Title

Decoding the Epigenetic Landscape of Apicomplexan Parasites: Mechanisms Mediating Virulence and Host-Pathogen Interactions

Permalink

https://escholarship.org/uc/item/1kq2211v

Author

Lenz, Todd Jeffery

Publication Date

2024

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Decoding the Epigenetic Landscape of Apicomplexan Parasites: Mechanisms Mediating Virulence and Host-Pathogen Interactions

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

Doctor of Philosophy

in

Genetics, Genomics and Bioinformatics

by

Todd J. Lenz

December 2024

Dissertation Committee:
- Dr. Karine Le Roch, Chairperson
- Dr. Stefano Lonardi
- Dr. Frances Sladek

The Dissertation of Todd J. Lenz is approved by:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without support, guidance, and encouragement from many individuals to whom I am deeply indebted.

First and foremost, I want to express my sincere gratitude to my advisor, Dr. Karine Le Roch, Her mentorship has been instrumental in shaping both this work and my development as a scientist. Her patience and belief in my potential as a researcher have enabled me to persevere and I am profoundly grateful for the countless hours she has invested in my growth.

Time spent with my lab family has been invaluable, offering fresh perspectives and welcome respite from the rigors of research. I am especially grateful to Dr. Zeinab Chahine, an accomplished scientist and dear friend whose generous support, both inside and outside the lab, made much of my work possible.

I also extend my appreciation to my collaborators—Dr. Ulf Ribacke, Dr. Kirk Deitsch, Dr. Emma Wilson, Dr. Stefano Lonardi, and numerous others—for generously sharing their expertise and providing critical research opportunities.

To my remarkable support system of family and friends who have aided my journey: I am profoundly grateful to my wife Amanda and our precious boys—Leo, Baltimore, and Shadow—who bring joy and purpose to each day. My grandfather's faith in my abilities gave me confidence to pursue my dreams, while my brother Trevor kept me grounded and laughing through the most challenging times.

To everyone mentioned here and to those whose support has touched my life in countless other ways, I am eternally grateful.

DEDICATION

To my wife, Amanda, you are the heart and soul of everything that I have accomplished. Your unconditional love and support have been my source of strength throughout this journey. You believed in me when I was at my lowest. This dissertation represents endless hours of research and writing, but it also represents your sacrifice and patience awaiting me to achieve my goals. This dissertation exists because of you.

ABSTRACT OF THE DISSERTATION

Decoding the Epigenetic Landscape of Apicomplexan Parasites: Mechanisms Mediating Virulence and Host-Pathogen Interactions

by

Todd J. Lenz

Doctor of Philosophy, Graduate Program in Genetics, Genomics and Bioinformatics
University of California, Riverside, December 2024
Dr. Karine G. Le Roch, Chairperson

Apicomplexan parasites like *Plasmodium falciparum* and *Toxoplasma gondii* are responsible for over a half million deaths each year globally. They employ sophisticated epigenetic mechanisms to regulate gene expression and evade host immune responses. This dissertation investigates the epigenetic and chromatin structural features that enable these parasites to coordinate a complex profile of virulence factors and survive within their hosts. Using complementary molecular and genomics approaches including RNA-seq, ChIP-seq, Hi-C, and proteomics, I characterized several key regulatory factors and mechanisms. In *P. falciparum*, I demonstrated that depletion of repressive H3K9me3 marks and small, localized changes in chromatin structure are required for expression of *var2csa*, a critical virulence gene involved in establishing the placental binding phenotype and pregnancy-associated malaria. I also showed that deletion of *var2csa* disrupts coordinated *var* gene switching by increasing heterochromatin compaction.

