's approach used stereospecific Curtius rearrangements to install tertiary carbinolamines as isonitrile precursors(13), and the Shenvi group's synthesis(14) featured their method for invertive displacement of tertiary trifluoroacetates (TMSCN, $Sc(OTf)_3$) on the bis-trifluoroacetate derived from **8**, which provided a stereoselective counterpoint to the endgame from **7**(15).



Scheme 1-1. Our previous synthesis via intermediate 9 enabled a lengthy formal synthesis by producing the Corey dione (7), for which stereocontrolled conversion to DICA is not straightforward. Our improved synthesis proceeds via 10, permitting a concise and stereocontrolled synthesis. TFAA: trifluoroacetic anhydride.

Our previous formal synthesis was problematic because it required over two dozen steps to arrive at the Corey dione, from which no satisfactory stereocontrolled endgame could be envisaged. Improving on this earlier work required addressing its critical problems: **1)** underfunctionalized early intermediate **9** required late-stage introduction of the C17 and C19 methyl groups; **2)** the THF ring in **9**, introduced as a protecting group for the corresponding 1,4-diol, proved challenging to open, and required a half dozen steps to appropriately elaborate; and **3)** the end-game from the Corey dione to DICA would be unselective. The reconciliation of these problems has led to the total synthesis of DICA

itself described herein, which was completed in fewer than half the number of operations and captures the spirit of our general ICT strategy.



Scheme 1-2. A short synthesis of DICA from dehydrocryptone.
HMPA: hexamethylphosphoramide; DMDO: dimethyldioxirane; DMSO: dimethyl sulfoxide.

Scheme 1-2 reveals how dehydrocryptone **(11)** is converted to DICA in 10 steps (13 steps from its chiral pool precursor perillaldehyde (16)). Conjugate arylation and *in situ* enolate trapping with ethyl bromoacetate provides the three-component coupling product in high yield with complete stereochemical control. Selective nucleophilic methylation of the ketone leads to lactone **12** (17). Equatorial methyl addition to the C7 ketone delivers an axial C–O bond that will later—upon displacement with TMSCN—result in the C7 equatorial isonitrile found in the natural product. Notably, the conversion of **11** to **12** introduces all of the skeletal carbon atoms of DICA except for the C16 methyl group. Sequential epoxidation and Lewis-acid-induced Meinwald rearrangement/Friedel–Crafts cyclodehydration affords tetracyclic styrene **13**(18).

Although dihydronapthalene **13** appears to be only moderately related to DICA, it is readily elaborated to the target's perhydropyrene core in only four carefully orchestrated transformations, with which five stereogenic centers are installed by reduction, each with greater than 10:1 diastereoselectivity. Birch reduction of **13** proceeds with the addition of six molar equivalents of electrons, reducing the styrene, the electron-rich aromatic ring, and—critically—the lactone. After hydrolysis of the initial reduction product, **14** is obtained with high control at C1 and C3 (C15 is easily epimerizable in a subsequent step, and the acetal center is irrelevant). Heterogeneous hydrogenation of the enone is completely selective for the *trans*-ring juncture and provides the ketone that serves as the pronucleophile for an aldol condensation with the acetal carbon. Controlling the reduction to deliver a lactol ensured the correct oxidation state for the subsequent proline-mediated aldol condensation, which directly provided **15**. A second heterogeneous enone hydrogenation—performed in basic methanol—introduces the C11 stereogenic center with high efficiency and selectivity and equilibrates the C15 center such that the methyl group resides equatorially, delivering **16** as a single diastereomer.

Given our focus on brevity to this point, we were compelled to find a way to directly introduce the C16 methyl group in an axial orientation. The Shenvi group reported that their efforts to do so were unsuccessful, opting instead for a multi-step but highly diastereoselective sequence of Peterson olefination and oxymercuration/reduction (14). Presented with this opportunity for discovery, we embarked upon a screen of combinations of different methyl organometallic reagents, Lewis acids additives, solvents, and co-

solvents, while simultaneously studying classical literature on carbonyl additions. In this way, and with guidance from the critical studies of Ashby and co-workers (20), we found that an excess of trimethylaluminum in solvents of low polarity tended to favor the generation of the desired axially methylated adduct **8**.

To date, we have managed selectivities of 2–2.5:1, but the trivial separation of the resulting diastereomers, the avoidance of stoichiometric mercury reagents, and the identical yield to Shenvi's multi-step solution makes this an attractive alternative. Conversion of **8** to **1** by application of Shenvi's invertive isocyanation protocol afforded the natural product exactly as described(14,15). At this stage, we have made over 30 mg of DICA via this sequence, without any particular efforts to scale up the chemistry; indeed, all indications to date suggest that significant quantities of the natural product should be available when needed.

While DICA itself has been tested many times, the activities of its isonitrile-bearing-carbon stereoisomers have not been reported. The general preference for equatorial methylation of cyclohexanones renders **16** particularly well suited to making the unnatural epimer of DICA with both isonitriles equatorial (C20-*epi*-DICA, **17**) (**Figure 1-2).** Treatment of **16** with methylmagnesium chloride was diastereoselective for equatorial methylation, and further processing afforded **17**(20). The potent activity of DICA is largely retained in its stereoisomer **17**, with only a five-fold decrease in potency toward a drug-resistant (Dd2) strain of *Plasmodium falciparum* and a 25-fold decrease against a drug-sensitive (3D7) strain. This outcome is consistent with the model put forth by *Wright, Tilley* and co-workers

pertaining to the antiplasmodial activity of many polycyclic ICTs(21); their modeling analysis suggests that the C7 equatorial isonitrile correlates with significant antiplasmodial activity, and that C20 axial substitution provides a supportive role in slightly improving potency. The model appears to embrace either electron-poor or alkyl groups axial at C20, which is further supported by **17**'s only modest loss in activity. Given this outcome, evaluation of the antiplasmodial activity of the other C7/C20 diastereomers of DICA is warranted—especially given the pseudo-symmetry of these compounds—and will be reported in due course.

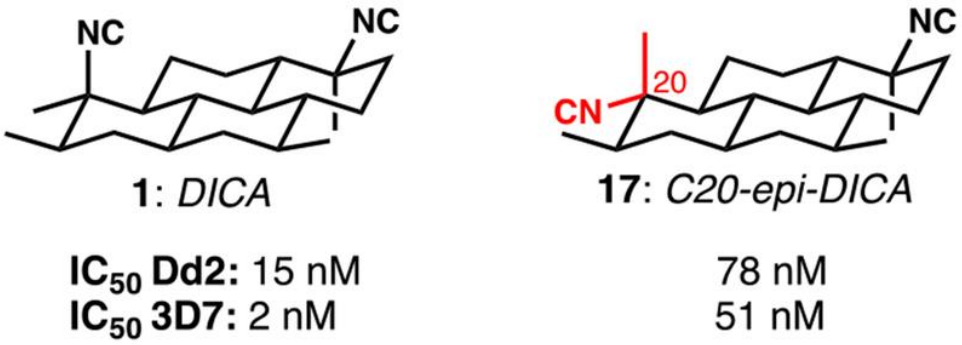


Figure 1-2. The antiplasmodial activity of DICA and C20-epi-DICA.

DISCUSSION

In conclusion, we have developed a particularly concise synthesis of DICA—based on our general strategy—that features a rapid buildup of much of its carbon scaffold, followed by efficient introduction of key stereogenic centers by reduction. We have thereby reduced the problem posed by DICA to a sequence of ten chemical steps from simple, known starting materials, obviating all of the unproductive steps that plagued our previous

effort. This synthesis further showcases the power of the conjugate addition/enolate trapping strategy as a valuable means of accessing ICT natural products. We have capitalized on this short, stereoselective synthesis to produce C20-*epi*-DICA, which was tested for antiplasmodial activity for the first time, the results of which support the current working hypothesis on the activity of polycyclic ICTs.

Our achievement provides a reminder that efficient syntheses need not necessarily rely on the development of new chemical methods. Indeed, with the wealth of chemical tools currently at our disposal, strategic advances will continue to be major contributors to efficiency.

MATERIALS AND METHODS

Plasmodium falciparum Culture

Parasites were cultured at 2.5% hematocrit (Type O^+) in RPMI 1640 w/L-Glutamine, w/o Phenol Red (Gibco), 0.043 mg/mL Gentamicin (Gibco), 0.014 mg/mL Hypoxanthine (Acros), 38.5 mM HEPES (Promega), 0.18% sodium bicarbonate (Gibco), 0.2% glucose (MP Biomedical), 2.6 mM NaOH (Sigma), 0.20% Albumax (Gibco), 5% human serum as described previously.4 Parasites were maintained at 37 °C under an atmosphere of 3% CO_2, 1% O_2, and 96% N_2

SYBR Green Fluorescent Based Drug Assay

The anti-malarial activity of individual compounds was evaluated in-vitro against *P. falciparum* 3D7 (MRA-102, drug sensitive) and Dd2 (MRA-156, multidrug resistant) strains (ATCC® Manassas, VA) via SYBR green-I based fluorescence assay. Parasite growth rate and stages of development were determined via microscopy in Giemsa stained smears of the cultures. Serial dilutions of the compounds were prepared in 96-well plates (Corning, Costar 3904) in 40 µL complete media followed by the introduction of 80 µL asynchronous culture of infected erythrocytes with 1–1.5% parasitemia and 2.5% hematocrit added to each well (120 µL-final volume). Eight wells were treated as positive control (with parasite, without drug) and 8 wells as negative controls (pure media with 2.5% hematocrit with neither parasite nor drug). Plates were incubated in a modular incubation chamber (Billups-Rothenberg, Del Mar, CA) maintained at 37 °C for 72 hr. in a low oxygen environment (96% N_2, 3% CO_2, 1% O_2). After 72 hr., plates were removed from incubation and stored at –80˚C for 24 hr. Plates were then thawed followed by incubation with 120 µL lytic buffer containing SYBR Green 1X for 1–6 h at 37˚C in darkness. Plates were read with a Molecular Devices SpectraMAX Gemini EM at Ex. 495 nm, Em. 525 nm. Assessment of anti-malarial activity of compounds was made on the basis of fifty percent inhibitory concentration values (IC_{50}) determined by DNA content of the parasite (SigmaPlot 10 (Systat)).

ACKNOWLEDGEMENTS

Early studies were supported by the National Science Foundation (CHE-1564340). This work is currently funded by the National Institutes of Health (AI-138139).

REFERENCES

1. Schnermann MJ, Shenvi RA, Nat. Prod. Rep 2015, 32, 543–577. [PMC free article] [PubMed] [Google Scholar].

2. Baker JT, Wells RJ, J. Am. Chem. Soc 1976, 98, 4010–4012. [Google Scholar].

3. (a) Angerhofer CK, Pezzuto JM, König GM, Wright AM, Sticher O, J. Nat. Prod 1992, 55, 1787–1789; [PubMed] [Google Scholar](b) König GM, Wright AD, Angerhofer CK, J. Org. Chem 1996, 61, 3259–3267; [Google Scholar]For a review, see:(c) Wright AD, König GM, Angerhofer CK, Greenidge P, Linden A, Desqueyroux-Faúndez R, J. Nat. Prod 1996, 59, 710–716. [PubMed] [Google Scholar].

4. Kalihinol B: Daub ME, Prudhomme J, Le Roch K, Vanderwal CD, J. Am. Chem. Soc 2015, 137, 4912–4915. [PMC free article] [PubMed] [Google Scholar]

5. Kalihinol analogues:Daub ME, Prudhomme J, Ben Mamoun C, Le Roch KG, Vanderwal CD, ACS Med. Chem. Lett 2017, 8, 355–360. [PMC free article] [PubMed] [Google Scholar]

6. DICA:(a) Roosen PC, Vanderwal CD, Angew. Chem. Int. Ed 2016, 55, 7180–7183 [PubMed] [Google Scholar]Angew. Chem 2016, 128, 7296–7299; [Google Scholar](b) Earlier routes were unsuccessful: Roosen PC, Vanderwal CD, Org. Lett 2014, 16, 4368–4371. [PubMed] [Google Scholar]

7. White AM, Dao K, Vrubliauskas D, Könst ZA, Pierens GK, Mándi A, Andrews KT, Skinner-Adams TS, Clarke ME, Narbutas P, Sim DC-M, Cheney KL, Kurtán T, Garson MJ, Vanderwal CD, J. Org. Chem 2017, 82, 13313–13323. [PubMed] [Google Scholar]

8. Such substructures played intermediary roles in other ICT syntheses, most notably in the Wood and Miyaoka groups' kalihinol work, wherein they were generated by intramolecular Diels–Alder reactions. See(a) White RD, Keaney GF, Slown CD, Wood JL, Org. Lett 2004, 6, 1123–1126. [PubMed] [Google Scholar](b) Miyaoka H, Shida H, Yamada N, Mitome H, Yamada Y, Tetrahedron Lett. 2002, 43, 2227–2230. [Google Scholar](c) Miyaoka H, Abe Y, Sekiya N, Mitome H, Kawashima E, Chem. Commun 2012, 48, 901–903. [PubMed] [Google Scholar](d) Miyaoka H, Abe Y, Kawashima E, Chem. Pharm. Bull 2012, 60, 1224–1226. [PubMed] [Google Scholar]

9. For a review of recurring themes in ICT synthesis strategies, see:Daub ME, Roosen PC, Vanderwal CD, J. Org. Chem 2017, 82, 4533–4541. [PubMed] [Google Scholar]

10. Corey EJ, Magriotis PA, J. Am. Chem. Soc 1987, 109, 287–288. [Google Scholar]

11. Miyaoka H, Okubo Y, Muroi M, Mitome H, Kawashima E, Chem. Lett 2011, 40, 246–247. [Google Scholar]

12. Robinson ER, Thomson RJ, J. Am. Chem. Soc 2018, 140, 1956–1965. [PubMed] [Google Scholar]

13. Fairweather KA, Mander LNA, Org. Lett 2006, 8, 3395–3398. [PubMed] [Google Scholar]

14. Lu H-H, Pronin SV, Antonova-Koch Y, Meister S, Winzeler EA, Shenvi RA, J. Am. Chem. Soc 2016, 138, 7268–7271. [PMC free article] [PubMed] [Google Scholar]

15. (a) Pronin SV, Shenvi RA, J. Am. Chem. Soc 2012, 134, 19604–19606. [PubMed] [Google Scholar](b) Pronin SV, Reiher CA, Shenvi RA, Nature 2013, 501, 195–199. [PubMed] [Google Scholar]

16. We use a slightly modified version of the procedure of Razdan and co-workers to make dehydrocryptone, in which the three-step sequence is performed with a single purification by distillation of 11:Siegel C, Gordon PM, Uliss DB, Handrick GR, Dalzell HC, Razdan RK, J. Org. Chem 1991, 56, 6865–6872. [Google Scholar]See ref. 6a for details.

17. The two reactions can also be combined itno a one-pot four-component coupling process, albeit with diminished yields (ca. 30%).

18. (a) Meinwald J, Labana SS, Chadha MS, J. Am. Chem. Soc 1963, 85, 582–585. [Google Scholar]For a recent application, see:(b) Liao X, Stanley LM, Hartwig JF, J. Am. Chem. Soc 2011, 133, 2088–2091. [PMC free article] [PubMed] [Google Scholar]

19. (a) Ashby EC, Yu SH, Roling PV, J. Org. Chem 1972, 37, 1918–1925. [Google Scholar](b) Laemmle J, Ashby EC, Roling PV, J. Org. Chem 1973, 38, 2526–2534. [Google Scholar]

20. See supporting information for details.

21. Wright AD, Wang H, Gurrath M, König GM, M. G; Kocak G, Neumann G, Loria P, Foley M, Tilley L, J. Med. Chem 2001, 44, 873–885. [PubMed] [Google Scholar]

CHAPTER 2

MED6-189: An Effective Antimalarial Kalihinol Analogue Targets *Plasmodium falciparum* Apicoplast and Membrane Biogenesis

Chahine Z[1], Abel S[1], Hollin T[1], Barnes GL[3], Chung JH[3], Daub ME[3], Renard I[2], Choi JY[2], Vydyam P[2], Pal A[2], Alba-Argomaniz M[9-11], Banks CAS[5], Kirkwood J[4], Saraf A[5$], Camino I[8], Castaneda P[8], Cuevas MC[8], De Mercado-Arnanz J[8], Fernandez-Alvaro E[8], Garcia-Perez A[8], Ibarz N[8], Viera-Morilla S[8], Prudhomme J[1], Joyner CJ[9-11], Bei AK[6], Florens L[5], Ben Mamoun C[2]*, Vanderwal CD[3,7]* and Le Roch KG[1]*.

[1]Department of Molecular, Cell and Systems Biology, University of California Riverside, CA, USA
[2]Department of Internal Medicine, Section of Infectious Diseases, Yale School of Medicine, New Haven, Connecticut, USA
[3]Department of Chemistry, University of California, Irvine, California, 92617, USA
[4]Metabolomics Core Facility, University of California, Riverside, CA 92521, USA
[5]Stowers Institute for Medical Research, Kansas City, MO 64110, USA
[6]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, Connecticut, USA
[7]Department of Pharmaceutical Sciences, University of California, Irvine, California 92697, USA
[8]GSK, C/ Severo Ochoa, 2 PTM, 28760 Tres Cantos (Madrid), Spain
[9]Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia,
Athens, GA, USA
[10]Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA, USA
[11]Center for Vaccines and Immunology, University of Georgia, Athens, GA, USA
[$]Current address: Shankel Structural Biology Center, The University of Kansas, Lawrence, KS,
USA

*Correspondence: Karine Le Roch (karine.leroch@ucr.edu), Choukri Ben Mamoun (choukri.benmamoun@yale.edu) and Christopher Vanderwal (cdv@uci.edu).

A version of this chapter has been published in *Science, 2023*.

PREFACE

The unicellular pathogen *Plasmodium falciparum*, responsible for malaria, poses a significant global health threat. Regrettably, *P. falciparum* has developed formidable mechanisms to resist conventional treatments, including the ability to develop resistance to many administered drugs. Possessing a uniquely high A/T rich genome and enigmatic mechanisms of genome regulation, the pathogen has adapted impressive mechanisms that evade all our currently approved drug therapies. Thus, there is an urgent need for innovative drug candidates possessing the complexity sufficient to impede the pathogen's evasion strategies. In this study, we employ a comprehensive systems-based approach to elucidate the mode of action of Kalihinols, a novel compound derived from the marine sponge, *Acanthella cavernosa*. Systems biology encompasses a multifaced vantage point of study. In doing so, we can identify how individual avenues of research provide insight into the cell's structure, and behavioral responses to our tested compound in an objective and global scale. Here, I applied major omics-based approaches that serve as important elements for a more rounded and impartial understanding of the molecular pathways affected by the kalihinols. I performed all molecular, *in vitro* assays on *P. falciparum* as well as nearly all the statistical analysis for this project. I was also tasked with writing the manuscript which is presented in this chapter. Collectively, the efforts made in work provides greater insight on this promising candidate in our quest to eliminate the spread of malaria.

ABSTRACT

We report the discovery of MED6-189, a new analogue of the kalihinol family of isocyanoterpene natural products, effective against drug-sensitive and -resistant *P. falciparum* strains, blocking both asexual replication and sexual differentiation. *In vivo* studies using a humanized mouse model of malaria confirms strong efficacy of the compound in animals with no apparent hemolytic activity or toxicity. Complementary chemical, molecular and genomics analyses revealed that MED6-189 targets the parasite apicoplast and acts by inhibiting lipid biogenesis and cellular trafficking. Genetic analyses revealed that a mutation in *PfSec13*, which encodes a component of the parasite secretory machinery, reduced susceptibility to the drug. The high potency of MED6-189, excellent therapeutic profile, and unique mode of action make it an excellent addition to the antimalarial drug pipeline.

INTRODUCTION

In 2022, an estimated 247 million clinical cases and 619,000 global deaths were due to malaria worldwide, most of which were caused by *Plasmodium falciparum* (*1*). Because of widespread resistance to commonly used antimalarials, artemisinin-based combination therapies (ACTs) represent the last resort in the antimalarial armamentarium for management of drug resistant malaria (*2*, *3*). Unfortunately, there have been reports of partial resistance to ACTs, particularly in the Greater Mekong region of Southeast Asia and most recently, emerging cases in parts of Africa (*4*-*6*). WHO reports describe resistance to ACT partner drugs as well as clinical treatment failure documented in Southeast Asia.

There are few antimalarials with efficacy comparable to that of ACTs, and current vaccines have only limited protective efficacy (*113*, *114*). There is therefore an urgent need to identify novel therapeutics to combat the ever-growing threat of drug resistance and to suppress the spread of the disease. Ideally, such molecules should target novel pathways not previously targeted by approved antimalarials or those in clinical development.

The isocyanoterpene (ICT) family of sponge-derived natural products encompasses a large number of biosynthetically related, isonitrile-, isothiocyanate-, isocyanate-, and formamide-containing diterpenoids, many of which have been shown to have potent antibacterial, antifungal, and antimalarial activity (*115*-*117*). The presence of the isonitrile functional group is critical for potent antimalarial activity. Among the ICTs, the kalihinol subfamily counts some of the most active compounds, with kalihinols A (*115, 116, 118*) and B (Fig. 1A, left) (*118*-*121*) exhibiting potent activity against drug-sensitive and -resistant *P. falciparum* isolates (*120, 121*). Our labs previously completed the first synthesis and testing of kalihinol B. We demonstrated that its activity was similar to that of kalihinol A and learned that the oxygen heterocycle motif was likely not critical for activity. We therefore dramatically truncated this group, which ultimately led to the discovery of analogue MED6-189 **(Fig. 2-1A, right),** whose simplified structure allowed for an easier synthesis.

Efforts to understand the mechanism of action of compounds in the ICT family go back more than two decades (*122, 123*). Wright and co-workers showed that some ICTs can

bind free heme and inhibit the formation of its crystalline form, -hematin (*123*), suggesting that inhibition of heme detoxification could be an important cause of ICT antimalarial activity (*123*). This study, however, focused on a specific subgroup of ICTs, namely the tetracyclic cycloamphilectanes and tricyclic amphilectanes; whereas the structurally distinct kalihinols, which inspired the development of MED6-189, were not investigated. More recently, another study showed that compounds in these same two structural classes inhibit both intraerythrocytic replication and liver stage development of malaria parasites (*124*), indicating that heme detoxification might not be the primary mechanism by which the larger group of ICTs exert their antiplasmodial activity.

We show that MED6-189, one of our most accessible synthetic analogues of kalihinols A and B, inhibits drug-sensitive and -resistant *P. falciparum* strains *in vitro*, blocks parasite sexual differentiation, and eliminates infection in a humanized mouse model of *P. falciparum* malaria. Closely related chemical probe molecules that retain antimalarial activity were incorporated into a broad systems biology strategy that integrates various multi-omics platforms, leading to an unbiased understanding of both the drug mechanism of action and possible routes of resistance. Our findings demonstrate that MED6-189 (and by extension, kalihinols and presumably other ICTs) disrupts the apicoplast, an organelle essential for the synthesis of parasite fatty acids and isoprenoids leading to *delayed death* in the pathogen. Additional metabolomic and proteomic analyses show that the compound also interferes with lipid metabolism and trafficking during the parasite asexual stage. Moreover, we found that mutations in the *PfSec13* gene (*PF3D7_1230700*), an important

component of the protein secretory machinery, is associated with altered susceptibility to the drug. The compound was also found to be potent against other zoonotic/human *Plasmodium* parasites, *P. knowlesi* and *P. cynomolgi*. Overall, our results show that MED6-189 has an efficient and sophisticated mode of action to which it has proven challenging for the pathogen to develop resistance. These combined characteristics make MED6-189 an optimal candidate for use in combination with fast-acting compounds or as a preventative, with reduced susceptibility for emergence of drug resistance.

RESULTS

Kalihinol Analogue Found to Impair Apicoplast Function

Generation of synthetic kalihinol analogues

Because large-scale biological production of ICT natural products is not sustainable, we chemically synthesized kalihinol B **(Fig. 2-1A, left),** a tetrahydrofuran isomer of the known, potent ICT kalihinol A (a tetrahydropyran). We also demonstrated its activity against *P. falciparum* Dd2 isolate with IC_{50} values of 4.6nM (*120*). Because this compound was challenging to produce on a sufficiently large scale for in-depth mechanism of action studies, several less complex synthetic analogues of kalihinols A and B were generated and their antiparasitic activity evaluated *in vitro* **(See Supporting Information 2-S1 and Fig. 2-S1)** (*120*). Among these compounds, MED6-189 emerged as a highly potent, more accessible drug with IC_{50} values below 50 nM against several *P. falciparum* strains **(Fig. 2-1A, right)** (*120*). Through a robust sequence of ten chemical steps, the

synthesis of MED6-189 was scaled to support the in-depth studies detailed in this work **(See Supporting Information 2-S1 and Fig. 2-S1)** (*125, 126*).

MED6-189 affects *P. falciparum* intraerythrocytic life cycle

To assess the effectiveness of MED6-189, **(Fig. 2-1A, right)** on the intraerythrocytic development (IED) of *P. falciparum*, we performed growth inhibition assays on several isolates with established sensitivity or resistance to various classes of antimalarials **(Fig. 2-1B)**. MED6-189 showed excellent activity against both drug-sensitive strains (3D7 and NF54), with IC_{50} values of 14 ± 2nM and 28 ± 5nM, respectively, and pyrimethamine- and chloroquine-resistant strains (HB3, Dd2 and W2) with IC_{50} values of 23 ± 2nM, 47 ± 7nM, and 27 ± 5nM, respectively (Sigmoidal, 4PL, X is concentration, n≥3, nonlinear regression, CI:95%) **(Fig. 2-1B, table 2-S1A).** To identify the specific stage of the *P. falciparum* asexual intraerythrocytic cycle (IEC) that was predominantly inhibited by the compound, we carried out phenotypic analyses on synchronized cultures that were incubated with either MED6-189 or a control vehicle (DMSO) at different time intervals following synchronization **(Fig. 2-1C).** Whereas no significant impact of the compound on parasite development could be observed during the first IEC, significant changes including cell cycle arrest and cellular stress were evident during the trophozoite stage of subsequent IEC **(Fig. 2-1D).** The treated parasites were unable to fully undergo schizogony or re-invasion ($p< 0.05$, n=3, 2-way ANOVA, Tukey t-test) **(Fig. 2-1E).** This phenotype is reminiscent of that of compounds that target the apicoplast, such as fosmidomycin (*127-129*). To assess whether MED6-189 could inhibit the parasite's ability to undergo sexual

differentiation, a critical step in malaria transmission, we examined *P. falciparum* gametocyte development in the absence or presence of the compound. Incubation with 300 nM of MED6-189 inhibited early gametocyte development (Stage III -Days 4 to 6) **(Fig. 2-1F)** by ~62% compared to controls ($p<0.05$, n=3, 2-way ANOVA). Furthermore, when exposed to high doses of MED6-189 (1M) during the early stages of gametocyte development, parasite growth was abolished, and after 48h of drug treatment, no quantifiable gametocytes were observed **(Fig. 2-S2A).** In contrast, the compound had little to no effect on mature-stage gametocytes (Stage V) **(Fig. 2-1F, 2-S2A, table 2-S1B).** These results suggest a delayed inhibitory activity of MED6-189 on parasite development.

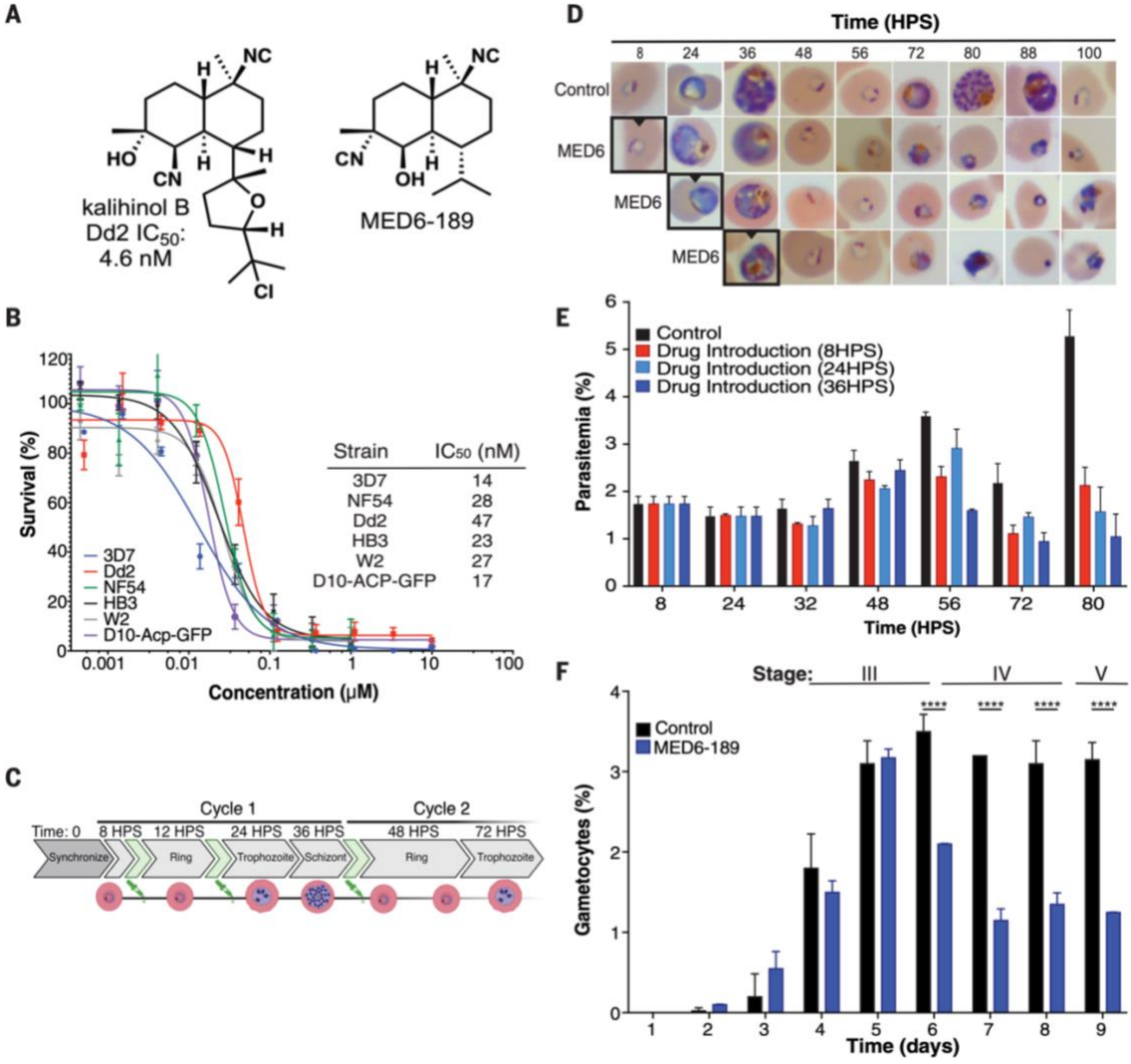

A
kalihinol B
Dd2 IC50:
4.6 nM
MED6-189
B
Survival (%)
Concentration (μM)
3D7
Dd2
NF54
HB3
W2
D10-Acp-GFP
Strain IC50 (nM)
3D7 14
NF54 28
Dd2 47
HB3 23
W2 27
D10-ACP-GFP 17
C
Cycle 1
Cycle 2
Time: 0
8 HPS
12 HPS
24 HPS
36 HPS
48 HPS
72 HPS
Synchronize
Ring
Trophozoite
Schizont
Ring
Trophozoite
D
Time (HPS)
Control
MED6
E
Control
Drug Introduction (8HPS)
Drug Introduction (24HPS)
Drug Introduction (36HPS)
Parasitemia (%)
Time (HPS)
F
Stage:
Control
MED6-189
Gametocytes (%)
Time (days)

Figure 2-1. Effect of MED6-189 on *P. falciparum* intraerythrocytic development A. Chemical structures of the natural product kalihinol B (left) and its analog MED6-189 (right) (*120, 121*). **B.** SYBR Green-based dose response assays were conducted on early-ring stage parasites (6 hours post invasion). The parasites were exposed to serial dilutions of MED6-189 for 72 hours, after which parasite growth was assessed. 3D7 WT (blue), NF54 (green) and drug-resistant strains Dd2 (red), HB3 (black), W2 (grey) and D10-Acp-GFP (purple) lines (Sigmoidal, 4PL, X is concentration, $n \geq 3$, nonlinear regression, CI:95%). **C.** Schematic diagram of the development of *P. falciparum* following two consecutive erythrocytic cycles. The time points at which MED6-189 was introduced are depicted in green. **D.** Giemsa-stained images of synchronized 3D7 parasites that were incubated with either DMSO or MED6-189 (at its IC_{80} concentration). The images depict various developmental stages of the parasite's intraerythrocytic life cycle. Bordered images represent timepoints when drug is first introduced ($n=3$, $p<0.05$) **E.** (%) parasitemia following exposure of 3D7 parasites to either DMSO (control) or MED6-189 at various stages of the parasite life cycle within erythrocytes ($p< 0.05$, $n=3$, 2-way ANOVA, Tukey t-test). **F.** Inhibition of *P. falciparum* gametocyte development following MED6-189 treatment (300 nM) (blue) during early gametocytogenesis compared to the control (black) ($p<0.05$, $n=3$, 2-way ANOVA).

MED6-189 targets apicoplast metabolism

To gain deeper insights into the cellular metabolic target of MED6-189, we synthesized a fluorescently labeled kalihinol analogue, MED6-131, featuring a JF549 fluorophore attachment **(see MED6-131 structure in Fig. 2-S2B)** and assessed its activity and cellular localization. Cell growth assays confirmed its activity against the *P. falciparum* D10-ACP-GFP strain (*130*), exhibiting an IC_{50} of 17 ± 0.9 nM. Fluorescence microscopy analyses were conducted using D10-Acp-GFP transgenic parasites, which express a fusion of the Acyl Carrier Protein (Acp) with GFP in the parasite apicoplast (*131*). These analyses revealed the co-localization of MED6-131 with Acp-GFP during the trophozoite and schizont stages of the parasite life cycle **(Fig. 2-2A).** These findings affirm localization of the compound within the apicoplast.

To explore whether MED6-189 targets the apicoplast non-mevalonate MEP/DOXP pathway, which is essential for parasite survival (*129*), we initiated drug-drug interaction studies. Our data showed that MED6-189 and fosmidomycin, a known inhibitor of the MEP/DOXP pathway, exhibited antagonistic effects, as indicated by an estimated Fractional Inhibitory Concentration (FIC) index of 2.4 ± 0.36 **(Fig. 2-2B-C).** Interactions of MED6-189 with drugs unrelated to apicoplast functions, namely chloroquine, (inhibits heme polymerization in parasite food vacuole) and atovaquone (a mitochondrial electron transport inhibitor), were additive, with FIC scores of 1.71 ± 0.4 for chloroquine and 0.91 ± 0.2 for atovaquone **(Fig. 2-2B-C, table 2-S2A and B)** *(132, 133).* The killing rate (k) of MED6-189- and fosmidomycin-treated parasites were comparable at 48- and 72-hours

post-treatment. Interestingly, the k value for MED6-189 was significantly higher during the third IEC **(Table 2-S2c),** consistent with the morphological alterations observed following treatment with this compound.

To verify targeting of the apicoplast by MED6-189, we investigated the effect of isopentenyl pyrophosphate (*29*) supplementation on MED6-189 antimalarial activity. IPP is a product of the MEP/DOXP pathway, and its supplementation reduces susceptibility to MEP inhibitors. Synchronized parasites were treated with either vehicle control (DMSO), MED6-189, fosmidomycin, or other known apicoplast targeting compounds, such as clindamycin and azithromycin that target protein synthesis in the apicoplast, along with atovaquone, a drug not known to affect the apicoplast, (*134-136*). Parasites were drugged at IC_{80} concentrations for one intraerythrocytic cycle prior to IPP (200 µM) supplementation. In contrast to non-supplemented samples, parasites supplemented with IPP survived the second IEC and initiated a third cycle in the presence of MED6-189 or fosmidomycin **(Fig. 2-2C).** Our results show that MED6-189 inhibition can only be rescued by IPP for one additional cycle. Together, these findings indicate that, in addition to inhibition of the apicoplast, MED6-189 likely hinders one or more additional targets within the parasite.

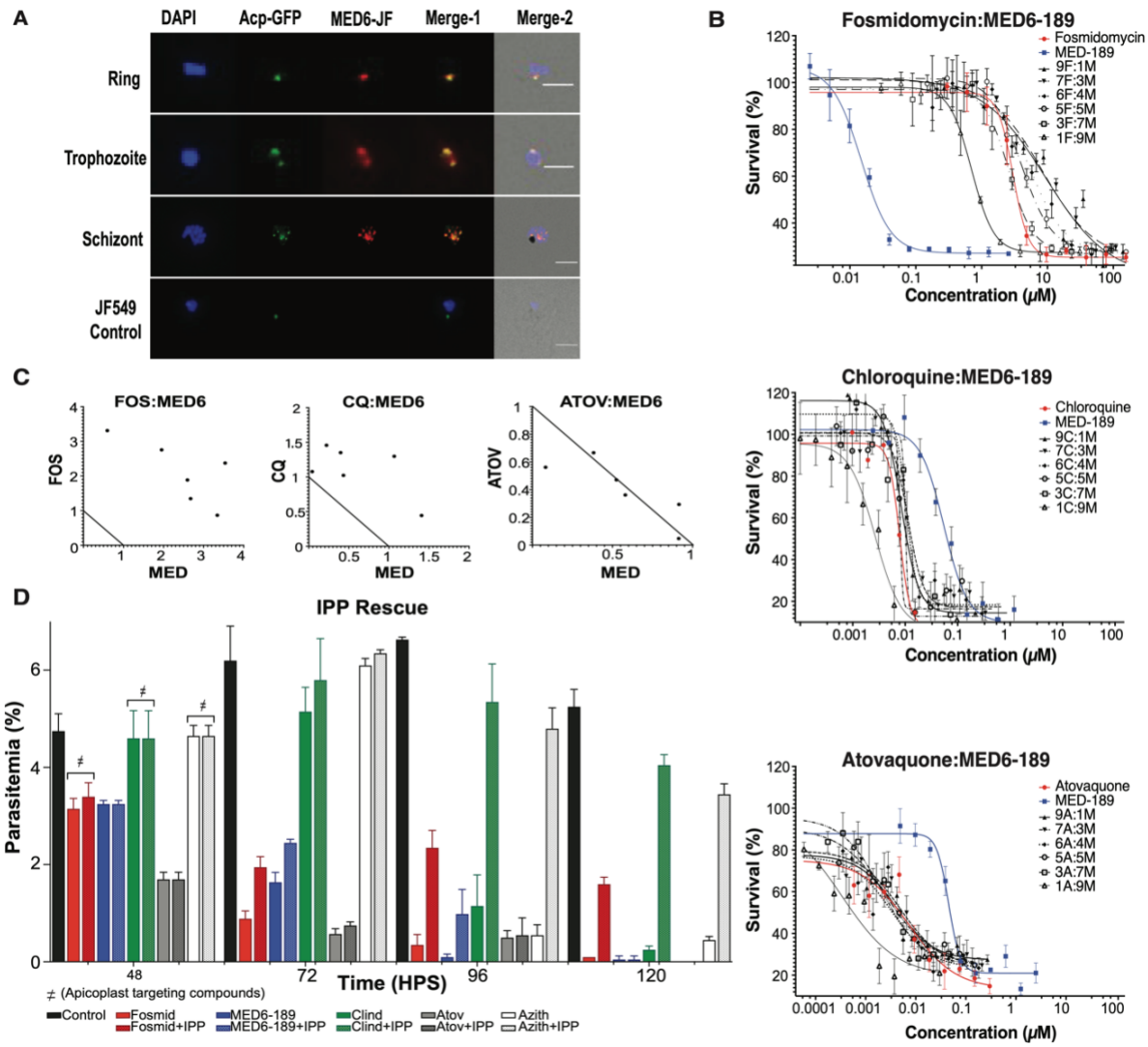

A
DAPI
Acp-GFP
MED6-JF
Merge-1
Merge-2
Ring
Trophozoite
Schizont
JF549 Control
B
Fosmidomycin:MED6-189
Fosmidomycin
MED-189
9F:1M
7F:3M
6F:4M
5F:5M
3F:7M
1F:9M
Survival (%)
Concentration (μM)
C
FOS:MED6
CQ:MED6
ATOV:MED6
FOS
CQ
ATOV
MED
Chloroquine:MED6-189
Chloroquine
MED-189
9C:1M
7C:3M
6C:4M
5C:5M
3C:7M
1C:9M
D
IPP Rescue
Parasitemia (%)
Time (HPS)
≠ (Apicoplast targeting compounds)
Control
Fosmid
Fosmid+IPP
MED6-189
MED6-189+IPP
Clind
Clind+IPP
Atov
Atov+IPP
Azith
Azith+IPP
Atovaquone:MED6-189
Atovaquone
MED-189
9A:1M
7A:3M
6A:4M
5A:5M
3A:7M
1A:9M

Figure 2-2. MED6-189 localization and activity in combination with other antimalarials. A. Cellular localization of MED6-131 (Red) in D10-Acp-GFP *P. falciparum* transgenic parasites. Nuclei are stained with DAPI (Blue). Overlap between ACP-GFP (Green) and MED6-131 can be seen during the trophozoite and schizont stages of the cell cycle. **B.** Dose-dependent interactions between MED6-189 (Blue) and various antimalarials with known mechanisms of action (Red). Logarithmic growth of parasites (Y-axis) is plotted as a function of drug concentrations for MED6-189 (M), fosmidomycin (F), chloroquine (C) or atovaquone (A) (X-axis). The regression line represents a nonlinear regression (Variable slope with four parameters), with significant differences considered if $p<0.05$. Activity correlations between each compound and MED6-189 were analyzed using Pearson correlation (r) using GraphPad Prism 9 (GraphPad Software, Inc.), n=3 (See table S2). **C.** Normalized isobolograms demonstrating drug interaction (synergism, indifference, or antagonism) according to the Loewe additivity model between apicoplast inhibitors and MED6-189. Isobologram curves are expected to be parallel to the diagonal for additive drug pairs, concave for synergistic drug pairs, and convex for antagonistic drug pairs. **D.** Rescue of 3D7 parasites exposed to DMSO (black), fosmidomycin (Red), MED6-189 (Blue), along with several other known antimalarials supplemented with IPP 48hrs post-synchronization (dotted). The analysis was performed using a 2-way ANOVA, n=3, with $p<0.05$ significance.

MED6-189 Impacts Secretory and Trafficking Systems

Multi-omics approaches unveil the mechanism of action of MED6-189

To elucidate the cellular and metabolic pathways affected by MED6-189, we employed a comprehensive multi-omics strategy. First, we analyzed transcriptional changes by conducting RNA-seq on *P. falciparum* (3D7 strain) in the absence or presence of MED6-189 at various time points during the parasite IEC. Whereas no significant change in transcript levels were detected between the control and drug-treated samples during the first IEC **(table 2-S3A),** significant changes (log_2 fold-change (FC) >0.5 or <-0.5) were manifest during the late ring and trophozoite stages of the second IEC, consistent with the delayed inhibitory activity of MED6-189 on parasite development **(able 2-S3A, 2-S3B).** Gene Ontology (GO) enrichment analyses identified several classes of transcripts that were significantly downregulated following exposure to MED6-189, including those known to be involved in parasite invasion and egress, such as *SUB1/SUB2* (PF3D7_0507500/ PF3D7_1136900). These findings indicate that the compound induces cell cycle arrest **(Fig. 2-S3A).** Among the upregulated genes, we identified those involved in the isoprenoid catabolic process and apicoplast function, such as G3PAT, autophagy, and stress responses **(Fig. 2-3A, table 2-S3C, Fig. 2-S3B)** (*137*). Although most changes detected in transcripts in the second IEC were likely induced by the drug in a nonspecific manner (cell cycle arrest and stress induction), our results also indicate a potential compensatory mechanism of the pathways affecting isoprenoid catabolic process and the apicoplast.

To assess the effect of MED6-189 on cellular metabolism, highly synchronized parasite cultures were maintained in the absence or presence of the compound and their metabolomic profiles were examined during the second IEC (72 hours post-treatment). A total of 1178 metabolites were detected, of which 40 were significantly affected by MED6-189 treatment **(Fig. 2-3B).** These included two polar metabolites (thiamine pyrophosphate (TPP) and S-adenosyl methionine (SAM)) and 38 lipids, including components of the Kennedy pathway for the synthesis of phosphatidylcholine and phosphatidylethanolamine (PC and PE, respectively) *(138, 139)***(Fig. 2-3B, and 2-S4).** TPP is essential for the activity of several enzymes including the apicoplast-associated pyruvate dehydrogenase (PDH) and vitamin B1 biosynthesis (*140-143*), and SAM is a major precursor for the synthesis of PC from the serine-decarboxylase-phosphoethanolamine-methyltransferase (SDPM) pathway (*138, 144*). This points to MED6-189 affecting apicoplast biogenesis and membrane biosynthesis.