Additionally, I identified *Pf*MORC as an essential chromatin-associated factor that maintains heterochromatin integrity and transcriptional repression during blood-stage development. In *T. gondii*, I investigated ROP55, a novel rhoptry protein that promotes parasite survival by preventing lytic cell death. I also revealed how infection reorganizes host cell chromatin architecture through weakening of topologically associating domains and disruption of A/B compartments, providing new insights into host-pathogen interactions. Together, this work advances our understanding of the complex interplay between nuclear organization, chromatin structure, and gene regulation in apicomplexan parasites. The findings highlight the importance of epigenetic mechanisms in parasite virulence and survival, with broad implications for developing novel intervention strategies against these major human pathogens.

TABLE OF CONTENTS

CHAPTER 4
The *Toxoplasma* rhoptry protein ROP55 is a major virulence factor that prevents lytic host cell death

CHAPTER 5
Toxoplasma gondii reorganizes the host chromatin structure and epigenetic features

FIGURES

INTRODUCTION

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

INTRODUCTION

The fascinating phylum of Apicomplexa

The phylum Apicomplexa comprises a vast group of single-celled, obligate intracellular protozoan parasites that exhibit a highly specialized parasitic lifestyle. Apicomplexans are among the most ecologically and medically significant eukaryotic pathogens due to their effect on human health and agriculture, infecting millions of humans and livestock animals each year resulting in widespread socioeconomic impact, morbidity and mortality. Notable genera include *Plasmodium*, the causative agent of malaria—one of the most devastating infectious diseases globally[1]; *Toxoplasma*, which infects virtually all warm-blooded animals and is the causative agent of toxoplasmosis[2,3]; Babesia, a tick-borne parasite that causes babesiosis—a malaria-like illness that shares many of the same symptoms[4]; and Eimeria, a major infectious agent of livestock that causes significant agricultural losses[5].

Apicomplexans are distinct from other pathogens, such as bacteria, viruses, and fungi, due to their complex biology as intracellular eukaryotes that inhabit the cells of other eukaryotic organisms. Though the phylum is morphologically diverse, all Apicomplexans contain a collection of specialized structures and secretory organelles that facilitate host cell invasion called the apical complex[6]. This structure contains secretory organelles—rhoptries, micronemes, and dense granules—that release proteins enabling parasite entry into host cells. This invasive strategy sets them apart from bacteria, which often use extracellular mechanisms, and viruses, which rely on host machinery for replication. Another distinguishing feature found in most Apicomplexans is a nonphotosynthetic plastid known as the apicoplast, which is critical for amino acid metabolism, fatty acid biosynthesis, and isoprenoid precursor synthesis[7,8]. Its evolutionary origin and essential functions make the apicoplast a promising drug target, as it is absent in host cells[9].

Apicomplexans exhibit a highly complex life cycle, often involving multiple hosts and intricate developmental stages **(Fig. I.1)**. This complexity contrasts sharply with the simpler life cycles of viruses and many bacteria, making Apicomplexans more challenging to study and control. Their heteroxenous life cycle and strict host cell dependency, requiring very specific growth conditions, prevents or severely impedes in vitro studies of many Apicomplexan species and raises ethical and logistical concerns for in vivo research[10–16]. Furthermore, because they are eukaryotic pathogens that share some similar molecular processes as their host cells, designing therapeutic interventions to target the parasites that also avoid damaging host cells is particularly challenging.

The scientific interest in Apicomplexans extends beyond their ability to cause disease as their eukaryotic interactions provides a unique perspective on host-parasite coevolution, intracellular survival strategies, and immune evasion. Their ability to manipulate host cell function to ensure their survival, such as altering host gene expression and immune signaling, makes them compelling subjects for molecular and cellular biology research. Moreover, understanding their biology provides critical insights into developing new therapeutic strategies, particularly in light of rising drug resistance in diseases like malaria[17,18].

Apicomplexans are a fascinating group of organisms that differentiate from other pathogens due to their unique intracellular parasitism, complex life cycles, and specialized organelles. They are not only central to understanding parasitic diseases but also offer a window into fundamental biological processes, making them a priority for both basic and applied research.

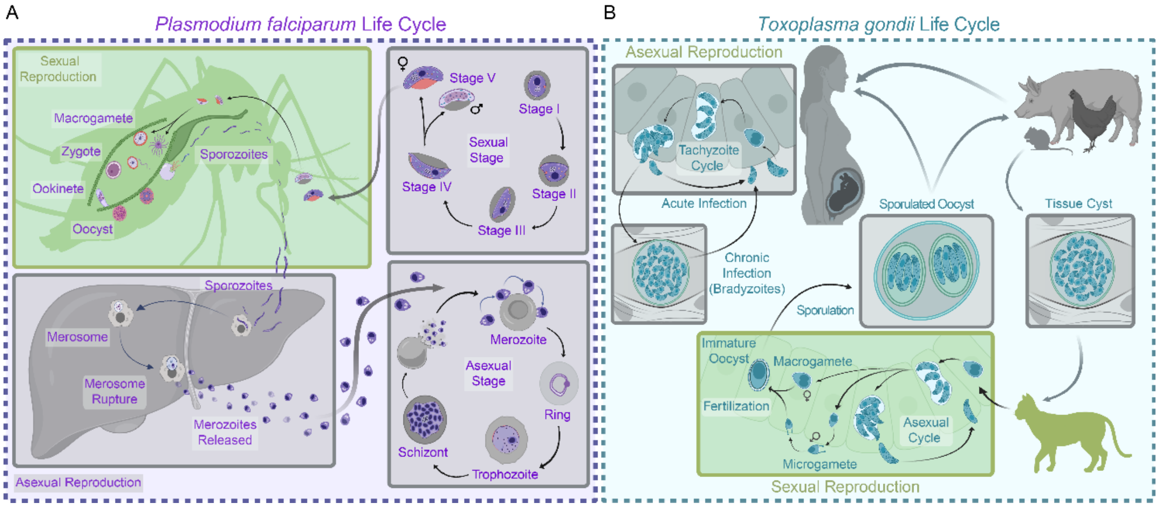

Fig. I.1. The *Plasmodium falciparum* and *Toxoplasma gondii* life cycles. (A) The life cycle of *P. falciparum* includes a multi-stage asexual reproduction cycle (purple) that occurs within the liver and circulating erythrocytes. Under environmental pressure the parasite transitions to the sexual stages and complete their sexual reproduction cycle inside the mosquito midgut (green). **(B)** The *T. gondii* definitive host (felids) sheds oocysts in their feces, which upon sporulation are ingested by warm-blooded animals to begin the asexual reproduction cycle (gray). Tissue cysts inside an intermediate host rupture when ingested by a new definitive host and thus the parasite can complete the sexual reproduction cycle (green).

The Plasmodium falciparum life cycle

The *Plasmodium falciparum* asexual life cycle **(Fig. I.1A)** begins when sporozoites are introduced into a human host through the bite of an infected female *Anopheles* mosquito. These sporozoites quickly enter the bloodstream and within one hour of inoculation they reach the liver where they invade hepatocytes[19,20]. The parasite undergoes a growth phase of differentiation and replication, generating thousands of merozoites, which are subsequently released into circulation where they can invade red blood cells (RBCs) and travel throughout the body. Notably, the liver stage varies among *Plasmodium* species; for instance, *P. vivax* and *P. ovale* can form dormant hypnozoites that remain in the liver for extended periods before activation[19,21,22].