MED6-189 binds to *P. falciparum* proteins involved in membrane trafficking

To identify proteins that interact with MED6-189, a biotinylated derivative with a similarly potent antimalarial activity (IC_{50}= 92nM ± 2.0), MED6-118 **(Fig. 2-S2C),** was synthesized and used in pull-down assays using *P. falciparum* protein extracts. The purified proteins were analyzed through Multidimensional Protein Identification Technology (MudPIT) and quantified using dNSAF spectral (*145-148*) counting **(table 2-S4).** Over 450 proteins were detected in at least two of the five biological replicates **(Fig. 2-3C).** Among them, 30 proteins were either absent in the negative controls or significantly enriched, as

determined by Log_2 (FC) ≥1.5 and Z-statistic ≥1.645, calculated using QPROT (*145*, *148*). These enriched proteins are involved in vesicular trafficking, lipid biogenesis and signaling **(Fig. 2-3C and table 2-S4).** Specifically, three proteins known to be involved in the endoplasmic reticulum (ER) Coat Protein Complex II (COPII) trafficking systems Sec62 (PF3D7_1438100), syntaxin SYN13 (PF3D7_1104100), a member of the Qa-SNARE family, and a conserved protein recently annotated as a putative translocation protein SEC66 (PF3D7_0204200) were found to be significantly enriched in the pull-down experiments from the MED6-118-treated but not the vehicle control samples(*149-153*).

To further investigate protein interactions with MED6-189, we conducted Thermal Proteome Profiling (TPP) (*154*) using whole protein extracts purified from either vehicle or MED6-189 treated parasite cultures (*154*). Following the addition of the ligand, the samples were incubated at temperatures ranging between 37˚C and 73˚C, and soluble proteins were analyzed via liquid chromatography coupled to tandem mass spectrometry **(table 2-S5),** as described previously (*155*). The data were processed using the mineCETSA R package (*155*), from which melting curves were generated for over 800 *P. falciparum* proteins **(table 2-S6 and Supporting Information 2-S2)** (*156, 157*). Melting curves were observed for 14 of the 30 proteins enriched in the MED6-118 affinity purification **(Fig. 2-3D and 2-S5A, table 2-S5).** Notably, SYN13 exhibited the most significant shift in melting temperature **(Fig. 2-3D),** while the other three proteins, namely the apicoplast localized bacterial histone-like protein (HU | PF3D7_0904700), lipocalin

(LCN | PF3D7_0925900), and 6-cysteine protein (P41 | PF3D7_0404900), displayed slight differences between the untreated and MED6-189 treated samples.

Other proteins of interest with substantial thermal shifts on MED6-189 treatment included UIS2 (PF3D7_1464600), a serine/threonine protein phosphatase localized to the parasitophorous vacuole, and PyKII (PF3D7_1037100) an apicoplast pyruvate kinase **(Fig. 2-S5B).** While these proteins were detected in the affinity purification dataset, they did not meet the significance cut-off **(table 2-S4).** Finally, a few proteins were not detected in the MED6-118 pull-downs but exhibited significant stabilization (PF3D7_0724100, PF3D7_1466900), or destabilization in the case of BOP1 (PF3D7_1405800), a protein involved in ribosome biogenesis **(Fig. 2-3E and 2-S5C).**

Overall, all but one (PF3D7_0721100) of the proteins we highlighted for their slight to marked changes in thermal profiles in the presence of MED6-189 **(Fig. 2-3D-E and 2-S5B-C)** did not show any difference in their melting curves in the DMSO and pyrimethamine-treated TPP samples reported by Dziekan *et al. (155)*, which likely indicates these changes in stability are specific to MED6-189 **(Fig. 2-S6 and Table 2-S6).** Collectively, these two complementary proteomics approaches suggest a possible mechanism of action of MED6-189 involving the association with, and potential disruption of the membrane trafficking apparatus between the ER-Golgi or ER-apicoplast systems.

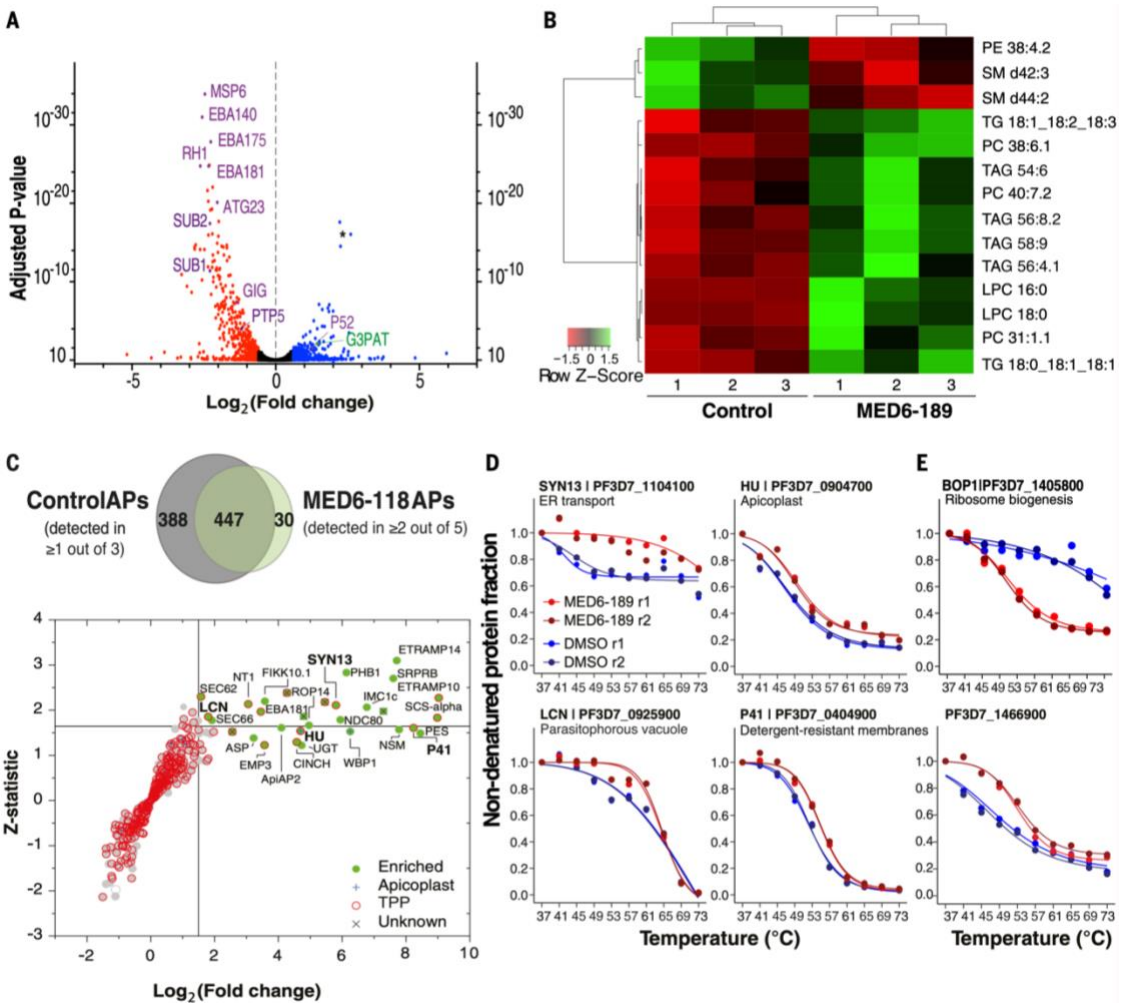

A
MSP6
EBA140
EBA175
RH1
EBA181
ATG23
SUB2
SUB1
GIG
PTP5
P52
G3PAT
Adjusted P-value
Log2(Fold change)
B
PE 38:4.2
SM d42:3
SM d44:2
TG 18:1_18:2_18:3
PC 38:6.1
TAG 54:6
PC 40:7.2
TAG 56:8.2
TAG 58:9
TAG 56:4.1
LPC 16:0
LPC 18:0
PC 31:1.1
TG 18:0_18:1_18:1
Row Z–Score
Control
MED6-189
C
ControlAPs
(detected in ≥1 out of 3)
388
447
30
MED6-118APs
(detected in ≥2 out of 5)
Z-statistic
Log2(Fold change)
Enriched
Apicoplast
TPP
Unknown
D
SYN13 | PF3D7_1104100
ER transport
HU | PF3D7_0904700
Apicoplast
LCN | PF3D7_0925900
Parasitophorous vacuole
P41 | PF3D7_0404900
Detergent-resistant membranes
MED6-189 r1
MED6-189 r2
DMSO r1
DMSO r2
Non-denatured protein fraction
Temperature (°C)
E
BOP1|PF3D7_1405800
Ribosome biogenesis
PF3D7_1466900
Temperature (°C)

Figure 3-3. Omics-based profiling of MED6-189 treated parasites. A. Volcano plot representing transcriptomic changes induced by MED6-189 treatment. A total of 5712 transcripts were identified with an adjusted p-value cut-off of 0.05. Transcripts associated with invasion and stress responses are highlighted in purple and those related to apicoplast function in green (***** represent ncRNAs of unknown function). **B.** Heatmap depicting the regulation of lipid metabolism in response to MED6-189. Metabolites significantly upregulated in response to MED6-189 treatment are shown in green and those downregulated in red. We used a Log_2 transformation of the data for the calculation of q-values (Benjamini-Hochberg adjusted p-values) and p-values using *Welch's t-test* or *ANOVA*. **C.** Protein pulldown assays using biotinylated kalihinol analogue, MED6-118. The significance plot displays all proteins detected in at least two of the five independent MED6-118-based affinity purifications (APs). Scatter plots with gray dots depict QPROT-derived log_2(FC) and Z-statistic values between MED6-118 APs and negative controls (table S4). Significantly enriched proteins with a Log_2 (FC) $\geq$1.5 and a Z score $\geq$1.645 or those not detected in controls are highlighted in green. Proteins for which thermal profiles are available are shown in red (table S5 and Supporting Information S2). Proteins localized to the apicoplast are indicated with a blue cross, while proteins of unknown function are marked with a gray "X". A Venn diagram shows the protein overlap between the MED6-118 APs and controls. **D.** TPP melting-curve analysis of *P. falciparum* lysates treated with MED6-189. The thermal profiles for four *P. falciparum* proteins significantly enriched in the MED6-118-based pull-downs are shown. **E.** Thermal profiles for 2 proteins not detected in MED6-118 pull-down but with significant changes in stability in the presence of MED6-189. For both panels D and E, stabilization is assessed on the relative amount of soluble protein remaining (Y-axis) after thermal challenge (X-axis). Sample replicates are color-coded in shades of red for MED6-189-treated samples and blue for DMSO controls. Differences in melting temperatures (ΔTm) along with arrow trends are reported when available (Table S6).

Selection of MED6-189 refractory parasites

To gain further insights into the cellular machineries impacted by MED6-189, we conducted extended *in vitro* resistance selection experiments (*158-160*). An initial inoculum of 10^8 3D7 parasites were first cloned through serial dilution (*161*) and isolated clones were split into controls or experimental lines and pulsed with DMSO or MED6-189 at 3X IC_{80} or 5XIC_{80} concentrations for 24-48 hours before allowing for parasite recovery. After several failed attempts to select resistance lines we sought an alternative approach. An inoculum of 10^8 cloned control and 10^8 experimental parasites were cultured in the presence of DMSO or MED6-189 at IC_{50} values before gradually escalating the drug dosage **(Fig. 2-4A).** After approximately 36 months of continuous *in vitro* propagation, parasites exhibiting noticeable drug tolerance emerged. Drug inhibition assays of these selected cloned lines revealed a 2 to 4-fold decrease in susceptibility to MED6-189 compared to wild type clones (3D7). The IC_{50} values for the drug-sensitive parental lines were approximately 14± 1 nM, while those for the resistant lines were approximately 60 ± 6 nM (Sigmoidal, 4PL, X is concentration, n=3, nonlinear regression, CI:95%) **(table 2-S1A).** Whole genome sequencing (WGS) identified single-nucleotide polymorphisms (SNPs) or Indels in two coding genes: *PfSec13* (PF3D7_1230700) and *Pfendoplasmin* (PF3D7_1222300) **(table 2-S7)** in the mutant lines, but not in the parental lines. However, endoplasmin showed no change in stability compared to DMSO controls in our TPP experiment **(Fig. 2-S5D).** Although *PfSec13* appeared to be the most promising target, owing to its established role in protein trafficking (*80*, *106*), the protein was detected in our TPP assay in presence of MED6-189 but not in our DMSO replicates, likely due to its low

abundance. To assess whether we could use the DMSO melting curves previously established for SEC13 by Dziekan *et al. (155)*, we calculated the difference in melting temperatures (ΔTm) between our DMSO dataset and Dziekan's for 364 proteins for which Tms were confidently calculated by mineCETSA **(Table 2-S6).** Considering the density distribution of the ΔTms between our dataset and Dziekan's **(Fig. 2-S5E),** a ΔTm greater than +15°C between these two independently acquired DMSO datasets (i.e., such as no denaturation is observed for the DMSO data points over the monitored range of temperatures, like for the MED6-189 data points) has less than 1% chance of occurring (two recorded instances in 364 proteins). Therefore, Dziekan's DMSO dataset can be considered a valid proxy and infers that the presence of MED6-189 stabilizes SEC13 beyond its previously observed melting temperature.

The role of the *PfSec13* gene in susceptibility to MED6-189 was confirmed in *Saccharomyces cerevisiae*. Since the yeast *sec13* gene is essential for viability, we assessed the effect of MED6-189 following overexpression of Sec13p using a multi-copy plasmid or its repression using a tetracycline-regulated (Tet-off) promoter. Sec13p overexpression significantly increased resistance to the drug by 2.4-fold, whereas its repression was accompanied with a ~5.5-fold increase in susceptibility to MED6-189 **(Fig. 2-4B).** Alteration of the expression levels of the endoplasmin in yeast did not result in discernible changes in susceptibility to MED6-189 (Data not shown).

To validate the role of SEC13 in drug resistance in *P. falciparum*, we used the CRISPR-Cas9 gene editing tool. The seven amino acid deletions detected in the *PfSec13* gene in the resistant clones formed a tandem repeat sequence, which could potentially enhance protein structural integrity or be involved in protein-protein interactions *(162)*. These tandem repeat sequences were targeted for deletion in the 3D7 wild type strain (Pf3D7_12_v3: 1261090- 1261110 (+)) **(Fig. 2-4C-D)** using CRISPR-Cas9 **(table 2-S8).** The targeted deletion was verified through Sanger sequencing and WGS **(Fig. 2-4E, table 2-S7)**. Four *PfSec13-mut* transgenic clones were tested for their sensitivity to MED6-189. All *PfSec13-mut* clones exhibited tolerance levels to MED6-189 similar to those of the MED6-189 resistant lines obtained after drug selection (IC_{50} range: 50-73nM, SEM ± 9.0-10nM) compared to the wild-type strain **(Fig. 2-4F, table 2-S1A).** Collectively, these genetic and pharmacological studies in both *P. falciparum* and *S. cerevisiae* underscore the crucial role of SEC13 in susceptibility to MED6-189.

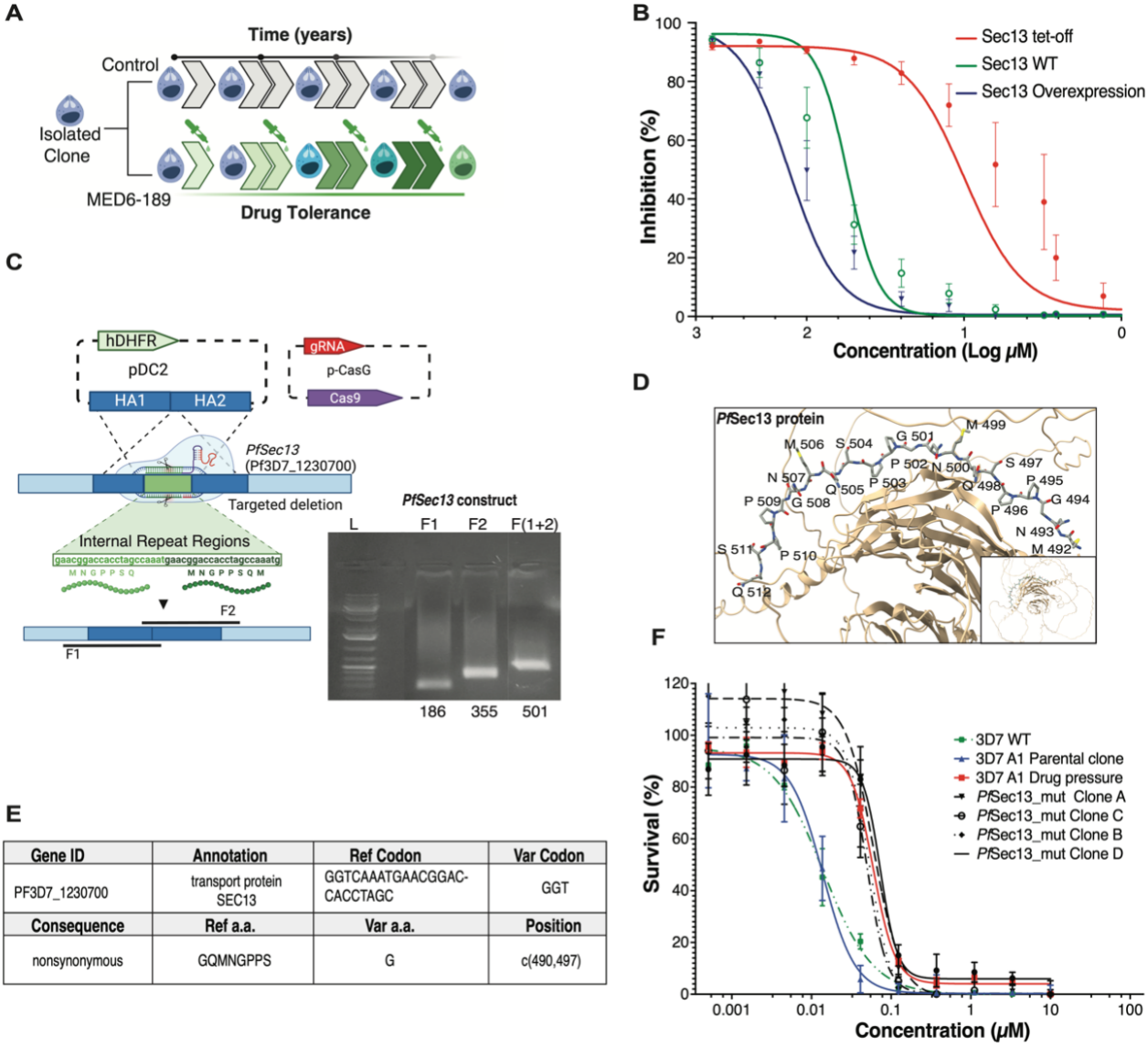

A
Time (years)
Control
Isolated Clone
MED6-189
Drug Tolerance
B
Sec13 tet-off
Sec13 WT
Sec13 Overexpression
Inhibition (%)
Concentration (Log μM)
C
hDHFR
pDC2
HA1
HA2
gRNA
p-CasG
Cas9
PfSec13 (Pf3D7_1230700)
Targeted deletion
Internal Repeat Regions
PfSec13 construct
L
F1
F2
F(1+2)
186
355
501
D
PfSec13 protein
E
Gene ID | Annotation | Ref Codon | Var Codon
PF3D7_1230700 | transport protein SEC13 | GGTCAAATGAACGGAC-CACCTAGC | GGT
Consequence | Ref a.a. | Var a.a. | Position
nonsynonymous | GQMNGPPS | G | c(490,497)
F
3D7 WT
3D7 A1 Parental clone
3D7 A1 Drug pressure
PfSec13_mut Clone A
PfSec13_mut Clone C
PfSec13_mut Clone B
PfSec13_mut Clone D
Survival (%)
Concentration (μM)

Figure 2-4. Evidence for a role of *Sec13* in susceptibility to MED6-189. A. Graphical illustration of the methodology employed to isolate MED6-189 resistant parasites. **B.** Comparison of the growth rates of wild type *S. cerevisiae* and transgenic clones with either overexpressed or down-regulated Sec13p following treatment with a vehicle (DMSO) or increasing concentrations of MED6-189. **C.** Schematic representation of the CRISPR-Cas9-based replacement strategy used to introduce a deletion of the targeted repeat regions in the *PfSec13* gene. The insertion was achieved through overlap extension PCR of fragments directly upstream and downstream of the target segment and subsequently formed by primer overlap extension PCR to replicate the desired deletion. The insertion was validated through whole genome sequencing. **D.** Predicted structure of the SEC13 protein highlighting the seven amino acid tandem repeat regions targeted for deletion (red). **E.** Results of the sequencing analysis, confirming the successful deletion of the tandem repeat region of *PfSec13* using CRISPR-Cas9 in an isolated clone. **F.** 3D7 WT and parental lines (green and blue), resistant lines (red) maintained in the presence of DMSO or MED6-189 and transgenic *PfSec13*-mut clones (dashed) were subjected to a parasite survival assay. The curves depict parasite survival (y-axis) in response to serial drug dilution of MED6-189 (x-axis). Data was analyzed using a Sigmoidal, 4PL (X represents concentration, n=3, nonlinear regression, CI:95%).

MED6-189 Demonstrates Promising Results *In Vivo*

MED6-189 exhibits a favorable safety and tolerability profile

The potency of MED6-189 against *P. falciparum in vitro* coupled with a comprehensive understanding of its mechanism of action, prompted us to explore its efficacy in an animal model of *P. falciparum* malaria. The *in vitro* safety profile of MED6-189 was first determined by assessing possible effects on the growth and metabolic activity of five human cell lines (HeLa, THP1, HEK293, HepG2 and hTERT) at concentrations ranging between 48 nM and 100 µM. No inhibitory activity could be found, resulting in an estimated *in vitro* therapeutic index of > 500, surpassing that of several approved antimalarial drugs **(table 2-S8).** MED6-189 was further evaluated in 15 *in vitro* assays using the Enhanced Cross Screen Panel (eXP), which provides pharmacological information and toxicology alerts for the target compound. Of these, hPXR (Human Pregnane-X-receptor) IC_{50} 7.4 µM & BSEP inhibition (Bile Salt Export Protein) IC_{50} 50 µM were highlighted as factors to be considered in future drug optimization studies aimed at identifying ideal partner drugs and mitigating potential hepatotoxicity. Given the susceptibility of isonitrile compounds to hydrolysis to formamides in acidic aqueous solution, we synthesized the formamide derivatives GB209-2 and GB209-3 by acidic hydrolysis of MED6-189 **(see Supporting Information 2-S1).** These derivatives exhibited significantly decreased antiplasmodial activities (IC_{50} values of 3.8 and 3.9 µM, respectively), indicating that these potential degradation products are not responsible for the observed antimalarial effects of MED6-189 administration. We also showed that MED6-189 had reasonable stability to aqueous acidic conditions that mimic the gut **(See**

Supporting Information 2-S1). Furthermore, MED6-189, GB209-2 and GB209-3 showed no hemolytic activity at concentrations up to 10 µM **(See Supporting information 2-S1 and Fig. 2-S7, S8).** To guide *in vivo* efficacy studies, we investigated the *in vivo* tolerability profile of MED6-189. Our data showed that animals treated with MED6-189 at doses up to 50 mg/kg did not exhibit any adverse events, and there were no significant changes in hematology, clinical chemistry analyses, nor necropsy observations **(table 2-S9).** Importantly, the data also indicated tolerable exposure levels and compound absorption.

In vivo MED6-189 efficacy in a humanized mouse model of *P. falciparum* malaria

Owing to the favorable safety and tolerability profile exhibited by MED6-189, we chose a dosage of 50 mg/kg to assess the *in vivo* efficacy of the compound using a humanized mouse model of *P. falciparum* malaria (*163*). NOD scid gamma (NSG) mice were intravenously engrafted with human red blood cells (hRBCs) daily until the percentage of hRBCs reached ~50%, at which point the mice were infected with $2x10^7$ *P. falciparum*-infected erythrocytes. At the third day post-infection, mice were administered either the vehicle alone or MED6-189 at 50 mg/kg orally, once a day for four days. Whereas the control group showed a rapid increase in parasite load with parasitemia reaching ~6% by day 6 post-infection, mice treated with MED6-189 showed a significant clearance of parasites achieving more than 90 percent reduction in parasitemia in the peripheral blood by the seventh day post-treatment **(Fig. 2-5A, table 2-S10).**

Pan-antimalarial activity of MED6-189

Given the favorable biological activity of MED6-189 against *P. falciparum*, we also assessed its effectiveness against other *Plasmodium* species that infect humans. Owing to the lack of robust *in vitro* culturing systems for *P. vivax*, we examined the efficacy of MED6-189 against *P. knowlesi* and *P. cynomolgi*, which infect human erythrocytes and are commonly used as model systems for *P. vivax* infection (*164*). *P. knowlesi*, an Asian Old World monkey parasite with a robust *in vitro* culture system, is known to be zoonotic for humans. *P. cynomolgi*, a simian parasite capable of infecting humans experimentally, is phylogenetically closely related to *P. vivax* and shares similar biological and genetic properties with *P. vivax* (*164*). MED6-189 demonstrated significant activity against both parasites, with IC_{50} values of 309 ± 42 nM and 276 ± 22 nM against *P. knowlesi* in human and rhesus erythrocytes, respectively and (136 ± 56 nM, n=3) against *P. **cynomolgi*** **(Fig. 2-5B and 2-S9,S10A-B)** (Sigmoidal, 4PL, X is concentration, n=3, nonlinear regression, CI:95%). Despite the increased IC_{50} values against *P. cynomolgi* and *P. knowlesi* compared to those obtained for *P. falciparum*, the values fall well within the low nM range. Differences in drug efficacy is not uncommon amongst *Plasmodium* species and is most likely caused by the ability of *P. falciparum* to transport and concentrate nutrients and other molecules, including antimalarials, at a high level inside the parasite. Together, the data validate MED6-189 as a potent antimalarial with a broad-spectrum activity that targets both *falciparum* and non-*falciparum* human malaria.

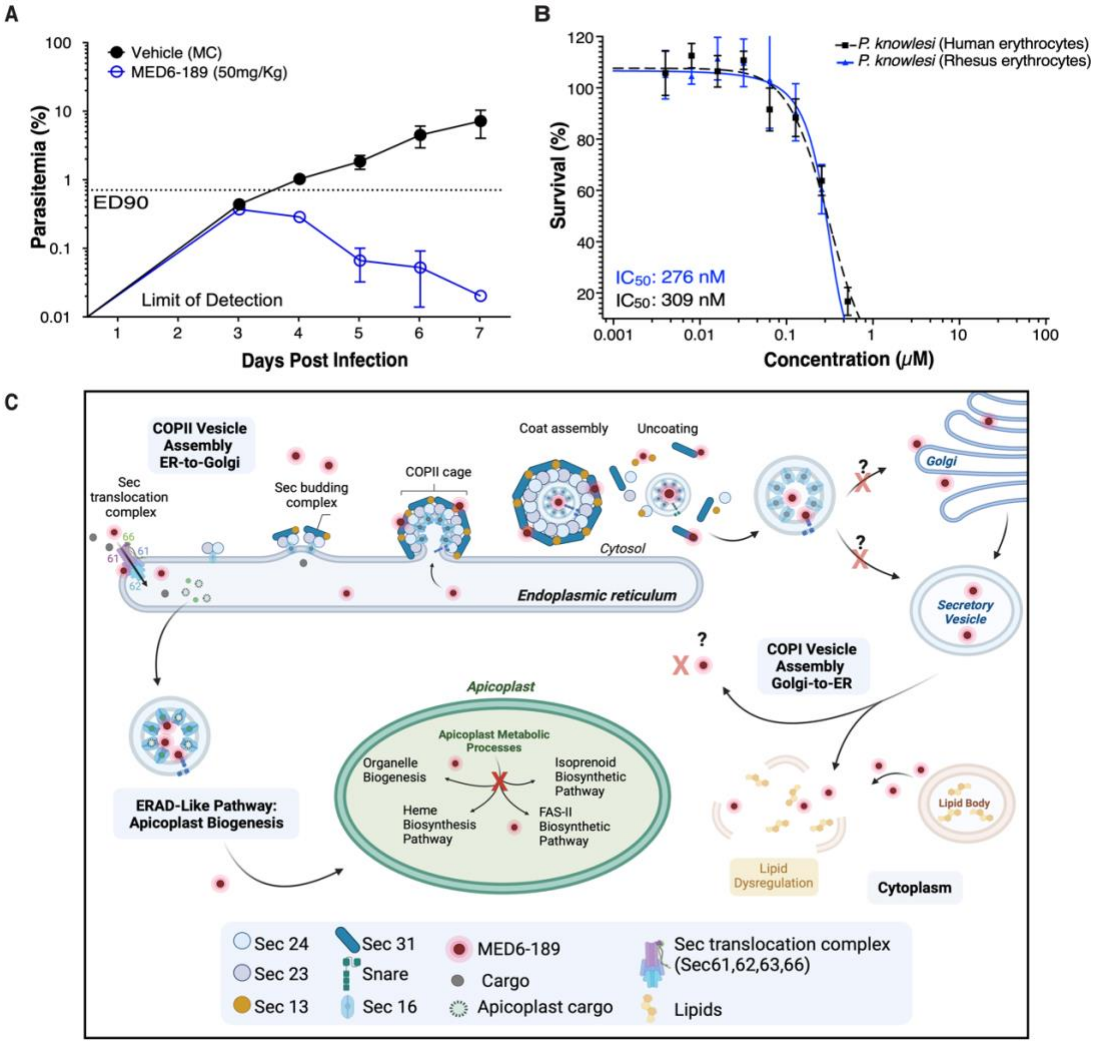

A
Vehicle (MC)
MED6-189 (50mg/Kg)
Parasitemia (%)
ED90
Limit of Detection
Days Post Infection
B
P. knowlesi (Human erythrocytes)
P. knowlesi (Rhesus erythrocytes)
Survival (%)
IC50: 276 nM
IC50: 309 nM
Concentration (μM)
C
COPII Vesicle Assembly ER-to-Golgi
Sec translocation complex
Sec budding complex
COPII cage
Coat assembly
Uncoating
Cytosol
Endoplasmic reticulum
Golgi
Secretory Vesicle
COPI Vesicle Assembly Golgi-to-ER
Apicoplast
Apicoplast Metabolic Processes
Organelle Biogenesis
Isoprenoid Biosynthetic Pathway
Heme Biosynthesis Pathway
FAS-II Biosynthetic Pathway
ERAD-Like Pathway: Apicoplast Biogenesis
Lipid Body
Lipid Dysregulation
Cytoplasm
Sec 24
Sec 23
Sec 13
Sec 31
Snare
Sec 16
MED6-189
Cargo
Apicoplast cargo
Sec translocation complex (Sec61,62,63,66)
Lipids

Figure 2-5. *In vivo* and broad-spectrum antimalarial efficacy of MED6-189. A. The *in vivo* efficacy of MED6-189 was evaluated in a humanized mouse model infected with *P. falciparum* (blue) compared to untreated controls (black). **B.** Dose-dependent response of MED6-189 on *P. knowlesi* YH1 human erythrocyte infecting (black) and rhesus erythrocyte infecting (blue) parasites. The graphs illustrate the logarithmic growth of parasites (Y-axis) in response to varying drug concentrations (X-axis). Error bars represent standard deviations from two independent experiments conducted in triplicate. The regression line is derived from a nonlinear regression analysis (Variable slope with four parameters, least squares fit). **C.** Proposed mode of action of MED6-189 in *P. falciparum*-infected erythrocytes. The compound is imported into the endoplasmic reticulum (ER) via the Sec translocation complex (SEC61,62,63,66), where it interacts with components of the ER transport machinery. The compound is translocated into the apicoplast where it directly interacts with proteins involved in crucial apicoplast function, ultimately disrupting this vital organelle. The interactions of MED6-189 with components of the apicoplast function and trafficking systems lead to dysregulation of lipids, resulting in the disruption of key biological processes within the *Plasmodium* parasite.

DISCUSSION

Using a multifaceted approach, we have investigated the *in vitro* and *in vivo* efficacy, mode of action, resistance mechanisms, and preclinical safety profile of the kalihinol analogue MED6-189. Our findings establish MED6-189 as a multi-target compound with limited capacity for the parasite to develop drug resistance, making a promising new lead in the fight against malaria. Our studies further highlighted the pivotal role of the apicoplast in the biological activity of the drug. By employing targeted and untargeted analyses, we have confirmed the apicoplast as a major target of its mode of action. Furthermore, our investigations have unveiled a *delayed death* mode of operation for MED6-189, reminiscent of other small molecules that target the integrity and biogenesis of the apicoplast.

To localize MED6-189 within the parasites, we used a fluorescent analogue, which confirmed the apicoplast as a major site of MED6-189 accumulation. Consistent with this finding, the compound was found to have a slow killing mode of action, like that of several classes of apicoplast-targeting molecules (*83, 127, 128*). Moreover, *in vitro* efficacy studies show that the activity of MED6-189 was antagonistic to that of fosmidomycin, a drug previously shown to disrupt the MEP pathway in the apicoplast. Interestingly, metabolic supplementation with IPP led to a partial rescue of parasite replication. Unlike fosmidomycin, MED6-189-treated parasites did not progress past the third intraerythrocytic cycle, suggesting that the compound may have multiple mechanisms of action, which may also account for the challenges in obtaining resistant parasites through standard drug selection methods.

Our metabolite profiling identified significant alterations in lysophosphatidylcholine (LPC), several triglycerides/triacylglycerides (TG/TAGs) and a limited subset of phosphatidylcholine (PC) forms in MED6-189-treated parasites. Notably, LPC exhibited the most pronounced change, with an estimated 3.5-fold increase in its steady-state levels. LPC, a precursor for the biosynthesis of PC, the most abundant phospholipid (50% of phospholipids in *P. falciparum* membranes), plays a key role in the regulation of sexual differentiation (*165*). Our data also indicated a significant reduction in the biosynthesis of sphingomyelin (*143, 166, 167*), sterol-esters and phosphatidylethanolamine (PE) in MED6-189-treated samples, indicating a significant disruption in lipid metabolism and turnover caused by the compound. Our metabolomic analysis identified thiamine pyrophosphate (TPP) as a major altered metabolite in *P. falciparum* following treatment with MED6-189. TPP is essential for the activity of several enzymes and biosynthetic pathways, including the pyruvate dehydrogenase PDH complex found in the apicoplast (*142*). The apicoplast PDH converts pyruvate into acetyl-CoA, the major fatty acid precursor, whereas a second distinct PDH fuels the tricarboxylic acid cycle in the mitochondria (*168*). Altogether, our data indicate that MED6-189 operates by modulating *P. falciparum* lipid metabolism and apicoplast biogenesis.

Proteomic analyses further supported the link between MED6-189 activity and *P. falciparum* membrane biogenesis and trafficking pathways. Several components of the SEC and SNARE secretory machinery, namely (PF3D7_1230700), SEC62

(PF3D7_1438100), SEC66 (PF3D7_0204200), and syntaxin SYN13 (PF3D7_1104100), were found either to be enriched in the drug pull-down experiments and/or showed differential melting profiles between MED6-189 treated versus vehicle-treated control samples. However, it is worth noting that not all targets identified in the pull-down assays were also identified by TPP. With relatively fewer proteins of significance found stabilized in our TPP dataset, one may hypothesize that MED6-189 may have greater impact on the ribosomal RNA, which is not captured in TPP results. Other factors can account for the differences seen between TPP and affinity purification findings including the level of abundance of specific proteins. For example, the SEC13 protein melting curves were not identified in our DMSO control samples yet consistently detected in our MED6-189 samples **(Fig. 2-S5D).** The problem with low abundance target proteins not generating TPP curves has been noted previously by Mateus and coworkers (*169*). However, the melting curves of SEC13 in the presence of DMSO and pyrimethamine have been previously established by Dziekan *et al* 2020 (*155*) and we showed that using this DMSO thermal profile **(Fig. 2-S5D)** was a reasonable comparison **(Fig. 2-S5E).** Combining these TPP analyses indicated that MED6-189 did indeed greatly stabilize SEC13 beyond 49°C, its previously observed melting temperature **(Fig. 2-S5D)** *(155)*, indicating a likely direct interaction with the drug.

SEC13 plays a major role in COPII-mediated vesicular transport between the nuclear pore complex, ER and Golgi membranes, whereas SEC62 and SEC66 (SEC71) are important components of protein translocation machinery in the ER *(80, 106, 151, 152, 170)*. The

ER-associated degradation system (ERAD) is a quality control mechanism that retro-translocates misfolded secretory proteins across the ER. A similar system, the ERAD-like system, is believed to play a critical role in protein transport into the apicoplast (*25*, *171*). Both the thermal proteomic profiling and pull-down assays identified significant interactions between MED6-189 and the SNARE protein family. Interestingly, syntaxin SYN13, a Qa-SNARE family protein, was also identified in our TPP assay as *P. falciparum* proteins specifically stabilized by MED6-189. The role of these proteins in COPI and COPII vesicle trafficking has previously been reported (*172-174*). These proteins play a key role in membrane identification and mediate membrane fusions across various organelles including the endoplasmic reticulum, mitochondria, apicoplast, and other double-membrane-bound vesicles (*172*, *173*, *175-177*). Most recently, SNARE proteins have been shown to play an important role in the biogenesis and maintenance of the apicoplast organelle in *Toxoplasma gondii* (*178*).

Other evidence for the disruption of apicoplast biogenesis stems from the thermal shift assay that revealed that MED6-189 significantly affected the stabilization of proteins such as 2C-methyl-d-erythritol 2,4-cyclodiphosphate enzyme (MECP | PF3D7_0209300), ssDNA binding protein (SSB | PF3D7_0508800), and acetyl-coA binding proteins (PF3D7_0810000) **(Fig. 2-S5C, bottom row).** All these proteins are integral to the biosynthesis of terpenoids (*179*), fatty acids, and apicoplast biogenesis systems (*25*, *171*, *179*).

Utilizing omics-based strategies, we have gained an understanding of how MED6-189 is trafficked into the apicoplast and disrupts its biogenesis, ultimately leading to parasite cell death consistent with other delayed acting apicoplast-disrupting therapies such as fosmidomycin **(Fig. 2-5C).** These studies were further complemented by genetic approaches aimed at selecting parasite clones with reduced susceptibility to MED6-189. Initial attempts to generate resistant variants through the one-step selection approach had failed (*180, 181*), yet we were ultimately able to achieve heightened drug tolerance in culture by incrementally escalating drug dosage over time through a stepwise selection method (*182, 183*). Unlike many other antimalarials for which resistant parasites can be selected within a small number of intraerythrocytic cycles, parasites with enhanced tolerance to MED6-189 required an estimated 36 consecutive months of drug pressure. The IC_{50} for MED6-189 in these resistant parasites was determined to be 3-4-fold that of the parent 3D7 strain. Whole genome sequencing of the resistant clones identified SNP and Indel mutations in genes involved in vesicular trafficking including *PfSec13,* and *Pfendoplasmin*. Because *Pfendoplasmin* showed no change in stability compared to DMSO controls in our TPP experiment, we validated *PfSec13* as a potential drug target. The importance of SEC13 in MED6-189 cellular response was therefore further validated using yeast as a model system. This finding led us to apply CRISPR-cas-9 genetic editing tool in *P. falciparum* to validate the role of SEC13 in MED6-189 tolerance. *PfSec13-mut* lines were subsequently cloned and survival assays confirmed the role of SEC13 as an adaptive mechanism of resistance to MED6-189 treatment. Our proteomics analyses pointed to a direct interaction between SEC13 and MED6-189 and indicate that both the mode of action

and mechanism of resistance of the drug are most likely linked. One of our stronger limitations from our studies is the fact that we were unable to obtain a fully resistant line, limiting our capability to fully understand the mechanism of action and resistance of MED6-189.

Altogether, our studies provide a comprehensive investigation into the activity and mode of action of MED6-189, a highly effective antimalarial compound targeting both the asexual and sexual stages of *P. falciparum*. The array of complex molecular pathways hindered by MED6-189 and the difficulty of selection of parasites that are refractory to the compound, render it an exciting candidate for preventing parasite progression with reduced susceptibility for emergence and spread of drug resistance. While our studies show that MED6-189 exhibits polypharmacology by targeting apicoplast biogenesis, and fatty acid biosynthesis and trafficking, other possible contributing mechanisms of action, including inhibition of heme detoxification, remain possible. Despite our comprehensive investigation we have yet to fully understand explicit pathways impacted by MED6-189. This would warrant further greater exploration onto the pathogen's secretory pathways and the roles they play in drug trafficking. We establish that MED6-189, and by extension, other similar kalihinols and ICTs more broadly, may serve as promising antimalarial agents, especially when coupled with more fast-acting drug therapies. Finally, the ability of MED6-189 to inhibit the growth of *P. falciparum in vitro* and in humanized mice as to inhibit the growth of *P. knowlesi* and *P. cynomolgi* makes it a promising antimalarial lead drug.

MATERIALS AND METHODS

Detailed information on materials and methods is available as Supplementary Materials and Methods.

In vivo cultures

In brief, the anti-malarial activity of individual analogues was evaluated *in-vitro* against *P. falciparum* 3D7 (MRA-102, drug sensitive) W2 (MRA-157, Chloroquine resistant) strains (ATCC® Manassas, VA), Dd2 (MRA-156), NF54 (MRA-1000), HB3 (MRA-155), and D10-Acp-GFP (MRA-568) (*161, 184, 185*). *P. knowlesi* assays were quantitatively measured by flow cytometry using SYBR Green I, and Mitrotracker Deep Red as previously described (*186, 187*) on a Beckman Coulter CytoFLEX. Nonhuman primate infections were required to generate *P. cynomolgi* M/B strain parasites for *in vitro* testing (*184, 188*).

RNA-sequencing

Differential expression analysis was done by use of R package DESeq2 with an adjusted P-value cutoff of 0.05. Volcano plots were made using R package Enhanced Volcano or GraphPad Prism 9 (GraphPad Software, Inc.).

LC-MS Metabolomics-lipids

LC-MS metabolomics analysis was performed on a Synapt G2-Si quadrupole time-of-flight mass spectrometer (Waters) coupled to an I-class UPLC system (Waters).

LC-MS Metabolomics-polar metabolites

Targeted metabolomics of polar, primary metabolites was performed on a TQ-XS triple quadrupole mass spectrometer (Waters) coupled to an I-class UPLC system (Waters).

Metabolomic data processing and analysis

Untargeted data processing (peak picking, alignment, deconvolution, integration, and spectral matching) was performed in Progenesis Qi software (Nonlinear Dynamics). Data were normalized to total ion count. Similar metabolites features were assigned a cluster ID using RAMClust (*147*). An extension of the metabolomics standard initiative guidelines was used to assign annotation level confidence (*100, 189*). Targeted data processing was performed in Skyline software (*190*).

P. falciparum PfSec13-mut strains

Gene editing of the *P. falciparum Sec13* gene *(PF3D7_1230700)* performed using a two-plasmid-based strategy (*2, 8, 191*). The pCasG-Cas9-sgRNA plasmid vector plasmid contains the site to express the sgRNA, along with the yDHODH gene as the positive selection marker. The sgRNA was selected from the database generated by Desai and Ribeiro *et al. 2018 (191-193).*

TPP mass spectrometry data

Data was analyzed using mineCETSA R-language package as described previously *(154, 155)*.

Proteomics data processing and analysis

Tandem mass (MS/MS) spectra were interpreted using ProluCID v.1.3.3(*194*). DTASelect v.1.9 (*195*) and swallow v.0.0.1, an in-house developed software (*196*) were used to control FDRs at less than 1.2%. All datasets were contrasted against their merged data set, respectively, using Contrast v1.9 (*195*) and in-house developed sandmartin. An in-house developed software, NSAF7 v.0.0.1, was used to generate spectral count-based label free quantitation results (*146*). QSPEC/QPROT (*145, 148*) was used to calculate $\log_2$ Fold-Changes and Z-statistics (*146, 184, 185, 188, 192, 197, 198*).

Yeast strain Sec13p expression analysis

Overexpression of *S. cerevisiae* Sec13p, was generated by an episomal *E. coli*/yeast shuttle vector, transformed into the BY4741 yeast strain. *SEC13* gene repression utilized the tet-off system.

Toxicological Profiles

Blood samples for analysis were carried out in the Abacus5 junior Vet → (Practice CVM S.L.L). Clinical chemistry was performed by mean analysis of whole blood in Vetscan→ (ABAXIS) & F560 (Menarini→) analyzer and by mean analysis of plasma.

In vivo Efficacy studies

MED6-189 oral efficacy was tested in female NOD-*scid IL-2R*γ^{null} (NSG) mice engrafted with human erythrocytes and infected *P. falciparum*-infected erythrocytes. Peripheral

blood from *P. falciparum*-infected mice were stained with TER-119-Phycoerythrine (marker of murine erythrocytes) and SYTO-16 (nucleic acid dye) and then analyzed by flow cytometry (FACSCalibur, BD).

Bioanalysis and pharmacokinetics analysis: LC-MS analysis

An Acquity Ultra-Performance liquid chromatography (UPLC) system (Waters Corp., Milford, MA, USA) couple to a triple quadrupole mass spectrometer (API 4000™, AB Sciex, Foster City, CA, USA) was used for the analysis(*199*).

Pharmacokinetic analysis

Blood concentration time data were analyzed by Non-Compartmental PK analysis using Phoenix WinNonlin software (Certara NY, US) to calculate PK parameters.

Statistical analyses

Parasitemia and proportion of asexual stages were analyzed using a two-way ANOVA with Tukey's test for multiple comparisons. Significant differences were indicated as following: * for $p< 0.05$; ** for $p< 0.01$, *** for $p<0.001$ and **** for $p<0.0001$. Statistical tests were performed with GraphPad Prism. Figures were generated with GraphPad Prism 9, Biorender, Adobe illustratorv27.3.1. Putative Sec13 protein was created by Alphafold monomer V2.0 prediction for protein transport protein SEC13 (*Q8I5B3*) with PDB reference *AF-Q8I5B3-F1-model_v4 (1).pdb* and protein structure was formed through ChimeraX.

ACKNOWLEDGEMENTS

We thank Matthias Göhl and Ryan Kozlowski for early efforts to scale up the synthesis of MED6-189. Alphafold monomer V2.0 prediction for protein transport protein SEC13 (Q8I5B3) and protein structure was formed through ChimeraX. All data used in this analysis is free to access.