Once the merozoites have parasitized circulating erythrocytes they commence the intraerythrocytic developmental cycle (IDC), a cycle encompassing several developmental stages that include the dormant ring stage, the metabolically active trophozoite stage, and the mature schizont stage where the parasite divides several times to produce new merozoites in a process referred to as schizogony. Ring stage infected erythrocytes are the predominant form detected in peripheral circulation and ring-infected cells are particularly difficult to detect and specifically target due to their unaltered morphology and low metabolic activity[23,24]. Although parasites at this stage are relatively quiescent, expression and trafficking of variant surface antigens to mediate cytoadherence provides the foundation for further parasite development[25,26]. This is followed by the metabolically active trophozoite stage, marked by rapid uptake of hemoglobin and expansion of the normally condensed heterochromatin that prevents widespread expression of genes involved in nutrient acquisition, hemoglobin catabolism and energy production[27–29]. Due to the export of various structural and adhesive proteins and growth of the parasite, the parasitized RBCs (pRBCs) begin to swell and present the more obvious hallmarks of infected cells[23]. Once the parasite transitions to the schizont stage, transcription decreases, and DNA is repackaged into tight heterochromatin clusters ready for DNA replication and division into daughter merozoites. Depending on the species of *Plasmodium*, the asexual cycle takes up to 72 hours and produces as many as 64 daughter merozoites. The production of new merozoites eventually causes the host RBC to lyse, releasing merozoites back into circulation to infect new cells. The characteristic flu-like symptoms of acute malaria are due to the presentation of the variant surface antigens that cause pRBCs to sequester from circulation, resulting in damage to the microvasculature and surrounding tissue[30–32]. Although cytoadhesion is harmful to the human host, removal from circulation enables immune evasion by shielding surface antigens from host phagocytes and ultimately prevents splenic clearance[33,34].

Environmental cues, including nutrient deprivation or stress signals, may prompt a transition from asexual replication to sexual development, though the exact triggers for gametocytogenesis remain elusive[35–38]. Sexual commitment is marked by expression of an epigenetically silenced gene that produces the DNA-binding protein PfAP2-G, which serves as a transcription factor (TF) for early gametocyte genes[39,40]. During sexual development, *P. falciparum* parasites differentiate into male and female gametocytes, a process that includes five distinct morphological stages, taking approximately 2–3 weeks for gametocytes to fully mature. Mature gametocytes circulate within the bloodstream of the human host, awaiting ingestion by an *Anopheles* mosquito during a blood meal[35,41]. Female gametocytes undergo cytoplasmic reorganization to form immobile macrogametes, while male gametocytes transform into microgametes. Microgametes actively seek out macrogametes, leading to fertilization in the mosquito midgut and the formation of a diploid zygote[35]. The zygote develops into a motile ookinete, which migrates through and penetrates the mosquito gut wall. Encystation follows, with the ookinete transforming into an oocyst within the gut epithelium. Inside the oocyst, sporozoites develop and eventually rupture out, migrating to the mosquito salivary glands, ready to infect a new human host. This intricate life cycle is tightly regulated by networks of proteins at the genomic, transcriptional, and translational levels, enabling parasite survival and propagation across hosts.

The Toxoplasma gondii life cycle

The life cycle of *Toxoplasma gondii* **(Fig. I.1B)** is similarly complex as it also involves both sexual and asexual stages, alternating between definitive and intermediate hosts. Felids, including domestic cats and other members of the cat family, serve as the definitive hosts for *Toxoplasma gondii* due to the unique biochemical and cellular environment of their intestinal epithelium, which supports the parasite's sexual reproduction[2]. The intermediate host range of *Toxoplasma gondii* is extraordinarily broad, encompassing virtually all warm-blooded animals, including mammals and birds. This wide host range is facilitated by the versatile asexual stages—tachyzoites and bradyzoites—which are capable of infecting nearly any nucleated cell[42].