FUNDING

This work was supported by the National Institutes of Allergy and Infectious Diseases of the National Institutes of Health (Grant R01 AI-138139 to C.D.V., C.B.M, and K.L.R) and the University of California, Riverside (NIFA-NIFAHatch- 225935 to K.G.L.R.). Proteomics datasets were generated in support of the NIH NIAID Grant no. (R01 AI-136511) and NIH R01 Grant No. (R01 AI-142743) at the Stower's Institute of Health. C.B.M. research is supported by NIH grants AI138139, AI123321, AI152220 and AI153100 and AI136118; the Steven and Alexandra Cohen Foundation (Lyme 62 2020); the Global Lyme Alliance and The Blavatnik Fund for Innovation. A.K.B. is supported by an International Research Scientist Development Award (K01 TW010496) from the Fogarty International Center of the National Institutes of Health. *P. k.* studies were supported by the Emory National Primate Research Center (Grant No. ORIP/OD P51OD011132) and Emory National Primate Research Center (Grant No. U42PDP11023). *P. cynomolgi* studies were funded in part by funding for INV-031788 granted to C.J.J. from the Bill and Melinda Gates Foundation. This publication includes data generated at the UC

San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (#S10 OD026929). The additional support of Tres Cantos Open Lab Foundation is gratefully acknowledged.

AUTHOR CONTRIBUTIONS

K.L.R., C.B.M., and C.D.V. conceived and designed all experiments. Z.C. performed all experiments with *P. falciparum* and S.A. contributed to the bioinformatics data analyses. M.E.D. and J.H.C. synthesized the kalihinol derivatives, including the biotinylated and fluorescent analogues. G.L.B. performed hydrolytic stability studies on MED6-189 and synthesized and characterized the formamide hydrolysis products. J.S.K. performed the metabolomic experiments while A.S., C.B and L.F. performed the proteomics mass-spectrometry analyses. J.Y.C. conducted the studies in yeast. I.R., M.C.C. and S.V-M. evaluated the *in vivo* efficacy of MED6-189 in murine malaria models. I.C., P. C., J. D. M-A., N. I. and A. G-P. evaluated the safety, pharmacokinetic and tolerability profiles of MED6-189. A.K.B. evaluated the *in vitro* efficacy of MED6-189 in *P. knowlesi* malaria parasites. C.J.J and M.A.A conducted the studies in *P. cynomolgi.* Z.C., T.H., K.L.R., C.B.M., C.D.V., J.C. and P.V. contributed to the writing of the manuscript. All authors reviewed, edited, and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests nor conflicts of interest.

Correspondence and requests for materials should be addressed to Karine Le Roch.

DATA AVAILABILITY

The MED6-189 purifications and TPP mass spectrometry datasets have been deposited to the ProteomeXChange (with PXD038457 and PXD038053 accessions, respectively) via the MassIVE repository (MSV000090812 and MSV000090667) and may also be accessed from the Stowers Original Data Repository (http://www.stowers.org/research/-publications/libpb-1759)(*100, 101*). WGS and RNA-seq datasets generated in this study have been deposited in the NCBI BioProject database under SubmissionID: SUB12241156 with BioProject ID: PRJNA930408(*102*). Metabolomics datasets generated in this study are available at Panorama database and can be accessed through the following the doi**:** https://doi.org/10.6069/5a05-5x26(*103*).

CODE AVAILABILITY

The entire in-house software suite (Kite) used for the MudPIT mass spectrometry analysis is supplied by Zenodo:
(https://zenodo.org/record/5914885#.Y9q0juzMKjg)(*96*).

REFERENCES

1. WHO, World malaria report 2023. (2023).

2. X. Wang *et al.*, Molecular Epidemiology of Drug Resistance Genes in Plasmodium falciparum Isolates Imported from Nigeria between 2016 and 2020: Continued Emergence of Fully Resistant Pfdhfr-Pfdhps Alleles. *Microbiol Spectr*, e0052822 (2022).

3. W. H. Organization, Global report on antimalarial drug efficacy and drug resistance, 2000–2010. *WHO*, (2010).

4. A. Uwimana *et al.*, Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda. *Nat Med* **26**, 1602-1608 (2020).

5. A. Uwimana *et al.*, Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study. *Lancet Infect Dis* **21**, 1120-1128 (2021).

6. N. Noreen *et al.*, New insights into the spread of resistance to artemisinin and its analogues. *J Glob Antimicrob Resist* **27**, 142-149 (2021).

7. M. V. Andrade *et al.*, The economic burden of malaria: a systematic review. *Malar J* **21**, 283 (2022).

8. A. M. Dondorp *et al.*, Artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med* **361**, 455-467 (2009).

9. A. Haakenstad *et al.*, Tracking spending on malaria by source in 106 countries, 2000-16: an economic modelling study. *Lancet Infect Dis* **19**, 703-716 (2019).

10. U. Bohme *et al.*, Complete avian malaria parasite genomes reveal features associated with lineage-specific evolution in birds and mammals. *Genome Res* **28**, 547-560 (2018).

11. R. Thomson-Luque *et al.*, Plasmodium falciparum transcription in different clinical presentations of malaria associates with circulation time of infected erythrocytes. *Nat Commun* **12**, 4711 (2021).

12. W. Liu *et al.*, Multigenomic Delineation of Plasmodium Species of the Laverania Subgenus Infecting Wild-Living Chimpanzees and Gorillas. *Genome Biol Evol* **8**, 1929-1939 (2016).

13. J. R. Poespoprodjo, N. M. Douglas, D. Ansong, S. Kho, N. M. Anstey, Malaria. *Lancet* **402**, 2328-2345 (2023).

14. I. Mueller, P. A. Zimmerman, J. C. Reeder, Plasmodium malariae and Plasmodium ovale--the "bashful" malaria parasites. *Trends Parasitol* **23**, 278-283 (2007).

15. C. J. Merrick, Hypnozoites in Plasmodium: Do Parasites Parallel Plants? *Trends Parasitol* **37**, 273-282 (2021).

16. S. Mehra, E. Stadler, D. Khoury, J. M. McCaw, J. A. Flegg, Hypnozoite dynamics for Plasmodium vivax malaria: The epidemiological effects of radical cure. *J Theor Biol* **537**, 111014 (2022).

17. T. B. Vieira, T. P. Astro, R. R. de Moraes Barros, Genetic Manipulation of Non-Falciparum Human Malaria Parasites. *Front Cell Infect Microbiol* **11**, 680460 (2021).

18. L. Tilley, M. McConville, Malaria: Sensing when it's time for sex. *Nature* **499**, 38-40 (2013).

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

20. S. Eksi *et al.*, Plasmodium falciparum gametocyte development 1 (Pfgdv1) and gametocytogenesis early gene identification and commitment to sexual development. *PLoS Pathog* **8**, e1002964 (2012).

21. J. Chawla, J. Oberstaller, J. H. Adams, Targeting Gametocytes of the Malaria Parasite Plasmodium falciparum in a Functional Genomics Era: Next Steps. *Pathogens* **10**, (2021).

22. J. D. Dvorin, D. E. Goldberg, Plasmodium Egress Across the Parasite Life Cycle. *Annu Rev Microbiol* **76**, 67-90 (2022).

23. G. I. McFadden, E. Yeh, The apicoplast: now you see it, now you don't. *Int J Parasitol* **47**, 137-144 (2017).

24. K. Keegstra, K. Cline, Protein import and routing systems of chloroplasts. *Plant Cell* **11**, 557-570 (1999).

25. D. W. Chung, N. Ponts, J. Prudhomme, E. M. Rodrigues, K. G. Le Roch, Characterization of the ubiquitylating components of the human malaria parasite's protein degradation pathway. *PLoS One* **7**, e43477 (2012).

26. J. D. Fellows, M. J. Cipriano, S. Agrawal, B. Striepen, A Plastid Protein That Evolved from Ubiquitin and Is Required for Apicoplast Protein Import in Toxoplasma gondii. *mBio* **8**, (2017).

27. G. I. McFadden, M. E. Reith, J. Munholland, N. Lang-Unnasch, Plastid in human parasites. *Nature* **381**, 482 (1996).

28. J. I. MacRae, E. Marechal, C. Biot, C. Y. Botte, The apicoplast: a key target to cure malaria. *Curr Pharm Des* **18**, 3490-3504 (2012).

29. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation. *Science* **295**, 1306-1311 (2002).

30. A. M. Guggisberg, R. E. Amthor, A. R. Odom, Isoprenoid biosynthesis in Plasmodium falciparum. *Eukaryot Cell* **13**, 1348-1359 (2014).

31. H. Jomaa *et al.*, Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285**, 1573-1576 (1999).

32. G. G. van Dooren, L. M. Stimmler, G. I. McFadden, Metabolic maps and functions of the Plasmodium mitochondrion. *FEMS Microbiol Rev* **30**, 596-630 (2006).

33. R. M. Simao-Gurge *et al.*, Biosynthesis of heme O in intraerythrocytic stages of Plasmodium falciparum and potential inhibitors of this pathway. *Sci Rep* **9**, 19261 (2019).

34. N. Thomsen-Zieger, J. Schachtner, F. Seeber, Apicomplexan parasites contain a single lipoic acid synthase located in the plastid. *FEBS Lett* **547**, 80-86 (2003).

35. C. Bisanz *et al.*, Toxoplasma gondii acyl-lipid metabolism: de novo synthesis from apicoplast-generated fatty acids versus scavenging of host cell precursors. *Biochem J* **394**, 197-205 (2006).

36. R. M. Packard, The origins of antimalarial-drug resistance. *N Engl J Med* **371**, 397-399 (2014).

37. N. Ma, Z. Zhang, F. Liao, T. Jiang, Y. Tu, The birth of artemisinin. *Pharmacol Ther* **216**, 107658 (2020).

38. X. Z. Su, L. H. Miller, The discovery of artemisinin and the Nobel Prize in Physiology or Medicine. *Sci China Life Sci* **58**, 1175-1179 (2015).

39. S. Oguche *et al.*, Efficacy of artemisinin-based combination treatments of uncomplicated falciparum malaria in under-five-year-old Nigerian children. *Am J Trop Med Hyg* **91**, 925-935 (2014).

40. K. Mahotorn *et al.*, In Vitro Sensitivity of Pyronaridine in Thai Isolates of Plasmodium falciparum. *Am J Trop Med Hyg* **98**, 51-56 (2018).

41. M. Madamet *et al.*, The Plasmodium falciparum chloroquine resistance transporter is associated with the ex vivo P. falciparum African parasite response to pyronaridine. *Parasit Vectors* **9**, 77 (2016).

42. P. Hodoameda, N. O. Duah-Quashie, N. B. Quashie, Assessing the Roles of Molecular Markers of Antimalarial Drug Resistance and the Host Pharmacogenetics in Drug-Resistant Malaria. *J Trop Med* **2022**, 3492696 (2022).

43. O. A. Folarin *et al.*, In vitro amodiaquine resistance and its association with mutations in pfcrt and pfmdr1 genes of Plasmodium falciparum isolates from Nigeria. *Acta Trop* **120**, 224-230 (2011).

44. A. Djimde *et al.*, A molecular marker for chloroquine-resistant falciparum malaria. *N Engl J Med* **344**, 257-263 (2001).

45. K. J. Wicht, S. Mok, D. A. Fidock, Molecular Mechanisms of Drug Resistance in Plasmodium falciparum Malaria. *Annu Rev Microbiol* **74**, 431-454 (2020).

46. T. Hollin, Z. Chahine, K. G. Le Roch, Epigenetic Regulation and Chromatin Remodeling in Malaria Parasites. *Annu Rev Microbiol* **77**, 255-276 (2023).

47. Malaria eradication: benefits, future scenarios and feasibility: a report of the Strategic Advisory Group on Malaria Eradication. *World Health Organization*, (2021).

48. WHO, Global Malaria Programme. (2024).

49. L. S. Garcia, Malaria. *Clin Lab Med* **30**, 93-129 (2010).

50. WHO. (2021).

51. M. Prof Christopher JL Murray *et al.*, Global malaria mortality between 1980 and 2010: a systematic analysis. *The Lancet Journal* **379**, 413-431 (2012).

52. A. Hammond *et al.*, Gene-drive suppression of mosquito populations in large cages as a bridge between lab and field. *Nat Commun* **12**, 4589 (2021).

53. P. E. Duffy, J. Patrick Gorres, Malaria vaccines since 2000: progress, priorities, products. *NPJ Vaccines* **5**, 48 (2020).

54. B. Mordmuller *et al.*, A PfSPZ vaccine immunization regimen equally protective against homologous and heterologous controlled human malaria infection. *NPJ Vaccines* **7**, 100 (2022).

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

56. S. Z. Davis, T. Hollin, T. Lenz, K. G. Le Roch, Three-dimensional chromatin in infectious disease-A role for gene regulation and pathogenicity? *PLoS Pathog* **17**, e1009207 (2021).

57. G. Batugedara, K. G. Le Roch, Unraveling the 3D genome of human malaria parasites. *Semin Cell Dev Biol* **90**, 144-153 (2019).

58. T. Hollin, K. G. Le Roch, From Genes to Transcripts, a Tightly Regulated Journey in Plasmodium. *Front Cell Infect Microbiol* **10**, 618454 (2020).

59. T. Hollin, M. Gupta, T. Lenz, K. G. Le Roch, Dynamic Chromatin Structure and Epigenetics Control the Fate of Malaria Parasites. *Trends Genet* **37**, 73-85 (2021).

60. T. Hollin, Z. Chahine, K. G. Le Roch, Epigenetic Regulation and Chromatin Remodeling in Malaria Parasites. *Annu Rev Microbiol*, (2023).

61. M. T. Watzlowik, S. Das, M. Meissner, G. Langst, Peculiarities of Plasmodium falciparum Gene Regulation and Chromatin Structure. *Int J Mol Sci* **22**, (2021).

62. X. Shang *et al.*, PfAP2-EXP2, an Essential Transcription Factor for the Intraerythrocytic Development of Plasmodium falciparum. *Front Cell Dev Biol* **9**, 782293 (2021).

63. T. L. Campbell, E. K. De Silva, K. L. Olszewski, O. Elemento, M. Llinas, Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. *PLoS Pathog* **6**, e1001165 (2010).

64. X. Shang *et al.*, Genome-wide landscape of ApiAP2 transcription factors reveals a heterochromatin-associated regulatory network during Plasmodium falciparum blood-stage development. *Nucleic Acids Res* **50**, 3413-3431 (2022).

65. R. M. Martins *et al.*, An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite Plasmodium falciparum. *Sci Rep* **7**, 14042 (2017).

66. E. Carrington *et al.*, The ApiAP2 factor PfAP2-HC is an integral component of heterochromatin in the malaria parasite Plasmodium falciparum. *iScience* **24**, 102444 (2021).

67. C. Flueck *et al.*, A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in chromosome end biology. *PLoS Pathog* **6**, e1000784 (2010).

68. M. Sierra-Miranda *et al.*, PfAP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds to Plasmodium falciparum telomeres. *Cell Microbiol* **19**, (2017).

69. Z. Chahine, K. G. Le Roch, Decrypting the complexity of the human malaria parasite biology through systems biology approaches. *Front Syst Biol* **2**, (2022).

70. W. A. Hoeijmakers *et al.*, Plasmodium falciparum centromeres display a unique epigenetic makeup and cluster prior to and during schizogony. *Cell Microbiol* **14**, 1391-1401 (2012).

71. J. Miao *et al.*, The malaria parasite Plasmodium falciparum histones: organization, expression, and acetylation. *Gene* **369**, 53-65 (2006).

72. W. J. Sullivan, Jr., Histone H3 and H3.3 variants in the protozoan pathogens Plasmodium falciparum and Toxoplasma gondii. *DNA Seq* **14**, 227-231 (2003).

73. W. A. Hoeijmakers *et al.*, H2A.Z/H2B.Z double-variant nucleosomes inhabit the AT-rich promoter regions of the Plasmodium falciparum genome. *Mol Microbiol* **87**, 1061-1073 (2013).

74. M. Petter *et al.*, H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy var gene promoters in the malaria parasite Plasmodium falciparum. *Mol Microbiol* **87**, 1167-1182 (2013).

75. M. Petter *et al.*, Expression of P. falciparum var genes involves exchange of the histone variant H2A.Z at the promoter. *PLoS Pathog* **7**, e1001292 (2011).

76. N. Coetzee *et al.*, Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual Plasmodium falciparum parasites. *Sci Rep* **7**, 607 (2017).

77. A. M. Salcedo-Amaya *et al.*, Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. *Proc Natl Acad Sci U S A* **106**, 9655-9660 (2009).

78. A. Saraf *et al.*, Dynamic and Combinatorial Landscape of Histone Modifications during the Intraerythrocytic Developmental Cycle of the Malaria Parasite. *J Proteome Res* **15**, 2787-2801 (2016).

79. M. B. Trelle, A. M. Salcedo-Amaya, A. M. Cohen, H. G. Stunnenberg, O. N. Jensen, Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum. *J Proteome Res* **8**, 3439-3450 (2009).

80. S. Siniossoglou *et al.*, A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. *Cell* **84**, 265-275 (1996).

81. C. Epp, F. Li, C. A. Howitt, T. Chookajorn, K. W. Deitsch, 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**, 116-127 (2009).

82. J. J. Lopez-Rubio, L. Mancio-Silva, A. Scherf, Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* **5**, 179-190 (2009).

83. S. A. Ralph *et al.*, Tropical infectious diseases: metabolic maps and functions of the Plasmodium falciparum apicoplast. *Nat Rev Microbiol* **2**, 203-216 (2004).

84. J. J. Lopez-Rubio *et al.*, 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol Microbiol* **66**, 1296-1305 (2007).

85. L. H. Freitas-Junior *et al.*, Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121**, 25-36 (2005).

86. S. Shrestha *et al.*, Distinct Histone Post-translational Modifications during Plasmodium falciparum Gametocyte Development. *J Proteome Res* **21**, 1857-1867 (2022).

87. J. Connacher *et al.*, H3K36 methylation reprograms gene expression to drive early gametocyte development in Plasmodium falciparum. *Epigenetics Chromatin* **14**, 19 (2021).

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

89. A. Kanyal *et al.*, Genome-wide survey and phylogenetic analysis of histone acetyltransferases and histone deacetylases of Plasmodium falciparum. *FEBS J* **285**, 1767-1782 (2018).

90. B. K. Chaal, A. P. Gupta, B. D. Wastuwidyaningtyas, Y. H. Luah, Z. Bozdech, Histone deacetylases play a major role in the transcriptional regulation of the Plasmodium falciparum life cycle. *PLoS Pathog* **6**, e1000737 (2010).

91. L. Cui *et al.*, PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in Plasmodium falciparum. *Eukaryot Cell* **6**, 1219-1227 (2007).

92. N. Coetzee *et al.*, Epigenetic inhibitors target multiple stages of Plasmodium falciparum parasites. *Sci Rep* **10**, 2355 (2020).

93. M. V. C. Greenberg, D. Bourc'his, The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* **20**, 590-607 (2019).

94. N. Ponts *et al.*, Genome-wide mapping of DNA methylation in the human malaria parasite Plasmodium falciparum. *Cell Host Microbe* **14**, 696-706 (2013).

95. E. Hammam *et al.*, Discovery of a new predominant cytosine DNA modification that is linked to gene expression in malaria parasites. *Nucleic Acids Res* **48**, 184-199 (2020).

96. J. Baum *et al.*, Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Res* **37**, 3788-3798 (2009).

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

98. Q. Jing *et al.*, Plasmodium falciparum var Gene Is Activated by Its Antisense Long Noncoding RNA. *Front Microbiol* **9**, 3117 (2018).

99. I. Amit-Avraham *et al.*, Antisense long noncoding RNAs regulate var gene activation in the malaria parasite Plasmodium falciparum. *Proc Natl Acad Sci U S A* **112**, E982-991 (2015).

100. E. L. Schymanski *et al.*, Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* **48**, 2097-2098 (2014).

101. M. Filarsky *et al.*, GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* **359**, 1259-1263 (2018).

102. S. D. Boltryk *et al.*, CRISPR/Cas9-engineered inducible gametocyte producer lines as a valuable tool for Plasmodium falciparum malaria transmission research. *Nat Commun* **12**, 4806 (2021).

103. F. Ay *et al.*, Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. *Genome Res* **24**, 974-988 (2014).

104. E. M. Bunnik *et al.*, DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite Plasmodium falciparum. *BMC Genomics* **15**, 347 (2014).

105. A. Weiner *et al.*, 3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite Plasmodium falciparum. *Cell Microbiol* **13**, 967-977 (2011).

106. N. Dahan-Pasternak *et al.*, PfSec13 is an unusual chromatin-associated nucleoporin of Plasmodium falciparum that is essential for parasite proliferation in human erythrocytes. *J Cell Sci* **126**, 3055-3069 (2013).

107. J. Kehrer *et al.*, Nuclear Pore Complex Components in the Malaria Parasite Plasmodium berghei. *Sci Rep* **8**, 11249 (2018).

108. L. H. Freitas-Junior *et al.*, Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature* **407**, 1018-1022 (2000).

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

110. J. E. Lemieux *et al.*, Genome-wide profiling of chromosome interactions in Plasmodium falciparum characterizes nuclear architecture and reconfigurations associated with antigenic variation. *Mol Microbiol* **90**, 519-537 (2013).

111. E. M. Bunnik *et al.*, Comparative 3D genome organization in apicomplexan parasites. *Proc Natl Acad Sci U S A* **116**, 3183-3192 (2019).

112. E. M. Bunnik *et al.*, Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. *Nat Commun* **9**, 1910 (2018).

113. A. M. Samuels *et al.*, Efficacy of RTS,S/AS01E malaria vaccine administered according to different full, fractional, and delayed third or early fourth dose regimens in children aged 5-17 months in Ghana and Kenya: an open-label, phase 2b, randomised controlled trial. *Lancet Infect Dis* **22**, 1329-1342 (2022).

114. K. E. Ward, D. A. Fidock, J. L. Bridgford, Plasmodium falciparum resistance to artemisinin-based combination therapies. *Curr Opin Microbiol* **69**, 102193 (2022).

115. C. C. W. J. P. A. R. D. M. S. P. J. M. G. K. C. J., Kalihinol-A, a Highly Functionalized Diisocyano Diterpenoid Antibiotic from a Sponge. *J. Am. Chem. Soc.* **106**, 4644–4646 (1984).

116. L. H. Yang, O. O. Lee, T. Jin, X. C. Li, P. Y. Qian, Antifouling properties of 10beta-formamidokalihinol-A and kalihinol A isolated from the marine sponge Acanthella cavernosa. *Biofouling* **22**, 23-32 (2006).

117. H. Miyaoka, Y. Abe, E. Kawashima, Synthesis of marine diterpene isocyanide (-)-kalihinol Y and diterpene isothiocyanate (-)-10-epi-kalihinol I. *Chem Pharm Bull (Tokyo)* **60**, 1224-1226 (2012).

118. D. Wolf, F. J. Schmitz, New diterpene isonitriles from the sponge Phakellia pulcherrima. *J Nat Prod* **61**, 1524-1527 (1998).

119. M. E. Daub, J. Prudhomme, K. Le Roch, C. D. Vanderwal, Synthesis and potent antimalarial activity of kalihinol B. *J Am Chem Soc* **137**, 4912-4915 (2015).

120. M. E. Daub, J. Prudhomme, C. Ben Mamoun, K. G. Le Roch, C. D. Vanderwal, Antimalarial Properties of Simplified Kalihinol Analogues. *ACS Med Chem Lett* **8**, 355-360 (2017).

121. A. D. Wright *et al.*, Antimalarial activity: the search for marine-derived natural products with selective antimalarial activity. *J Nat Prod* **59**, 710-716 (1996).

122. M. J. Schnermann, R. A. Shenvi, Syntheses and biological studies of marine terpenoids derived from inorganic cyanide. *Nat Prod Rep* **32**, 543-577 (2015).

123. A. D. Wright *et al.*, Inhibition of heme detoxification processes underlies the antimalarial activity of terpene isonitrile compounds from marine sponges. *J Med Chem* **44**, 873-885 (2001).

124. H. H. Lu *et al.*, Synthesis of (+)-7,20-Diisocyanoadociane and Liver-Stage Antiplasmodial Activity of the Isocyanoterpene Class. *J Am Chem Soc* **138**, 7268-7271 (2016).

125. D. F. Taber, B. P. Gunn, Control elements in the intramolecular Diels-Alder reaction. Synthesis of (.+-.)-torreyol. *Journal of the American Chemical Society* **101**, 3992-3993 (1979).

126. S. V. Pronin, C. A. Reiher, R. A. Shenvi, Stereoinversion of tertiary alcohols to tertiary-alkyl isonitriles and amines. *Nature* **501**, 195-199 (2013).

127. E. L. Dahl, P. J. Rosenthal, Multiple antibiotics exert delayed effects against the Plasmodium falciparum apicoplast. *Antimicrob Agents Chemother* **51**, 3485-3490 (2007).

128. K. Kennedy *et al.*, Delayed death in the malaria parasite Plasmodium falciparum is caused by disruption of prenylation-dependent intracellular trafficking. *PLoS Biol* **17**, e3000376 (2019).

129. S. C. Nair *et al.*, Apicoplast isoprenoid precursor synthesis and the molecular basis of fosmidomycin resistance in Toxoplasma gondii. *J Exp Med* **208**, 1547-1559 (2011).

130. C. J. Tonkin, N. S. Struck, K. A. Mullin, L. M. Stimmler, G. I. McFadden, Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites. *Mol Microbiol* **61**, 614-630 (2006).

131. R. F. Waller, M. B. Reed, A. F. Cowman, G. I. McFadden, Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. *EMBO J* **19**, 1794-1802 (2000).

132. J. Meletiadis, S. Pournaras, E. Roilides, T. J. Walsh, Defining fractional inhibitory concentration index cutoffs for additive interactions based on self-drug additive combinations, Monte Carlo simulation analysis, and in vitro-in vivo correlation data for antifungal drug combinations against Aspergillus fumigatus. *Antimicrob Agents Chemother* **54**, 602-609 (2010).

133. T. C. Chou, Derivation and properties of Michaelis-Menten type and Hill type equations for reference ligands. *J Theor Biol* **59**, 253-276 (1976).

134. S. Steinbacher *et al.*, Structural basis of fosmidomycin action revealed by the complex with 2-C-methyl-D-erythritol 4-phosphate synthase (IspC). Implications for the catalytic mechanism and anti-malaria drug development. *J Biol Chem* **278**, 18401-18407 (2003).

135. E. Yeh, J. L. DeRisi, Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage Plasmodium falciparum. *PLoS Biol* **9**, e1001138 (2011).

136. T. Uddin, G. I. McFadden, C. D. Goodman, Validation of Putative Apicoplast-Targeting Drugs Using a Chemical Supplementation Assay in Cultured Human Malaria Parasites. *Antimicrob Agents Chemother* **62**, (2018).

137. S. E. Lindner *et al.*, Enzymes involved in plastid-targeted phosphatidic acid synthesis are essential for Plasmodium yoelii liver-stage development. *Mol Microbiol* **91**, 679-693 (2014).

138. G. Pessi, G. Kociubinski, C. B. Mamoun, A pathway for phosphatidylcholine biosynthesis in Plasmodium falciparum involving phosphoethanolamine methylation. *Proc Natl Acad Sci U S A* **101**, 6206-6211 (2004).

139. N. Kilian, J. Y. Choi, D. R. Voelker, C. Ben Mamoun, Role of phospholipid synthesis in the development and differentiation of malaria parasites in the blood. *J Biol Chem* **293**, 17308-17316 (2018).

140. X. W. Chan *et al.*, Chemical and genetic validation of thiamine utilization as an antimalarial drug target. *Nat Commun* **4**, 2060 (2013).

141. J. Knockel *et al.*, Filling the gap of intracellular dephosphorylation in the Plasmodium falciparum vitamin B1 biosynthesis. *Mol Biochem Parasitol* **157**, 241-243 (2008).

142. E. E. Reid, P. Thompson, C. R. Lyttle, D. T. Dennis, Pyruvate dehydrogenase complex from higher plant mitochondria and proplastids. *Plant Physiol* **59**, 842-848 (1977).

143. J. P. Musabyimana *et al.*, Plasmodium falciparum S-Adenosylmethionine Synthetase Is Essential for Parasite Survival through a Complex Interaction Network with Cytoplasmic and Nuclear Proteins. *Microorganisms* **10**, (2022).

144. Y. Ouyang, Q. Wu, J. Li, S. Sun, S. Sun, S-adenosylmethionine: A metabolite critical to the regulation of autophagy. *Cell Prolif* **53**, e12891 (2020).

145. H. Choi, D. Fermin, A. I. Nesvizhskii, Significance analysis of spectral count data in label-free shotgun proteomics. *Mol Cell Proteomics* **7**, 2373-2385 (2008).

146. Y. Zhang, Z. Wen, M. P. Washburn, L. Florens, Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem* **82**, 2272-2281 (2010).

147. C. D. Broeckling, F. A. Afsar, S. Neumann, A. Ben-Hur, J. E. Prenni, RAMClust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data. *Anal Chem* **86**, 6812-6817 (2014).

148. H. Choi, S. Kim, D. Fermin, C. C. Tsou, A. I. Nesvizhskii, QPROT: Statistical method for testing differential expression using protein-level intensity data in label-free quantitative proteomics. *J Proteomics* **129**, 121-126 (2015).

149. M. Deponte *et al.*, Wherever I may roam: protein and membrane trafficking in P. falciparum-infected red blood cells. *Mol Biochem Parasitol* **186**, 95-116 (2012).

150. R. Tuteja, Unraveling the components of protein translocation pathway in human malaria parasite Plasmodium falciparum. *Arch Biochem Biophys* **467**, 249-260 (2007).

151. A. S. Mamidi, A. Ray, N. Surolia, Structural Analysis of PfSec62-Autophagy Interacting Motifs (AIM) and PfAtg8 Interactions for Its Implications in RecovER-phagy in Plasmodium falciparum. *Front Bioeng Biotechnol* **7**, 240 (2019).

152. H. Fang, N. Green, Nonlethal sec71-1 and sec72-1 mutations eliminate proteins associated with the Sec63p-BiP complex from S. cerevisiae. *Mol Biol Cell* **5**, 933-942 (1994).

153. A. Tripathi, E. C. Mandon, R. Gilmore, T. A. Rapoport, Two alternative binding mechanisms connect the protein translocation Sec71-Sec72 complex with heat shock proteins. *J Biol Chem* **292**, 8007-8018 (2017).

154. D. Martinez Molina *et al.*, Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84-87 (2013).

155. J. M. Dziekan *et al.*, Cellular thermal shift assay for the identification of drug-target interactions in the Plasmodium falciparum proteome. *Nat Protoc* **15**, 1881-1921 (2020).

156. A. Florentin, D. W. Cobb, H. M. Kudyba, V. Muralidharan, Directing traffic: Chaperone-mediated protein transport in malaria parasites. *Cell Microbiol* **22**, e13215 (2020).

157. M. M. Savitski *et al.*, Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784 (2014).

158. C. Nagaraj, N. K. Chakrabarati, M. V. Narasimham, A longitudinal study of P. falciparum response to chloroquine in relation to the epidemiology. *J Commun Dis* **21**, 229-240 (1989).

159. T. Masserey *et al.*, The influence of biological, epidemiological, and treatment factors on the establishment and spread of drug-resistant Plasmodium falciparum. *Elife* **11**, (2022).

160. A. L. Malisa *et al.*, Drug coverage in treatment of malaria and the consequences for resistance evolution--evidence from the use of sulphadoxine/pyrimethamine. *Malar J* **9**, 190 (2010).

161. T. J. Green *et al.*, Cloning and characterization of Plasmodium falciparum FCR-3/FMG strain. *Am J Trop Med Hyg* **34**, 24-30 (1985).

162. M. V. Katti, R. Sami-Subbu, P. K. Ranjekar, V. S. Gupta, Amino acid repeat patterns in protein sequences: their diversity and structural-functional implications. *Protein Sci* **9**, 1203-1209 (2000).

163. M. B. Jimenez-Diaz *et al.*, Improved murine model of malaria using Plasmodium falciparum competent strains and non-myelodepleted NOD-scid IL2Rgammanull mice engrafted with human erythrocytes. *Antimicrob Agents Chemother* **53**, 4533-4536 (2009).

164. S. Tachibana *et al.*, Plasmodium cynomolgi genome sequences provide insight into Plasmodium vivax and the monkey malaria clade. *Nat Genet* **44**, 1051-1055 (2012).

165. N. M. B. Brancucci *et al.*, Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite Plasmodium falciparum. *Cell* **171**, 1532-1544 e1515 (2017).

166. P. K. Sheokand *et al.*, A Plasmodium falciparum lysophospholipase regulates host fatty acid flux via parasite lipid storage to enable controlled asexual schizogony. *Cell Rep* **42**, 112251 (2023).

167. G. G. Holz, Jr., Lipids and the malarial parasite. *Bull World Health Organ* **55**, 237-248 (1977).

168. B. J. Foth *et al.*, The malaria parasite Plasmodium falciparum has only one pyruvate dehydrogenase complex, which is located in the apicoplast. *Mol Microbiol* **55**, 39-53 (2005).

169. A. Mateus, T. A. Maatta, M. M. Savitski, Thermal proteome profiling: unbiased assessment of protein state through heat-induced stability changes. *Proteome Sci* **15**, 13 (2016).

170. Y. Harada, H. Li, J. S. Wall, H. Li, W. J. Lennarz, Structural studies and the assembly of the heptameric post-translational translocon complex. *J Biol Chem* **286**, 2956-2965 (2011).

171. D. Prusty *et al.*, Single-stranded DNA binding protein from human malarial parasite Plasmodium falciparum is encoded in the nucleus and targeted to the apicoplast. *Nucleic Acids Res* **38**, 7037-7053 (2010).

172. R. Prekeris, J. Klumperman, Y. A. Chen, R. H. Scheller, Syntaxin 13 mediates cycling of plasma membrane proteins via tubulovesicular recycling endosomes. *J Cell Biol* **143**, 957-971 (1998).

173. L. Ayong, G. Pagnotti, A. B. Tobon, D. Chakrabarti, Identification of Plasmodium falciparum family of SNAREs. *Mol Biochem Parasitol* **152**, 113-122 (2007).

174. L. A. Parish, J. C. Rayner, Plasmodium falciparum secretory pathway: characterization of PfStx1, a plasma membrane Qa-SNARE. *Mol Biochem Parasitol* **164**, 153-156 (2009).

175. J. Y. van der Meer, A. K. Hirsch, The isoprenoid-precursor dependence of Plasmodium spp. *Nat Prod Rep* **29**, 721-728 (2012).

176. C. Ungermann, D. Kummel, Structure of membrane tethers and their role in fusion. *Traffic* **20**, 479-490 (2019).

177. T. Wang, L. Li, W. Hong, SNARE proteins in membrane trafficking. *Traffic* **18**, 767-775 (2017).

178. J. Fu *et al.*, Apicoplast biogenesis mediated by ATG8 requires the ATG12-ATG5-ATG16L and SNAP29 complexes in Toxoplasma gondii. *Autophagy*, 1-19 (2022).

179. F. Rohdich *et al.*, Biosynthesis of terpenoids. 2C-Methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspF) from Plasmodium falciparum. *Eur J Biochem* **268**, 3190-3197 (2001).

180. E. L. Flannery, D. A. Fidock, E. A. Winzeler, Using genetic methods to define the targets of compounds with antimalarial activity. *J Med Chem* **56**, 7761-7771 (2013).

181. G. LaMonte *et al.*, Mutations in the Plasmodium falciparum Cyclic Amine Resistance Locus (PfCARL) Confer Multidrug Resistance. *mBio* **7**, (2016).

182. S. Menard *et al.*, Induction of Multidrug Tolerance in Plasmodium falciparum by Extended Artemisinin Pressure. *Emerg Infect Dis* **21**, 1733-1741 (2015).

183. F. Ariey *et al.*, A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* **505**, 50-55 (2014).

184. C. Lambros, J. P. Vanderberg, Synchronization of Plasmodium falciparum erythrocytic stages in culture. *J Parasitol* **65**, 418-420 (1979).

185. W. Trager, J. B. Jensen, Human malaria parasites in continuous culture. 1976. *J Parasitol* **91**, 484-486 (2005).

186. A. K. Bei, M. T. Duraisingh, Measuring Plasmodium falciparum Erythrocyte Invasion Phenotypes Using Flow Cytometry. *Methods Mol Biol* **1325**, 167-186 (2015).

187. C. Amaratunga, A. T. Neal, R. M. Fairhurst, Flow cytometry-based analysis of artemisinin-resistant Plasmodium falciparum in the ring-stage survival assay. *Antimicrob Agents Chemother* **58**, 4938-4940 (2014).

188. C. J. Joyner *et al.*, Humoral immunity prevents clinical malaria during Plasmodium relapses without eliminating gametocytes. *PLoS Pathog* **15**, e1007974 (2019).

189. L. W. Sumner *et al.*, Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 211-221 (2007).

190. B. MacLean *et al.*, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966-968 (2010).

191. K. Rajaram, H. B. Liu, S. T. Prigge, Redesigned TetR-Aptamer System To Control Gene Expression in Plasmodium falciparum. *mSphere* **5**, (2020).

192. J. M. Ribeiro *et al.*, Guide RNA selection for CRISPR-Cas9 transfections in Plasmodium falciparum. *Int J Parasitol* **48**, 825-832 (2018).

193. S. H. Adjalley, M. C. Lee, D. A. Fidock, A method for rapid genetic integration into Plasmodium falciparum utilizing mycobacteriophage Bxb1 integrase. *Methods Mol Biol* **634**, 87-100 (2010).

194. T. Xu *et al.*, ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J Proteomics* **129**, 16-24 (2015).

195. D. L. Tabb, W. H. McDonald, J. R. Yates, 3rd, DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J Proteome Res* **1**, 21-26 (2002).

196. Z. Wen, kite: a software suite for processing and analysis of tandem mass spectrometry data. *Zenodo* (2022).

197. D. A. Fidock, T. Nomura, T. E. Wellems, Cycloguanil and its parent compound proguanil demonstrate distinct activities against Plasmodium falciparum malaria parasites transformed with human dihydrofolate reductase. *Mol Pharmacol* **54**, 1140-1147 (1998).

198. L. Florens, M. P. Washburn, Proteomic analysis by multidimensional protein identification technology. *Methods Mol Biol* **328**, 159-175 (2006).

199. S. Mnaimneh *et al.*, Exploration of essential gene functions via titratable promoter alleles. *Cell* **118**, 31-44 (2004).

CHAPTER 3

PfMORC protein regulates chromatin accessibility and transcriptional repression in the human malaria parasite, *P. falciparum*

Chahine Z[1#], Gupta M[1†#], Lenz T[1#], Hollin T[1#], Abel S[1], Banks CAS[2], Saraf A[2$], Prudhomme J[1], Bhanvadia S[1], Florens L[2] and Le Roch KG[1*].

[1] Department of Molecular, Cell and Systems Biology, University of California Riverside, CA, USA

[2] Stowers Institute for Medical Research, 1000 E. 50th Street, Kansas City, MO 64110, USA

[†] Current address: Cancer Early Detection Advanced Research Center, Oregon Health and Science University, S Moody Avenue, Portland, OR 97201, USA

[$] Current address: Shankel Structural Biology Center, The University of Kansas, Lawrence, KS, USA

[#] Contributed equally to this work

[*] Correspondence: Karine Le Roch (karine.leroch@ucr.edu).

A version of this chapter has been published in *eLife.,2023*.

PREFACE

Our understanding of the mechanisms used by the eukaryotic pathogen *Plasmodium falciparum* to regulate gene and protein expression, as well as protein complexes throughout its life cycle, remains incomplete and poses a significant barrier in malaria research. In model eukaryotes, gene expression is regulated through a complex network of interactions and protein complexes that play crucial roles in transcription regulation, mRNA stabilization/turnover, transport, and translation efficiency. Microrchidia (MORC) proteins have been implicated in DNA compaction and gene silencing across plant and animal kingdoms. However, the functional impact of this protein on *P. falciparum* is not yet fully understood. In this chapter, I work in collaboration with

several colleagues to investigate the role of the *Pf*MORC protein. In this project we performed all experimentation, analyzed all datasets, and wrote the finalized manuscript presented herein. We provide significant evidence of the importance of this protein in regulating heterochromatin stability and gene silencing in *Plasmodium*.

ABSTRACT

The environmental challenges the human malaria parasite, *Plasmodium falciparum*, faces during its progression into its various life cycle stages warrant the use of effective and highly regulated access to chromatin for transcriptional regulation. Microrchidia (MORC) proteins have been implicated in DNA compaction and gene silencing across plant and animal kingdoms. Accumulating evidence has shed light into the role MORC protein plays as a transcriptional switch in apicomplexan parasites. In this study, using CRISPR/Cas9 genome editing tool along with complementary molecular and genomics approaches, we demonstrate that *PfMORC* not only modulates chromatin structure and heterochromatin formation throughout the parasite erythrocytic cycle, but is also essential to the parasite survival. Chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) experiments suggest that *PfMORC* binds to not only sub-telomeric regions and genes involved in antigenic variation but is also most likely a key modulator of stage transition. Protein knockdown experiments followed by chromatin conformation capture (Hi-C) studies indicate that downregulation of *PfMORC* impact histone marks and induces the collapse of the parasite heterochromatin structure leading to its death. All together these

findings confirm that *PfMORC* plays a crucial role in chromatin structure and gene regulation, validating this factor as a strong candidate for novel antimalarial strategies.

INTRODUCTION

Malaria is a mosquito-borne infectious disease that is caused by protozoan parasites of the genus *Plasmodium*. Among the five human-infecting species, *Plasmodium falciparum* is the deadliest, with over 619,000 deaths in 2023 (1). To adapt to extreme environmental challenges, *P. falciparum* possesses unique strategies that direct the tightly coordinated changes in gene expression and control transcriptional switching in genes encoded by multigene families to ensure antigenic variation and immune evasion. Changes in gene expression throughout the parasite life cycle are controlled by a surprisingly low repertoire of transcription factors (TFs) encoded in the *Plasmodium* genome(2-9). The 27 apicomplexan APETALA2 (ApiAP2) DNA-binding proteins are the only well documented parasite TFs known to contribute to the modulation of gene expression throughout various stages of the parasite's development (3,5-8,10-13). Since the discovery of the AP2 gene family (3), evidence has alluded to their role as master regulators of gene expression throughout transitory phases of the parasite life cycle. The most well documented are the gametocyte specific TFs (AP2-G, AP2-G2) (7,14) with subsequent discovery of those associated with transition to sexual differentiation 13 as well as sporozoite (AP2-SP) and liver stages (AP2-L) (6,1). The *Plasmodium* genome, however, encompasses well over 5,000 protein-coding genes suggesting there are most likely other molecular components responsible for gene expression. It is believed that, to offset the relatively low TF range,

Plasmodium has evolved additional mechanisms regulating gene expression including mechanisms that use epigenetics factors, RNA binding proteins or regulate chromatin structure(15-20). All together these components work in combination to regulate the dynamic organization of DNA and the cascade of gene expression required for the parasite life cycle progression (21-28). Despite their importance, the identification and functional characterization of regulatory players controlling chromatin remains a challenge in this intractable organism. However, with the advent of sensitive technologies capable of capturing important chromatin associated regulatory complexes such as chromatin immunoprecipitation (ChIP), chromatin conformation capture Hi-C technologies, and chromatin enrichment for proteomics (ChEP), we can now solve important facets of molecular components controlling epigenetics and chromatin structure.

Microrchidia (MORC) belong to a highly conserved nuclear protein superfamily with widespread domain architectures that link MORCs with signaling-dependent chromatin remodeling and epigenetic regulation across plant and animal kingdoms including the apicomplexan parasites (29-32). In all organisms, MORC contains several domains that form a catalytically active ATPase. Apicomplexan MORC ATPases are encircled by Kelch-repeat β propellers as well as a CW-type zinc finger domain functioning as a histone reader (33). In higher eukaryotes, MORCs were first identified as epigenetic regulators and chromatin remodelers in germ cell development. Currently, these proteins are shown to be involved in various human diseases including cancers and are expected to serve as important biomarkers for diagnosis and treatment (34). In the apicomplexan parasite,

*Toxoplasma gondii, Tg*MORC was shown to recruit the histone deacetylase HDAC3 to particular genome loci to regulate chromatin accessibility, restricting sexual commitment (33,35). *Tg*MORC-depleted cells also resulted in change in gene expression with up regulation of secondary AP2 factors and a significant shift from asexual to sexual differentiation (32,33,36-38). Having multiple homologs with *T. gondii*, *Plasmodium* AP2 protein conservation is primarily restricted to their AP2 DNA-binding domains.(39,40). Recently, ChEP experiments done in *P. falciparum* identified *Pf*MORC as one of the highest enriched chromatin-bound proteins at different stages of the parasite intraerythrocytic development cycle (IDC)(28). *Pf*MORC was also detected at a relatively high level throughout the parasite life cycle, including sporozoites and liver stages(33,41), and has been identified in several protein pull down experiments as targeting both AP2 TFs and epigenetic factors(32,33,38,42-44). Immunoprecipitation experiments demonstrated that *Pf*MORC seems to interact with AP2-G2(32), a TF that plays a critical role in the maturation of gametocytes. However, the genome wide distribution and exact function of this protein throughout the parasite life cycle remained elusive.

In this study, we apply CRISPR/Cas9 genomic editing technologies to determine the function, genome distribution, and indispensability of *Pf*MORC throughout the parasite's IDC. Immunoprecipitation of an HA-tagged parasite line validates the role of *Pf*MORC in heterochromatin structure maintenance. Using downregulation of *Pf*MORC induced through TetR-DOZI system, we demonstrate the functional significance of *Pf*MORC in heterochromatin stability and gene repression. Immunofluorescence based

assays and ChIP-seq experiments show that *Pf*MORC localizes to heterochromatin clusters at or near *var* genes with significant overlap with well-known signatures of the parasite heterochromatin post-translational modification (PTM) H3K9-trimethylation (H3K9me3) marks. When *Pf*MORC was downregulated, level of H3K9me3 was detected at a lower level, demonstrating a possible role of this protein *PfMORC* in epigenetic regulation and gene repression. Finally, Hi-C analyses demonstrate that downregulation of *Pf*MORC results in significant dysregulation of chromatin architectural stability resulting in parasite death. All together our work provides significant insight into the role of *Pf*MORC in the maintenance of the pathogens' epigenetics and chromosomal architectural integrity and validates this protein as a promising target for novel therapeutic interventions.

RESULTS

Generation of *Pf*MORC-HA transgenic line shows nuclear localization in heterochromatin clusters

To characterize the role of *Pf*MORC in *P. falciparum*, we applied CRISPR/Cas9 genome editing tool to add a 3X-HA tag at the C-terminal coding region of the *Pf*MORC locus (*PF3D7_1468100*) in an NF54 line (45**) (Fig. 3-1a, Sup. data 3-1).** Recovered transgenic parasites were cloned using limiting dilution and the correct incorporation of the tag into the parasite genome was validated via PCR and whole genome sequencing (WGS) **(Fig. 3-1b, Sup. data 3-2).** WGS analysis confirmed the presence of the HA tag at the expected locus with no obvious off-target effect but uncovered an additional nonsense mutation in the gametocyte development protein 1 (*gdv1*) gene (*PF3D7_0935400*) at amino acid

position 561 out of 599 total. Mutations in *gdv1* have been detected in the past and seem to emerge relatively often indicating a fitness benefit in the *in vitro* culture system used in the laboratory (8). The involvement of GDV1 as essential in sexual commitment is well substantiated and further study involving the role of *Pf*MORC in sexual commitment could not be fully addressed (7,8,14,46-48) **(Sup. data 3-2).** Despite this finding, we were able to design experiments to determine the role of *Pf*MORC throughout the IDC as no other major detrimental mutations were detected in the genome. We first validated *Pf*MORC protein expression through Western blot analysis **(Fig. 3-S1).** Regardless of potential protein degradation, results showed an expected band size within the ~290 kDa range that was absent in NF54 control compared to our tagged line **(Fig. 3-S1).** Immunofluorescence (IFA) analysis of intracellular parasites revealed that *Pf*MORC is localized in the nucleus throughout the parasite IDC in punctate patterns **(Fig. 3-1c).** A single foci per parasite was detected in the nuclear periphery at the ring stage. An increased number of foci with a more diffuse signal could be detected as the number of DNA copies increased at the schizont stage. Moreover, punctuates were shown to have strong colocalization signals with histone H3K9me3 marks present at the different parasite developmental stages analyzed. This was most evident during the early ring stage of the parasite IDC where the chromatin organization is more compact with a single DNA copy. Altogether, our findings indicated that *Pf*MORC is a true nuclear protein that is most likely associated with *P. falciparum* heterochromatin cluster(s).

***Pf*MORC interacting partners include proteins involved in heterochromatin maintenance, transcription regulation and chromatin remodeling**

To confirm the association of *Pf*MORC with proteins involved in heterochromatin cluster(s), we performed immunoprecipitation (IP) followed by mass-spectrometry (MS) analysis using mature stages of *P*fMORC-HA and parental line as control. Our IP-MS analysis matched an average of 1109 spectra to 105 unique peptide sequences corresponding to the *Pf*MORC protein consistent with HA-affinity enrichment of *Pf*MORC. MORC was detected with highest peptides (97 and 113) and spectra (1041 and 1177) counts compared to control conditions (5 and 7 peptides; 16 and 43 spectra) confirming the efficiency of our pull-down. However, considering the relatively large size of the MORC protein (295kDa) and its weak detection in the control, the $\log_2$ FC and Z-statistic after normalization are minimal when compared to smaller proteins that were not identified in the control samples. To define a set of *Pf*MORC associated proteins, we used the QPROT statistical framework(49) to compare proteins detected in *Pf*MORC samples and controls. We identified 73 *Pf*MORC associated proteins ($\log_2$ FC ≥ 2 and Z statistic > 5) (**Fig. 3-1d, Sup. data 3**)**.** Additional chromatin remodeling proteins were detected such as ISWI chromatin-remodeling complex ATPase (PF3D7_0624600) and SWIB/MDM2 domain-containing protein (PF3D7_0611400)(50), as well as the epigenetic readers BDP3 (PF3D7_0110500) and PHD2 (PF3D7_1433400). We also detected a significant enrichment of histone erasers and writers such as a N-acetyltransferase (PF3D7_1020700), HDAC1 (PF3D7_0925700), and SET3 (PF3D7_0827800). Several of which have been

detected to interact with MORC in previous studies including the AP2-G5, HDAC1, ELM2 and ApiAp2 proteins(33,42-44)(**Fig. 3-1e**). Notably, we observed that many of the detected proteins were enriched at levels comparable to those of the bait protein such as AP2-P and HDAC1. This may be a result of proteins forming complexes in a one-to-one ratio, leading to similar enrichment levels. Notably, two of these three proteins have been reported to interact with MORC in several studies, further supporting a strong interaction between them. Gene Ontology (GO) enrichment analysis revealed that these proteins are nuclear and are associated with heterochromatin, DNA binding, and transcription factor activity (**Fig. 3-1f**). Detailed analysis showed the detection of SIP2, involved in heterochromatin formation and chromosome end regulation(51), and 8 AP2 transcription factors, including AP2-G2, AP2-O2, AP2-G5. Among the other AP2s detected, 5 were previously identified by ChIP-seq as *Plasmodium falciparum* AP2 Heterochromatin-associated Factors (*Pf*AP2-HFs)(52). Our findings, corroborated with several published works, indicate that *Pf*MORC interacts with multiple ApiAP2 TFs, chromatin remodelers, and epigenetic players associated with heterochromatin regions.

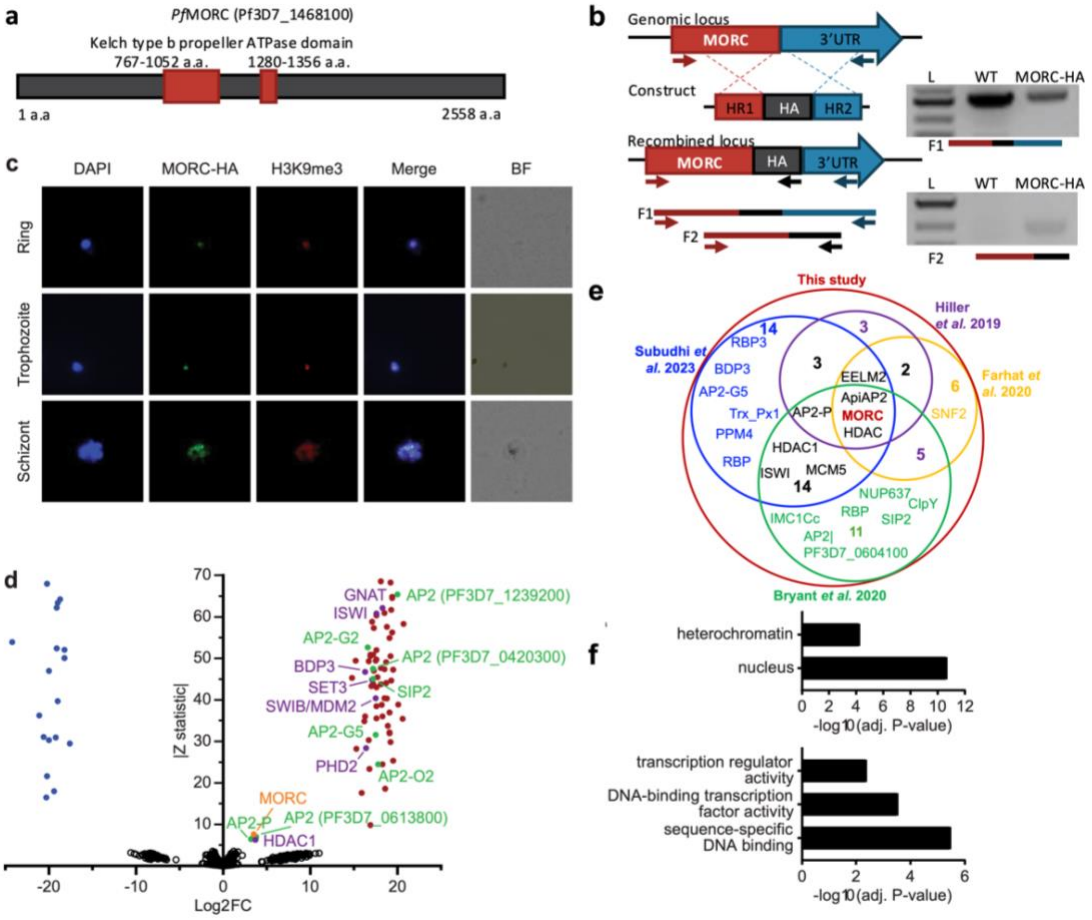


Fig. 3-1. *Pf*MORC-HA is associated with heterochromatin. (a) Illustration of the *Pf*MORC containing domains including Kelch type b propeller and ATPase domains using InterProScan. **(b)** Design strategy applied for *Pf*MORC C-terminal HA tagging. PCR amplification of the genomic C-terminus end of *Pfmorc* region extending towards the 3'UTR (F1) as well as extension from the C-terminus towards the HA flanking sequence (F2) verifies correct insertion site. NF54 genomic DNA was used as negative control. **(c)** IFA experiment: *Pf*MORC foci 22 expressing colocalization with H3K9me3 marks (red). Cell nuclei are stained with DAPI (blue). BF: brightfield **(d)** Protein immunoprecipitation: Significance plot representing *Pf*MORC interactome recovered through immunoprecipitation followed by mass spectrometry (IP-MS). Graph list MORC (orange) bindings partners of highest affinity associated with TF regulation (green) and chromatin remodelers, erasers, and writers (purple). Proteins enriched in the *Pf*MORC-HA samples compared with controls were filtered with log2 FC $\geq$ 2 and Z statistic > 5. **(e)** Venn diagram representing overlapping proteins identified among five publications. numbers represent the total number of significant proteins identified between two overlapping subsets**. (f)** Gene ontology enrichment analysis of the significantly enriched proteins. The top 2 terms of Cellular Component (top) and top 3 terms of Molecular Function (bottom) are represented as -Log_{10} (adjusted P-value) (Fisher's exact test with Bonferroni adjustment).

Genome-wide distribution of *Pf*MORC displays functional roles in stage-specific gene silencing and antigenic variation.

Given the multifunctional role of MORC proteins throughout apicomplexan parasites, we further explored the association of *Pf*MORC with chromatin accessibility and stage-specific gene regulation by assessing the genome-wide distribution of *Pf*MORC using chromatin immunoprecipitation followed by deep sequencing (ChIP–seq) in duplicate at the ring, trophozoite and schizont stages. Inspection of *Pf*MORC binding sites across each chromosome revealed a strong signal for telomeric and subtelomeric regions of the parasite genome, as well as the internal clusters of genes responsible for antigenic variation such as *var* and *rifin* (**Fig. 3-2a-c, Sup. data 4**). We found that while only 47% of reads mapped to antigenic genes in the ring stage, 95% of reads mapped to the same genes during the trophozoite stage and 82% during the schizont stage. Although a vast majority of the genome is actively transcribed during the trophozoite stage, these antigenic gene families are under tight regulatory control to ensure mutually exclusive expression and participate in immune evasion(53,54). This could explain the disparity in the level of *Pf*MORC binding of the coding regions between the ring and trophozoite stages (**Fig.3- 2c, Fig. S2b**).

We also evaluated *Pf*MORC binding of stage-specific gene families including gametocyte-related genes, and merozoite surface proteins (*msp*)(55). In trophozoite and schizont stages, gametocyte-associated genes contain a mean of <0.5 RPKM normalized reads per nucleotide of *PfMORC* binding within their promoter region, whereas antigenic gene families such as *var* and *rifin* contain ~1.5 and 0.5 normalized reads, respectively **(Fig. 3-2b).** The difference is even greater within the gene body with gametocyte genes displaying almost no reads mapped more than 200 bp downstream of the TSS and antigenic gene families containing 0.5-1.5 RPKM normalized reads. However, the gametocyte-specific transcription factor AP2-G, known to be repressed during the IDC and required for sexual commitment, deviates from this trend and contains similar levels of *PfMORC* binding to *var* genes **(Fig. 3-2c and S2c, d).** These results indicate a major role of *Pf*MORC in controlling AP2-G and sexual differentiation. For some of the *msp* genes that are usually expressed later during the IDC to prepare the parasite for egress and invasion of new erythrocytes, *Pf*MORC binding was detected in the gene bodies at the trophozoite stage. We also detected a small switch in *Pf*MORC binding sites, moving from their gene bodies to their intergenic and promoter regions at the trophozoite stage (**Fig. S2d**). *Pf*MORC most likely moved away from the gene body to the regulatory regions surrounding the transcription start site (TSS) to guide RNA Polymerase and transcription factors and aid in

activating expression of genes at the schizont stage that are crucial for egress and invasion. These results provide strong evidence for the direct effects that *Pf*MORC binding has on tightly controlled antigenic and stage-specific genes including crucial gametocyte genes.

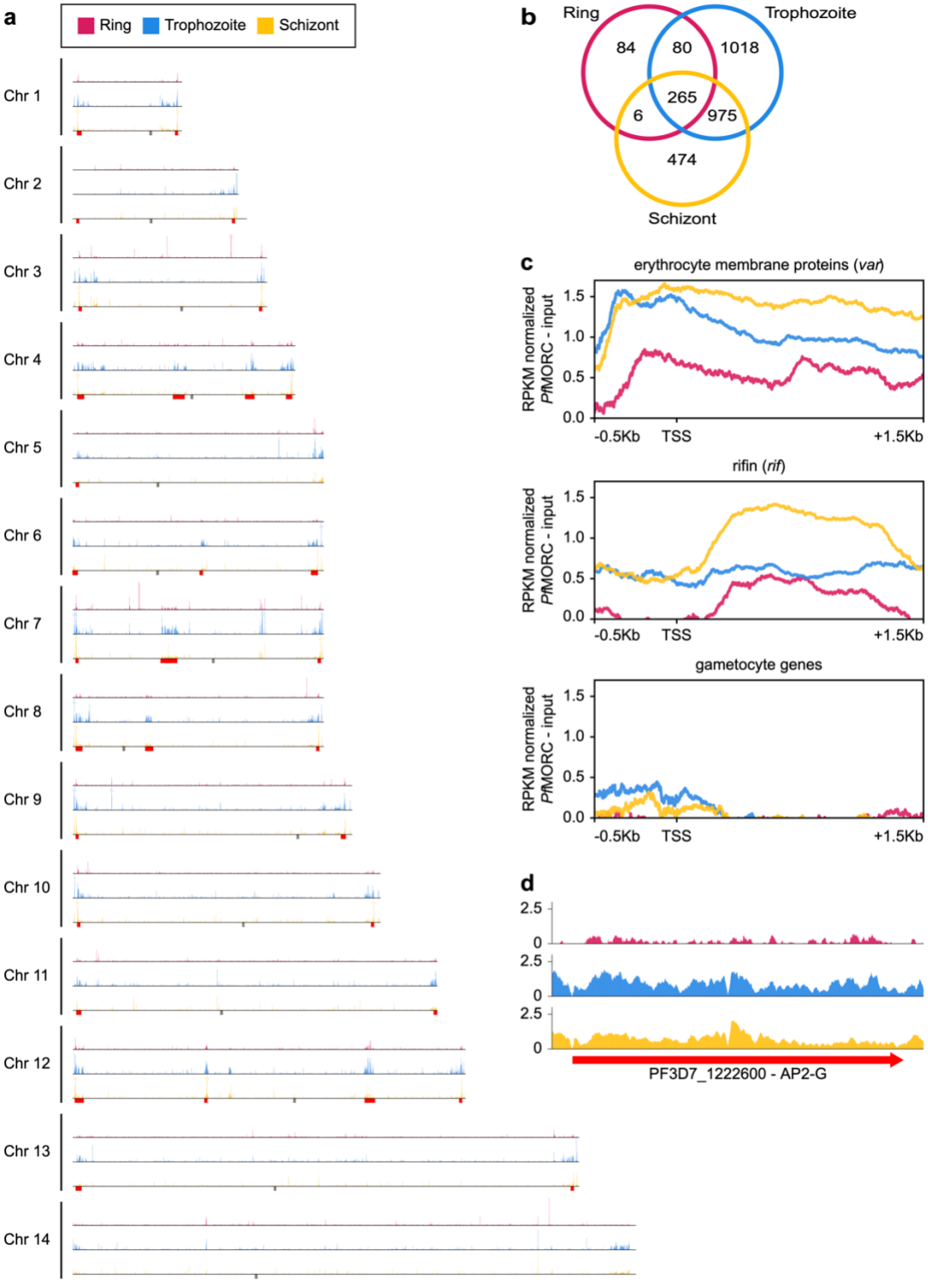

a
Ring
Trophozoite
Schizont
Chr 1
Chr 2
Chr 3
Chr 4
Chr 5
Chr 6
Chr 7
Chr 8
Chr 9
Chr 10
Chr 11
Chr 12
Chr 13
Chr 14
b
Ring
Trophozoite
Schizont
84
80
1018
265
6
975
474
c
erythrocyte membrane proteins (var)
RPKM normalized PfMORC - input
1.5
1.0
0.5
0.0
-0.5Kb
TSS
+1.5Kb
rifin (rif)
gametocyte genes
d
2.5
0
PF3D7_1222600 - AP2-G

Fig. 2. Genome wide distribution of *Pf*MORC proteins. (a) Chromosome distribution plots of *Pf*MORC binding showing a predisposition for subtelomeric and internal *var* gene regions (red). Each track is input subtracted, and per-million read count normalized before normalizing the track height to allow for direct comparison between stages. Grey boxes indicate the position of centromeres. **(b)** Overlap of called peaks between time points. **(c)** Profile plots showing *Pf*MORC coverage from 0.5 kb 5' of the transcription start site (TSS) to 1.5 kb 3' of the TSS in the ring, trophozoite, and schizont stages. Each plot includes per-million read count normalized coverage at 1 bp resolution for all genes within the *var* and *rifin* gene families as well as gametocyte-specific genes. **(d)** *Pf*MORC coverage of the gametocyte-specific transcription factor *ap2-g*.

*Pf*MORC is essential for *P. falciparum* survival

We next sought to confirm the functional relevance of *Pf*MORC protein in parasite survival. While partial *Pf*MORC knockdown using the glmS-ribozyme system had previously shown to have no significant effect on the parasite survival (41), another study using transposon mutagenesis (piggyBac) identified *Pf*MORC as likely essential (56). To resolve these conflicting results, we applied a complementary approach using the CRISPR/Cas9 gene editing strategy to incorporate an inducible TetR-DOZI system to knockdown (KD) *Pf*MORC (57) through administration of anhydrotetracycline (aTC) **(Fig. 3-3a)** (45,58). The protein was also modified to include a C-terminal 3x-HA tag. Parental and transgenic clones were validated via PCR **(Fig. 3-3b)** and WGS to confirm the correct insertion of the inducible system **(Sup. data 2)** as well as the absence of major mutation that could explain some of the phenotypes observed. While results from our WGS validated our editing strategy without the detection of any obvious deleterious off-target effect, we identified a nonsense mutation in AP2-G gene (*Pf3D7_1222600*) explaining our inability to obtain mature gametocytes. Protein expression and successful KD was confirmed via western blot with ~ 58% and ~ 88.2% decrease in *Pf*MORC expression at 24h and 36h after aTC removal, respectively **(Fig. 3-S3).** This observed downregulation of *Pf*MORC was slightly above the level observed by *Singh* and colleagues (41).

We then performed a phenotypic assay on (+/-) aTC *Pf*MORC-HA-TetR-DOZI parasite cultures. The assay was conducted in replicates of synchronized cultures that were then split into (+) aTC and (-) aTC conditions either at the ring or trophozoite stages on

*Pf*MORC-HA-TetRDOZI clones and WT control lines **(Fig. 3-3c-f, Sup. data 5).** Sequential morphological screens were performed through Giemsa-stained smears and monitored by microscope imaging. As opposed to what was observed previously with the glmS-ribozyme system (41), aTC removal at the ring stage induced clear signs of stress and cell cycle arrest in mid trophozoite and schizont stages of the first intraerythrocytic cycle **(Fig. 3-3c-d)** compared to (+) aTC *Pf*MORC and WT parasites. Our data showed an ~53%, 77%, and 84% (n=3, $p \leq 0.0001$) drop in parasitemia at 48-, 72- and 120-hours post invasion (hpi), respectively, when compared to control conditions. aTC supplemented and WT cultures showed unperturbed cell cycle progression, reinvasion, and morphological development. Interestingly, when aTC was withheld at the late trophozoite stages (24 to 30 hpi), parasites could complete the first cycle and successfully reinvade into new red blood cells (RBCs**) (Fig. 3-3e-f).** The evident hindered phenotypic response may either be associated with the delay time for the protein to be completely depleted or may suggest a more dispensable role of *Pf*MORC in the schizont stage. However, clear indications of stress and cell cycle arrest were ultimately detected at trophozoite and early schizont stages of the second cell cycle. Quantitative analysis of (-) aTC *Pf*MORC cultures revealed significantly decreased parasitemia of ~37% and 66% (n=3, $p \leq 0.0001$) at 96 hpi and 120 hpi, respectively, compared to (+) aTC *Pf*MORC and WT control lines **(Fig. 3-3e-f, Sup.**

data 3-5), confirming the importance of the protein at the trophozoite and early schizont stages.

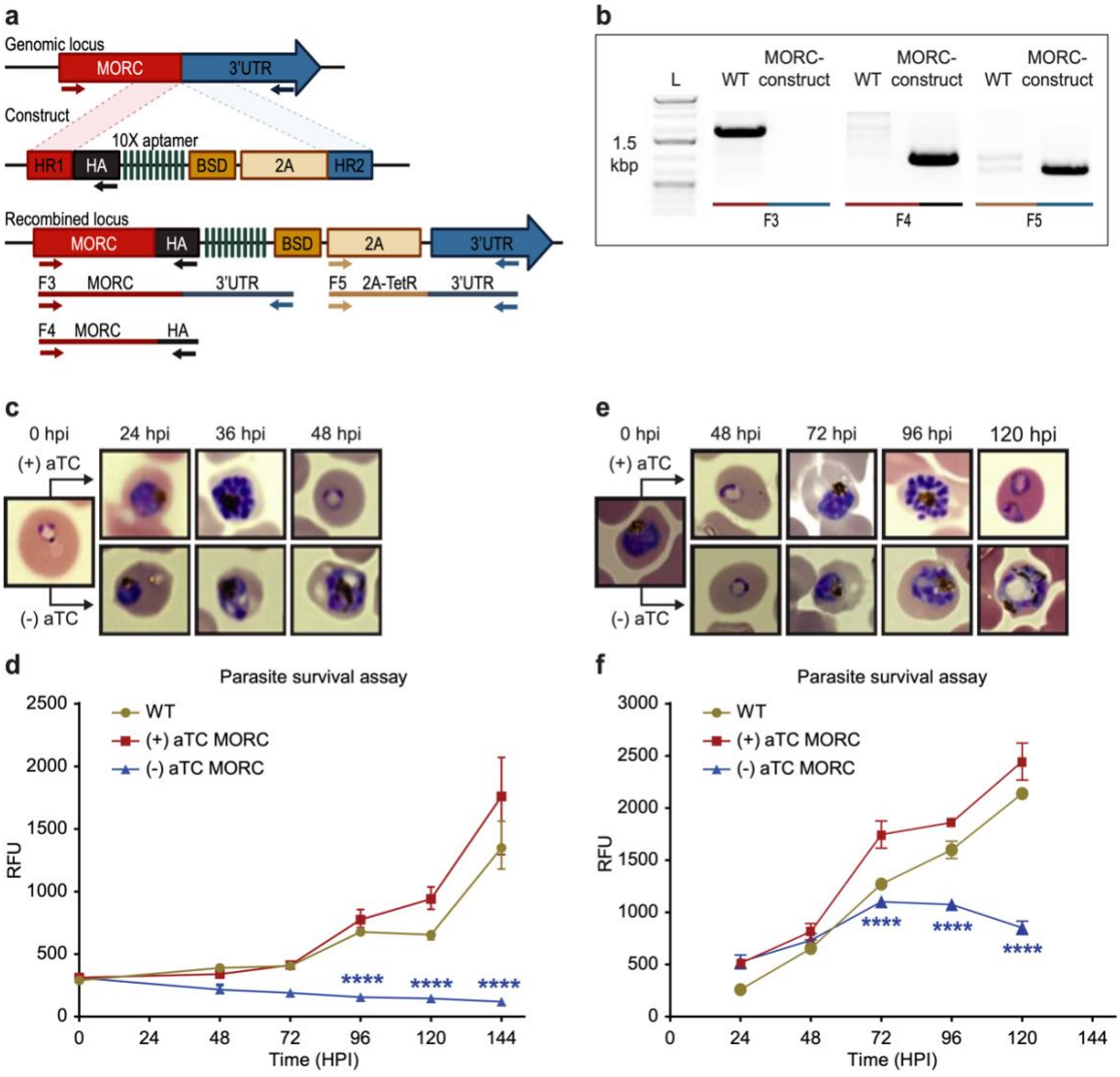


Fig. 3-3. *Pf*MORC is essential for cell survival. (a) Diagram representation of *Pf*MORC-HA-TetRDOZI plasmid. **(b).** PCR amplification is used to verify genomic insertion using primers sets targeting 1.5 kbp of WT *Pfmorc* genome locus absent in transgenic line (MORC construct) (F3) as well as verification of HA insertion (F4) and TetR-DOZI system extending along 3' UTR of the construct (F5). **(c)** Phenotypic and quantitative **(d)** analysis of parasite cell progression after aTC withdrawal at the ring stage (0-6 hpi) (2-way ANOVA, n=3, $p \leq 0.0001$). **(e)** Phenotypic and quantitative **(f)** analysis of parasite cell progression after aTC removal at the trophozoite stage of cell cycle progression (24 hpi) (2-way ANOVA, n=3, $p \leq 0.0001$).

Effect of *Pf*MORC Knockdown on parasite transcriptome

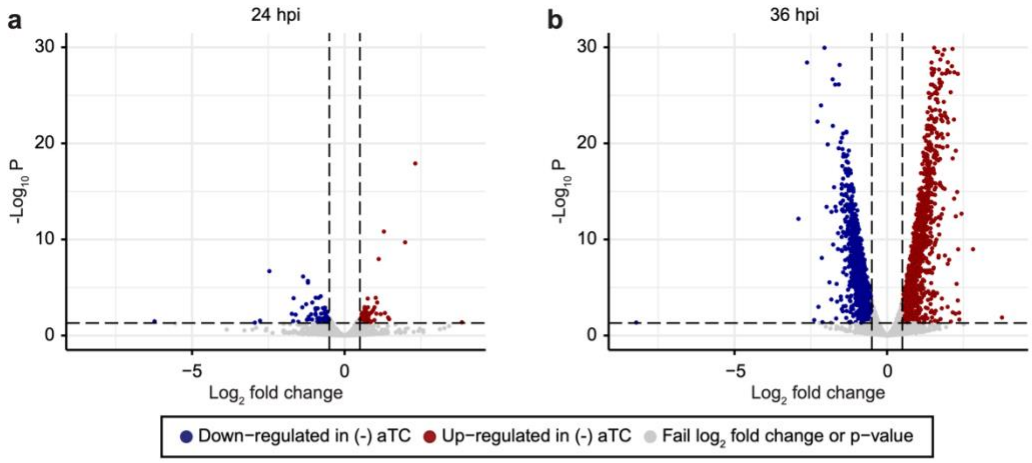
To define the effects of *Pf*MORC KD on transcription, we conducted detailed time-course measurements of mRNA levels using RNA-sequencing (RNA-seq) throughout the parasite asexual cycle on (+/-) aTC *Pf*MORC lines. Parasites were first synchronized and aTC was removed from one set of samples. Total RNA was then extracted at the trophozoite (24 hpi) and the schizont (36 hpi) stages to allow for detection of early and late changes in gene expression after aTC removal. Pairwise correlations analysis (**Fig. S4a**) between (+/-) aTC *Pf*MORC treated lines at the two different time points confirmed appropriate reproducibility of our RNA-seq experiments.

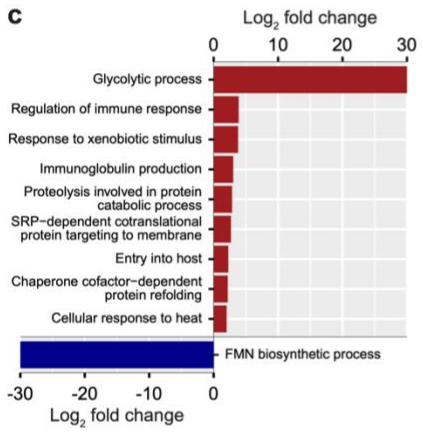
Our results identified a relatively low number of differentially expressed genes at 24 hpi with 96 and 93 upregulated and downregulated genes, respectively (FDR < 0.05 and $\log_2$ FC > 0.5 or < -0.5) (**Fig 3-4a, Sup. data 6**). This data is in agreement with the absence of major phenotypic changes observed in (-) aTC *Pf*MORC parasite cultures after 24 hours. GO enrichment analysis indicated upregulation of genes in response to xenobiotic stimulus including several phosphatases, hydrolases and heat shock proteins suggesting stress sensing of the parasites at this stage (**Fig. 3-4c**). GO analyses of downregulated genes, on the other hand, were found to be closely associated with regular metabolic processes and intracellular transport mechanisms suggesting dysregulation of regular cellular activity. Although, at this stage it is probable that *Pf*MORC depletion is limited and the effects at the transcriptional level are quite restricted **(See Fig. S3)**. It is highly possible that the observed down-regulated genes are likely due to an indirect effect of cell cycle arrest. At 36 hpi the number of genes that exhibit changes in gene expression were significantly

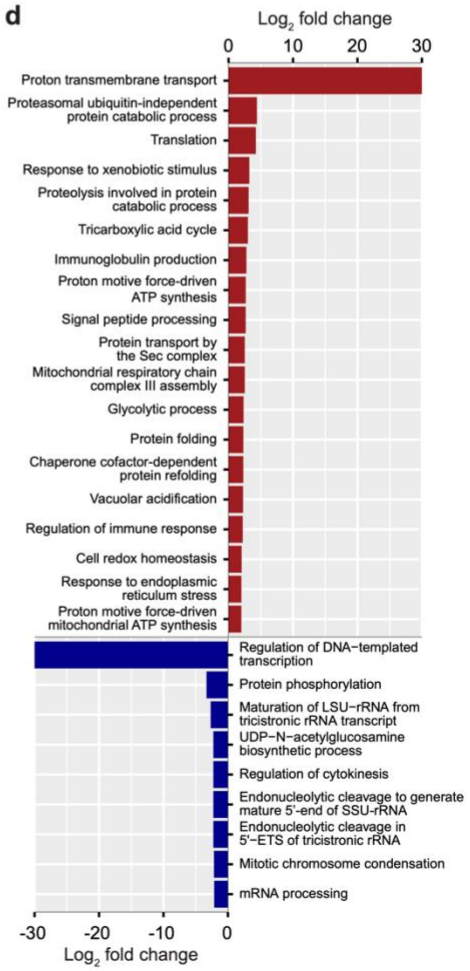
higher with 1319 upregulated and 1150 downregulated (FDR < 0.05 and $\log_2$ FC > 0.5 or < -0.5) in (-) aTC *Pf*MORC conditions, here also confirming the significant changes observed at the phenotypic level (**Fig. 3-4b**). These results emphasize the cascading effects of *Pf*MORC KD as the parasite progressed throughout its cell cycle. GO analysis indicated an enrichment of upregulated genes involved in multiple pathways including ATP metabolic process, mitochondria, translation, food vacuole and protein folding (**Fig. 3-4d**); characteristic of not only stress, but also a clear signal of cell cycle arrest in (-) aTC *Pf*MORC samples. Among upregulated genes, we also observed several *var* genes and genes exported at the surface of the red blood cell that could be linked to significant decreased in *Pf*MORC binding of these genes families as well as a disorganization of the heterochromatin cluster(s) at the trophozoite and schizont stages. GO enrichment analysis for genes that were downregulated included many genes involved in DNA replication, chromosome organization and mitotic spindle further emphasizing a strong cell cycle arrest and absence of potential major compensatory mechanisms for cell division and chromatin organization (**Fig. 3-4d**). Additional downregulated genes were required for invasion such as several merozoite surface proteins. Our results clearly indicate stress sensing and an arrest in cell cycle progression at the trophozoite and schizont stages (**Sup. data 5**)(59).

Although these results are compelling, the small overlap observed between the ChIP-seq signals and RNA-seq results indicate that major changes observed in gene expression at the schizont stages are not simply the result of reduced *Pf*MORC binding in targeted gene bodies but a combination of direct and indirect effects of the degradation of

*Pf*MORC that leads to cell cycle arrest and potential collapse of the heterochromatin **(Fig. 3-4e).**


a
24 hpi
b
36 hpi
-Log10 P
Log2 fold change
Down−regulated in (-) aTC
Up−regulated in (-) aTC
Fail log2 fold change or p−value


c
Log2 fold change
Glycolytic process
Regulation of immune response
Response to xenobiotic stimulus
Immunoglobulin production
Proteolysis involved in protein catabolic process
SRP−dependent cotranslational protein targeting to membrane
Entry into host
Chaperone cofactor−dependent protein refolding
Cellular response to heat
FMN biosynthetic process


d
Log2 fold change
Proton transmembrane transport
Proteasomal ubiquitin-independent protein catabolic process
Translation
Response to xenobiotic stimulus
Proteolysis involved in protein catabolic process
Tricarboxylic acid cycle
Immunoglobulin production
Proton motive force-driven ATP synthesis
Signal peptide processing
Protein transport by the Sec complex
Mitochondrial respiratory chain complex III assembly
Glycolytic process
Protein folding
Chaperone cofactor-dependent protein refolding
Vacuolar acidification
Regulation of immune response
Cell redox homeostasis
Response to endoplasmic reticulum stress
Proton motive force-driven mitochondrial ATP synthesis
Regulation of DNA−templated transcription
Protein phosphorylation
Maturation of LSU−rRNA from tricistronic rRNA transcript
UDP−N−acetylglucosamine biosynthetic process
Regulation of cytokinesis
Endonucleolytic cleavage to generate mature 5′-end of SSU-rRNA
Endonucleolytic cleavage in 5′−ETS of tricistronic rRNA
Mitotic chromosome condensation
mRNA processing

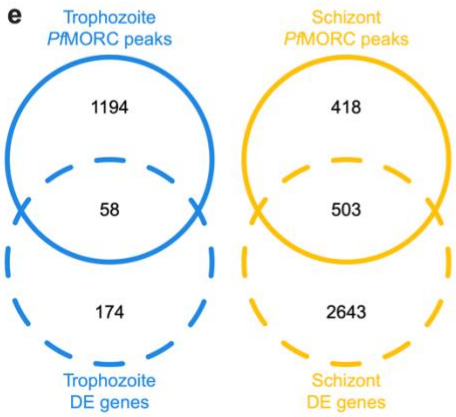

e
Trophozoite PfMORC peaks
Schizont PfMORC peaks
1194
58
174
418
503
2643
Trophozoite DE genes
Schizont DE genes

Fig. 3-4. *Pf*MORC KD on parasite transcriptome. Volcano plots denoting upregulated (red), and downregulated (blue) genes discovered through differential expression analysis following *Pf*MORC knockdown at **(a)** 24 hpi **(b)** 36 hpi. Gene ontology enrichment analysis for upregulated (red) and downregulated (blue) genes at **(c)** 24 hpi and **(d)** 36 hpi. **(e)** Overlap of differentially expressed genes and genes containing significant peaks called by *PfMORC* ChIP-seq analysis at the trophozoite and schizont stages.

*Pf*MORC knockdown erodes antigenic-gene silencing framework

We next sought to analyze the effect of *Pf*MORC down-regulation on the global chromatin landscape. We therefore performed a ChIP-seq experiment against histone H3K9me3 and H3K9ac marks in response to *Pf*MORC depletion. Synchronized parasites were equally split between permissive and repressive conditions. At trophozoite (24 hpi) and schizont (36 hpi) stages of development, both control (+aTC) and experiment (-aTC) samples were fixed and parasites collected for chromatin immunoprecipitation followed by sequencing. The experimental procedure was performed in duplicate between (+/-) aTC *Pf*MORC treated lines. Correlation analysis confirms the reproducibility of our ChIP–seq experiments.

Results revealed no significant changes in histone H3K9ac marks across the genome (data not shown) but a reduction in the heterochromatin landscape in the *Pf*MORC depleted conditions, specifically in the telomere regions of the chromosomes **(Fig. 3-5a)**. *Pf*MORC KD resulted in significantly reduced (Mann-Whitney *U* test, $p < 0.05$) H3K9me3 marks within *var* gene promoters from 200-800 bp upstream of the TSS, as well as 400-800 bp downstream of the TSS **(Fig. 3-5b-c)**. This result coincides with our transcriptomic profile and the upregulation of *var* genes at 36 hpi of (-)aTC samples **(Fig 3-4d, Sup. data 6)**. Further analyses revealed similar trends in other gene families associated with parasite reinvasion and immune response typically silenced under WT conditions. Most noticeable was the reduced H3K9me3 coverage across the *rifin* gene family, which displayed a significant reduction ($p < 0.05$) in all bins from 1000 bp upstream of the TSS to 1000 bp

downstream from the end of the gene **(Fig. 3-5b-c)** in response to *Pf*MORC depletion. There is a clear reduction of H3K9me3 coverage of virulence genes shown to be upregulated following *Pf*MORC knockdown **(Fig. 3-5c)**. Overall, these findings confirm that *Pf*MORC protein depletion not only affects chromatin landscape but has significant impact on heterochromatin. *Pf*MORC downregulation leads to the erosion of heterochromatin integrity most likely required for mutually exclusive gene expression and immune evasion within heterochromatin clusters.

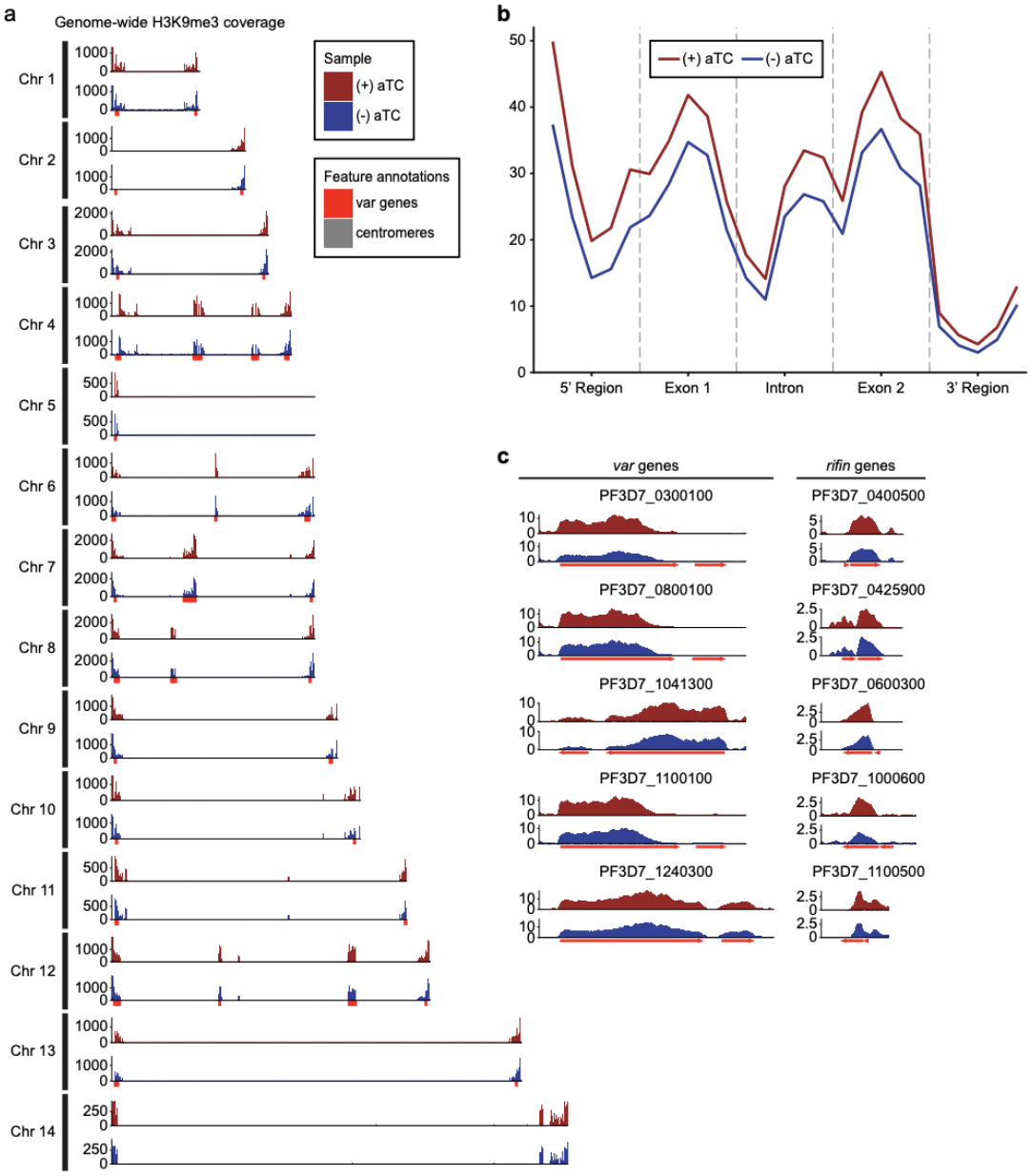

a
Genome-wide H3K9me3 coverage
Sample
(+) aTC
(-) aTC
Feature annotations
var genes
centromeres
Chr 1
Chr 2
Chr 3
Chr 4
Chr 5
Chr 6
Chr 7
Chr 8
Chr 9
Chr 10
Chr 11
Chr 12
Chr 13
Chr 14
b
(+) aTC
(-) aTC
5' Region
Exon 1
Intron
Exon 2
3' Region
c
var genes
rifin genes
PF3D7_0300100
PF3D7_0400500
PF3D7_0800100
PF3D7_0425900
PF3D7_1041300
PF3D7_0600300
PF3D7_1100100
PF3D7_1000600
PF3D7_1240300
PF3D7_1100500

Fig. 3-5. Impact of *Pf*MORC KD on heterochromatin markers. (a) Genome-wide H3K9me3 coverage of IGG subtracted and per-million normalized (+/-) aTC show similar distribution and concentration within telomeres and antigenic gene clusters highlighted in red. Replicates (n = 2) are merged using the mean of the normalized read coverage per base pair. **(b)** Binned coverage of *var* and *rifin* genes from 1 kb upstream of the TSS to 1 kb downstream of the end. The exons and intron of all genes within these families are split into five equal sized bins and the 5' and 3' regions are binned into five 200 bp bins. Read counts within each bin are per-million and bin length normalized prior to plotting. **(c)** Coverage of the five most upregulated *var* and *rifin* genes as determined by the transcriptomic analysis which shows elevated coverage in (+) aTC cells.

Knockdown of *Pf*MORC expression results in the loss of tightly regulated heterochromatin structures

To better understand the effect that downregulation of *Pf*MORC has on the chromatin structure and investigate whether changes in chromatin accessibility may explain large changes in gene expression observed using RNA-seq, we performed Hi-C on (+/-) aTC *Pf*MORC cultures at 24 hpi (trophozoite) and 36 hpi (schizont). Biological replicates for each sample were collected and used to generate Hi-C libraries with >37 million reads per replicate/sample, ensuring comprehensive coverage of both intrachromosomal and interchromosomal interactions. After processing (pairing, mapping, and quality filtering) the raw sequences via HiC-Pro (60), there were approximately 15 million ($\sigma = \sim 10$ million) high quality interaction pairs per sample. Due to the high number of reads and relatively small size of the *P. falciparum* genome (23.3 Mb) compared to higher eukaryotes, we elected to bin our Hi-C data at 10-kb resolution, which allowed the identification of genome-wide patterns while not introducing much noise by binning at too high of a resolution. A high stratum-adjusted correlation **(Fig. S4)**, especially at the 24 hpi time point, suggested that the chromatin structure was consistent between biological replicates, therefore we chose to combine replicates for downstream analyses and visualization.

Because of variation in the number of valid interaction pairs between (+/-) aTC *Pf*MORC samples, 100 iterations of random sampling were performed on samples with higher read count to obtain ~35 million and ~9 million interaction pairs at 24 hpi and 36 hpi, respectively. ICED normalized intrachromosomal heatmaps displayed a high proportion of

interactions at distances less than 10% the length of each chromosome as well as strong subtelomeric interactions and internal regions containing genes involved in antigenic variation **(Fig. 3-6a-b, Fig. S6-S9)** This pattern was similar to those observed previously at various stages of the life cycle (27,60,61) and confirm that genes involved in antigenic variation are usually confined within dense heterochromatin rich regions to aid the tight control of *var* genes necessary for their mutually exclusive expression and immune evasion (53,54). Data generated at 36 hpi in (-) aTC *Pf*MORC parasite cultures were however tumultuous, with a decrease in defined heterochromatin borders and fewer long-range intrachromosomal interactions across the genome. This may indicate a significant loss of chromatin maintenance in (-) aTC *Pf*MORC parasites **(Fig. 3-6a-b).**

To pinpoint which regions were most strongly affected by knockdown of *Pf*MORC, we used Selfish(62) to identify differential intrachromosomal and interchromosomal interactions in the (+/-) aTC *Pf*MORC samples. Although the 24 hpi and 36 hpi time points shared similarities in loci most highly affected by *Pf*MORC KD on many chromosomes, the changes at 36 hpi were more significant with a higher $\log_2$ FC (FDR < 0.05) at most loci **(Fig. 3-6c-d, Fig. S10-S11).** We also observed consistent loss of interactions between most regions containing var genes. These results indicate a failed attempt of the (-) aTC *Pf*MORC parasites to maintain their overall chromatin structure with a significant weakening of the tightly controlled heterochromatin cluster.

We further validated our results using PASTIS to generate coordinate matrices and subsequently visualize a consensus three-dimensional model of the chromosome folding and their overall organization within the nucleus (63). This can indicate changes in spatial distance that best describe the interaction data. The overall change in the chromatin 3D structure was clearly shown by our models of (+/-) aTC *Pf*MORC samples **(Fig. 3-6e).** While the co-localization of centromeres and telomeres in different regions of the nucleus were conserved, the 3D chromatin structure of (-) aTC *Pf*MORC at 36 hpi displayed clear opening of the chromatin and loss of interactions in most regions. This could explain the large changes in gene expression, including increased expression of all *var* and *rifin* genes in the (-) aTC *Pf*MORC line **(Sup. data 3-6).** While heterochromatin maintenance may be essential for tight control of *var* gene expression, preservation of the overall structure of the chromatin may be necessary to regulate the accurate expression of *most P. falciparum* transcripts.

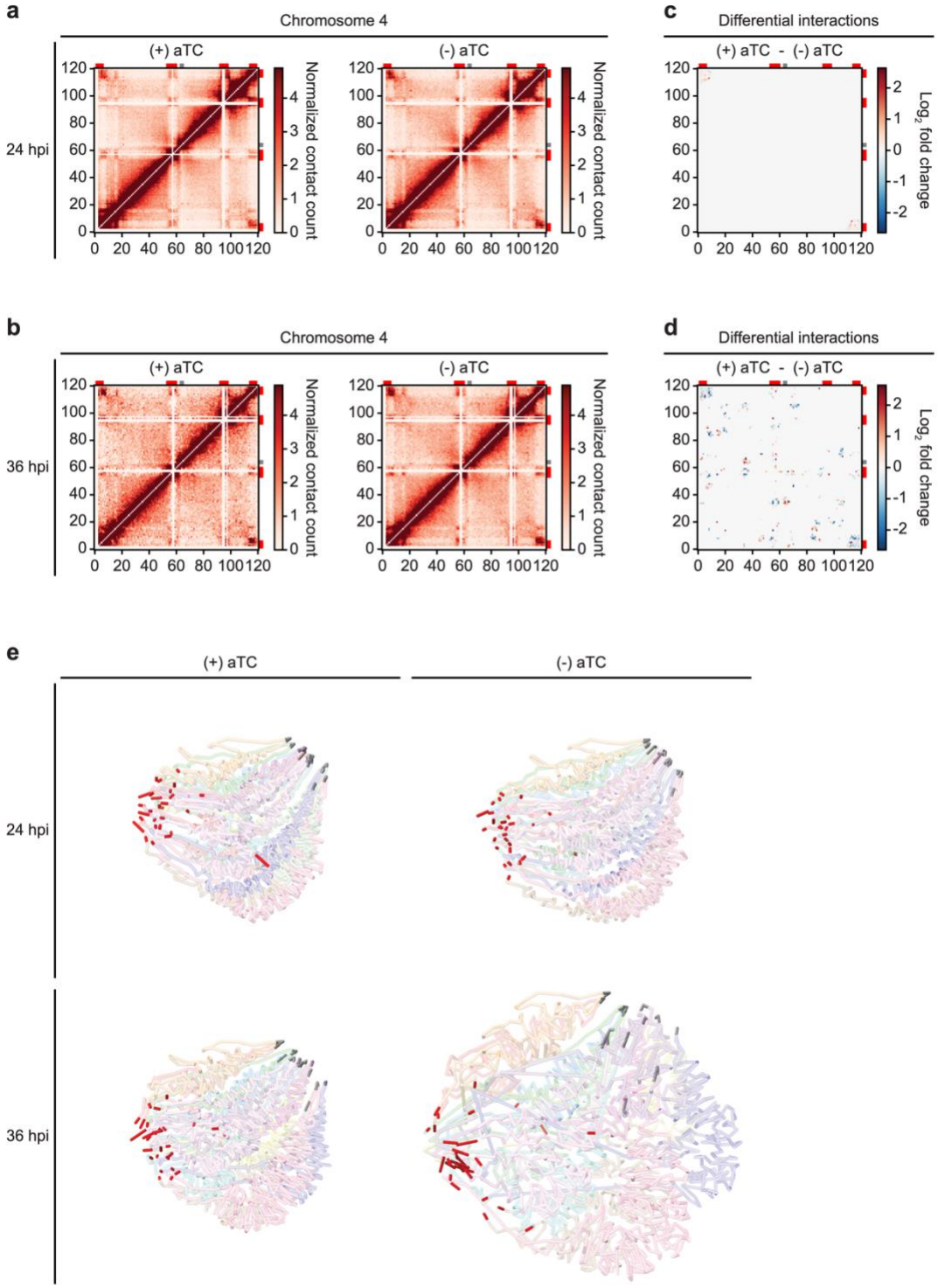

a
Chromosome 4
(+) aTC
(-) aTC
24 hpi
Normalized contact count
c
Differential interactions
(+) aTC - (-) aTC
Log2 fold change
b
Chromosome 4
(+) aTC
(-) aTC
36 hpi
Normalized contact count
d
Differential interactions
(+) aTC - (-) aTC
Log2 fold change
e
(+) aTC
(-) aTC
24 hpi
36 hpi

Fig. 3-6. Loss of *Pf*MORC expression correlates with heterochromatin expansion. Intrachromosomal interaction heatmaps of (+/-) aTC *Pf*MORC for chromosome 4 at **(a)** 24 hpi and **(b)** 36 hpi displaying heterochromatin clustering within antigenic (*var*, *rifin* and *stevor*) gene-dense regions (red). Differential interaction heatmaps highlight changes in chromatin structure following removal of aTC and subsequent *Pf*MORC knockdown at **(c)** 24 hpi and **(d)** 36 hpi. **(e)** Whole genome 3D models of the chromatin structure at both time points (24 hpi and 36 hpi) and (+/-) aTC.

DISCUSSION

Despite intensive investigations on the dynamic nature of the *Plasmodium* chromatin, progress has been slow in capturing the regulatory factors that maintain and control chromatin structure and gene expression throughout parasite development. Preliminary studies performed by our lab and others have identified *Pf*MORC as one of the most abundant chromatin-bound proteins (32,35). MORC proteins from plant to animal cells have a broad range of binding sites within the genome and participate in either DNA methylation establishment or heterochromatin formation (30,64-69). They have also been shown to co-localize with TFs in unmethylated promoter regions to alter chromatin accessibility and regulate TF binding and gene expression (64-69). For instance, AP2-P (PF3D7_1107800), AP2-G5, ISW1, NUP2, MCM5 and HDAC1 have been shown to interact with MORC in several published works (33,42,43), including this work. Protein pull-down experiments in *Plasmodium spp.* and *T. gondii* studies have detected MORC in complexes with several AP2 TFs. In *T. gondii*, *Tg*MORC was characterized as an upstream transcriptional repressor of sexual commitment (33,35). In *Plasmodium*, the role of *Pf*MORC was still poorly understood. Conflicting results on the essentiality of *Pf*MORC using either KD or KO strategies remained unresolved (32,56). Using the CRISPR/Cas9 genome editing tool, we have determined the nuclear localization, genome-wide distribution, and regulatory impacts of *Pf*MORC on the parasite chromatin, transcriptome, and cell cycle development. We now have powerful evidence demonstrating that *Pf*MORC is not only critical for parasite cell cycle progression and its survival but has a direct role in heterochromatin formation and gene silencing, including regulating immune evasion

through antigenic variation. Disparities observed in the essentiality of *Pf*MORC for the parasite survival in previous studies (32) are most likely the results of weak protein disruption highlighting the need for a significant downregulation of *Pf*MORC for true functional analysis. Using IFA and protein pull-downs, we have confirmed that *Pf*MORC localizes to the nucleus and interacts with multiple ApiAP2 TFs, chromatin remodelers, and epigenetic players associated with heterochromatin regions including H3K9me3, SIP2, HDAC1 and the ISWI chromatin-remodeling complex (SWIB/MDM2) (52) **(Fig. 3-1).** Although some discrepancies have been observed with other studies, mainly because MORC protein was not used as the bait protein, our study validated the interaction of MORC with these partners, supporting that *Pf*MORC is in complex with key heterochromatin regulators. Using a combination of ChIP-seq, protein knock down, RNA-seq and Hi-C experiments, we also demonstrated that the MORC protein is essential for the tight regulation of gene expression. We can speculate that lack of MORC impacts heterochromatin and chromatin compaction, preventing access to gene promoters from TFs and the general transcriptional machinery in a stage specific manner. Although additional experiments will be required, our hypothesis is reinforced by the fact that downregulation of the *Pf*MORC significantly reduced H3K9me3 coverage in the heterochromatin cluster(s) and 5' flanking regions of *var* and *rifin* genes. While we were unable to confirm a direct role of *Pf*MORC in sexual conversion due to a nonsense mutation in AP2-G gene in our transfected lines, its strong interaction with AP2-G during the asexual cell cycle indicates that *Pf*MORC in combination with other epigenetic factors may most likely control AP2-G expression and sexual differentiation. It is also important to recognize that the many

pathways affected at the transcriptional level throughout the asexual stages in (-) aTC *Pf*MORC lines are not only the direct results of *Pf*MORC downregulation and decrease of its targeted DNA binding sites. They are most likely a combination of direct and indirect effects of *Pf*MORC KD warranted by the cell cycle arrest observed at the phenotypic level and the collapse of the chromatin organization confirmed using chromatin conformation capture experiment. These direct and indirect effects should be carefully considered, and a combination of functional genomic studies should be completed when interpreting changes in gene expression in mutant generated lines.

All together our work demonstrated that in addition to its direct role in heterochromatin formation and antigenic variation, *Pf*MORC may act as a potential repressor to control a set of parasite specific genes including genes involved in parasite egress and invasion, and antigenic variation between the trophozoite to schizont stage transition (52). As our data confirm the importance of *Pf*MORC and the parasite specificity of several interacting partners, it is tempting to speculate that drugs targeting these protein complexes could lead to novel antiparasitic strategies.

MATERIALS AND METHODS

Asexual parasite Culture and maintenance

Asexual *P. falciparum* strain NF54 parasites (MRA-1000) were propagated in 5% of human O+ erythrocytes and 10 mL of RPMI-1640 medium containing 0.5% Albumax II (Invitrogen), 2 mM L-glutamine, 50 mg/L hypoxanthine, 25 mM HEPES, 0.225% NaHCO3 and 10 mg/mL gentamicin. They were maintained at 37°C and gassed with a

sterile mixture of 5% O2, 5% CO2 and 90% N2. Parasite synchronization was achieved using two 5% D-sorbitol treatments 8 hours apart (70,71).

Plasmids Construction

Flagging of the *P. falciparum* MORC (*Pf3D7_1468100*) gene spanning position (Ch14: 2785386 - 2797934 (-)) was performed using a two-plasmid design to insert a 3x HA tag. The pCasG-Cas9- sgRNA plasmid vector (gifted by Dr. Sean Prigge) contains the site to express the sgRNA, along with the *yDHODH* gene as the positive selection marker. The gRNA oligos (Sup. data 1) were ligated after digestion of the plasmid with BsaI. The pDC2-cam-Cas9-U6 plasmid (gifted by Dr. Marcus Lee) was digested with BamHI and Apal to remove the eGFP tag from the backbone. 453 bp of C-terminal region of *Pf*MORC and 469 bp of 3'UTR were amplified from *P. falciparum* genomic DNA with their respective primers **(Sup. data 3-1).** A 3x HA-tag was fused to the Cterminal region of the amplified product along with the 3'UTR formed through Gibson assembly master mix (NEB, E2611S). To generate the *Pf*MORC-HA knockdown constructs, a pKD*Pf*AUBL plasmid (gifted by Dr. Sean Prigge) (72) was digested with AscI and AatII. Homology arms HA1 and HA2 of PfMORC were amplified with respective primers **(Sup. data 1)** having 20 bp overhang and inserted into the digested plasmid using Gibson assembly mix. The resulting pKD-MORCHA-TetR-DOZI construct was linearized with EcoRV enzyme prior to transfection. All constructs were confirmed through restriction enzyme digestion and Sanger sequencing.

Plasmids were isolated from 250 mL cultures of *Escherichia coli* (XL10-Gold Ultracompetent Cells, Agilent Cat. 200314) and 60 µg of pDC2-MORC-HA or linearized pKD-MORC-HA-TetRDOZI, were used with 60 µg of pCasG-plasmid containing gRNA to transfect 200 µl of fresh red blood cells (RBCs) infected with 3-5% early ring stage parasites. After one erythrocytic cycle, transfected cultures were supplemented with 1.5 µM WR99210 (2.6nM) (provided by the (Jacobus Pharmaceuticals, Princeton, NJ) and 2.5 µg/mL Blasticidin (RPI Corp B12150-0.1). *Pf*MORC-HA-TetR-DOZI transfected parasites were maintained with 500 nM anhydrotetracycline (aTC) (45,58). Media and drug selection was replenished every 24 hours for 7 consecutive days after which DSM-1 drug selection was halted Once parasites were detected by microscopy, integration of the insert was confirmed by PCR amplification. To generate genetically homogenous parasite lines, the transfected parasites were serially diluted to approximately 0.5% parasite/well, into 96 well plates.

Molecular analysis of the transgenic lines

Genomic DNA (gDNA) was extracted and purified using the DNeasy Blood & Tissue kit (Qiagen) following instructions from the manufacturer. The diagnostic PCR analysis was used to genotype the transfected lines using the primers listed in Supplementary data 1. The PCR amplification was conducted using KAPA HiFi HotStart ReadyMix (Roche) and amplicons were analyzed by gel electrophoresis followed by sequencing. For whole genome sequencing, genomic DNA was fragmented using a Covaris S220 ultrasonicator and libraries were generated using KAPA LTP Library Preparation Kit (Roche, KK8230).

To verify that the insertion was present in the genome at the correct location in both transfected lines, reads were mapped using Bowtie2 (v2.4.4) to the *P. falciparum* 3D7 reference genome (PlasmoDB, v48), edited to include the insertion sequence in the intended location. Integrative Genomic Viewer (IGV, Broad Institute) was used to verify that reads aligned to the modified sequence.

Variant analysis by genome-wide sequencing

To call variants (SNPs/indels) in the transfected lines compared to a previously sequenced control 3D7 line, genomic DNA reads were first trimmed to adapters and aligned to the Homo sapiens genome (assembly GRCh38) to remove human-mapped reads. Remaining reads were aligned to the *P. falciparum* 3D7 genome using bwa (version 0.7.17) and PCR duplicates were removed using PicardTools (Broad Institute). GATK HaplotypeCaller (https://gatk.broadinstitute.org/hc/en-us) was used to call variants between the sample and the 3D7 reference genome for both the transfected lines and the NF54 control. Only variants that were present in both transfected lines but not the NF54 control line were kept. We examined only coding-region variants and removed those that were synonymous variants or were located in *var, rifin,* or *stevor* genes. Quality control of variants was done by hard filtering using GATK guidelines.

Immunofluorescence assays

A double-staining immunofluorescence assay was used on NF54 control and transgenic parasite lines of mixed parasite population growth stages. Parasites were washed in

incomplete medium prior to fixing onto coverslips with 4% paraformaldehyde for 20 min at RT under darkness. After fixation, samples were washed 3-5 times with 1x PBS followed by permeabilization with 0.5%Triton X-100 in PBS for 25 min at RT. Subsequently, samples were subjected to PBS washes and then incubated overnight at 4°C in blocking buffer (2mg/ml BSA solution in PBS containing 0.05% Tween-20 solution). Following overnight blocking, samples were washed and incubated for 1 hr at RT with Anti-HA Rb Ab (1:500, Abcam, ab9110) in blocking buffer. After primary Ab incubation, samples were subject to 3x washes with wash buffer (1x PBS containing 0.05% Tween-20) followed by incubation with anti-rabbit DyLight 550 (Abcam ab98489; 1:500) secondary antibody for 1 h at room temperature. After incubation and series of washes with wash buffer, slides were incubated with anti-H3K9me3 antibody, Alexa Fluor 488 conjugate (Millipore 07–442-AF488; 1:100) for 1 hr at RT. Slides are then washed and mounted in Vectashield Antifade Mounting Medium with DAPI (Vector Laboratories, H-1200). Images were acquired using a Keyence BZ-X810 Fluorescence Microscope and were processed through ImageJ.

Western blotting

Parasites were synchronized twice at 8 hr. intervals between synchronizations. After one cycle, culture was washed 3x with complete media to remove residual aTC followed by dividing the cultures into two equal conditions. Parasites were grown with or without 500 nM aTC for 24 and 36 hrs. after which RBCs were lysed using 0.15 % saponin and parasites were collected after being washed with ice cold 1X PBS. Proteins were recovered from the

lysed parasites after 30 min of incubation in lysis buffer (150 mM NaCl, 0.5 % NP40, 50 mM Tris pH 8, 1 mM EDTA and protease inhibitors) and 10 sec of sonication. Proteins were quantified with Pierce™ BCA Protein Assay Kit (Thermo Fisher 23227). 20 µg of proteins were loaded onto 3–8% Criterion XT Tris-Acetate Midi Protein Gels (Bio-Rad, 3450129). After migration, proteins were transferred onto a PVDF membrane and the membrane was blocked and then probed overnight with anti-HA tag antibody (1:2500, Abcam, ab9110) as well as anti-aldolase antibody as a loading control (1:10000, abcam, ab252953). After primary Ab incubation, blots were subsequently washed 3x with a washing buffer followed by HRP-labeled Goat anti-Rabbit IgG (H + L) (1:10,000, Novex™, A16104). Clarity™ Western ECL Substrate (Bio-Rad, 1705060) was applied to reveal the blots. Relative abundance of (+/-) aTC *Pf*MORC was calculated by Bio-Rad ChemiDoc imagelab software.

Immunoprecipitation followed by MudPIT mass spectrometry

Mid- to late-stage asexual parasites were collected following saponin treatment and purified samples were then resuspended into fresh IP buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 2 mM AEBSF 0.5% Triton X-100, and EDTA-free protease inhibitor cocktail (Roche)). Post cell lysis solution was homogenized via sonication for 6-9 rounds. The soluble extracts were centrifuged at 13,000 x g for 15 min at 4°C. The lysates were precleared with Dynabeads™ Protein A (Invitrogen) for 1h at 4°C. Anti-HA tag antibody (1:2500, Abcam, ab9110) are added to control and HA tagged PfMORC precleared protein extract samples for 1hr at 4°C under constant rotation

followed by the addition of fresh Dynabeads™ Protein A beads to each sample and incubated overnight at 4°C. Dynabeads were washed 3 times with 500 µL of buffer (PBS, 0.05% Tween-20). Proteins were eluted into an elution buffer (50 mM Tris-HCl pH 6.7, 100 mM DTT and 2% SDS). The eluent was subsequently precipitated overnight in 20% TCA followed by cold acetone washes. The urea-denatured, reduced, alkylated, and digested proteins were analyzed by Multidimensional Protein Identification Technology (MudPIT) on a Orbitrap Elite mass spectrometer coupled to an Agilent 1260 series HPLC, as described previously (73).

Proteomics data processing and analysis

Tandem mass (MS/MS) spectra were interpreted using ProluCID v.1.3.3 (74) against a database consisting of 5527 non-redundant (NR) *Plasmodium falciparum* 3D7 proteins (PlasmoDB, v42), 36661 NR human proteins (NCBI, 2018-03-30 release), 419 common contaminants (human keratins, IgGs, and proteolytic enzymes), together with shuffled versions of all of these sequences. DTASelect v.1.9 (75) and swallow v.0.0.1, an in-house developed software (https://github.com/tzwwen/kite) were used to control FDRs resulting in protein FDRs less than 1.86%. All datasets were contrasted against their merged data set, respectively, using Contrast v1.9 (75) and in-house developed sandmartin v.0.0.1 (https://github.com/tzw-wen/kite/tree/master/ kitelinux). Our inhouse developed software, NSAF7 v.0.0.1

(https://github.com/tzwwen/kite/tree/master/windowsapp/NSAF7x64), was used to generate spectral count-based label free quantitation results (76). QPROT (49,77) was used

to calculate values of $\log_2$ fold change and Z statistic to compare two replicate PfMORC affinity purifications to two negative controls. Proteins enriched in the *Pf*MORC-HA samples with values of $\log_2$ fold change ≥ 2 and Z-statistic ≥ 5 were considered significantly enriched.

ChIP assay

*Pf*MORC-HA and *Pf*MORC KD asexual stage parasites were supplemented with and without aTC along with NF54 parasites (as a control) and were harvested at 24 and 36 HPS and cross linked with 1% formaldehyde for 10 min at 37°C followed by quenching with 150 mM glycine and 3 x washing with 1 x PBS. The pellets were resuspended in 1 mL of nuclear extraction buffer (10 mM HEPES, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 0.5 mM AEBSF and 1x Protease inhibitor cocktail), and incubated on ice for 30 min before addition of NP-40/Igepal to a final of 0.25%. After lysis, the parasites are subject to homogenization by passing through a 26.5G 3/8 needle/syringe to burst the nucleus. After centrifugation for 20 min at 2,500 x g at 4°C, samples were resuspended in 130 µL of shearing buffer (0.1% SDS, 1 mM EDTA, 10 mM Tris-HCl pH 7.5 and 1X Protease inhibitor cocktail) and transferred to a 130 µL Covaris sonication microtube. The samples were then sonicated using a Covaris S220 Ultrasonicator for 8 min (Duty cycle: 5%, intensity peak power:140, cycles per burst: 200, bath temperature: 6°C). The samples were transferred to ChIP dilution buffer (30 mM Tris-HCl pH 8.0, 3 mM EDTA, 0.1% SDS, 30 mM NaCl, 1.8% Triton X-100, 1X protease inhibitor tablet, 1X phosphatase inhibitor tablet) and centrifuged for 10 min at 13,000 rpm at 4°C, retaining the supernatant.

For each sample, 13 µL of protein A agarose/salmon sperm DNA beads were washed three times with 500 µL ChIP dilution buffer by centrifuging for 1 min at 1,000 rpm at room temperature. For pre-clearing, the diluted chromatin samples were added to the beads and incubated for 1 hour at 4°C with rotation, then pelleted by centrifugation for 1 min at 1,000 rpm. Before adding antibody, ~10% of each sample was taken as input. 2 µg of anti-HA tag antibodies (1:2500, Abcam, ab9110) or rabbit polyclonal anti-H3K9me3 (Millipore no. 07–442) antibodies were added to the sample and incubated overnight at 4°C with rotation. For each sample, rabbit IgG antibody (Cat.689 No. 10500c, Invitrogen) was used as a negative-control library. Per sample, 25 µL of protein A agarose/salmon sperm DNA beads were washed with ChIP dilution buffer (no inhibitors), blocked with 1 mg/mL BSA for 1 hour at 4°C, and then washed three more times with buffer. 25 µL of washed and blocked beads were added to the sample and incubated for 1 hour at 4°C with continuous mixing to collect the antibody/protein complex. Beads were pelleted by centrifugation for 1 min at 1,000 rpm at 4°C. The bead/antibody/protein complex was then washed with rotation using 1 mL of each buffer twice; low salt immune complex wash buffer (1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl), high salt immune complex wash buffer (1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 500 mM NaCl), and TE wash buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Complexes were eluted from antibodies by adding 250 µL of freshly prepared elution buffer (1% SDS, 0.1 M sodium bicarbonate). 5 M NaCl were added to the eluate and cross-linking was reversed by heating at 45°C overnight followed by addition of 15 µL of 20 mg/mL RNase A with 30 min incubation at 37°C. After this, 10 µL of 0.5 M EDTA, 20 µL of 1 M Tris-HCl pH

7.5, and 2 µL of 20 mg/mL proteinase K were added to the eluate and incubated for 2 hours at 45°C. DNA was recovered by phenol/chloroform extraction and ethanol precipitation. DNA was purified using Agencourt AMPure XP beads. Libraries were then prepared from this DNA using KAPA LTP Library Preparation Kit (Roche, KK8230) and sequenced on a NovaSeq 6000 machine.

ChIP-seq analysis

FastQC (version 0.11.8, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to analyze raw read quality. Any adapter sequences were removed using Trimmomatic (version 0.39, http://www.usadellab.org/cms/?page=trimmomatic). Bases with Phred quality scores below 20 were trimmed using Sickle (version 1.33, https://asexual stage parasitesb.com/najoshi/sickle). The resulting reads were mapped against the *P. falciparum* genome (v48) using Bowtie2 (version 2.4.4). Using Samtools (version 1.11), only properly paired reads with mapping quality 40 or higher were retained and reads marked as PCR duplicates were removed by PicardTools MarkDuplicates (version 2.18.0, Broad Institute). Genome-wide read counts per nucleotide were normalized by dividing millions of mapped reads for each sample (for all samples including input) and subtracting IgG read counts from the anti-HA IP counts. Peak calling was performed using MACS3 (version 3.0.0a7) (78) with the upper and lower limits for fold enrichment set to 2 and 50, respectively.

Phenotypic analyses

P. falciparum NF54 line along with transgenic *Pf*MORC-HA-TetR-DOZI line were synchronized, grown for one cycle post-synchronization and subject to 3 consecutive washes with 1X PBS before being supplemented either with or without (aTC) for 48 hrs. or 72 hrs. (depending on which stage the parasites were washed). Culture media were replaced daily with or without aTC supplementation.

Quantitative Growth assay

Synchronous cultures either ring or trophozoite stage at 0.5 % parasitemia were grown with or without aTC (500 nM) into a 96 well plate in triplicates for each time point. Parasites were collected at every 24 hrs. time for 3 cycles and subjected to SYBR green assay (Thermo Fisher, S7523).

RNA-seq library preparation

Parasites at ring, trophozoite or schizont stage were extracted following saponin treatment before flash freezing. Two independent biological replicates were generated for each time point, culture condition and line. Total RNA was extracted with TRIzol® LS Reagent (Invitrogen) followed by incubation for 1 hr with 4 units of DNase I (NEB) at 37°C. RNA samples were visualized by RNA electrophoresis and quantified on Synergy™ HT (BioTek). mRNA was then purified using NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB, E7490) according to the manufacturer's instructions. Libraries were prepared using NEBNext® Ultra™ Directional RNA Library Prep Kit (NEB, E7420) and

amplified by PCR with KAPA HiFi HotStart Ready Mix (Roche). PCR conditions consisted of 15 min at 37°C followed by 12 cycles of 98°C for 30 sec, 55°C for 10 s and 62°C for 1 min 15 sec, then finally one cycle for 5 min at 62°C. The quantity and quality of the final libraries were assessed using a Bioanalyzer (Agilent Technology Inc). Libraries were sequenced using a NovaSeq 6000 DNA sequencer (Illumina), producing paired end 100-bp reads.

RNA-seq data processing and differential expression analysis

FastQC [https://www.bioinformatics.babraham.ac.uk/projects/fastqc/] was used to analyze raw read quality and thus 11 bp of each read and any adapter sequences were removed using Trimmomatic (v0.39) [http://www.usadellab.org/cms/?page=trimmomatic]. Bases were trimmed from reads using Sickle with a Phred quality threshold of 25 (v1.33) [https://github.com/najoshi/sickle] and reads shorter than 18 bp were removed. The resulting reads were mapped against the *P. falciparum* 3D7 genome (PlasmoDB, v53) using HISAT2 (v2-2.2.1) with default parameters. Uniquely mapped, properly paired reads with mapping quality of 40 or higher were retained using SAMtools (v1.11) [http://samtools.sourceforge.net/]. Raw read counts were determined for each gene in the *P. falciparum* genome using BedTools [https://bedtools.readthedocs.io/en/latest/#] to intersect the aligned reads with the genome annotation. Differential expression analysis was performed using DESeq2 to call up- and downregulated genes (FDR < 0.05 and $\log_2$ FC > 0.5). Volcano plots were made using the R package EnhancedVolcano.

Hi-C library preparation

(+/-) aTC parasites at 24 hpi and 36 hpi were cross linked using 1.25% formaldehyde for 15 minutes at 37°C in 10 mL total volume followed by quenching with 150 mM glycine. Parasites were then washed three times with chilled 1x PBS on a rocking platform. Parasite nuclei were released using lysis buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 2 mM AEBSF, 0.25% Igepal CA-630, and 1X EDTA-free protease inhibitor cocktail (Roche)) and 15 syringe passages through a 26.5-gauge needle. After releasing the crosslinked chromatin with 0.1% SDS, chromatin was digested using MboI restriction enzyme overnight at 37°C. Thereafter Hi-C libraries were prepared as previously described (27,79).

Hi-C data processing, differential interaction analysis and generation of 3D models

Paired-end Hi-C libraries were processed (pairing, mapping, quality filtering, binning, and normalization) using HiC-Pro (60). A mapping quality cutoff of 30 was set while aligning to the *P. falciparum* genome (PlasmoDB, v58) and the resulting reads were binned at 10 kb resolution and ICED-normalized. Stratum-adjusted correlation coefficients were calculated using HiCRep (80), then replicates were merged to generate a single consensus sample for each time point and condition. The normalized matrices were per-million read count normalized and maximum values for the heatmap color scale on each chromosome was set to the maximum value for all samples to allow for direct comparisons between each condition and time point. Furthermore, due to the overrepresentation of short-range interactions, the maximum values for each heatmap were also set to the 90th percentile of each chromosome to slightly compress highly interacting regions and enhance

visualization of interactions. Significant intrachromosomal interactions were identified with FitHiC (81) to identify the log-linear relationship between contact probability and genomic distance. Differential interchromosomal and intrachromosomal interactions between (+) aTC and (-) aTC were identified using Selfish (62) with default parameters (FDR < 0.05). The maximum and minimum values for the color scales in the differential heatmaps were set to the absolute value of the largest $\log_2$ fold change within each chromosome. PASTIS (63) was used to generate coordinate matrices from the raw read count matrices output by HiC-Pro, and then visualized as 3D chromatin models in ChimeraX (82), highlighting regions containing var genes, telomeres, and centromeres.

Statistical analyses

Parasitemia and proportion of asexual stages were analyzed using a two-way ANOVA with Tukey's test for multiple comparisons. Significant differences were indicated as following: * for $p < 0.05$; ** for $p < 0.01$, *** for $p < 0.001$ and **** for $p < 0.0001$. Statistical tests were performed with GraphPad Prism. Figures were generated with GraphPad Prism and BioRender.

ACKNOWLEDGEMENTS

This work was supported by NIH grants to KGLR (nos. 1R01 AI136511 and R21 AI142506-01) and by the University of California, Riverside to KGLR (no. NIFA-Hatch-225935). This publication includes data generated at the UC San Diego IGM Genomics

Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (#S10 OD026929).

AUTHOR CONTRIBUTIONS

Conceptualization was the responsibility of KGLR that also supervised the project together with LF. MG generated all the constructs and cell lines used in the study. Immunoprecipitations, Growth assay were performed by MG and ZC. ChIP seq, RNA seq libraries were generated by MG and analyzed by TL and SA. Hi-C experiments were performed by MG and the data was analyzed by TL and generated all the figures. Methodologies were performed by MG, ZC, TH, SA, TL, CB, and AS; TL, SA, TH, CB, LF and KGLR carried out the analysis. Software, formal analysis, and data curation were provided by TL, SA, and CB. ZC, TH, TL and KGLR wrote the original draft. All authors reviewed and edited the final version of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY

Sequence reads for all sequencing experiments have been deposited in the NCBI Sequence Read Archive with accession PRJNA994684. Original data underlying this manuscript generated at the Stowers Institute can be accessed from the Stowers Original Data Repository at http://www.stowers.org/research/publications/LIBPB-2404. The mass

spectrometry dataset generated for this study is available from the MassIVE data repository (ftp://MSV000092353@massive.ucsd.edu) using the identifier MSV000092353. All other data are available in the main text, methods, and supplementary data.

REFERENCES

1. WHO, World malaria report 2023. (2023).

2. R. M. Coulson, N. Hall, C. A. Ouzounis, Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. *Genome Res* **14**, 1548-1554 (2004).

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

4. E. K. De Silva *et al.*, Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc Natl Acad Sci U S A* **105**, 8393-8398 (2008).

5. M. Yuda *et al.*, Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. *Mol Microbiol* **71**, 1402-1414 (2009).

6. S. Iwanaga, I. Kaneko, T. Kato, M. Yuda, Identification of an AP2-family protein that is critical for malaria liver stage development. *PLoS One* **7**, e47557 (2012).

7. A. Sinha *et al.*, A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. *Nature* **507**, 253-257 (2014).

8. B. F. Kafsack *et al.*, A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507**, 248-252 (2014).

9. K. M. Lesage *et al.*, Cooperative binding of ApiAP2 transcription factors is crucial for the expression of virulence genes in Toxoplasma gondii. *Nucleic Acids Res* **46**, 6057-6068 (2018).

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

11. K. Modrzynska *et al.*, A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle. *Cell Host Microbe* **21**, 11-22 (2017).

12. C. Gu *et al.*, Multiple regulatory roles of AP2/ERF transcription factor in angiosperm. *Bot Stud* **58**, 6 (2017).

13. M. Yuda, I. Kaneko, S. Iwanaga, Y. Murata, T. Kato, Female-specific gene regulation in malaria parasites by an AP2-family transcription factor. *Mol Microbiol* **113**, 40-51 (2020).

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

15. Z. Bozdech *et al.*, Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray. *Genome Biol* **4**, R9 (2003).

16. K. G. Le Roch *et al.*, Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503-1508 (2003).

17. Z. Bozdech *et al.*, The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. *PLoS Biol* **1**, E5 (2003).

18. P. Srinivasan *et al.*, Analysis of the Plasmodium and Anopheles transcriptomes during oocyst differentiation. *J Biol Chem* **279**, 5581-5587 (2004).

19. F. Silvestrini *et al.*, Genome-wide identification of genes upregulated at the onset of gametocytogenesis in Plasmodium falciparum. *Mol Biochem Parasitol* **143**, 100-110 (2005).

20. M. Yang *et al.*, Full-Length Transcriptome Analysis of Plasmodium falciparum by Single-Molecule Long-Read Sequencing. *Front Cell Infect Microbiol* **11**, 631545 (2021).

21. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation. *Science* **295**, 1306-1311 (2002).

22. J. Dostie *et al.*, Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* **16**, 1299-1309 (2006).

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

24. J. R. Dixon *et al.*, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380 (2012).

25. L. Dembele *et al.*, Persistence and activation of malaria hypnozoites in long-term primary hepatocyte cultures. *Nat Med* **20**, 307-312 (2014).

26. X. Deng *et al.*, Bipartite structure of the inactive mouse X chromosome. *Genome Biol* **16**, 152 (2015).

27. E. M. Bunnik *et al.*, Comparative 3D genome organization in apicomplexan parasites. *Proc Natl Acad Sci U S A* **116**, 3183-3192 (2019).

28. G. Batugedara *et al.*, The chromatin bound proteome of the human malaria parasite. *Microb Genom* **6**, (2020).

29. Z. J. Lorkovic, MORC proteins and epigenetic regulation. *Plant Signal Behav* **7**, 1561-1565 (2012).

30. D. Q. Li, S. S. Nair, R. Kumar, The MORC family: new epigenetic regulators of transcription and DNA damage response. *Epigenetics* **8**, 685-693 (2013).

31. N. E. Weiser *et al.*, MORC-1 Integrates Nuclear RNAi and Transgenerational Chromatin Architecture to Promote Germline Immortality. *Dev Cell* **41**, 408-423 e407 (2017).

32. S. Singh *et al.*, The PfAP2-G2 transcription factor is a critical regulator of gametocyte maturation. *Mol Microbiol* **115**, 1005-1024 (2021).

33. D. C. Farhat *et al.*, A MORC-driven transcriptional switch controls Toxoplasma developmental trajectories and sexual commitment. *Nat Microbiol* **5**, 570-583 (2020).

34. H. Wang, L. Zhang, Q. Luo, J. Liu, G. Wang, MORC protein family-related signature within human disease and cancer. *Cell Death Dis* **12**, 1112 (2021).

35. Z. Zhong *et al.*, MORC proteins regulate transcription factor binding by mediating chromatin compaction in active chromatin regions. *Genome Biol* **24**, 96 (2023).

36. N. Saksouk *et al.*, Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite Toxoplasma gondii. *Mol Cell Biol* **25**, 10301-10314 (2005).

37. A. Bougdour *et al.*, Drug inhibition of HDAC3 and epigenetic control of differentiation in Apicomplexa parasites. *J Exp Med* **206**, 953-966 (2009).

38. C. Hillier *et al.*, Landscape of the Plasmodium Interactome Reveals Both Conserved and Species-Specific Functionality. *Cell Rep* **28**, 1635-1647 e1635 (2019).

39. M. D. Jeninga, J. E. Quinn, M. Petter, ApiAP2 Transcription Factors in Apicomplexan Parasites. *Pathogens* **8**, (2019).

40. T. Hiyoshi, J. A. Wada, Feline amygdaloid kindling and the sleep-waking pattern: observations on daily 22-hour polygraphic recording. *Epilepsia* **31**, 131-138 (1990).

41. M. K. Singh *et al.*, A nuclear protein, PfMORC confers melatonin dependent synchrony of the human malaria parasite P. falciparum in the asexual stage. *Sci Rep* **11**, 2057 (2021).

42. A. K. Subudhi *et al.*, DNA-binding protein PfAP2-P regulates parasite pathogenesis during malaria parasite blood stages. *Nat Microbiol* **8**, 2154-2169 (2023).

43. J. M. Bryant *et al.*, Exploring the virulence gene interactome with CRISPR/dCas9 in the human malaria parasite. *Mol Syst Biol* **16**, e9569 (2020).

44. B. V. A. Singh Maneesh Kumar, Da Silva Israel Tojal, Santiago Verônica Feijoli, Moraes Miriam S., Adderley Jack, Doerig Christian, Palmisano Giuseppe, Llinás Manuel, Garcia Célia R. S., A Plasmodium falciparum MORC protein complex modulates epigenetic control of gene expression through interaction with heterochromatin *eLIfe* **12**, (2023).

45. S. M. Ganesan, A. Falla, S. J. Goldfless, A. S. Nasamu, J. C. Niles, Synthetic RNA-protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nat Commun* **7**, 10727 (2016).

46. M. Filarsky *et al.*, GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* **359**, 1259-1263 (2018).

47. M. Yuda, I. Kaneko, Y. Murata, S. Iwanaga, T. Nishi, Mechanisms of triggering malaria gametocytogenesis by AP2-G. *Parasitol Int* **84**, 102403 (2021).

48. R. S. Kent *et al.*, Inducible developmental reprogramming redefines commitment to sexual development in the malaria parasite Plasmodium berghei. *Nat Microbiol* **3**, 1206-1213 (2018).

49. H. Choi, S. Kim, D. Fermin, C. C. Tsou, A. I. Nesvizhskii, QPROT: Statistical method for testing differential expression using protein-level intensity data in label-free quantitative proteomics. *J Proteomics* **129**, 121-126 (2015).

50. W. A. M. Hoeijmakers *et al.*, Epigenetic reader complexes of the human malaria parasite, Plasmodium falciparum. *Nucleic Acids Res* **47**, 11574-11588 (2019).

51. C. Flueck *et al.*, A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in chromosome end biology. *PLoS Pathog* **6**, e1000784 (2010).

52. X. Shang *et al.*, Genome-wide landscape of ApiAP2 transcription factors reveals a heterochromatin-associated regulatory network during Plasmodium falciparum blood-stage development. *Nucleic Acids Res* **50**, 3413-3431 (2022).

53. Q. Chen *et al.*, Developmental selection of var gene expression in Plasmodium falciparum. *Nature* **394**, 392-395 (1998).

54. A. Scherf *et al.*, Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. *EMBO J* **17**, 5418-5426 (1998).

55. L. Meerstein-Kessel *et al.*, Probabilistic data integration identifies reliable gametocyte-specific proteins and transcripts in malaria parasites. *Sci Rep* **8**, 410 (2018).

56. M. Zhang *et al.*, Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, (2018).

57. S. J. Goldfless, J. C. Wagner, J. C. Niles, Versatile control of Plasmodium falciparum gene expression with an inducible protein-RNA interaction. *Nat Commun* **5**, 5329 (2014).

58. A. S. Nasamu *et al.*, An integrated platform for genome engineering and gene expression perturbation in Plasmodium falciparum. *Sci Rep* **11**, 342 (2021).

59. T. Kiss, Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* **109**, 145-148 (2002).

60. N. Servant *et al.*, HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol* **16**, 259 (2015).

61. E. M. Bunnik *et al.*, Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. *Nat Commun* **9**, 1910 (2018).

62. A. R. Ardakany, F. Ay, S. Lonardi, Selfish: discovery of differential chromatin interactions via a self-similarity measure. *Bioinformatics* **35**, i145-i153 (2019).

63. N. Varoquaux, F. Ay, W. S. Noble, J. P. Vert, A statistical approach for inferring the 3D structure of the genome. *Bioinformatics* **30**, i26-33 (2014).

64. A. H. Tencer *et al.*, Molecular mechanism of the MORC4 ATPase activation. *Nat Commun* **11**, 5466 (2020).

65. Y. Shao *et al.*, Involvement of histone deacetylation in MORC2-mediated down-regulation of carbonic anhydrase IX. *Nucleic Acids Res* **38**, 2813-2824 (2010).

66. F. He *et al.*, Structural insight into the zinc finger CW domain as a histone modification reader. *Structure* **18**, 1127-1139 (2010).

67. Y. Zhang *et al.*, MORC3 Forms Nuclear Condensates through Phase Separation. *iScience* **17**, 182-189 (2019).

68. J. Luo, S. Zeng, C. Tian, MORC4 Promotes Chemoresistance of Luminal A/B Breast Cancer via STAT3-Mediated MID2 Upregulation. *Onco Targets Ther* **13**, 6795-6803 (2020).

69. Y. Mimura, K. Takahashi, K. Kawata, T. Akazawa, N. Inoue, Two-step colocalization of MORC3 with PML nuclear bodies. *J Cell Sci* **123**, 2014-2024 (2010).

70. W. Trager, J. B. Jensen, Human malaria parasites in continuous culture. 1976. *J Parasitol* **91**, 484-486 (2005).

71. D. A. Fidock, T. Nomura, T. E. Wellems, Cycloguanil and its parent compound proguanil demonstrate distinct activities against Plasmodium falciparum malaria parasites transformed with human dihydrofolate reductase. *Mol Pharmacol* **54**, 1140-1147 (1998).

72. K. Rajaram, H. B. Liu, S. T. Prigge, Redesigned TetR-Aptamer System To Control Gene Expression in Plasmodium falciparum. *mSphere* **5**, (2020).

73. L. Florens, M. P. Washburn, Proteomic analysis by multidimensional protein identification technology. *Methods Mol Biol* **328**, 159-175 (2006).

74. T. Xu *et al.*, ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J Proteomics* **129**, 16-24 (2015).

75. D. L. Tabb, W. H. McDonald, J. R. Yates, 3rd, DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J Proteome Res* **1**, 21-26 (2002).

76. Y. Zhang, Z. Wen, M. P. Washburn, L. Florens, Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem* **82**, 2272-2281 (2010).

77. H. Choi, D. Fermin, A. I. Nesvizhskii, Significance analysis of spectral count data in label-free shotgun proteomics. *Mol Cell Proteomics* **7**, 2373-2385 (2008).

78. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

79. M. K. Gupta, T. Lenz, K. G. Le Roch, Chromosomes Conformation Capture Coupled with Next-Generation Sequencing (Hi-C) in Plasmodium falciparum. *Methods Mol Biol* **2369**, 15-25 (2021).

80. T. Yang *et al.*, HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Res* **27**, 1939-1949 (2017).

81. F. Ay, T. L. Bailey, W. S. Noble, Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res* **24**, 999-1011 (2014).

82. T. D. Goddard *et al.*, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).

CHAPTER 4

Novel insights into the role of long non-coding RNA in the human malaria parasite, *Plasmodium falciparum*

Gayani Batugedara[1]*, Xueqing M. Lu[1]*, Borislav Hristov[2], Steven Abel[1], **Zeinab Chahine[1],** Thomas Hollin[1], Desiree Williams[1], Tina Wang[1], Anthony Cort[1], Todd Lenz[1], Trevor A. Thompson[1], Jacques Prudhomme[1], Abhai K. Tripathi[3], Guoyue Xu[3], Juliana Cudini[4], Sunil Dogga[4], Mara Lawniczak[4], William Stafford Noble[2], Photini Sinnis[3] and Karine G. Le Roch[1✧]

[1]Department of Molecular Cell and Systems Biology, University of California Riverside, Riverside, CA 92521, USA

[2] Department of Genome Sciences, University of Washington, Seattle, WA 98195-5065, USA

[3]Department of Molecular Microbiology and Immunology and the Johns Hopkins Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA

[4]Wellcome Sanger Institute, Hinxton, CB10 1SA, UK

* These authors contributed equally to this work.

A version of this chapter has been published in *Nature Communications*, 2023.

PREFACE

One of the most lethal pathogens transmitted by mosquitoes is the *Plasmodium falciparum* parasite, responsible for human malaria. Extensive research on this parasite focuses on identifying gene regulation mechanisms and targets for potential antimalarial drugs. While protein targets have been the primary focus, recent findings suggest that long non-coding RNAs (lncRNAs) may also play a significant role in regulating parasite genes. Employing computational methods, we predicted the location and orientation of lncRNAs across the parasite genome and conducted their experimental validation, with a particular focus on lncRNA-14. This comprehensive approach revealed various properties of the identified

lncRNAs, including their expression during different stages of the parasite life cycle, subcellular localization, size, GC content, and genomic location, notably in telomeric regions.

My role in this work consisted of characterizing the impact of LncRNAs on parasite sexual commitment and gametocyte development. I executed the successful knock-down of the LncRNA-14 on *P. falciparum* through insertion of a drug selection marker into the gene body of the target region. This was achieved through CRISPR-Cas9 genetic manipulation. With a successful clone obtained I proceeded with comparing the parasite asexual stage progression and gametocyte development through multiple avenues of study. I performed phenotypic analysis and quantification of the gametocyte development as well as applied molecular investigations of parasite genomic, transcriptomic comparisons. My efforts helped determine the changes in RNA-expression post LncRNA-14 perturbation and determined how, in this case, LncRNA disruption has a significant impact on the pathogens asexual development upon initiation of gametocytogenesis, drastically affecting sexual stage development important for transmission.

ABSTRACT

The complex life cycle of *Plasmodium falciparum* requires coordinated gene expression regulation to allow host cell invasion, transmission, and immune evasion. Increasing evidence now suggests a major role for epigenetic mechanisms in gene expression in the parasite. In eukaryotes, many lncRNAs have been identified to be pivotal

regulators of genome structure and gene expression. To investigate the regulatory roles of lncRNAs in *P. falciparum* we explore the intergenic lncRNA distribution in nuclear and cytoplasmic subcellular locations. Using nascent RNA expression profiles, we identify a total of 1,768 lncRNAs, of which 718 (~41%) are novels in *P. falciparum*. The subcellular localization and stage-specific expression of several putative lncRNAs are validated using RNA-FISH. Additionally, the genome-wide occupancy of several candidate nuclear lncRNAs is explored using ChIRP. The results reveal that lncRNA occupancy sites are focal and sequence-specific with a particular enrichment for several parasite-specific gene families, including those involved in pathogenesis and sexual differentiation. Genomic and phenotypic analysis of one specific lncRNA demonstrate its importance in sexual differentiation and reproduction. Our findings bring a new level of insight into the role of lncRNAs in pathogenicity, gene regulation and sexual differentiation, opening new avenues for targeted therapeutic strategies against the deadly malaria parasite.

INTRODUCTION

Malaria, a mosquito-borne infectious disease, is caused by protozoan parasites of the genus *Plasmodium*. Among the human-infecting species, *Plasmodium falciparum* is the most prevalent and deadly, with an estimated 627 000 deaths in 2020(1). The parasite has a complex life cycle involving multiple biological stages in both human and mosquito hosts. As sporozoites are transmitted from a *Plasmodium*-infected mosquito to the human bloodstream, they migrate to the liver to invade hepatocytes and initiate parasite amplification. After this pre-erythrocytic cycle, which can last 7 to 10 days, tens of

thousands of infectious merozoites are released into the bloodstream to invade red blood cells. Within the erythrocyte, the parasite matures from the ring to the trophozoite and to the multinucleated schizont stages. After 48 h, the newly formed merozoites burst out of the erythrocyte to reinfect new red blood cells. This rupture is usually associated with clinical symptoms. During this intraerythrocytic developmental cycle, a subset of the parasites can differentiate into male and female gametocytes. Once ingested by a female *Anopheles* mosquito during a blood meal, these gametocytes undergo sexual replication inside the mosquito's gut to form a zygote that can differentiate into a mobile ookinete and an oocyst. The oocyst will grow and produce thousands of new sporozoites that will migrate to the mosquito's salivary glands ready to infect a new human host during a subsequent blood meal. This multi-stage developmental life cycle leads to distinct morphological and physiological changes in response to altered environmental conditions and is tightly regulated by coordinated changes in gene expression.

Gene expression profiling(2,3) including bulk RNA-seq experiments(2,4-7), nascent RNA expression profiles(8), as well as single cell sequencing(9) has revealed that a majority of the genes in the parasite are transcribed in a cascade of gene expression throughout the parasite life cycle but the exact molecular mechanisms regulating these events are largely unknown.

Compared to other eukaryotes with a similar genome size, *P. falciparum* has an extremely AT-rich genome and a relatively low number of sequence-specific transcription factors

(TFs), approximately two-thirds of the TFs expected based on the size of the genome. Only 27 apicomplexan apetala2 (ApiAP2) DNA-binding proteins have been identified as specific TFs in the parasite genome. These ApiAP2 are unique to *Apicomplexa* (10)and have been demonstrated to have a major role as activators or repressors of transcription(11). Our understanding of the regulation of these TFs, and how various TFs could act together to organize transcriptional networks, is still limited but the patterns of gene expression observed are likely the result of a combination of transcriptional(9,12-14) and post-transcriptional regulatory events(15-18). Additionally, epigenetic studies(19-25) and chromosome conformation capture methods (Hi-C)(26-28) have suggested that the chromatin state and the three-dimensional (3D) genome structure of *P. falciparum* are strongly connected with transcriptional activity of gene families(28). Machine learning algorithms have also suggested that the ApiAP2 TFs may indeed work in conjunction with epigenetic factors(29). However, how all the regulators of transcription are recruited to their DNA binding motifs and their chromatin regions remains to be elucidated. Understanding the exact mechanisms regulating the parasite replication life cycle is essential if we want to identify novel therapeutic targets.

With advances in biotechnology and next generation sequencing technologies, huge strides have been made in genomics studies revealing that the transcriptome of an organism is much larger than expected. In eukaryotes spanning from yeast to human, many non-coding RNAs (ncRNAs) have been detected and linked to diseases ranging from cancers to neurological disorders and are now actively studied for their potential as novel therapeutic

and diagnostic agents(30). Over the past few years, ncRNAs been recognized as key regulators of chromatin states and gene expression(31-33). One class of ncRNAs, the long noncoding RNAs (lncRNAs), are defined as non-protein coding RNA molecules which are ≥ 200 nucleotides in length. Many lncRNAs share features with mature mRNAs including 5' caps, polyadenylated tails as well as introns(34). LncRNAs are expressed and functionally associated in a cell-type specific manner. Based on their genomic localization, lncRNAs are categorized as sense, antisense, bidirectional, intronic and intergenic(35). because lncRNAs can bind DNAs, RNAs and proteins, their functions are diverse(34). LncRNAs enriched in the nuclear fraction often associate with regulation of transcription(36-39). By tethering genomic DNA, lncRNAs can control long-range interaction. They can also regulate promoter accessibility by recruiting, guiding or enhancing either TFs or chromatin remodeling enzymes including histone acetyltransferases and methyltransferases. LncRNA have also been shown to interact with spliceosomal factors to affect the frequency and efficiency of mRNA splicing. In the cytosol, lncRNAs can regulate gene expression by mediating mRNA export, RNA stability and translation.

In mammalian systems, the X inactive specific transcript (Xist) is a well-studied example of a lncRNA mediating X-chromosome inactivation during zygotic development(40). Deposition of Xist on the X-chromosome recruits histone-modifying enzymes that place repressive histone marks, such as H3K9 and H3K27 methylation, leading to gene silencing and the formation of heterochromatin. Two other lncRNAs, the HOX transcript antisense

(HOTAIR) and the antisense lncRNA in the INK4 locus (ANRIL), have also been shown to interact with multiprotein Polycomb PRotein Complexes (PRC1 and PRC2) to catalyze histone marks and silence gene expression(30). Similarly, long telomeric repeat-containing lncRNAs (TERRA) have been recently identified as a major component of telomeric heterochromatin(41-42).With thousands of lncRNAs transcribed in mammalian cells we are only starting to grasp their role in regulating major biological processes.

Although the role of lncRNA in malaria parasite has only been studied more recently, they are emerging as new players in the development of parasite life cycle stages. To date, several studies have already explored the presence of ncRNAs in *P. falciparum*(43-54). Technological advances including strand-specific and long read sequencing platforms have identified >2500 lncRNA candidates, including 1300 circular lncRNAs(51-53,55,56). These initial studies confirmed that parasite lncRNAs are developmentally regulated but only a few of these annotated ncRNAs have been functionally characterized. Some have been linked to regulation of virulence genes(57-62). It has also been established that GC-rich ncRNAs serve as epigenetic regulatory elements that play a role in activating *var* gene transcription as well as several other clonally variant gene families(63). In addition, a family of twenty-two lncRNAs transcribed from the telomere-associated repetitive elements (TAREs) has been identified in the parasite(45,47,59). These TARE-lncRNAs show functional similarities to the eukaryotic family of non-coding RNAs involved in telomere and heterochromatin maintenance(64) and could have a role in regulating virulence factors. More recently, the functional characterization of two lncRNAs, *gdv1*-as-

lncRNA and *md1*-lncRNA that were detected during gametocytogenesis, has revealed that sexual differentiation and sex determination in *P. falciparum* is at least partially regulated by lncRNAs(65,66). While it is becoming evident that lncRNAs serve as an integral part of the mechanisms regulating gene expression in *Plasmodium*, the localization and function of most of the identified lncRNAs remain a mystery.

Here, to investigate the localization and subsequently the potential role of lncRNAs in *P. falciparum*, we explore the intergenic lncRNA distribution separately in nuclear and cytoplasmic subcellular locations. Using deep sequencing and nascent RNA expression profiles(8), we identify a total of 1768 lncRNAs, of which 41% are novels in *P. falciparum*. We further validate the subcellular localization and stage-specific expression of several putative lncRNAs using RNA fluorescence in situ hybridization (RNA-FISH) and single-cell RNA sequencing (scRNA-seq). Additionally, the genome-wide occupancy of 7 candidate nuclear lncRNAs is explored using Chromatin Isolation by RNA Purification followed by deep sequencing (ChIRP-seq). Our ChIRP-seq experiments on our candidate lncRNAs reveal that lncRNA occupancy sites within the parasite genome are sequence-specific with a particular enrichment for several parasite-specific gene families, including those involved in pathogenesis, remodeling of the red blood cell, and regulation of sexual differentiation. We also demonstrate that the presence of some of these lncRNAs correlates with changes in gene expression demonstrating that these lncRNAs can possibly work in cooperation with TFs and epigenetic factors. We further validate the role of a lncRNA identified as enriched in gametocytes. Using the CRISPR-Cas9 editing tool, we

functionally characterize lncRNA-ch14 and validate its role during sexual differentiation and development, particularly affecting female gametocytes. Transmission studies demonstrate that even partial deletion of this lncRNA significantly affects parasite development throughout all mosquito stages. Collectively, our results provide evidence that in addition to being developmentally regulated, lncRNAs are distributed in distinct cellular compartments in *P. falciparum*. Depending on their nuclear or cytoplasmic localization, they may play important roles in gene regulation at the transcriptional or translational levels respectively, ultimately regulating the malaria parasite life cycle progression.

RESULTS

Identification of lncRNAs

To comprehensively identify lncRNA populations in *P. falciparum*, we extracted total RNA from both nuclear and cytoplasmic fractions using synchronized parasite cultures at early ring, early trophozoite, late schizont, and gametocyte stages **(Fig. 4-1a).** The samples collected here allow for gene expression profiling during the critical processes of parasite egress, invasion, and sexual differentiation. In brief, extracted parasites were subjected to a modified cell fractionation procedure described in the PARIS kit (ThermoFisher) (see "Methods"). Successful isolation of both subcellular fractions was validated using western blot with an anti-histone H3 antibody as a nuclear marker and an anti-aldolase antibody as a cytoplasmic marker **(Fig. 4-1b).** After separation of nuclear material from the cytoplasmic material, total RNA and subsequent polyadenylated mRNA

was isolated from both fractions. Strand-specific libraries were then prepared and sequenced **(see methods for details, Supplementary data 4-1a)**. For verification, Spearman correlations in gene expression levels were calculated among nuclear samples and cytoplasmic samples(67) **(Fig. 4-S1).**

Once validated, a computational pipeline was implemented for the identification of lncRNAs. Briefly, all nuclear and cytoplasmic RNA libraries were merged, resulting in one nuclear and one cytoplasmic merged file, then assembled into nuclear and cytosol transcriptomes independently using cufflinks. Subsequently, transcripts were filtered based on length, expression level, presence of nascent transcript from previously published GRO-seq dataset[8], and sequence coding potential **(Fig. 4-1a).** To specifically identify lncRNA candidates within the intergenic regions and avoid any potential artefacts introduced by PCR amplification of the AT rich genome, we removed any predicted transcripts that have at least 30% overlap with annotated genes. Our goal was to select transcripts that are ≥200 bp in length, consistently expressed in both published nascent RNA and steady-state RNA expression profiles, and that are likely to be non-protein-coding genes. As a result, we identified a total of 1768 intergenic lncRNAs in *P. falciparum* irrespective of the developmental stage **(Supplementary data 4-1b).** Nine hundred fifty-one lncRNAs have no overlap with any UTR regions. Overall, 1050 lncRNAs (~59%) overlapped with previously identified intergenic lncRNAs(49,51-53,55,56) and 718 lncRNAs were identified as novel in *P. falciparum* **(Fig. 4-1c).**

To evaluate the essentiality of the lncRNAs identified in this study, we used *piggyBac* insertion sites from Zhang and colleagues(68). In this work, the authors used a high-throughput transposon insertional mutagenesis method to distinguish essential and dispensable genes in the *P. falciparum* genome during the asexual stages of the parasite life cycle. We focused our analysis on the integration of the transposon that occurred within each of the identified lncRNAs. The *piggyBac* insertion site coordinates were overlapped with the genomic ranges for all detected lncRNAs. This was performed after accounting for differences in the parasite strains used in the two studies. Overall, we were unable to uncover *piggyBac* insertion for 558 lncRNAs (31.6%), suggesting that these lncRNAs are either difficult to disrupt or are potentially essential for the parasite asexual development **(Fig. 4-1d).** Because we observed an insertion in 292 lncRNAs (16.5%) that were specifically detected in gametocyte stage, it will be important to validate at the phenotypic level whether those lncRNAs are essential during sexual differentiation rather than the asexual cycle. Additionally, significantly fewer insertions per possible insertion site (TTAA sequence) were found for telomeric as compared to subtelomeric lncRNAs, and for subtelomeric as compared to other lncRNAs **(Fig. 4-S2a).** This suggests that *piggyBac* insertions in telomeric and sub-telomeric lncRNAs are also either difficult to disrupt due to their co-localization with heterochromatin or are more likely to be essential than others. It was also found that the 5' flanking regions of lncRNAs **(Fig. 4-S2b)** had more insertions per possible site as compared to the rest of the lncRNAs, suggesting that, as a general trend, these regions of the lncRNAs are the most disposable while the gene body and 3' flanking regions may be more important for their function. Further analysis of the 558 lncRNA with

zero *piggyBac* insertions illustrated that 158 of them (28.3%) overlapped by at least 1 bp with a gene, including UTRs, found to be essential in the Zhang study. For the remaining 400 lncRNAs (71.7%), the lack of insertions cannot be tied to a nearby gene rather than the lncRNA itself. In addition, novel lncRNAs were found to have fewer insertions (and thus be more likely essential) than previously known lncRNAs (38.6% of novel lncRNAs had zero insertions, whereas 26.8% of previously known lncRNAs had zero insertions, and novel lncRNAs had fewer insertions per TTAA site (0.184 vs. 0.204)). Thus, many lncRNAs identified in this study may be essential for parasite survival. Although additional experiments will be needed to validate these results, it is also possible that some of the genes found to be essential in the Zhang study including some of the *rifin*, *stevor*, and pseudogenes, could be due to extensive overlap with the identified lncRNAs.

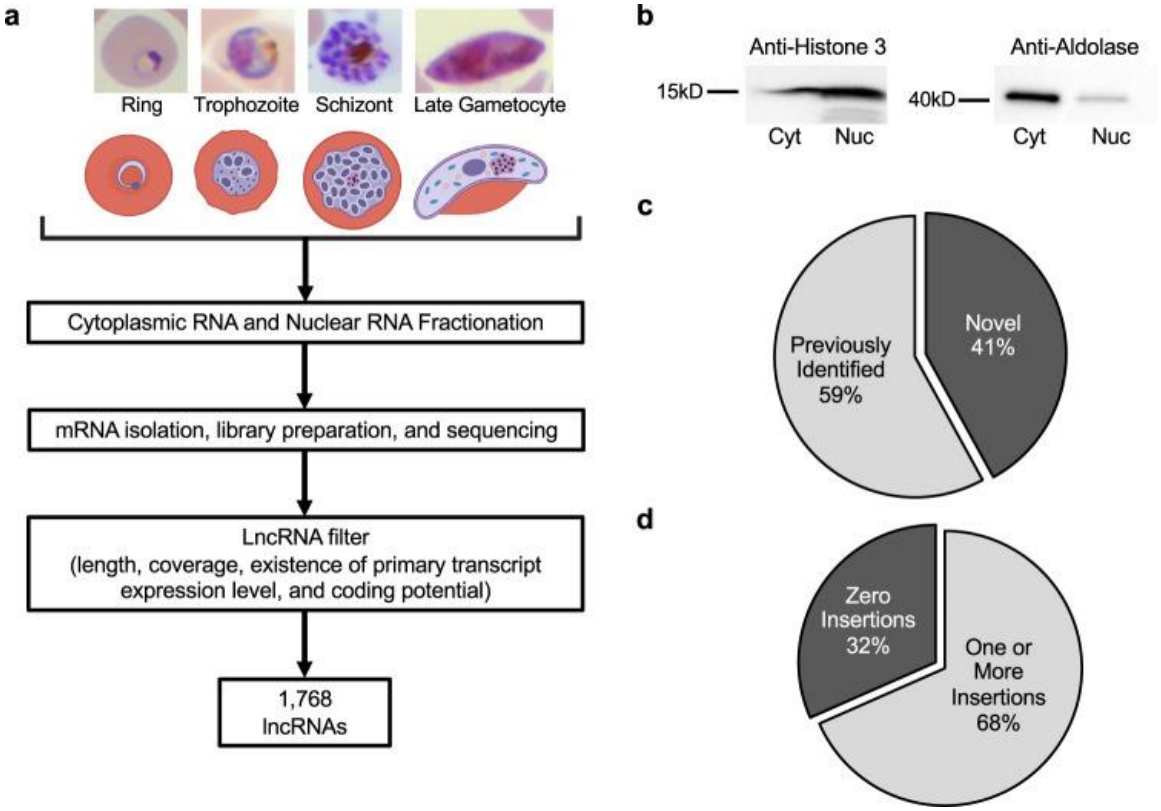


Figure 4-1. Nuclear and cytoplasmic lncRNA identification. a A general overview of the lncRNA identification pipeline. Created with BioRender.com. **b** Validation of cell fractionation efficiency using anti-histone H3 and anti-aldolase as nuclear (Nuc) and cytoplasmic (Cyt) markers. Blot is representative of two independent biological replicates. **c** Comparison of lncRNA candidates with lncRNAs identified from previous publications. **d** Essentiality of lncRNAs using *piggyBac* insertion (*Zhang et al., 2018*). LncRNAs that cannot be disrupted are more likely to be essential.

Length, GC content, and RNA stability of cytoplasmic and nuclear lncRNAs

LncRNAs exhibit diverse subcellular distribution patterns, ranging from nuclear foci to cytoplasmic localization. Their localization patterns are linked to their distinct regulatory effects at their site of action(69,70). Therefore, to better understand the potential function of the lncRNA in *Plasmodium*, we categorized the subcellular localization of our candidate lncRNAs into nuclear lncRNAs, cytoplasmic lncRNAs, or indistinguishable

lncRNAs that are equally distributed in both fractions. Among the total identified 1768 lncRNAs, 719 lncRNAs (41%) were enriched in the nuclear fraction, 204 lncRNAs (11%) were enriched in the cytoplasmic fraction, and 845 lncRNAs (48%) showed similar distribution between both subcellular fractions **(Fig. 4-2a).** Further, we explored the physical properties of these lncRNAs. We observed that lncRNAs are in general shorter in length and less GC-rich as compared to protein-encoding mRNAs **(Fig. 4-2b,c).** Using total steady-state RNA expression profiles and nascent RNA expression profiles (GRO-seq) **(Fig. 4-2d),** we then estimated the expression levels and stability of the lncRNAs. RNA stability was calculated as the ratio between steady-state RNA expression levels over nascent RNA expression levels. We discovered that, although the overall life cycle gene expression pattern of the lncRNAs is similar to the expression pattern of coding mRNAs, lncRNAs are less abundant and less stable than coding mRNAs; nuclear IncRNAs are particularly lowly expressed and unstable as compared to the other two groups of IncRNAs **(Fig. 4-2e).** These observations are consistent with previous lncRNA annotation studies in human breast cancer cells(71) and noncoding RNA stability studies in mammalian genomes(72). Our results suggest that the low expression level and the low stability of these lncRNAs may be the reason why they failed to be detected in previous identification attempts. By taking advantage of primary transcripts detected in our GRO-seq dataset, we significantly improved the sensitivity of lncRNA detection, especially for those localized in the nuclear fraction and expressed at a lower level.

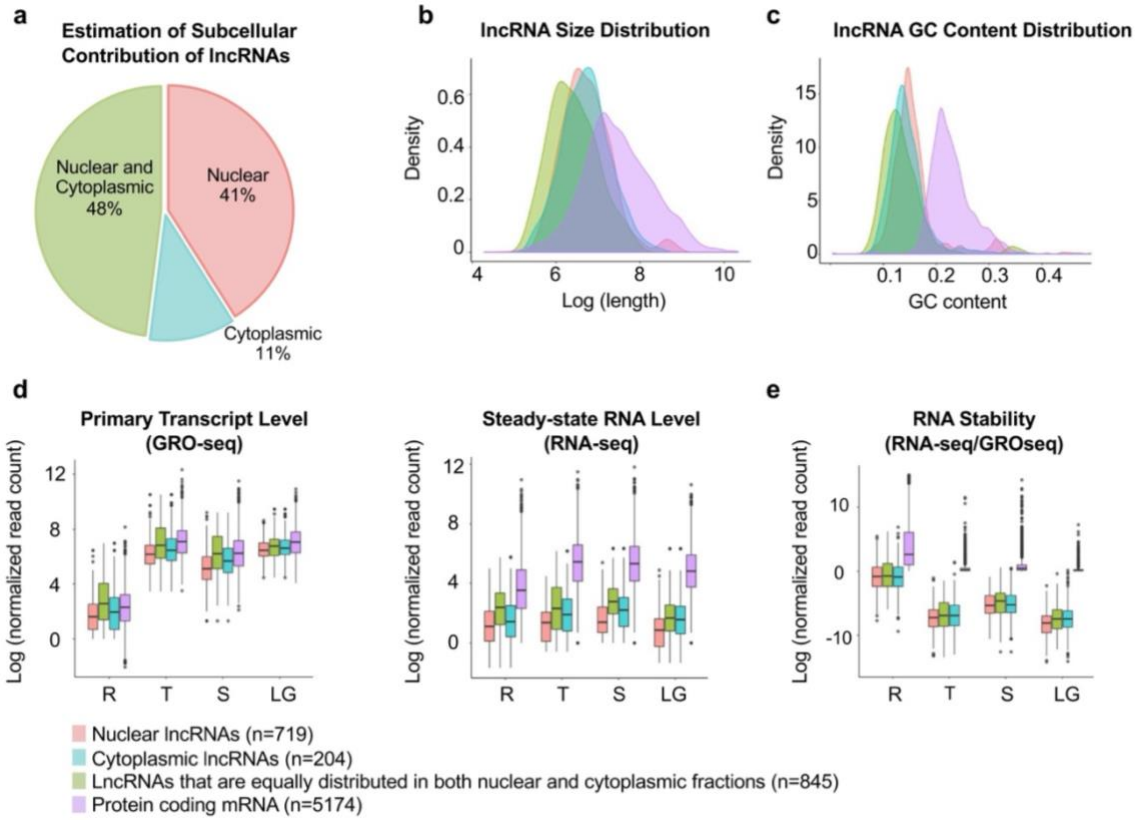


Figure 4-2. Candidate lncRNA categorization. a A total of 1768 lncRNA candidates were identified, covering 719 nuclear enriched lncRNAs (red), 204 cytoplasmic enriched lncRNAs (blue), and 845 lncRNAs found in both fractions (green). Cellular distribution of predicted lncRNAs is based on $\log_2$ fold change > or <0.5 of summed nuclear vs cytoplasmic expression level. Density plots of size (**b**) and GC content (**c**) of lncRNA candidates and annotated protein encoding mRNAs (purple). **d** Expression levels of primary transcripts (left), steady-state RNA (middle) of lncRNA candidates and annotated protein encoding mRNAs for Ring (R), Trophozoite (T), Schizont (S) and Late Gametocyte (LG) stages. **e** Relative stability of lncRNA candidates and annotated protein encoding mRNAs. The stability is based on the ratio of RNA-seq/GRO-seq transcript level. Each box represents the 25/75 percentiles, the line across the box represents the median, and the whiskers represent maximum and minimum values. Outliers are indicated with dots.

Stage-specific expression of cytosolic and nuclear lncRNAs

As lncRNAs often exhibit specific expression patterns in a tissue dependent manner, we investigated the stage specificity of identified candidate lncRNAs across the asexual and sexual life cycle stages. Using k-means clustering, we were able to group lncRNAs into 10 distinct clusters **(Fig. 4-3a).** Generally, nearly all lncRNAs showed a strong coordinated cascade throughout the parasite's life cycle. Similar to what was observed with mRNA, a large fraction of the lncRNAs was highly expressed at mature stages compared to the ring stages **(Fig. 4-3b).** Cluster 1 contains lncRNAs that are more abundantly expressed in the nuclear fraction of ring stage parasites and are lowly expressed in the nuclear fraction of schizont stage parasites. LncRNAs representative of this cluster are the lncRNA-TAREs. We observed that most lncRNA-TAREs identified in this study (19 out of 21) are clustered into this group with an average expression of 1.18 $\log_2$ fold change of nuclear to cytoplasmic ratio **(Fig. 4-3a).**

The remaining two identified lncRNA-TAREs were found in cluster 6, where transcription peaks at the schizont stage. This finding validates our approach and suggests that lncRNAs in this cluster may contribute to the maintenance and regulation of chromatin structure and telomere ends. Approximately 28% of the identified lncRNAs are more abundantly found in either the nuclear or cytoplasmic fraction at the schizont stage (cluster 6, 7 and 8), after DNA replication and the peak of transcriptional activity observed at the trophozoite stage. We observed a few lncRNAs that are solely expressed during the asexual cycle with distinct changes in heterochromatin marks **(Fig. 4-3c).**

Based on clustering analysis, we also found that 460 of our detected lncRNAs are exclusively expressed at a high level at the gametocyte stage (cluster 9 and 10). Interestingly, two unique lncRNAs in this cluster, lncRNA-ch9 (Pf3D7_09_v3:1,384,241-1,386,630) and lncRNA-ch14 (Pf3D7_14_v3:3,148,960 − 3,150,115), were identified in a previous study to be located within heterochromatin regions marked by repressive histone marks H3K9me3 at the trophozoite stage(28)(**Fig. 4-3d and 4-3e).** At the gametocyte stage however, the H3K9me3 was lost. Additionally, both lncRNAs are transcribed from regions adjacent to gametocyte-specific genes. To validate the expression of these gametocyte-specific lncRNAs, we performed RT-PCR **(Fig. 4-S3 and Supplementary data 4-1c)** as well as single cell RNA-seq (scRNA-seq) across key-stages of the parasite life cycle. LncRNA expression was visualized on the UMAP embedding generated from coding gene expression **(Fig. 4-3f).** LncRNA-chr9 and lncRNA-chr14 were expressed in sexual-stage parasites, with lncRNA-ch9 and lncRNA-ch14 were expressed in sexual-stage parasites, with specific enrichment in male and female gametocytes, respectively **(Fig. 4-3f, right panel).** Collectively, these results emphasize the stage-specific expression of parasite lncRNAs and the potential function of gametocyte-specific lncRNAs in regulating sexual development.

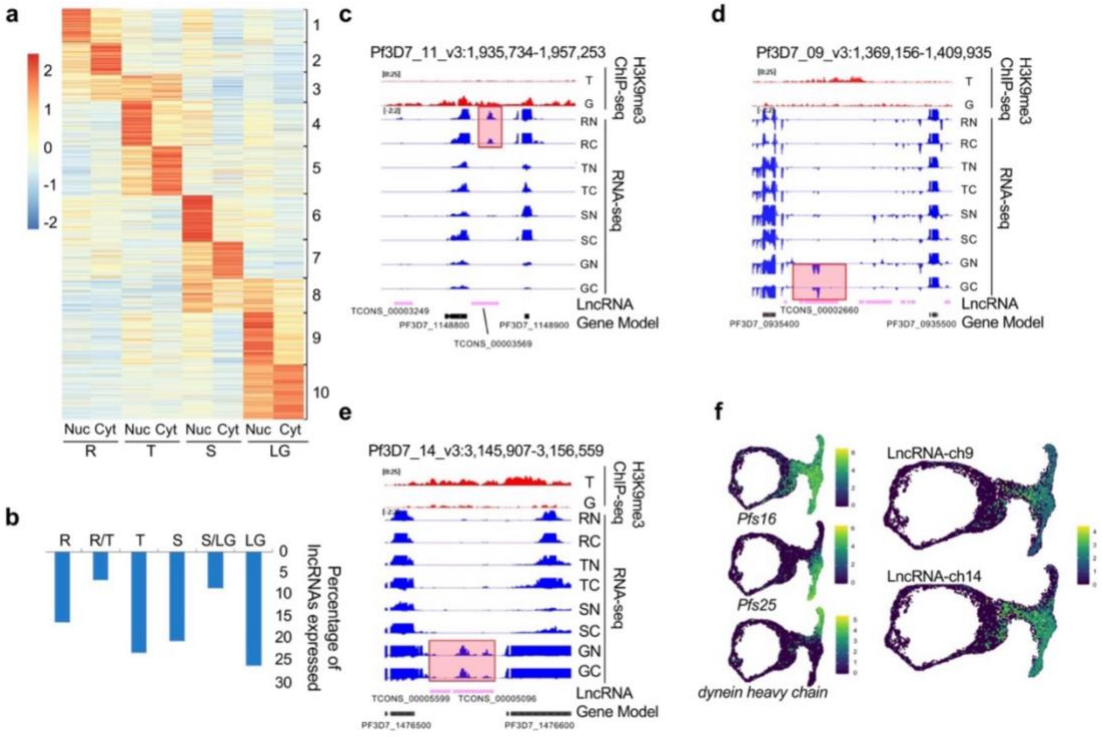


Figure 4-3. Gene expression pattern of lncRNAs. a lncRNAs are grouped into 10 clusters based on their life cycle expression patterns in the Nuclear (Nuc) or Cytoplasmic (Cyt) fraction for the Ring (R), Trophozoite (T), Schizont (S) and Late Gametocyte (LT) stages. **b** Percentage of lncRNAs that are highly expressed at ring, trophozoite, schizont, and late gametocyte stages. Genome browser views of H3K9me3 ChIP-seq and RNA-seq datasets in the region of a representative asexual stage-specific lncRNA on chromosome 11 (**c**), two gametocyte-specific lncRNAs located at the intergenic regions of chromosome 9, lncRNA-ch9 (**d**) and 14, lncRNA-ch14 (**e**). RNA-seq data of Nuclear (N) and Cytoplasmic (C) fraction at Ring (R), Trophozoite (T), Schizont (S) and Gametocyte (G) stages are shown (e.g., RN=nuclear fraction at ring stage). **f** scRNA-seq analysis. 2-dimensional UMAP projection of *P. falciparum* parasites, both asexual (ring) and sexual (T-shape). Each dot represents a single cell. Left panel: Cells colored according to log-normalized gene expression values for gametocytes (*Pfs16* (PF3D7_0406200), top), females (*Pfs25* (PF3D7_1031000), middle), and males (*dynein heavy chain* (PF3D7_0905300), bottom). Right panel: Log-normalized expression of lncRNA-ch9 (top) and lncRNA-ch14 (bottom) across the *P. falciparum* life cycle.

Validation of lncRNA localization and stage-specific expression

To validate the cellular localization of several candidate lncRNAs, we utilized RNA fluorescence in situ hybridization (RNA-FISH). We selected two candidates that were

detected as enriched in the cytoplasmic fraction (lncRNA-267 (Pf3D7_08_v3: 1382128-1382689) and lncRNA-643 (Pf3D7_14_v3:1606672-1607587)), four candidates detected as enriched in the nuclear fraction of asexual parasites (lncRNA-13 (Pf3D7_01_v3:491225-494291), lncRNA-178 (Pf3D7_06_v3:53758-54745), lncRNA-271 (Pf3D7_02_v3:590844-592940) and lncRNA-1494 (Pf3D7_06_v3:1311694-1312858)) and two candidates detected as enriched in the nuclear fraction of sexually mature gametocytes (lncRNA-ch9 and lncRNA-ch14).

Finally, we also selected a lncRNA that had been previously identified and known to be transcribed from the telomere region on chromosome 4, termed lncRNA-TARE4 (Pf3D7_04_v3:1194786-1199684)(47) as our positive control. Briefly, mixed stage parasites were fixed and hybridized to fluorescently labeled ~200-300 nucleotide antisense RNA probes **(Supplementary data 4-1c,d).** The hybridization images, representative of 15-20 parasites, demonstrate that the nuclear lncRNAs localize to distinct foci within or close to the DAPI-stained nuclei **(Fig. 4-4a),** while cytoplasmic lncRNAs are localized outside the DAPI-stained genomic DNA **(Fig. 4-4b).** Additionally, using RNA-FISH, we validated the stage-specific expression of our candidate lncRNAs. Specifically, expression of lncRNA-267 and lncRNA-13 were enriched at the ring and trophozoite stages; lncRNA-178 was expressed at the trophozoite and schizont stages; lncRNA-643 was expressed at the schizont stage only and lncRNA-TARE4 was expressed at all three asexual stages. LncRNA-ch9 and lncRNA-ch14 were only expressed at the gametocyte stage. These

results highlight that, similar to protein-coding transcripts, these candidate lncRNAs are developmentally regulated.

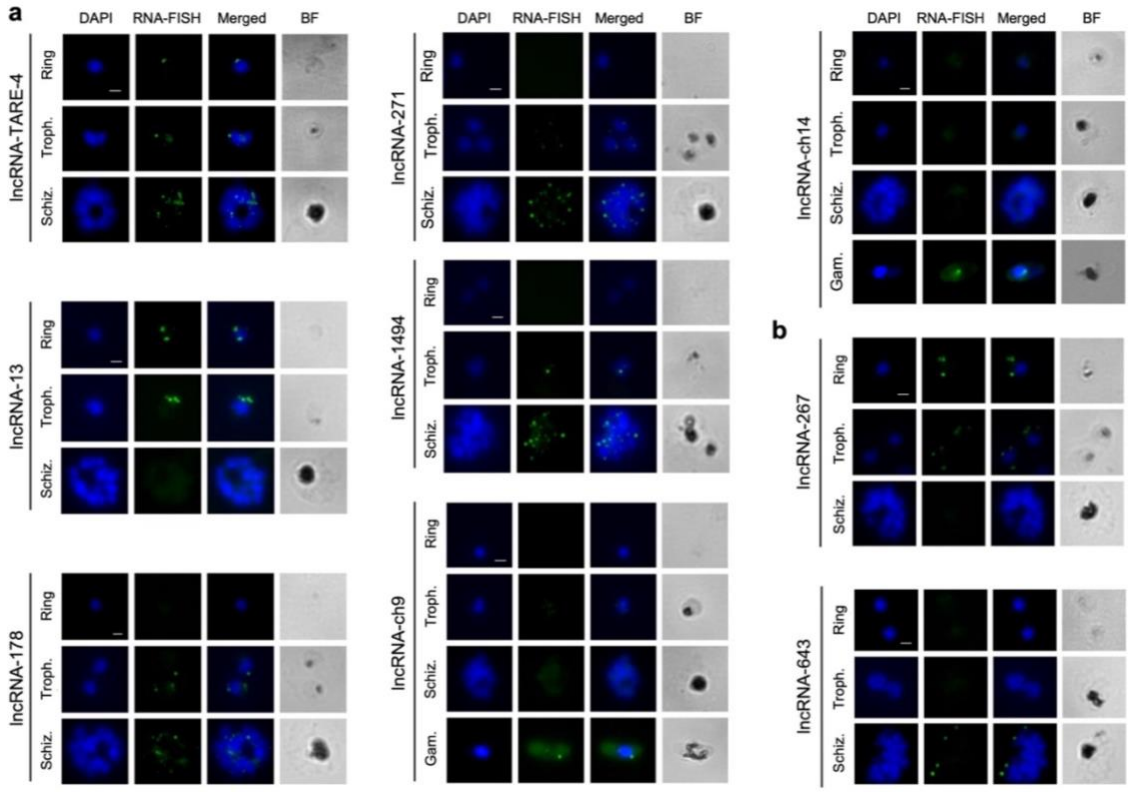


Figure 4-4. RNA-FISH show localization of several candidate lncRNAs. a Nuclear lncRNAs (lncRNA-TARE4, lncRNA-13, lncRNA-178, lncRNA-271, lncRNA-1494, lncRNA-ch9 and lncRNA-ch14) colocalize with nuclei stained with DAPI in Ring, Trophozoite (Troph.), Schizont (Schiz.) and Gametocyte (Gam.) stages. **b** Cytoplasmic lncRNA-267 and lncRNA-643 do not colocalize with the nuclei stained with DAPI. Scale bar indicates 2 μm. Hybridization images are representative of approximately 15 stained parasites from two independent experiments. BF brightfield.

Genomic maps of RNA-chromatin interactions

The locations of the binding sites of most lncRNAs remain unknown. Accordingly, the role of lncRNAs in chromatin and gene regulation in eukaryotes, including the malaria parasite, has been mostly deduced from the indirect effects of lncRNA perturbation. To

explore the role of lncRNAs in gene expression, we sought to identify occupancy sites of our selected candidate lncRNAs within the parasite genome. For an unbiased high-throughput discovery of RNA-bound DNA in *P. falciparum*, we adapted a method termed Chromatin Isolation by RNA Purification (ChIRP) **(Fig. 4-5a)**(73,74). ChIRP-seq is based on affinity capture of target lncRNA: chromatin complex by tiling antisense-biotinylated-oligos to allow the identification of lncRNA-DNA binding sites at single base-pair resolution with high sensitivity and low background(75,76). Such experiments can identify whether lncRNAs are working in *cis* on neighboring genes or in *trans* to regulate distant genes.

ChIRP-seq is applicable to all detected lncRNAs and requires no knowledge of the RNA's structure. This method has recently been used in *Plasmodium* to investigate the role of ncRNA RUF6 on heterochromatin formation(76). In our experiments, synchronized parasites were extracted and crosslinked. Parasite nuclei were then extracted, and chromatin was solubilized and sonicated. Biotinylated antisense oligonucleotides tiling our candidate lncRNAs **(Supplementary data 4-1e,f)** were hybridized to target RNAs and isolated using magnetic beads. These candidates correspond to the seven nuclear lncRNAs validated using RNA-FISH: lncRNA-TARE4, lncRNA-13, lncRNA-178, lncRNA-1494 and lncRNA-271 detected in the nuclear fraction of the asexual stages, as well as lncRNA-ch9 and lncRNA-ch14 detected in the nuclear fraction of the sexual stages. To validate the specificity of the biotinylated oligonucleotides to target our RNA of interest, we performed RT-PCR following our RNA pulldown. RT-PCR results confirmed that lncRNAs and control serine tRNA ligase probes retrieve the selected lncRNA and the serine tRNA ligase

RNA (PF3D7_0717700), respectively **(Fig. 4-5b).** No RNA was retrieved in the negative controls that were incubated with no probes or templates. These results confirm that the biotinylated probes target the RNA of interest with specificity. Purified DNA fragments were then sequenced using next-generation sequencing technology. An input control was used to normalize the signal from ChIRP enrichment.

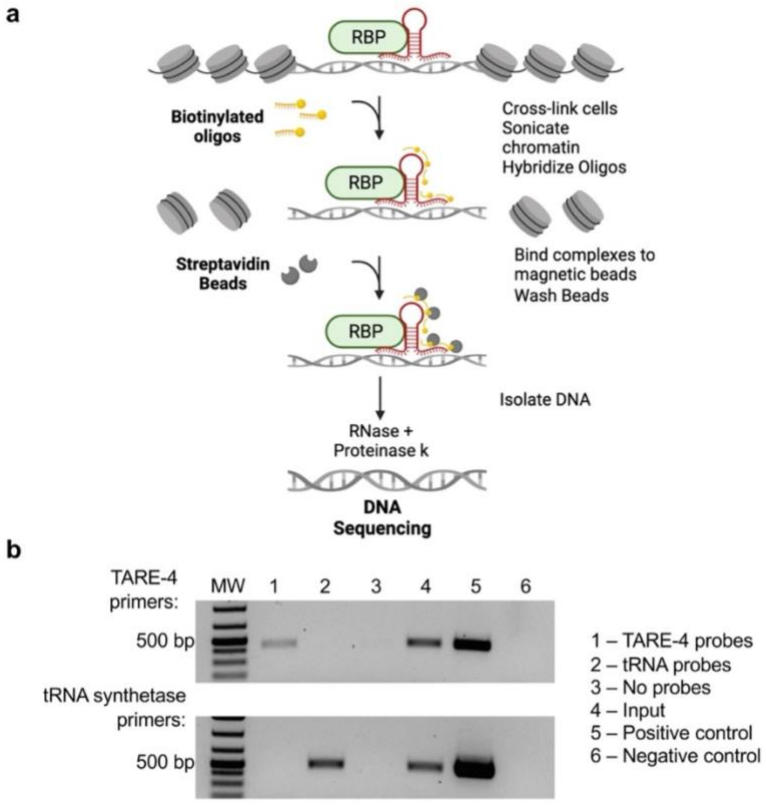


Figure 4-5. Chromatin Isolation by RNA Purification (ChIRP). a Schematic representation of the ChIRP methodology. Created with BioRender.com. RBP: RNA-binding protein. **b** RT-PCR following the ChIRP protocol validates the specificity of the biotinylated antisense probes. Following the ChIRP-pulldown, the RNA fraction was analyzed, and RT-PCR results confirm that lncRNA-TARE-4 probes retrieve the lncRNA-TARE-4 RNA (535 bp PCR product) and the control serine tRNA ligase probes retrieve the serine tRNA ligase RNA (505 bp PCR product), respectively. No RNA was retrieved in the no probe control. RNA from a ChIRP-input sample as well as WT 3D7 parasites (wells 4 and 5, respectively) was used to confirm the lncRNA-TARE-4 and serine tRNA ligase primers. The negative controls (well 6) represent no template controls. RT-PCRs are representative of two independent replicates.

For all seven lncRNAs, ChIRP-seq experiments were performed in duplicate at the stages when the lncRNA was either highly or lowly expressed. We also generated several input controls for each analyzed stage to demonstrate the consistency of our coverage in various heterochromatin and euchromatin regions **(Fig. 4-S4).** For all lncRNAs investigated here, ChIRP-seq performed at time points where the lncRNAs were least expressed retrieved low to no significant signals **(Fig. 4-6a,b and Fig. 4-S4).** The lncRNA-TARE4, transcribed from the telomere region on chromosome 4, and expressed throughout the parasite life cycle stages investigated in this study was found to strongly interact with most telomeres in a very specific manner **(Fig. 4-6a).** Interestingly, one focus per cell for lncRNA-TARE4 was detected by RNA-FISH experiment at the ring stage. One focus per cell was also observed by IFA using an antibody against histone H3K9me3, known to also localize in the telomeres, as well as the subtelomeric and internal *var* gene cluster of the parasite genome(28,65) **(Fig. 4-S5.** This data correlates nicely with the Hi-C data published previously(27,28,77) that demonstrate that most telomere ends, including *var* genes, interact with each other in a large heterochromatin cluster before and after DNA replication. LncRNA-13, transcribed on chromosome 1 and highly expressed at the trophozoite stage **(Fig. 4-S4),** was found to be around surface antigen genes, including PF3D7_0113100 (SURFIN4.1) and PF3D7_1149200 (RESA, ring-infected erythrocyte surface antigen). Both the SURFIN and RESA families of proteins have been implicated in erythrocyte invasion-related processes and are transcribed in mature-stage parasites(78,79). Given that a trophozoite stage lncRNA was identified adjacent to surface

antigen genes which are transcribed at the schizont stage, lncRNA-13 is possibly playing a role in recruiting chromatin modifying enzymes to edit the epigenetic state of the chromatin and allow the recruitment of transcription factor(s) needed to activate the transcription of these genes. To further investigate the possible link between lncRNAs and their role in epigenetics and regulation of gene expression, we developed a software pipeline in Python to identify all specific binding sites in the genome using Bowtie for mapping and PePr for peak calling(80). We then aligned lncRNA ChIRP-seq signals across all 5' and 3' UTRs as well as the gene bodies. Similar to what was detected in higher eukaryotes, we discovered that the lncRNA occupancy is enriched either in the gene bodies for lncRNA-13 and lncRNA-178 or near the end of the 5′ UTR of each gene (lncRNA-1494 and lncRNA-271) **(Fig. 4-S6).** While the data will need to be further validated at the molecular level, this pattern provides support for our candidate lncRNAs to promote either transcriptional elongation or transcriptional initiation, respectively. We next retrieved the genes closest to the identified lncRNA binding sites and calculated the $\log_2$ fold change of their expression from inactive to active stage. We compared the resultant information to the change in the expression profiles for all other genes in the *P. falciparum* genome **(Fig. 4-6c).** For each of the lncRNAs investigated, we detected a significant increase in the expression of the genes near the ChIRP signals. These results indicate that overall, the presence of lncRNA correlates with a significantly increased gene expression. To further demonstrate that lncRNAs interact specifically with DNA, we looked for motif enrichment **(Fig. 4-6d).** Motif analysis of ChIRP-seq data revealed one motif for lncRNA-13 (pval = $1.8e^{-3}$) occurring in 131 of the 138 retrieved lncRNA-13 sequences and two motifs

for lncRNA-TARE-4 (pval = $3.7e^{-7}$ and pval = $5.1e^{-5}$) occurring in 72% and 61% of the TARE-4 binding sites. These data demonstrate that our ChIRP-seq experiments were highly sensitive and specific, and that we were able to retrieve biological insights into their function.

We then focused our attention on ChIRP-seq data generated using probes against lncRNA-ch9 and lncRNA-ch14, two lncRNAs enriched in gametocytes. ChIRP-seq signals **(Fig. 4-6b, Fig. 4-S4, and Supplementary data 4-1k)** showed significant enrichment in the genomic regions where the lncRNAs are transcribed. The lncRNA-ch9 lie between genes that have been implicated in gametocyte differentiation. These include PF3D7_0935500(81), a *Plasmodium* exported protein of unknown function, PF3D7_0935600, a gametocytogenesis-implicated protein, and PF3D7_0935400, Gametocyte development protein 1. These three genes are known to be significantly up regulated in gametocytes and have been demonstrated to be essential to sexual commitment(65). For lncRNA-ch14, the genes are PF3D7_1476500, a probable protein of unknown function, PF3D7_1476600(82), a *Plasmodium* exported protein of unknown function and PF3D7_1476700, a lysophospholipase, three genes on chromosome 14 that are only detected either in gametocyte or ookinete stages. When overlaid with previous ChIP-seq data generated against histone H3K9me3 during the asexual and sexual stages of the parasite life cycle, we noticed that the presence of these lncRNAs correlate with a loss of H3K9me3 marks at the gametocyte stage. For lncRNA-ch14, 11 additional peaks were detected as statistically significant in the gametocyte stage **(Fig. 4-6b, and Supplementary data 4-1).** Most of these peaks were identified in the promotors of genes

that were described as conserved *Plasmodium* protein of unknown function but were also known to be expressed in gametocyte including PF3D7_1145400, a dynamin-like protein overexpressed in female gametocytes(83).While these data will need to be further validated, our results suggest that these lncRNAs may recruit histone demethylase and/or histone acetyl transferase to change the epigenetic state of the chromatin and activate the expression of these genes during sexual differentiation. Collectively, these experiments propose that lncRNAs in the parasite could be essential to recruit chromatin remodeling and modifying enzymes as well as sequence-specific transcription factors to regulate gene expression.

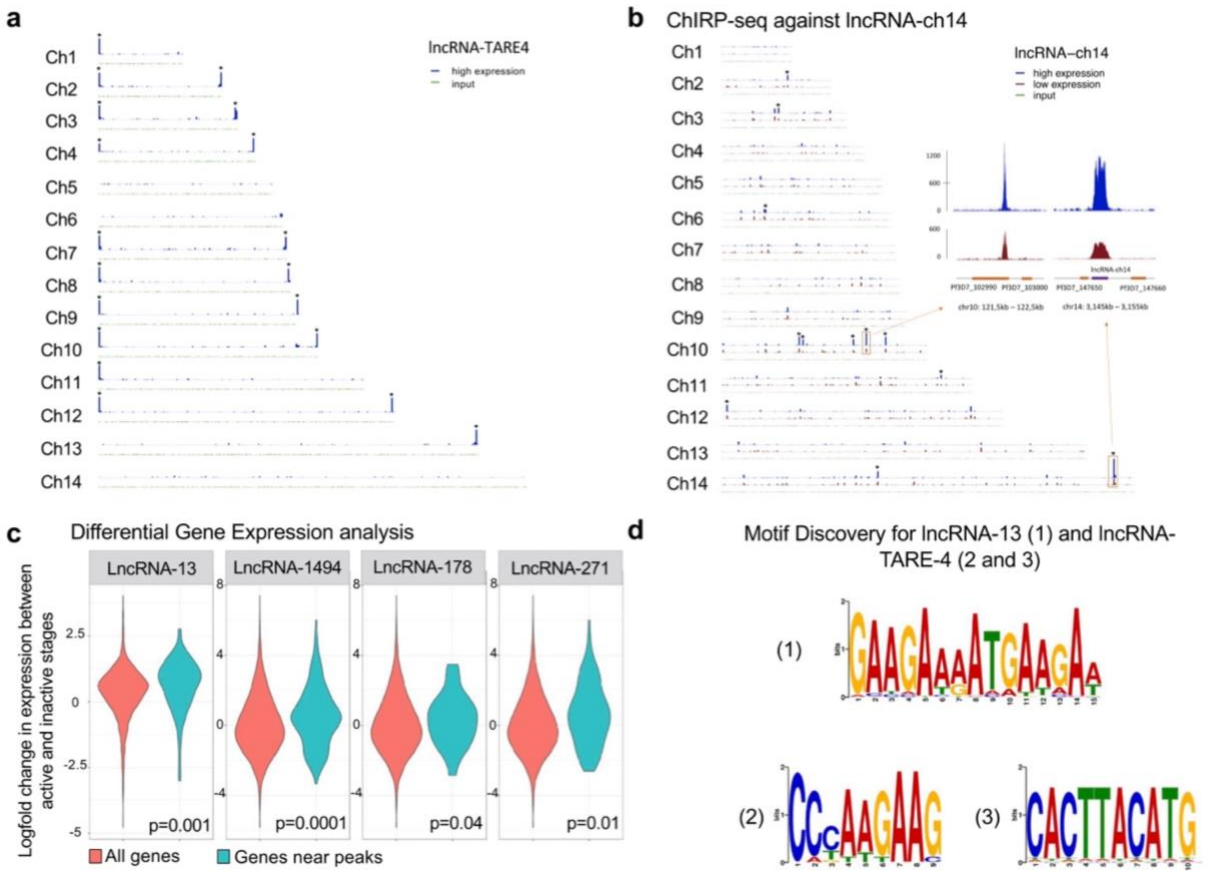


Figure 4-6. ChIRP-seq reveals candidate lncRNA binding sites. a Genome-wide binding sitesof lncRNA-TARE-4 and (**b**) lncRNA-ch14. Mapped, normalized reads from the active stage (top, blue track) and inactive stage (bottom, red track) are shown for each chromosome (Ch). Significant peaks are highlighted with an asterisk. **c** Differential gene expression analysis. The $\log_2$-fold change of gene expression was calculated for the genes closest to the lncRNA peaks between the inactive and active stage (right violin). This distribution was compared to the $\log_2$ fold change in expression for all other genes in the *Plasmodium* genome (left violin) using a two-sided *t* test, and the *p* values are reported at the bottom of each panel. **d** Motif identification. 100 bp sequences centered at the peaks' summits were extracted, and we used STREME specifying 2nd order Markov model and default for the rest of parameters to search for possible motifs. We identified one motif for lncRNA-13 ($p = 1.8e^{-3}$) occurring in 131 of the 138 (95%) lncRNA-13 sequences and two motifs for lncRNA-TARE-4, ($p = 3.7e^{-6}$ and $p = 5.1e^{-5}$), occurring in 72% and 61%, respectively, of the TARE binding areas.

Role of lncRNA-ch14 in sexual commitment and development

We next sought to validate the role of one lncRNA in the parasite development. We therefore selected lncRNA-ch14, which was detected as upregulated in female

gametocytes. We began by disrupting its full length via the CRISPR-Cas9 editing tool. After four unsuccessful attempts, we concluded that either the system was not able to target this genomic region, or this large chromosomal deletion was lethal to the parasite. However, we were able to successfully disrupt the lncRNA-ch14 gene through insertion of a resistance marker spanning the position (Ch14:3,148,960–3,150,115) of the gene **(Fig. 4-S7a).** Parasite lncRNA14 disruptive lines (two clones, named ΔlncRNA-ch14 B1 and F2) were recovered and validated via PCR and RT-PCR **(Fig. 4-S7b).** We also confirmed the absence of obvious off-target effects via whole genome sequencing (WGS). We then examined parasite growth using our two selected clones along with wild-type NF54 parasites during the erythrocytic cycle. Growth was monitored in triplicates using Giemsa-stained blood culture smears for two full cycles. No significant difference was observed in the asexual stages of the ΔlncRNA-ch14 clones compared to the WT **(Fig. 4-S8a).** This indicates that partial disruption of lncRNA-ch14 does not have a significant role in the asexual blood stage replication.

We subsequently aimed to analyze the effect of ΔlncRNA-ch14 in gametogenesis. Relative gametocyte numbers were determined by microscopic examination of Giemsa-stained blood smears prepared from day 16 gametocyte cultures. To consider the impact prolonged culturing times may have on gametogenesis, the assays were conducted between our two ΔlncRNA-ch14 clones as well as two NF54 strains; a NF54 WT lab strain as well as the NF54 parental line used for the initial transfection. The NF54 parental line was maintained in culture in parallel with our selected ΔlncRNA-ch14 clones. We detected a significant

decrease in mature stage V gametocytes in clones B1 and F2 compared to the WT NF54 line ($n = 4$, $p < 0.05$). This decrease was however not detected as significant between the parental NF54 line and ΔlncRNA-ch14 clones **(Fig. 4-7a).** To better understand discrepancies observed between the NF54 lines, we purified gDNA for WGS. While we confirmed successful disruption of ncRNA-ch14 in our two selected clones, we also identified a nonsense mutation in the gametocyte developmental protein 1 (*gdv1)* gene (PF3D7_0935400), a gene essential in sexual differentiation, in the parental line. The mutation detected indicated a premature stop codon leading to a C-terminal truncation of 39 amino acids (GDV39) similar to what was previously observed by *Tibúrcio* and colleagues(84,85). Of the 110 reads covering the mutation site, 43 were shown to retain the reference base while 67 displayed the *gdv1* mutation described. This result suggests the development of a spontaneous mutation after transfection and that the reduced number of gametocytes observed in the parental line were most likely a result of a significant portion (60%) of parasites with the *gdv1* mutation, not producing mature gametocytes. This mutation was however absent in both of our ΔlncRNA-ch14 clones as well as our NF54 WT **(Supplementary data 4-2),** explaining the discrepancies observed in our gametocyte induction assays with the NF54 parental line. Development of spontaneous mutations in culture attests the need of WGS to validate the phenotypes observed in the parasites for both the WT and genetically modified strains. As the parental line was still capable of generating healthy mature gametocytes, albeit at a lower frequency, and gametocytemia is normalized prior to mosquito feeds, the use of the NF54 parental line as our control to analyze the impacts of lncRNA-ch14 disruption on transmission to the mosquito was not

considered to be a major issue because GDV1 is only essential for early sexual commitment. Our reasoning was that the gametocytes produced from the parental line would still closely resemble the ΔlncRNA-ch14 clones at the genomic level.

We then analyzed the formation and ratio of mature male and female gametocytes between the parental line and ΔlncRNA-ch14 clones. To discriminate between male and female gametocytes, blood smears prepared from day 16 gametocyte cultures were stained with Giemsa and ≥100 mature stage V gametocytes were counted to determine sex ratio in each line. Male gametocytes can be distinguished from their female counterparts as they are less elongated with rounder ends and their cytoplasm is distinctly pink while the cytoplasm of female gametocytes, with their large stores of RNA, are dark blue. As shown in (**Fig. 4-7b),** the male to female ratio was significantly affected in the ΔlncRNA-ch14 clones compared to control parasites. In control lines the ratio of female to male gametocytes was approximately 2:1, with the expected larger number of females than males. In contrast, in the lncRNA-ch14 clones it was approximately 2:7, with significantly more males than females ($n = 2, p < 0.05$) **(Fig. 4-7b).** These data suggest that lncRNA-ch14 has a role in the ratio of male and female gametocytes produced under our culture conditions. Exflagellation assays revealed a dramatic drop in microgametocyte exflagellation with an average of a 65% decrease in exflagellation centers observed in the ΔlncRNA-ch14 clones compared to the parental lines, indicating a defect in male gametogenesis and microgamete formation in ΔlncRNA-ch14 parasites **(Fig. 4-7c).** All together these data demonstrate a role of lncRNA-ch14 in gametogenesis.

We next investigated the transmissibility of ΔlncRNA-ch14 gametocytes to mosquitoes. Infectious blood meals were prepared with parental and ΔlncRNA-ch14 stage V gametocytes. Mosquitoes were fed with 0.2% mature stage V gametocyte infected blood using membrane feeders as described earlier[86]. Mosquito midguts were dissected and analyzed on day 11 post blood feeding. As expected, our mutated parental line was able to produce a high number of oocysts and sporozoites. However, we detected a significant decrease in the number of oocysts per midgut in the ΔlncRNA-ch14 clones compared to the parental control. Both prevalence and intensity of infection were impacted in the ΔlncRNA-ch14 clones. While 90% of control infected mosquitoes had oocysts, only 17% and 37% of the ΔlncRNA-ch14 clones, B1 and F2, respectively were positive for oocysts. Additionally, the mosquitoes infected with the ΔlncRNA-ch14 clones had only 1 oocyst while the control had a median of 7 oocysts ($n = 2$, $p < 0.05$) **(Fig. 4-7d).** As expected, these lower oocyst numbers resulted in significantly lower salivary gland sporozoite numbers. On day 17 salivary glands were dissected from control and ΔlncRNA-ch14 infected mosquitoes and the average of number of salivary glands sporozoites from 2 biological replicates was 19,000 for the control line and 667 and 1993 for the ΔlncRNA-ch14 clones B2 and F1, respectively, a decrease of about 96% ($n = 2$, $p < 0.05$) **(Fig. 4-7e).** Altogether, though we could only partially disrupt our candidate lncRNA, we clearly demonstrate that lncRNA-ch14 has an important role in gametocyte development and in the infectivity of these gametocytes for mosquitoes.

Transcriptome perturbation in asexual and sexual stages

Based on our phenotypic assays, we predicted that perturbation of lncRNA-ch14 would affect parasite gene expression. We therefore performed RNA-seq analysis for two biological replicates each of our WT NF54 and ΔlncRNA-ch14 clones. For each parasite line, RNA was extracted at both schizont stage, before sexual commitment and at the late gametocyte stages after sexual commitment. For each respective stage, we observed up-regulation of 78 and 57 genes in the ΔlncRNA-ch14 lines compared to controls as well as downregulation of 383 and 18 genes **(Fig. 4-7f and Supplementary data 4-3).** Gene Ontology (GO) enrichment analysis revealed that up-regulated genes were involved in translation and cytoadherence ($p = 10e^{-14}$) that are known to be expressed in a stage specific manner between the human and vector hosts **(Fig. 4-S8b).** Downregulated genes were mostly involved in microtubule movement, cell signaling and oxidation-reduction processes that are known to be critical during sexual differentiation **(Fig. 4-7g).** Five specific genes known to be upregulated in female gametocytes were detected to be significantly downregulated in our ΔlncRNA-ch14 clones compared to our control line. These genes include PF3D7_1250100, the osmiophilic body protein G377; PF3D7_1407000, LCCL domain-containing protein; PF3D7_0719200, the NIMA related kinase 4; PF3D7_0525900, the NIMA related kinase 2, and PF3D7_1031000, the ookinete surface protein P25. Four male-specific genes were also detected as downregulated in the lncRNA-ch14 clones compared to the control. Those genes included PF3D7_1113900, the mitogen-activated protein kinase 2, PF3D7_1014200, the male gamete fusion factor HAP2;

PF3D7_1216700, the perforin-like protein 2 and PF3D7_1465800, the putative dynein beta chain coding gene. All together this data confirms that lncRNA-ch14 controls, at least partially, the regulation of the transcripts known to be critical in gametocyte development including several key kinases involved in cell signaling relevant to gametogenesis.

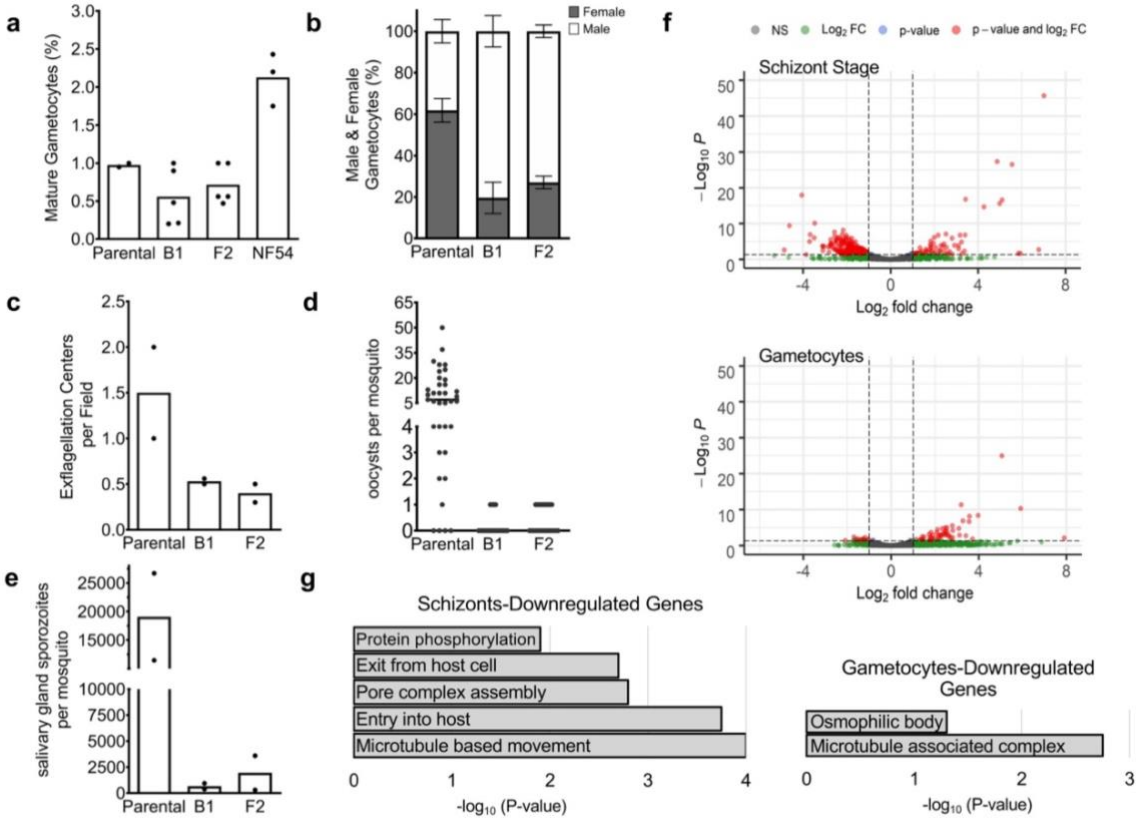


Figure 4-7. LncRNA-ch14 disruption design and characterization. a Percentage of mature gametocytes. Gametocyte cultures were sampled by Giemsa-stained blood smears to assess the percent of Stage V gametocytes. Assays were performed at least in duplicate and significance of the results was calculated using the one-way ANOVA with Dunnett's multiple comparison test (**$p < 0.01$, ***$p < 0.001$). **b** Percentage of gametocytes identified as male and female. Two independent biological experiments were performed, and data are presented as mean values ±SD ($n > 500$ mature gametocytes). Significance of the results was calculated using the two-way ANOVA with Holm-Šídák correction (**$p < 0.01$, ***$p < 0.001$). **c** Exflagellation assays were performed and the number of exflagellation centers per field ($n > 10$) were counted. Shown are results from two biological replicates. Bars indicate the mean. (d&e) Mosquito passage: Gametocyte cultures were fed to *Anopheles stephensi* mosquitoes. On day 11 post-infection, midguts were removed, and oocysts were counted (**d**) and on day 17 post-infection salivary glands were harvested and sporozoites were counted (**e**). Data are pooled from 2 biological replicates. Oocyst counts were performed with 15–25 mosquito midguts per experiment. Significance of the results was calculated using the Kruskal-Wallis test with Dunn's post-test (****$p < 0.0001$). For salivary gland sporozoite counts, salivary glands from 20 mosquitoes were harvested, homogenized and sporozoites counted using a hemocytometer. Shown is the average number of salivary gland sporozoites for each of two experiments. Significance was calculated using the Chi-Squared tests (****$p < 0.0001$). **f** Volcano plots for gene expression profile between the WT and ΔlncRNA-ch14 lines by $-\text{Log}_{10}$ P (y-axis) and $\log_2$ Fold change (x-axis) in asexual stage parasites (Top) and in mature gametocytes (Bottom). NS: Non-significant; FC: Fold Change. **g** Bar graph representations of selected Gene Ontology (GO) enrichment of downregulated genes between WT and ΔlncRNA-ch14 lines are presented by Log_{10} *P* (y-axis) in asexual mature (Left) and mature gametocyte (Right) stages. Exact *p* values and raw data are indicated in the Source data file.

DISCUSSION

Many lncRNAs are now recognized as essential regulators of chromatin structure and gene expression in eukaryotes. While some of the identified lncRNAs have been shown to work in *cis* on neighboring genes, others seem to work in *trans* to regulate distantly located genes. Specifically, functions of nuclear lncRNAs have been determined as either directly promoting or repressing gene expression activity(87,88) guiding or enhancing the functions of regulatory proteins(37,88-91) or assisting the alteration of chromatin structures by shaping 3D genome organization(39,92,93).

The extent of lncRNA regulation in the human malaria parasite is only now starting to emerge. A few lncRNAs have already been suggested to regulate *var* gene expression(57,63,94) or drive sexual commitment(65,66) confirming that at least some of the identified *P. falciparum* lncRNA candidates may have a functional role in the parasite life cycle progression.

In *P. falciparum*, emerging evidence has shown that chromatin structure and genome organization are of vital importance for the parasite's gene expression and regulation system(28,95). Depending on their localization and their specific interactions with DNA, RNA and proteins, lncRNAs can modulate chromatin and epigenetics in the nucleus or mRNA stability and translation in the cytosol, ultimately affecting gene expression. Therefore, identification and characterization of nuclear or cytoplasmic enriched lncRNAs may support the discovery of lncRNAs that are either chromatin-associated or translational-associated regulators of gene expression in the parasite. The dataset generated in this study presents the first global detection of lncRNAs from different subcellular

locations throughout several *P. falciparum* life cycle stages. By utilizing published total and nascent RNA expression profiles (GRO-seq)(8), we were able to significantly improve the sensitivity of lncRNA detection, especially for the identification of nuclear lncRNAs. Using both experimental and computational pipelines, we identified 1768 lncRNAs covering 204 cytoplasmic enriched, 719 nuclear enriched, and 845 lncRNAs that localized to both fractions. Our data suggest that nuclear and cytoplasmic lncRNAs are coordinately expressed but cytoplasmic lncRNAs are less abundant as compared to the number of nuclear lncRNAs in the parasite. In addition, we observed that a small group of cytoplasmic lncRNAs is highly expressed at the trophozoite stage, the stage where a large proportion of genes are transcribed(8). Though more in-depth studies will be required to confirm the functions of these trophozoite-expressed cytoplasmic lncRNAs, it is possible that some of these lncRNAs are involved in mRNA stability or translational regulation.

In our present work, we also observed that many lncRNAs enriched in the nuclear fraction, including the lncRNA-TAREs, are highly abundant at the ring and schizont stages. This finding suggests that some of these lncRNAs **(cluster 1, Fig. 4-3a)** are likely to be involved in heterochromatin maintenance or chromatin structure re-organization events, as previous ChIP-seq and Hi-C experiments have shown that epigenetics and chromatin are critical to gene expression at the initiation level(27). Additionally, ChIRP experiments mapping genome-wide binding sites of lncRNAs revealed that lncRNA-TARE4 binds to subtelomeric regions on multiple chromosomes as well as regulatory regions around genes involved in pathogenesis and immune evasion. Previous reports showed that subtelomeric

regions as well as virulence gene families cluster in perinuclear heterochromatin. Therefore, evidence suggests a role for lncRNA-TARE4 in transcriptional and/or epigenetic regulation of parasite telomeric and subtelomeric regions by interacting with or recruiting histone-modifying complexes to targeted regions to maintain them in a heterochromatin state, much like the case of X chromosome inactivation regulated via lncRNA Xist(96).

Genomic occupancy of other lncRNAs explored here, suggest that these lncRNAs bind around the gene regions. In all cases investigated, a positive correlation was observed between the lncRNA expression and the expression of genes around the lncRNA occupancy sites **(Fig. 4-6).** Given already existing evidence for lncRNA-associated epigenetic modification and transcriptional regulation in other eukaryotes(97-99), it is likely that the lncRNAs identified in the parasite nucleus are responsible for coordinated recruitment of distinct repressing proteins and/or histone-modifying complexes to target loci. Additionally, we uncovered that the lncRNA binding sites were situated upstream of the start codon of target genes **(Fig. 4-S6).** This pattern of lncRNA occupancy provides additional support for the idea that the lncRNAs explored here might have a role in recruiting protein complexes to promoter regions of target genes to regulate transcription, either by activating the formation of the pre-initiation complex or recruiting histone modifiers. However, while additional experiments are needed to confirm the roles of these nuclear lncRNAs in the parasite, using ChIRP-seq, we demonstrate that genome-wide collections of RNA binding sites can be used to discover the DNA sequence motifs

enriched by lncRNAs. These findings signify the existence of lncRNA target sites in the genome, an entirely new class of regulatory elements that could be essential for transcriptional regulation in the malaria parasite.

Genetic disruption of lncRNA-ch14, a transcript detected specifically in gametocytes, demonstrates that this lncRNA plays an important role in sexual differentiation and is required for onward transmission to the mosquito **(Fig. 4-7a,b).** This finding is supported by our transcriptomic analysis where we identified significant downregulation of genes involved in sexual differentiation including NEK and MAP kinases(110,101) ookinete/oocyst development(102) and microtubule function (i.e., dyneins and kinesins) most likely important in reshaping the parasite into sexual stages **(Fig. 4-7f,g and Fig. 4-S8b and Supplementary data 4-3).** Importantly, the skewed sex ratio of the ΔlncRNA-ch14 parasites does not completely account for the dramatic decrease in the ability of these parasites to be transmitted to mosquitoes. Indeed, our data suggest that the gametocytes of the ΔlncRNA-ch14 parasites are less infectious, a result that is also supported by the decrease in exflagellation of the male gametocytes **(Fig. 4-7c).** It is currently difficult to assess infectiousness of female gametocytes, but we would hypothesize that these are also impacted by the disruption of lncRNA-ch14. While further experiments will be needed to further validate the effect of the full deletion or downregulation of the lncRNA-ch14 transcript in the mosquitoes stage, the results presented here confirm that some of the lncRNAs identified in this study play a role in the parasite's sexual development and onward transmission to the mosquito.

Compared to the progress made in understanding lncRNA biology in higher eukaryotes, the field of lncRNAs in *Plasmodium* is still evolving. Analysis of promoter and gene body regions with available histone modification datasets (H3K9me3, H3K36me3, and H3K9ac) are still needed for further annotation of these candidate lncRNAs. It is clear that lncRNAs represent a new paradigm in chromatin remodeling and genome regulation. Therefore, this newly generated dataset will not only assist future lncRNA studies in the malaria parasite but will also help in identifying parasite-specific gene expression regulators that can ultimately be used as new anti-malarial drug targets.

MATERIALS AND METHODS

Parasite culture

P. falciparum 3D7 strain at ~8% parasitemia was cultured in human erythrocytes at 5% hematocrit in 25 mL of culture as previously described in (103). Two synchronization steps were performed with 5% D-sorbitol treatments at ring stage within eight hours. Parasites were collected at early ring, early trophozoite, and late schizont stages. Parasite developmental stages were assessed using Giemsa-stained blood smears.

Nuclear and cytosolic RNA isolation

Highly synchronized parasites were first extracted using 0.15% saponin solution followed by centrifugation at 1500 x *g* for 10 min at 4 °C. Parasite pellets were then washed twice with ice cold PBS and re-collected at 1500 × *g*. Parasite pellets were resuspended in 500 µL

ice cold Cell Fractionation Buffer (PARIS kit, ThermoFisher; AM1921) with 10 µL of RNAse Inhibitor (SUPERaseIn 20U/µL, Invitrogen; AM2694) and incubated on ice for 10 minutes. Samples were centrifuged at 500 x *g* for 5 min at 4 °C. After centrifugation, the supernatant containing the cytoplasmic fraction was collected. Nuclei were resuspended in 500 µL Cell fractionation buffer and 15 µL RNAse Inhibitor as described above. To obtain a more purified nuclear fraction, the pellet was syringed with a 26 G inch needle five times. The sample was incubated on ice for 10 min and centrifuged at 500 × *g* for 5 min at 4 °C. The nuclear pellet was resuspended in 500 µL of ice-cold Cell Disruption Buffer (PARIS kit, ThermoFisher; AM1921). For both cytoplasmic and nuclear fractions, RNA was isolated by adding 5 volumes of Trizol LS Reagent (Life Technologies, Carlsbad, CA, USA) followed by a 5 min incubation at 37 °C. RNA was then isolated according to manufacturer's instructions. DNA-free DNA removal kit (ThermoFisher; AM1906) was used to remove potential genomic DNA contamination according to manufacturer's instruction, and the absence of genomic DNA was confirmed by performing a 40-cycle PCR on the PfAlba3 gene using 200 to 500 ng input RNA.

mRNA isolation and library preparation

Messenger RNA was purified from total cytoplasmic and nuclear RNA samples using NEBNext Poly(A) nRNA Magnetic Isolation module (NEB; E7490S) with manufacturer's instructions. Once mRNA was isolated, strand-specific RNA-seq libraries were prepared using NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB; E7420S) with library amplification specifically modified to accommodate the high AT content of *P.*

falciparum genome: libraries were amplified for a total of 12 PCR cycles (45 s at 98 °C followed by 15 cycles of 15 s at 98 °C, 30 s at 55 °C, 30 s at [62 °C], 5 min 62 °C). Libraries were then sequenced on Illumina NExtSeq500 generating 75 bp paired-end sequence reads.

Sequence mapping

After sequencing, the quality of raw reads was analyzed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The first 15 bases and the last base were trimmed. Contaminating adaptor reads, reads that were unpaired, bases below 28 that contained Ns, and reads shorter than 18 bases were also filtered using Sickle (https://github.com/najoshi/sickle)(104). All trimmed reads were then mapped to the *P. falciparum* genome (v34) using HISAT2(105) with the following parameters: –t, –downstream-transcriptome assembly, –max-intronlen 3000, –no-discordant, –summary-file, –known-splicesite-infile, –rna-strandness RF, and –novel-splicesite-outfile. After mapping, we removed all reads that were not uniquely mapped, not properly paired (samtools v 0.1.19-44428 cd(106)) and are likely to be PCR duplicates (Picard tools v1.78, broadinstitute.github.io/picard/). The final number of working reads for each library is listed in Supplementary data 4-1. For genome browser tracks, read coverage per nucleotide was first determined using BEDTools(107) and normalized per million mapped reads.

Transcriptome assembly and lncRNA identification

To identify lncRNAs in the nuclear and cytoplasmic fractions, we first merged all nuclear libraries and cytoplasmic libraries for each replicate, resulting in one pair of nuclear and

cytoplasmic dataset per replicates. Next, we assembled the transcriptome (cufflinks v2.1.1) (108) for each of the datasets using the following parameters: -p 8 -b PlasmoDB-34_*Pfalciparum*3D7_Genome.fasta -M PlasmoDB-34_*Pfalciparum*3D7.gff –library-type first strand -I 5000. After transcriptome assembly, we filtered out transcripts that are less than 200 basepairs and were predicted to be protein-coding (CPAT, http://lilab.research.bcm.edu). We then merge transcripts in both replicates using cuffmerge and removed any transcripts that are on the same strand and have more than 30% overlap with annotated regions (BEDTools intersect). Lastly, we further selected transcripts that have both primary and steady-state transcriptional evidence. For primary transcription, we used GRO-seq dataset (GSE85478) and removed any transcript that has a read coverage below 15% of the median expression of protein-encoding genes, as well as transcript that has an FPKM count less than 10 at any given stage. The same filtering criteria were also applied to the steady-state RNA-seq expression profiles.

To estimate the cellular location for each predicted lncRNA, we first calculated the summed read count of all nuclear libraries and the summed read count of all cytoplasmic libraries. Then, we measured the $\log_2$-fold change ($\log_2$ FC) of the summed nuclear signal to the summed cytoplasmic signal. Any transcripts with a $\log_2$ FC value above 0.5 were classified as nuclear enriched lncRNAs, and any transcripts with a $\log_2$ FC value below −0.5 were classified as cytoplasmic enriched lncRNAs. In addition, lncRNAs with a $\log_2$ FC value between the above thresholds were classified as lncRNAs expressed equally in both fractions.

Western blot

Mixed-stage parasites were collected as described above. Parasite pellets were gently resuspended in 500 µL of ice-cold Cell Fractionation Buffer (PARIS kit, ThermoFisher; AM1921) and 50 µL of 10X EDTA-free Protease inhibitor (cOmplete Tablets, Mini EDTA-free, EASY pack, Roche; 05 892 791 001). Solution was incubated on ice for 10 min and the sample was centrifuged for 5 min at 4 °C and 500 × *g*. The supernatant containing cytoplasmic fraction was collected carefully and the nuclear pellet was resuspended in 500 µL Cell Fractionation Buffer followed by needle-lysis 5x using 26 G inch needle. Nuclei were collected again at 4 °C and 500 × g. The supernatant was discarded and the nuclei pellet in 500 µL of Cell Disruption Buffer (PARIS kit, ThermoFisher; AM1921) and incubated on ice for 10 min. The nuclear fraction was then sonicated 7x with 10 s on/30 s off using a probe sonicator. Extracted nuclear protein lysates were incubated for 10 min at room temperature and centrifuged for 2 min at 10,000 × *g* to remove cell debris. Seven micrograms of parasite cytoplasmic and nuclear protein lysates were diluted in a 2X laemmli buffer at a 1:1 ratio followed by heating at 95 °C for 10 min. Protein lysates are then loaded on an Any-KD SDS-PAGE gel (Bio-Rad, 569033) and run for 1 h at 125 V. Proteins were transferred to a PVDF membrane for 1 h at 18 V, then stained using commercial antibodies generated against histone H3 (1:3000 dilution, Abcam; ab1791) and PfAldolase (1:1000 dilution, Abcam; ab207494), and secondary antibody, Goat Anti-Rabbit IgG HRP Conjugate (1:25,000 dilution, Bio-Rad; 1706515).

Membranes were visualized using the Bio-Rad ChemiDoc MP Gel Imager and images were treated using Image Lab software. Uncropped blots are presented in Source data file.

PiggyBac insertion analysis

To analyze lncRNA essentiality, we used *piggyBac* insertion sites from(68), who performed saturation mutagenesis to uncover essential genes in *P. falciparum*. Since that study used an NF54 reference genome, we converted the coordinates to be applicable to the 3D7 reference genome (v38, PlasmoDB), using liftOver (Kent tools v427, UCSC Genome Bioinformatics group, https://github.com/ucscGenomeBrowser/kent). A chain file for the two genomes, needed for liftOver, was manually constructed as described here: http://genomewiki.ucsc.edu/index.php/LiftOver_Howto.

Custom Python scripts were used to overlap insertion site coordinates with lncRNA ranges to count the number of insertions that occurred in each lncRNA, as well as to locate TTAA sites (sites where *piggyBac* insertions could potentially occur) in the genome and count the number of TTAA sites in each lncRNA. These scripts were also used to determine the normalized location of each TTAA site and insertion site, in one of 50 windows either across the lncRNA range or also including the 5' and 3' flanking regions, which were each given 50% of the length of the lncRNA. The ratio of number of insertion sites to number of TTAA sites within a lncRNA was used as a loose measure of essentiality.

Estimation of transcript stability

Read coverage values were calculated from total steady-state RNA datasets (SRP026367, SRS417027, SRS417268, SRS417269) using BEDTools v2.25.0. The read counts were then normalized as described in the original publication, and ratios between RNA-seq and GRO-seq coverage values were calculated for each lncRNA and gene. This ratio reflects the relative abundance of the mature RNA transcript over its corresponding primary transcript and is a simple but convenient measurement for transcript stability.

Reverse transcription PCR

Total RNA was isolated from 10 mL of mixed-stage asexual *P. falciparum* culture and 25 mL of late gametocyte stage culture. Total RNA quality was checked on an agarose gel and genomic DNA contamination was removed using a DNA-free DNA removal kit (ThermoFisher; AM1906) according to manufacturer's instructions. The absence of genomic DNA was validated using a primer set targeting an intergenic region within *PfAlba3* (PF3D7_1006200). Approximately 1 µg of DNase I treated RNA from each sample was used in a 35-cycle PCR reaction to confirm the absence of genomic DNA contamination. DNase-treated total RNA was then mixed with 0.1 µg of random hexamers, 0.6 µg of oligo-dT (20), and 2 µL 10 mM dNTP mix (Life Technologies) in total volume of 10 µL, incubated for 10 minutes at 70 °C and then chilled on ice for 5 minutes. This mixture was added to a solution containing 4 µL 10X RT buffer, 8 µL 20 mM MgCl2, 4 µL 0.1 M DTT, 2 µL 20U/µl SuperaseIn and 1 µL 200 U/µL SuperScript III Reverse Transcriptase (Invitrogen, 18080044). First-strand cDNA was synthesized by incubating

the sample for 10 min at 25 °C, 50 min at 50 °C, and finally 5 minutes at 85 °C. First strand cDNA is then mixed with 70 µL of nuclease free water, 30 µL 5x second-strand buffer (Invitrogen, 10812014), 3 µL 10 mM dNTP mix (Life Technologies), 4 µL 10 U/µl *E. coli* DNA Polymerase (NEB, M0209), 1 µL 10 U/µL *E. coli* DNA ligase (NEB, M0205) and 1 µL 2 U/µL *E. coli* RNase H (Invitrogen, 18021014). Samples were incubated for 2 h at 16 °C and double stranded cDNA was purified using AMPure XP beads (Beckman Coulter, A63881). For testing transcription activity of predicted genes, 450 ng of double stranded cDNA was mixed with 10 pmole of both forward and reverse primers. DNA was incubated for 5 minutes at 95 °C, then 30 s at 98 °C, 30 s at 55 °C, 30 s at 62 °C for 25 cycles. All primers used for PCR validation are listed in Supplementary data.

Single-cell sequencing and data processing

P. falciparum strain NF54 was cultured in O+ blood in complete RPMI 1640 culture medium at 37 °C in a gas mixture of 5% O_2/5% CO_2/90% N_2, as described previously(9,109). Sexual commitment was induced at 1% parasitemia and 3% hematocrit and culture media were supplemented with 10% human serum. After 4, 6 and 10 days post sexual commitment, samples were taken from the culture for single cell sequencing. Cells from each day were loaded into separate inlets in a 10X chromium controller using the manufacturer's instructions for a 10,000-target cell capture. Libraries for the days 4 and 6 samples were obtained using Chromium 10X version 2 chemistry, whereas libraries for the day 10 sample were obtained using version 3 chemistry. Cells were sequenced on a single lane of a HighSeq4000 using 150-bp paired-end reads. Raw reads were mapped to a custom

gtf containing lncRNA coordinates appended to the *P. falciparum* 3D7 V3 reference genome (www.sanger.ac.uk/resources/downloads/protozoa/). Read mapping, deconvolution of cell barcodes and UMIs and the generation of single cell expression matrices were performed using the CellRanger pipeline v 3.0.0. LncRNA regions were labeled as 'protein coding' to be prioritized in STAR mapping in CellRanger. CellRanger was also run separately for each sample using the 3D7 reference genome that did not contain the appended non-coding regions for comparison. Resultant count matrices were loaded into the R package Seurat (v3.2.2) for pre-processing.

Quality control and lncRNA expression

Single-cell transcriptomes (SCTs) were log-normalized, and expression scaled using Seurat (v.3.2.2). Each cell was assigned a stage by mapping to the Malaria Cell Atlas[109] using scmap-cell (v1.8.0). Cells were assigned the stage of their closest neighbor in the Malaria Cell Atlas if they reached a cosine similarity of > 0.2. Cells identified as an early/late ring or late schizont containing <50 UMI/cell and <50 genes/cell were removed due to poor quality. Cells mapped to late stages, or cells not assigned to a stage in the Malaria Cell Atlas were removed if they contained <100 UMIs/cell or <80 genes/cell. Data from days 4, 6 and 10 were integrated together using Seurat's IntegrateData function using 2000 integration anchors and 10 significant principal components. A variance stabilizing transformation was performed on the integrated matrix to identify the 750 most highly variable coding genes, and these were used to perform a principal component (PC) analysis. Significant PCs were then used to calculate three-dimensional UMAP embeddings using

only coding genes. LncRNA expression was visualized on the UMAP embedding generated from coding gene expression using the package ggplot2 to assign stage-specific expression for the lncRNA.

RNA in situ hybridization (RNA-FISH)

RNA FISH was performed with slight modifications as described by Sierra-Miranda, 2012[59] on mixed-stage asexual and gametocyte stage parasites. Antisense RNA probes for seven nuclear lncRNAs; -TARE4, −178, −13, −1494, −271, −4076 -ch9, -ch14 and two cytoplasmic lncRNAs; −267, −643, were labeled by in vitro transcription in the presence of fluorescein. RNA FISH was also performed using sense RNA probes as controls. Briefly, fixed and permeabilized parasites were incubated with RNA probes overnight at 37 °C. Parasites were washed with 2x SSC three times for 15 min each at 45 °C followed by one wash with 1x PBS for 5 min at room temperature. The slides were mounted in a Vectashield mounting medium with DAPI and visualized using the Olympus BX40 epifluorescence microscope. Images were treated with ImageJ. Pictures are representative of 15-20 positive parasites examined.

Immunofluorescence assays

Parasites were fixed with 4% paraformaldehyde and 0.0075% glutaraldehyde for 15 min at 4˚C, and then sedimented on Poly-D-lysine coated coverslips for 1 hr. at room temperature. After PBS washes, parasites were permeabilized and saturated with 0.2% Triton X-100, 5% BSA, 0.1% Tween 20 in PBS for 30 min at room temperature. Anti-H3K9me3 mAb

(Abcam, ab184677) was diluted at 1:500 in 5% BSA, 0.1% Tween 20 and PBS, and applied for 1 hr. at room temperature. After PBS washes, Goat anti-Mouse Alexa Fluor 488 (Invitrogen, A11001) was diluted at 1:2000 and applied for 1 h at room temperature. Coverslips were mounted in Vectashield Antifade Mounting Medium with DAPI (Vector Laboratories, H-1200). Images were acquired using a Keyence BZ-X810 fluorescence microscope and treated with ImageJ.

Chromatin isolation by RNA purification (ChIRP)

ChIRP-seq experiments were performed in duplicate for all nuclear lncRNAs, at the time point of highest lncRNA expression and lowest expression. Synchronized parasite cultures were collected and incubated in 0.15% saponin for 10 min on ice to lyse red blood cells. Parasites were centrifuged at 3234 x g for 10 min at 4 °C and subsequently washed three times with PBS by resuspending in cold PBS and centrifuging for 10 min at 3234 x *g* at 4 °C. Parasites were cross-linked for 15 min at RT with 1% glutaraldehyde. Cross-linking was quenched by adding glycine to a final concentration of 0.125 M and incubating for 5 min at 37 °C. Parasites were centrifuged at 2500 x *g* for 5 min at 4 °C, washed three times with cold PBS and stored at −80 °C.

To extract nuclei, parasite was first incubated on in nuclear extraction buffer (10 mM HEPES, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 0.5 mM 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF), EDTA-free protease inhibitor cocktail (Roche) and phosphatase inhibitor cocktail (Roche)) on ice. After 30 min, Igepal CA-630 (Sigma-Aldrich, I8896) was added to a final concentration of 0.25% and

needle sheared seven times by passing through a 26 G ½ needle. Parasite nuclei were centrifuged at 2500 x *g* for 20 min at 4 °C and resuspended in shearing buffer (0.1% SDS, 1 mM EDTA, 10 mM Tris–HCl pH 7.5, EDTA-free protease inhibitor cocktail and phosphatase inhibitor cocktail). Chromatin was fragmented using the Covaris UltraSonicator (S220) to obtain 100-500 bp DNA fragments with the following settings: 5% duty cycle, 140 intensity incident power, 200 cycles per burst. Sonicated samples were centrifuged for 10 min at 17000 × *g* at 4 °C to remove insoluble material.

Fragmented chromatin was precleared using Dynabeads MyOne Streptavidin T1 (Thermo Fisher, 65601) by incubating for 30 min at 37 °C to reduce non-specific background. Per ChIRP sample using 1 mL of lysate, 10 µL each was removed for the RNA input and DNA input, respectively. Each sample was diluted in 2x volume of hybridization buffer (750 mM NaCl, 1% SDS, 50 mM Tris-Cl pH 7.5, 1 mM EDTA, 15% formamide, 0.0005x volume of AEBSF, 0.01x volume of SUPERase-In (Ambion, AM2694) and 0.01x volume of protease inhibitor cocktail). ChIRP probes used for each lncRNA (see Supplementary data [1]) were pooled, heated at 85 °C for 3 min and cooled on ice. ChIRP probes were added to each sample (2 µL of 100 µM pooled probes per sample) and incubated at 37 °C with end-to-end rotation for 4 h. Prior to completion of hybridization, Dynabeads MyOne Streptavidin T1 beads were washed three times on a magnet stand using lysis buffer (50 mM Tris-Cl pH 7, 10 mM EDTA, 1% SDS). After the hybridization, 100 µL of washed T1 beads were added to each tube and incubated for 30 min at 37 °C. Beads were washed with wash buffer (2x SSC, 0.5% SDS, 0.005x volume of AEBSF) and split evenly for isolation of DNA and RNA fractions.

For RNA isolation, the RNA input and chromatin-bound beads were resuspended in RNA elution buffer (100 mM NaCl, 10 mM Tris-HCl pH 7.0, 1 mM EDTA, 0.5% SDS, 1 mg/mL Proteinase K), incubated at 50 °C for 45 min, boiled at 95 °C for 15 min and subjected to trizol:chloroform extraction. Genomic DNA contamination was removed using a DNA-free DNA removal kit (ThermoFisher, AM1906) according to manufacturer's instructions. The absence of genomic DNA was validated using a primer set targeting an intergenic region within (PF3D7_1006200) in a 35-cycle PCR reaction. DNase-treated RNA was then mixed with 0.1 μg of random hexamers, 0.6 μg of oligo-dT (20), and 2 μL 10 mM dNTP mix (Life Technologies) in total volume of 10 μL, incubated for 10 min at 70 °C and then chilled on ice for 5 min. This mixture was added to a solution containing 4 μL 10X RT buffer, 8 μL 20 mM MgCl2, 4 μL 0.1 M DTT, 2 μL 20 U/μl SUPERase-In and 1 μL 200 U/μL SuperScript III Reverse Transcriptase. First-strand cDNA was synthesized by incubating the sample for 10 min at 25 °C, 50 min at 50 °C, and finally 5 min at 85 °C followed by a 20 min incubation with 1 μL 2 U/μL *E. coli* RNase H at 37 °C. Prepared cDNA was then subjected to quantitative reverse-transcription PCR for the detection of enriched TARE-4 and serine tRNA ligase transcripts with the following program: 5 min at 95 °C, 30 cycles of 30 s at 98 °C, 30 s at 55 °C, 30 s at 62 °C and a final extension 5 min at 62 °C. All primers used for PCR validation are listed in Supplementary data 4-1.

Libraries from the ChIRP samples were prepared using the KAPA Library Preparation Kit (KAPA Biosystems). Libraries were amplified for a total of 12 PCR cycles (12 cycles of 15 s at 98 °C, 30 s at 55 °C, 30 s at (62 °C) using the KAPA HiFi HotStart Ready Mix (KAPA Biosystems). Libraries were sequenced with a NextSeq500 DNA sequencer

(Illumina). Raw read quality was first analyzed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were mapped to the *P. falciparum* genome (v38, PlasmoDB) using Bowtie2 (v2.4.2). Duplicate, unmapped, and low quality (MAPQ < 20) reads were filtered out using Samtools (v1.9), and only uniquely mapped reads were retained. All libraries, including the input, were then normalized by dividing by the number of mapped reads in each of them. For each nucleotide, the signal from the input library was then subtracted from each of the ChIRP-seq libraries, and any negative value was replaced with a zero. Genome tracks were generated by the R package ggplot2.

Peak calling

Peaks were called using PePr v1.1. For a given lncRNA of interest, the tool was run in differential binding analysis mode using the filtered ChIRP-seq libraries for when the lncRNA was active versus non-active with the following parameters specified: -peaktype broad -threshold $1e^{-10}$. The top 25% of all reported peaks were selected because they exhibited the strongest signal (see Fig. 4-S4) and used in downstream analyses. For differential gene expression analysis, the closest gene to each peak was selected, and its expression in the two stages (active versus inactive) was obtained from(29,67,110).

LncRNA-ch14 gene disruption

Gene knockout (KO) for the long non-coding RNA 14 (lncRNA-ch14) spanning position (Ch14:3,148,960–3,150,115) on chromosome 14 was performed using a two-plasmid

design. The plasmid pDC2-Cas9-sgRNA-hdhfr(111), gifted from Marcus Lee (Wellcome Sanger Institute) contains the SpCas9, a site to express the sgRNA, and a positive selectable marker human dihydrofolate reductase (*hdhfr*). The sgRNA was selected from the database generated by(112) and cloned into pDC2-Cas9-sgRNA-hdhfr at the *BbsI* restriction site. The homology directed repair plasmid (modified pDC2-donor-*bsd* without eGFP) was designed to insert a selectable marker, blasticidin S-deaminase *(bsd)*, disrupting the lncRNA-ch14 region. The target specifying homology arm sequences were isolated through PCR amplification and gel purification. The right homology regions (RHR), and the left homology regions (LHR) of each gene were then ligated into the linearized vector via Gibson assembly. The final donor vectors were confirmed by restriction digests and Sanger sequencing. All primers are indicated in Supplementary data 4-1n Plasmids were isolated from 250 mL cultures of *Escherichia coli* (XL10-Gold Ultracompetent Cells, Agilent Cat. 200314) and 60 µg of each plasmid was used to transfect ring stage parasites. 24-hrs before transfection, mature parasite cultures (6–8% parasitemia) were magnetically separated using magnetic columns (MACS LD columns, Miltenyi Biotec) and diluted to 1% parasitemia containing 0.5 mL fresh erythrocytes(113). The next day, ~3% ring stage parasites were pelleted and washed in 4 mL of cytomix(114). 200 µL of the infected erythrocytes were resuspended with the two plasmids in cytomix to a total volume of 400 µL in a 0.2 cm cuvette. Electroporation was performed with a single pulse at 0.310 kV and 950 µF using the Biorad GenePulser electroporator. Cells were immediately transferred to a flask containing 12 mL media and 400 µL erythrocytes. The media was exchanged 5 hrs post electroporation with 12 mL of fresh media. The following day, fresh

culture media was added and supplemented with 2.5 nM WR99210 and 2.5 μg/mL blasticidin (RPI Corp, B12150-0.1). Media and drug selection were replenished every 48 h. After 14 days, the culture was split into two flasks and 50 μL of erythrocytes were added every two weeks. Once parasites were detected by microscopy, WR99210 was removed (selection for Cas9). Integration of the *bsd* gene was confirmed by gDNA extraction and PCR.

Isolation of ΔlncRNA-ch14 clone

To generate genetically homogenous parasite lines, the transfected parasites were serially diluted to approximately 0.5%, into 96 well plates. 200 μL final volume of cultured parasites were incubated with bsd drug selection for 1 month with weekly erythrocytic and media changes for the first 2 weeks of dilution followed by media changes every 2 days until parasite recovery is observed through Giemsa-stained smears.

Verification of ΔlncRNA-ch14 line

Genomic DNA (gDNA) was extracted and purified using DNeasy Blood & Tissue kit (Qiagen, 69504) following instructions from the manufacturer. The genotyping PCR analysis was used to genotype the KO lines using primer indicated in Supplementary data 4-1n. The PCR amplification was done using 2xKAPA master mix for thirty cycles with an annealing temperature of 50 °C and an extension temperature of 62 °C. The PCR amplicons were analyzed on a 1% agarose gel electrophoresis. For whole genome sequencing, genomic DNAs were fragmented using a Covaris S220 ultrasonicator and

libraries were generated using KAPA LTP Library Preparation Kit (Roche, KK8230). To verify that the insertion was present in the genome at the correct location in both transfected lines, reads were mapped using Bowtie2 (version 2.4.4) to the *P. falciparum* 3D7 reference genome (v48, PlasmoDB), edited to include the insertion sequence in the intended location. IGV (Broad Institute) was used to verify that reads aligned to the insertion sequence.

ΔlncRNA-ch14 line genome-wide sequencing and variant analysis

Libraries were sequenced using a NovaSeq 6000 DNA sequencer (Illumina), producing paired-end 100-bp reads. To verify that the insertion was present in the genome at the correct location in both transfected lines, reads were mapped using Bowtie2 (version 2.4.4) to the *P. falciparum* 3D7 reference genome (v48, PlasmoDB), edited to include the insertion sequence in the intended location. IGV (Broad Institute) was used to verify that reads aligned to the insertion sequence. To call variants (SNPs/indels) in the transfected lines compared to a previously sequenced NF54 control line, genomic DNA reads were first trimmed of adapters and aligned to the *Homo sapiens* genome (assembly GRCh38) to remove human-mapped reads.

Remaining reads were aligned to the *P. falciparum* 3D7 genome using bwa (version 0.7.17) and PCR duplicates were removed using PicardTools (Broad Institute). GATK HaplotypeCaller (https://gatk.broadinstitute.org/hc/en-us) was used to call variants between the sample and the 3D7 reference genome for both the transfected lines and the NF54 control. Only variants that were present in both transfected lines but not the NF54 control line were kept. We examined only coding-region variants and removed those that

were synonymous variants or were in *var*, *rifin*, or *stevor* genes. Quality control of variants was done by hard filtering using GATK guidelines.

Assessment of gametocyte development

Viability of gametocytes was assessed via microscopy in parasite laboratory strains NF54 and two of the ΔlncRNA-ch14 clones, F2 and B1. The morphology of parasite gametocytes was assessed in a Giemsa-stained thin blood smear. Gametocytes were classified either as viable (normal intact morphology of mature gametocytes) or dead (deformed cells with a decrease in width, a thin needle-like appearance or degraded cytoplasmic content).

Gametocyte cultures and mosquito feeding

This was performed as outlined previously [50]. Briefly, asexual stage cultures were grown in RPMI-1640 containing 2 mM L-glutamine, 50 mg/L hypoxanthine, 25 mM HEPES, 0.225% $NaHCO_3$, contained 10% *v/v* human serum in 4% human erythrocytes. Five mL of an asexual stage culture at 5% parasitemia was centrifuged at 500 × *g* for 5 min at room temperature. Gametocyte cultures were initiated at 0.5% asynchronous asexual parasitemia from low passage stock and maintained up to day 18 with daily media changes but without any addition of fresh erythrocytes. The culture medium was changed daily for 15–18 days, by carefully aspirating ~ 70-80% of the supernatant medium to avoid removing cells, and 5 mL of fresh complete culture medium was added to each well.

Giemsa-stained blood smears were made every alternate day to confirm that the parasites remained viable. Instead of a gas incubator, cultures were maintained at 37 °C in a candle

jar made of glass desiccators. On day 15 to 18, gametocyte culture, containing largely mature gametocytes, were used for mosquito feeds. Cultures were transferred to pre-warmed tubes and centrifuged at 500 x *g* for 5 min. The cells were diluted in a pre-warmed 50:50 mixture of uninfected erythrocytes and normal human serum to achieve a mature gametocytemia of 0.2% and the resulting 'feeding mixture' was placed into a pre-warmed glass feeder. Uninfected *Anopheles stephensi* mosquitoes, starved overnight of sugar water, were allowed to feed on the culture for 30 min. Unfed mosquitoes were removed, and the mosquito cups were placed in a humidified 26 °C incubator, with 10% sugar-soaked cotton pads placed on top of the mosquito cage.

Oocyst and salivary gland sporozoite quantification

On days 11 and 17 after the infective-blood meal, mosquitoes were dissected and midguts or salivary glands, respectively, were harvested for sporozoite counts. Day 11 midguts were stained with mercurochrome and photographed for oocyst counts by brightfield and phase microscopy using an upright Nikon E600 microscope with a PlanApo 4× objective. On day 17, salivary glands from ~20 mosquitoes were pooled, homogenized, and released sporozoites were counted using a haemocytometer.

Gametocyte quantification, sex determination and exflagellation assay

Between days 15 to 18, blood smears were prepared from gametocyte cultures, fixed with methanol and stained with Giemsa (Sigma, GS500), prepared as a 1:5 dilution in buffer (pH = 7.2) made using Gurr buffered tablets (VWR, 331942 F) and filtered. Slides were

stained for 20 min, washed with buffer, and allowed to dry before observation using a Nikon E600 microscope with a PlanApo 100× oil objective. For calculation of gametocytemias and male: female ratios, at least 500 mature gametocytes were scored per slide. To count exflagellation centers, 500 µL of mature gametocyte culture was centrifuged at 500 × *g* for 4 min and the resulting pellet was resuspended in equal volume of prewarmed normal human serum. Temperature was dropped to room temperature to activate gametogenesis and after 15 min incubation 10 µl of culture was transferred to a glass slide and covered with a cover slip. Exflagellation centers were counted at 40x objective in at-least ten fields for each gametocyte culture. To avoid bias, microscopic examination was performed in a blinded fashion by a trained reader.

ΔlncRNA-ch14 transcriptome analysis

Libraries were prepared from the extracted total RNA, first by isolating mRNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, E7490), then using the NEBNext Ultra Directional RNA Library Prep Kit (NEB, E7420). Libraries were amplified for a total of 12 PCR cycles (12 cycles of 15 s at 98 °C, 30 s at 55 °C, 30 s at [62 °C]) using the KAPA HiFi HotStart Ready Mix (KAPA Biosystems). Libraries were sequenced using a NovaSeq 6000 DNA sequencer (Illumina), producing paired-end 100-bp reads.

FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), was used to assess raw read quality and characteristics, and based on this information, the first 11 bp of each read and any adapter sequences were removed using Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic). Bases were trimmed from reads using Sickle with a Phred quality threshold of 20 (https://github.com/najoshi/sickle). These reads were mapped against the *Homo sapiens* genome (assembly GRCh38) using Bowtie2 (version 2.4.4) and mapped reads were removed. The remaining reads were mapped against the *Plasmodium falciparum* 3D7 genome (v48, PlasmoDB) using HISAT2 (version 2.2.1), using default parameters. Uniquely mapped, properly paired reads with mapping quality 40 or higher were retained using SAMtools (http://samtools.sourceforge.net/), and PCR duplicates were removed using PicardTools (Broad Institute). Genome browser tracks were generated and viewed using the Integrative Genomic Viewer (IGV) (Broad Institute).

Raw read counts were determined for each gene in the *P. falciparum* genome using BedTools (https://bedtools.readthedocs.io/en/latest/#) to intersect the aligned reads with the genome annotation. Differential expression analysis was done by use of R package DESeq2 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) to call up- and down-regulated genes with an adjusted *P*-value cutoff of 0.05. Gene ontology enrichment was done using PlasmoDB (https://plasmodb.org/plasmo/app). Volcano plots were generated using R package EnhancedVolcano (https://bioconductor.org/packages/release/bioc/html/EnhancedVolcano.html).

Statistical analysis

Descriptive statistics were calculated with GraphPad Prism version 9.1.2 (GraphPad Software, San Diego, CA, USA) for determining mean, percentages, standard deviation and plotting of graphs. Excel 2013 and GraphPad Prism 9.1.2 were used for the calculation of gametocytemia of microscopic data.

ACKNOWLEDGEMENTS

This work was supported by NIH grants to KGLR (nos. 1R01 AI136511 and R21 AI142506-01) and by the University of California, Riverside to KGLR (no. NIFA-Hatch-225935). We also thank the parasitology and insectary core facilities at the JHMRI and Bloomberg Philanthropies for their support of these facilities.

AUTHOR CONTRIBUTIONS

Conceptualization was the responsibility of KGLR that also supervised the project together with ML, WSN and PS. Methodology was performed by GB, XML, SA, ZC, TH, DW, TW, AC, TL, JP, AKT, GX, JC, TAT and SD; GB, XML, SA, BH and KGLR carried out the analysis. Software, formal analysis and data curation were provided by SA, BH and SD. GB, ZC, SA and KGLR wrote the original draft. All authors reviewed and edited the manuscript.

COMPETEING INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY

The GRO-seq data used in this study are available in the Gene Expression Omnibus database under accession code GSE85478. The steady-state RNA-seq data used in this study are available the NCBI Sequence Read Archive under accession code SRP026367, SRS417027, SRS417268, and SRS417269.

WGS, ChIRP-seq and RNA-seq data generated in this study (66 libraries) have been deposited in the NCBI Sequence Read Archive with accession PRJNA869073 and are available at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA869073/. All other data generated in this study are provided in the Supplementary Information, Supplementary data and Source Data file. Source data are provided with this paper.

REFERENCES

1. WHO. *World malaria report 2021*.

https://www.who.int/publications/i/item/9789240040496 (2021).

2. Le Roch KG, et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science.* 2003;301:1503–1508. doi: 10.1126/science.1087025. [PubMed] [CrossRef] [Google Scholar]

3. Bozdech Z, et al. Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray. *Genome Biol.* 2003;4:R9. doi: 10.1186/gb-2003-4-2-r9. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

4. Bozdech Z, et al. The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. *PLoS Biol.* 2003;1:E5. doi: 10.1371/journal.pbio.0000005. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

5. Young JA, et al. The Plasmodium falciparum sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol. Biochem. Parasitol.* 2005;143:67–79. doi: 10.1016/j.molbiopara.2005.05.007. [PubMed] [CrossRef] [Google Scholar]

6. Silvestrini F, et al. Genome-wide identification of genes upregulated at the onset of gametocytogenesis in Plasmodium falciparum. *Mol. Biochem. Parasitol.* 2005;143:100–110. doi: 10.1016/j.molbiopara.2005.04.015. [PubMed] [CrossRef] [Google Scholar]

7. Lemieux JE, et al. Statistical estimation of cell-cycle progression and lineage commitment in Plasmodium falciparum reveals a homogeneous pattern of transcription in ex vivo culture. *Proc. Natl. Acad. Sci. USA.* 2009;106:7559–7564. doi: 10.1073/pnas.0811829106. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

8. Lu XM, et al. Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite Plasmodium falciparum. *Nucleic Acids Res.* 2017;45:7825–7840. doi: 10.1093/nar/gkx464. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

9. Reid, A. J. et al. Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites. *Elife***7**, 10.7554/eLife.33105 (2018). [PMC free article] [PubMed]

10. Balaji S, Babu MM, Iyer LM, Aravind L. 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.* 2005;33:3994–4006. doi: 10.1093/nar/gki709. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

11. Shang X, et al. Genome-wide landscape of ApiAP2 transcription factors reveals a heterochromatin-associated regulatory network during Plasmodium falciparum blood-stage development. *Nucleic Acids Res.* 2022;50:3413–3431. doi: 10.1093/nar/gkac176. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

12. De Silva EK, et al. Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc. Natl. Acad. Sci. USA.* 2008;105:8393–8398. doi: 10.1073/pnas.0801993105. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

13. Gomez-Diaz E, et al. Epigenetic regulation of Plasmodium falciparum clonally variant gene expression during development in Anopheles gambiae. *Sci. Rep.* 2017;7:40655. doi: 10.1038/srep40655. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

14. Gupta AP, et al. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite Plasmodium falciparum. *PLoS Pathog.* 2013;9:e1003170. doi: 10.1371/journal.ppat.1003170. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

15. Bunnik EM, et al. The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum. *Genome Biol.* 2016;17:147. doi: 10.1186/s13059-016-1014-0. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

16. Lacsina JR, LaMonte G, Nicchitta CV, Chi JT. Polysome profiling of the malaria parasite Plasmodium falciparum. *Mol. Biochem. Parasitol.* 2011;179:42–46. doi: 10.1016/j.molbiopara.2011.05.003. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

17. Mair GR, et al. Regulation of sexual development of Plasmodium by translational repression. *Science.* 2006;313:667–669. doi: 10.1126/science.1125129. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

18. Shock JL, Fischer KF, DeRisi JL. Whole-genome analysis of mRNA decay in Plasmodium falciparum reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle. *Genome Biol.* 2007;8:R134. doi: 10.1186/gb-2007-8-7-r134. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

19. Freitas-Junior LH, et al. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria

parasites. *Cell.* 2005;121:25–36. doi: 10.1016/j.cell.2005.01.037. [PubMed] [CrossRef] [Google Scholar]

20. Petter M, et al. 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.* 2013;87:1167–1182. doi: 10.1111/mmi.12154. [PubMed] [CrossRef] [Google Scholar]

21. Lopez-Rubio JJ, et al. 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol. Microbiol.* 2007;66:1296–1305. doi: 10.1111/j.1365-2958.2007.06009.x. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

22. Lopez-Rubio JJ, Mancio-Silva L, Scherf A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe.* 2009;5:179–190. doi: 10.1016/j.chom.2008.12.012. [PubMed] [CrossRef] [Google Scholar]

23. Tonkin CJ, et al. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. *PLoS Biol.* 2009;7:e84. doi: 10.1371/journal.pbio.1000084. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

24. Ukaegbu UE, et al. Recruitment of PfSET2 by RNA polymerase II to variant antigen encoding loci contributes to antigenic variation in P. falciparum. *PLoS Pathog.* 2014;10:e1003854. doi: 10.1371/journal.ppat.1003854. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

25. Saxena, H. & Gupta, A. Plasmodium falciparum PfRUVBL proteins bind at the TARE region and var gene promoter located in the subtelomeric region. *Pathog Dis***80**, 10.1093/femspd/ftac018 (2022). [PubMed]

26. Lemieux JE, et al. Genome-wide profiling of chromosome interactions in Plasmodium falciparum characterizes nuclear architecture and reconfigurations associated with antigenic variation. *Mol Microbiol.* 2013;90:519–537. doi: 10.1111/mmi.12381. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

27. Ay F, et al. Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. *Genome Res.* 2014;24:974–988. doi: 10.1101/gr.169417.113. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

28. Bunnik EM, et al. Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. *Nat. Commun.* 2018;9:1910.

doi: 10.1038/s41467-018-04295-5. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

29. Read DF, Cook K, Lu YY, Le Roch KG, Noble WS. Predicting gene expression in the human malaria parasite Plasmodium falciparum using histone modification, nucleosome positioning, and 3D localization features. *PLoS Comput. Biol.* 2019;15:e1007329. doi: 10.1371/journal.pcbi.1007329. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

30. Abdi E, Latifi-Navid S. Emerging long noncoding RNA polymorphisms as novel predictors of survival in cancer. *Pathol Res. Pract.* 2022;239:154165. doi: 10.1016/j.prp.2022.154165. [PubMed] [CrossRef] [Google Scholar]

31. Rinn JL, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* 2007;129:1311–1323. doi: 10.1016/j.cell.2007.05.022. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

32. Corcoran AE. The epigenetic role of non-coding RNA transcription and nuclear organization in immunoglobulin repertoire generation. *Semin Immunol.* 2010;22:353–361. doi: 10.1016/j.smim.2010.08.001. [PubMed] [CrossRef] [Google Scholar]

33. Gupta RA, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature.* 2010;464:1071–1076. doi: 10.1038/nature08975. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

34. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* 2021;22:96–118. doi: 10.1038/s41580-020-00315-9. [PMC free article] [PubMed] [CrossRef] [Google Scholar

35. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. *RNA Biol.* 2013;10:925–933. doi: 10.4161/rna.24604. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

36. Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* 2018;19:143–157. doi: 10.1038/nrm.2017.104. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

37. Engreitz JM, Ollikainen N, Guttman M. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat. Rev. Mol. Cell Biol.* 2016;17:756–770. doi: 10.1038/nrm.2016.126. [PubMed] [CrossRef] [Google Scholar]

38. Quinodoz S, Guttman M. Long noncoding RNAs: an emerging link between gene regulation and nuclear organization. *Trends Cell Biol.* 2014;24:651–663.

doi: 10.1016/j.tcb.2014.08.009. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

39. Nakagawa S, Kageyama Y. Nuclear lncRNAs as epigenetic regulators-beyond skepticism. *Biochim. Biophys. Acta.* 2014;1839:215–222. doi: 10.1016/j.bbagrm.2013.10.009. [PubMed] [CrossRef] [Google Scholar]

40. Maclary E, Hinten M, Harris C, Kalantry S. Long nonoding RNAs in the X-inactivation center. *Chromosome Res.* 2013;21:601–614. doi: 10.1007/s10577-013-9396-2. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

41. Schoeftner S, Blasco MA. Chromatin regulation and non-coding RNAs at mammalian telomeres. *Semin Cell Dev. Biol.* 2010;21:186–193. doi: 10.1016/j.semcdb.2009.09.015. [PubMed] [CrossRef] [Google Scholar]

42. Feuerhahn S, Iglesias N, Panza A, Porro A, Lingner J. TERRA biogenesis, turnover and implications for function. *FEBS Lett.* 2010;584:3812–3818. doi: 10.1016/j.febslet.2010.07.032. [PubMed] [CrossRef] [Google Scholar]

43. Patankar S, Munasinghe A, Shoaibi A, Cummings LM, Wirth DF. Serial analysis of gene expression in Plasmodium falciparum reveals the global expression profile of erythrocytic stages and the presence of anti-sense transcripts in the malarial parasite. *Mol. Biol. Cell.* 2001;12:3114–3125. doi: 10.1091/mbc.12.10.3114. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

44. Gunasekera AM, et al. Widespread distribution of antisense transcripts in the Plasmodium falciparum genome. *Mol. Biochem. Parasitol.* 2004;136:35–42. doi: 10.1016/j.molbiopara.2004.02.007. [PubMed] [CrossRef] [Google Scholar]

45. Raabe CA, et al. A global view of the nonprotein-coding transcriptome in Plasmodium falciparum. *Nucleic Acids Res.* 2010;38:608–617. doi: 10.1093/nar/gkp895. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

46. Sorber K, Dimon MT, DeRisi JL. RNA-Seq analysis of splicing in Plasmodium falciparum uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. *Nucleic Acids Res.* 2011;39:3820–3835. doi: 10.1093/nar/gkq1223. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

47. Broadbent KM, et al. A global transcriptional analysis of Plasmodium falciparum malaria reveals a novel family of telomere-associated lncRNAs. *Genome Biol.* 2011;12:R56. doi: 10.1186/gb-2011-12-6-r56. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

48. Wei C, et al. Deep profiling of the novel intermediate-size noncoding RNAs in intraerythrocytic Plasmodium falciparum. *PLoS One.* 2014;9:e92946.

doi: 10.1371/journal.pone.0092946. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

49. Liao Q, et al. Genome-wide identification and functional annotation of Plasmodium falciparum long noncoding RNAs from RNA-seq data. *Parasitol Res.* 2014;113:1269–1281. doi: 10.1007/s00436-014-3765-4. [PubMed] [CrossRef] [Google Scholar]

50. Siegel TN, et al. Strand-specific RNA-Seq reveals widespread and developmentally regulated transcription of natural antisense transcripts in Plasmodium falciparum. *BMC Genomics.* 2014;15:150. doi: 10.1186/1471-2164-15-150. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

51. Broadbent KM, et al. Strand-specific RNA sequencing in Plasmodium falciparum malaria identifies developmentally regulated long non-coding RNA and circular RNA. *BMC Genomics.* 2015;16:454. doi: 10.1186/s12864-015-1603-4. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

52. Lee, V. V. et al. Direct nanopore sequencing of mRNA reveals landscape of transcript isoforms in apicomplexan parasites. *mSystems***6**, 10.1128/mSystems.01081-20 (2021). [PMC free article] [PubMed]

53. Chappell L, et al. Refining the transcriptome of the human malaria parasite Plasmodium falciparum using amplification-free RNA-seq. *BMC Genomics.* 2020;21:395. doi: 10.1186/s12864-020-06787-5. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

54. Mourier T, et al. Genome-wide discovery and verification of novel structured RNAs in Plasmodium falciparum. *Genome Res.* 2008;18:281–292. doi: 10.1101/gr.6836108. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

55. Yang M, et al. Full-Length Transcriptome Analysis of Plasmodium falciparum by Single-Molecule Long-Read Sequencing. *Front Cell Infect Microbiol.* 2021;11:631545. doi: 10.3389/fcimb.2021.631545. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

56. Hoshizaki J, et al. Correction: A manually curated annotation characterises genomic features of P. falciparum lncRNAs. *BMC Genomics.* 2023;24:189. doi: 10.1186/s12864-023-09164-0. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

57. Epp C, Li F, Howitt CA, Chookajorn T, Deitsch KW. Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite Plasmodium falciparum. *RNA.* 2009;15:116–127. doi: 10.1261/rna.1080109. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

58. Amit-Avraham I, et al. Antisense long noncoding RNAs regulate var gene activation in the malaria parasite Plasmodium falciparum. *Proc. Natl. Acad. Sci. USA.* 2015;112:E982–E991. doi: 10.1073/pnas.1420855112. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

59. Sierra-Miranda M, et al. Two long non-coding RNAs generated from subtelomeric regions accumulate in a novel perinuclear compartment in Plasmodium falciparum. *Mol. Biochem. Parasitol.* 2012;185:36–47. doi: 10.1016/j.molbiopara.2012.06.005. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

60. Simantov K, Goyal M, Dzikowski R. Emerging biology of noncoding RNAs in malaria parasites. *PLoS Pathog.* 2022;18:e1010600. doi: 10.1371/journal.ppat.1010600. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

61. Rovira-Graells N, et al. Deciphering the principles that govern mutually exclusive expression of Plasmodium falciparum clag3 genes. *Nucleic Acids Res.* 2015;43:8243–8257. doi: 10.1093/nar/gkv730. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

62. Jing Q, et al. Plasmodium falciparum var Gene Is Activated by Its Antisense Long Noncoding RNA. *Front Microbiol.* 2018;9:3117. doi: 10.3389/fmicb.2018.03117. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

63. Barcons-Simon, A., Cordon-Obras, C., Guizetti, J., Bryant, J. M. & Scherf, A. CRISPR Interference of a Clonally Variant GC-Rich Noncoding RNA Family Leads to General Repression of var Genes in Plasmodium falciparum. *mBio***11**, 10.1128/mBio.03054-19 (2020). [PMC free article] [PubMed]

64. Luke B, Lingner J. TERRA: telomeric repeat-containing RNA. *EMBO J.* 2009;28:2503–2510. doi: 10.1038/emboj.2009.166. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

65. Filarsky M, et al. GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science.* 2018;359:1259–1263. doi: 10.1126/science.aan6042. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

66. Gomes AR, et al. A transcriptional switch controls sex determination in Plasmodium falciparum. *Nature.* 2022;612:528–533. doi: 10.1038/s41586-022-05509-z. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

67. Bunnik EM, et al. DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite Plasmodium falciparum. *BMC Genomics.* 2014;15:347. doi: 10.1186/1471-2164-15-347. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

68. Zhang, M. et al. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science***360**, 10.1126/science.aap7847 (2018). [PMC free article] [PubMed]

69. Sun Q, Hao Q, Prasanth KV. Nuclear long noncoding RNAs: Key regulators of gene expression. *Trends Genet.* 2018;34:142–157. doi: 10.1016/j.tig.2017.11.005. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

70. Noh JH, Kim KM, McClusky WG, Abdelmohsen K, Gorospe M. Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip Rev. RNA.* 2018;9:e1471. doi: 10.1002/wrna.1471. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

71. Sun M, Gadad SS, Kim DS, Kraus WL. Discovery, annotation, and functional analysis of long noncoding RNAs controlling cell-cycle gene expression and proliferation in breast cancer cells. *Mol. Cell.* 2015;59:698–711. doi: 10.1016/j.molcel.2015.06.023. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

72. Clark MB, et al. Genome-wide analysis of long noncoding RNA stability. *Genome Res.* 2012;22:885–898. doi: 10.1101/gr.131037.111. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

73. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell.* 2011;44:667–678. doi: 10.1016/j.molcel.2011.08.027. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

74. Quinn JJ, et al. Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat. Biotechnol.* 2014;32:933–940. doi: 10.1038/nbt.2943. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

75. Chu, C., Quinn, J. & Chang, H. Y. Chromatin isolation by RNA purification (ChIRP). *J. Vis. Exp.*, 10.3791/3912 (2012). [PMC free article] [PubMed]

76. Fan, Y. *et al*. Rrp6 Regulates Heterochromatic Gene Silencing via ncRNA RUF6 Decay in Malaria Parasites. *mBio***11**, 10.1128/mBio.01110-20 (2020). [PMC free article] [PubMed]

77. Bunnik EM, et al. Comparative 3D genome organization in apicomplexan parasites. *Proc. Natl. Acad. Sci. USA.* 2019;116:3183–3192. doi: 10.1073/pnas.1810815116. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

78. . Winter G, et al. SURFIN is a polymorphic antigen expressed on Plasmodium falciparum merozoites and infected erythrocytes. *J. Exp. Med.* 2005;201:1853–1863. doi: 10.1084/jem.20041392. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

79. Pei X, et al. The ring-infected erythrocyte surface antigen (RESA) of Plasmodium falciparum stabilizes spectrin tetramers and suppresses further invasion. *Blood.* 2007;110:1036–1042. doi: 10.1182/blood-2007-02-076919. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

80. Zhang Y, Lin YH, Johnson TD, Rozek LS, Sartor MA. PePr: a peak-calling prioritization pipeline to identify consistent or differential peaks from replicated ChIP-Seq data. *Bioinformatics.* 2014;30:2568–2575. doi: 10.1093/bioinformatics/btu372. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

81. Silvestrini F, et al. Protein export marks the early phase of gametocytogenesis of the human malaria parasite Plasmodium falciparum. *Mol. Cell Proteomics.* 2010;9:1437–1448. doi: 10.1074/mcp.M900479-MCP200. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

82. Gardiner DL, et al. Implication of a Plasmodium falciparum gene in the switch between asexual reproduction and gametocytogenesis. *Mol. Biochem. Parasitol.* 2005;140:153–160. doi: 10.1016/j.molbiopara.2004.12.010. [PubMed] [CrossRef] [Google Scholar]

83. Lasonder E, et al. Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression. *Nucleic Acids Res.* 2016;44:6087–6101. doi: 10.1093/nar/gkw536. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

84. Tiburcio, M. et al. A 39-Amino-Acid C-Terminal Truncation of GDV1 Disrupts Sexual Commitment in Plasmodium falciparum. *mSphere***6**, 10.1128/mSphere.01093-20 (2021). [PMC free article] [PubMed

85. Usui M, et al. Plasmodium falciparum sexual differentiation in malaria patients is associated with host factors and GDV1-dependent genes. *Nat. Commun.* 2019;10:2140. doi: 10.1038/s41467-019-10172-6. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

86. Tripathi, A. K., Mlambo, G., Kanatani, S., Sinnis, P. & Dimopoulos, G. Plasmodium falciparum Gametocyte Culture and Mosquito Infection Through Artificial Membrane Feeding. *J. Vis. Exp*., 10.3791/61426 (2020). [PMC free article] [PubMed]

87. Guil S, Esteller M. Cis-acting noncoding RNAs: friends and foes. *Nat. Struct. Mol. Biol.* 2012;19:1068–1075. doi: 10.1038/nsmb.2428. [PubMed] [CrossRef] [Google Scholar]

88. Orom UA, et al. Long noncoding RNAs with enhancer-like function in human cells. *Cell.* 2010;143:46–58. doi: 10.1016/j.cell.2010.09.001. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

89. Ng SY, Bogu GK, Soh BS, Stanton LW. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. *Mol. Cell.* 2013;51:349–359. doi: 10.1016/j.molcel.2013.07.017. [PubMed] [CrossRef] [Google Scholar]

90. Prensner JR, et al. The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. *Nat. Genet.* 2013;45:1392–1398. doi: 10.1038/ng.2771. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

91. Tsai MC, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science.* 2010;329:689–693. doi: 10.1126/science.1192002. [PMC free article] [PubMed] [CrossRef] [Google Scholar

92. Mele M, Rinn JL. "Cat's Cradling" the 3D Genome by the Act of LncRNA Transcription. *Mol. Cell.* 2016;62:657–664. doi: 10.1016/j.molcel.2016.05.011. [PubMed] [CrossRef] [Google Scholar]

93. Rinn J, Guttman M. RNA Function. RNA and dynamic nuclear organization. *Science.* 2014;345:1240–1241. doi: 10.1126/science.1252966. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

94. Guizetti J, Barcons-Simon A, Scherf A. Trans-acting GC-rich non-coding RNA at var expression site modulates gene counting in malaria parasite. *Nucleic Acids Res.* 2016;44:9710–9718. [PMC free article] [PubMed] [Google Scholar]

95. Abel S, Le Roch KG. The role of epigenetics and chromatin structure in transcriptional regulation in malaria parasites. *Brief Funct. Genomics.* 2019;18:302–313. doi: 10.1093/bfgp/elz005. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

96. Cerase A, Pintacuda G, Tattermusch A, Avner P. Xist localization and function: new insights from multiple levels. *Genome Biol.* 2015;16:166. doi: 10.1186/s13059-015-0733-y. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

97. Hanly, D. J., Esteller, M. & Berdasco, M. Interplay between long non-coding RNAs and epigenetic machinery: emerging targets in cancer? *Philos. Trans. R Soc. Lond. B Biol. Sci*. **373**, 10.1098/rstb.2017.0074 (2018). [PMC free article] [PubMed]

98. Vance KW, Ponting CP. Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.* 2014;30:348–355. doi: 10.1016/j.tig.2014.06.001. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

99. Wierzbicki AT. The role of long non-coding RNA in transcriptional gene silencing. *Curr. Opin. Plant Biol.* 2012;15:517–522. doi: 10.1016/j.pbi.2012.08.008. [PubMed] [CrossRef] [Google Scholar]

100. Reininger L, et al. A NIMA-related protein kinase is essential for completion of the sexual cycle of malaria parasites. *J. Biol. Chem.* 2005;280:31957–31964. doi: 10.1074/jbc.M504523200. [PubMed] [CrossRef] [Google Scholar]

101. Reininger L, et al. An essential role for the Plasmodium Nek-2 Nima-related protein kinase in the sexual development of malaria parasites. *J. Biol. Chem.* 2009;284:20858–20868. doi: 10.1074/jbc.M109.017988. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

102. Tomas AM, et al. P25 and P28 proteins of the malaria ookinete surface have multiple and partially redundant functions. *EMBO J.* 2001;20:3975–3983. doi: 10.1093/emboj/20.15.3975. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

103. Trager W, Jensen JB. Human malaria parasites in continuous culture. *Science.* 1976;193:673–675. doi: 10.1126/science.781840. [PubMed] [CrossRef] [Google Scholar]

104. Joshi NA, F. J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. *[Software]*. **(Version 1.33)** (2011).

105. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37:907–915. doi: 10.1038/s41587-019-0201-4. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

106. Li H, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25:2078–2079. doi: 10.1093/bioinformatics/btp352. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

107. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26:841–842. doi: 10.1093/bioinformatics/btq033. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

108. Trapnell C, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 2012;7:562–578. doi: 10.1038/nprot.2012.016. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

109. Howick, V. M. et al. The Malaria Cell Atlas: Single parasite transcriptomes across the complete Plasmodium life cycle. *Science***365**, 10.1126/science.aaw2619 (2019). [PMC free article] [PubMed]

110. Jiang L, et al. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. *Nature.* 2013;499:223–227. doi: 10.1038/nature12361. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

111. Sonoiki E, et al. A potent antimalarial benzoxaborole targets a Plasmodium falciparum cleavage and polyadenylation specificity factor homologue. *Nat. Commun.* 2017;8:14574. doi: 10.1038/ncomms14574. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

112. Ribeiro JM, et al. Guide RNA selection for CRISPR-Cas9 transfections in Plasmodium falciparum. *Int. J. Parasitol.* 2018;48:825–832. doi: 10.1016/j.ijpara.2018.03.009. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

113. Ribaut C, et al. Concentration and purification by magnetic separation of the erythrocytic stages of all human Plasmodium species. *Malar J.* 2008;7:45. doi: 10.1186/1475-2875-7-45. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

114. Adjalley SH, Lee MC, Fidock DA. A method for rapid genetic integration into Plasmodium falciparum utilizing mycobacteriophage Bxb1 integrase. *Methods Mol. Biol.* 2010;634:87–100. doi: 10.1007/978-1-60761-652-8_6. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

PEER REVIEWED PUBLICATIONS

(Most recent listed first)

1. **Chahine Z**, Abel S, Hollin T, Chung JH, Barnes GL, Daub ME, Renard I, Choi JY, Vydyam P, Pal A, Kirkwood J, Saraf A, Banks CAS, Camino I, Castaneda P, Cuevas MC, De Mercado Arnanz J, Fernandez E, Garcia A, Ibarz N, Viera-Morilla S, Prudhomme J, Bei AK, Florens L, Ben Mamoun C, Vanderwal C and Le Roch KG. *A Kalihinol Analogue Disrupts Apicoplast Function and Vesicular Trafficking in P. falciparum. Science, 2024.* Manuscript Accepted.

2. **Chahine Z***, Gupta M*, Lenz T*, Hollin T*, Abel S, Banks C, Saraf A, Prudhomme J, Florens L, Roch KL*. *Pf*MORC protein regulates chromatin accessibility and transcriptional repression in the human malaria parasite, *P. falciparum*. eLife. Accepted. 2023. ***** Co-first Authors.

3. Batugedara G, Lu XM, Hristov B, Abel S, **Chahine Z**, Hollin T, Williams D, Wang T, Cort A, Lenz T, Thompson TA, Prudhomme J, Tripathi AK, Xu G, Cudini J, Dogga S, Lawniczak M, Noble WS, Sinnis P, Le Roch KG. Novel insights into the role of long non-coding RNA in the human malaria parasite, *Plasmodium falciparum*. Nat Commun. 2023;14(1):5086. Epub 20230822. PubMed PMID: 37607941.

4. Dwulet NC, **Chahine Z**, Le Roch KG, Vanderwal CD. An Enantiospecific Synthesis of Isoneoamphilectane Confirms Its Strained Tricyclic Structure. J Am Chem Soc. 2023;145(6):3716-26. Epub 20230202. PubMed PMID: 36730688.

5. Romano PS, Akematsu T, Besteiro S, Bindschedler A, Carruthers VB, **Chahine Z**, Coppens I, Descoteaux A, Alberto Duque TL, He CY, Heussler V, Le Roch KG, Li FJ, de Menezes JPB, Menna-Barreto RFS, Mottram JC, Schmuckli-Maurer J, Turk B, Tavares Veras PS, Salassa BN, Vanrell MC. Autophagy in protists and their hosts: When, how and why? Autophagy Rep. 2023;2(1). Epub 20230309. PubMed PMID: 37064813. Review.

6. Thompson TA, **Chahine Z**, Le Roch KG. The role of long noncoding RNAs in malaria parasites. Trends Parasitol. 2023;39(7):517-31. Epub 20230429. PubMed PMID: 37121862. Review.

7. Hollin T, **Chahine Z**, Le Roch KG. Epigenetic Regulation and Chromatin Remodeling in Malaria Parasites. Annu Rev Microbiol. 2023. Epub 20230602. PubMed PMID: 37268002. Review.
8. Zeeshan M, Rashpa R, Ferguson DJP, Abel S, **Chahine Z**, Brady D, Vaughan S, Moores CA, Le Roch KG, Brochet M, Holder AA, Tewari R. Genome-wide functional analysis reveals key roles for kinesins in the mammalian and mosquito stages of the malaria parasite life cycle. PLoS Biol. 2022;20(7):e3001704. Epub 20220728. PubMed PMID: 35900985.

9. **Chahine Z**, Le Roch KG. Decrypting the complexity of the human malaria parasite biology through systems biology approaches. Front Syst Biol. 2022;2. Epub 20220916. PubMed PMID: 37200864. Review.

10. Clifton BD, Jimenez J, Kimura A, **Chahine Z**, Librado P, Sanchez-Gracia A, Abbassi M, Carranza F, Chan C, Marchetti M, Zhang W, Shi M, Vu C, Yeh S, Fanti L, Xia XQ, Rozas J, Ranz JM. Understanding the Early Evolutionary Stages of a Tandem *Drosophila melanogaster*-Specific Gene Family: A Structural and Functional Population Study. Mol Biol Evol. 2020;37(9):2584-600. PubMed PMID: 32359138.

11. Karns AS, Ellis BD, Roosen PC, **Chahine Z**, Le Roch KG, Vanderwal CD. Concise Synthesis of the Antiplasmodial Isocyanoterpene 7,20-Diisocyanoadociane. Angew Chem Int Ed Engl. 2019;58(39):13749-52. Epub 20190827. PubMed PMID: 3127092.

CONCLUDING STATEMENTS

Malaria remains one of the most complex and challenging blood-borne diseases with a profound global impact, having affected millions of lives throughout the years. In this work, I bring attention to the *Plasmodium* parasite's unique features and sophisticated mechanisms it has adapted to evade host immune responses and develop resistance to therapeutic interventions. I illustrated the range of complex epigenetic mechanisms needed to regulate all aspects of cellular activity in *Plasmodium*. Elements such as transcription factors, heterochromatin associated proteins and non-coding RNAs, which work systematically and complementarily within nuclear organization dynamics to coordinate gene regulation during parasite development and transmission. The introduction of more precise sequencing platforms and more sensitive omics-based technologies, along with advanced genetic editing tools, has deepened our understanding of the pathogen. This improved understanding has offered new perspectives on the *Plasmodium* parasite, paving the way for alternative drug therapies and enabling a more accurate comprehension of the effects these potential antimalarials have, both on the parasite, and human biology. By further studying the molecular mechanics involved in drug mechanisms of action and drug resistance, we can more appropriately and responsibly administer drug therapies while reducing the potential for resistance outbreaks. My research journey has underscored the intricacies of these mechanisms, highlighting the need for innovative and comprehensive approaches to combat the disease. Through my work in the lab, I have gained essential skills and deepened my understanding of malaria’s biology, emphasizing the benefits of collaborative work, the importance of systems biology in advancing our knowledge and

developing effective treatments. I am profoundly grateful to all who have supported me throughout this endeavor—your guidance, encouragement, and the breadth of knowledge provided have equipped me with the confidence and expertise needed to excel as a scientist. As I move forward, I carry with me the lessons learned and the collaborative spirit that will drive my future contributions to the field.