The transmission of *T. gondii* to an intermediate host is most often facilitated through environmental contamination when a grazing animal ingests food or water contaminated by sporulated oocysts. Humans contract the parasite through consumption of unwashed produce, or inadvertently through ingestion of oocysts by consuming undercooked meat from infected animals. Once an oocyst is ingested, sporozoites are released in the digestive tract and quickly penetrate the intestinal epithelium. The initial invasion occurs when sporozoites use a gliding motility mechanism to secrete specific proteins that allow them to breach the host cell membrane[43,44]. This process involves the invagination of the host cell plasma membrane, forming the parasitophorous vacuole (PV), a protective microenvironment that allows for nutrient exchange and the export of parasite proteins while also preventing fusion with host cell lysosomes, ultimately protecting the parasite from the immune defenses of the host[43]. Once inside, sporozoites differentiate into rapidly replicating tachyzoites which then disseminate throughout the host, predominantly infecting neural tissue and muscle cells. Tachyzoites replicate through endodyogeny, a unique form of asexual reproduction wherein two daughter cells are formed within the plasma membrane of the parent[45]. This process allows for exponential population growth, with a single tachyzoite capable of producing multiple generations of parasites within a single host cell before the cell lyses. Under immune pressure or nutrient-limited conditions, tachyzoites can transition to bradyzoites, a hearty and slow-growing form that persists within tissue cysts. This chronic form of the parasite enables long-term persistence within the host while remaining metabolically quiescent and provides exceptional protection against the host immune system. While largely innocuous in a healthy host, the parasite can reactivate in the absence of environmental pressure and resume replication and further propagation in the tachyzoite form[3,46]. The ability to form bradyzoite-containing tissue cysts is particularly important for parasite transmission. When a cat consumes an intermediate host infected with tissue cysts, such as a rodent or bird, the bradyzoites can be released within the small intestine and initiate the sexual reproductive cycle.

After two days inside the definitive host, *T. gondii* undergoes its sexual phase in the epithelial cells of the small intestine. Initial stages involve the differentiation of parasites into male and female pre-sexual stages called microgamonts and macrogamonts, respectively. These precursor cells undergo complex developmental changes, with microgamonts producing numerous small, flagellated microgametes and macrogamonts developing into larger, immobile macrogametes[2,47–49]. The microgametes actively seek and fertilize the macrogametes, resulting in the formation of a zygote. Following fertilization, the zygote develops into an oocyst, which is then shed in the feces of the cat. A single infected cat can shed millions of oocysts over a short period, typically 1-2 weeks following initial infection, as this brief but intense shedding period is crucial for parasite transmission. The oocysts become infectious only after sporulation, at which point they can be ingested by intermediate hosts, restarting the complex life cycle.

The *Plasmodium falciparum* genome and diverse multicopy gene families

The *Plasmodium falciparum* genome maintains a complex organizational framework optimized for antigenic variation and immune evasion. Comprising approximately 23 Mb distributed across 14 chromosomes, the genome demonstrates a distinct architectural pattern with highly conserved central regions contrasting markedly with variable subtelomeric domains[50]. The subtelomeric regions exhibit a precise organizational hierarchy, beginning with six telomere-associated repeat elements

(TAREs) of variable size, followed by an ordered arrangement of variant antigen gene families[32]. This architecture places *var* genes proximal to the telomeres, followed by *rif* and *stevor* genes, along with other multigene families including *surfin*, *phist*, *clag*, and *fikk*[51]. This arrangement facilitates extensive genetic recombination, enabling the generation of millions of new antigenic structures within a single infected individual.

The complex var gene repertoire

The *var* gene family, comprising approximately 60 members per haploid genome[50], is controlled by a system of antigenic variation in *P. falciparum*. These genes demonstrate distinct architectural features with a conserved two-exon structure: exon 1 encoding the highly variable extracellular portion and cytoadhesive domains, and exon 2 encoding the more conserved transmembrane and intracellular components[31]. Each *var* gene contains two promoters—an upstream promoter controlling mutually exclusive expression and an intronic promoter generating regulatory lncRNAs[52,53]. The upstream promoters are categorized into three distinct classes based on their chromosomal location, with upsA exclusively in subtelomeric regions, while upsB and upsC show more varied positioning[31,50]. Transcriptional activation of *var* genes involves a multi-step process starting with initial low-level expression of multiple *var* genes, then selection and high-level activation of a single *var* gene. This aids in the evasion of the innate immune response by switching the primary *var* gene each generation, thus preventing recognition of parasitized cells via extracellular components of the protein[54–56].

Plasmodium falciparum erythrocyte membrane protein-1 (PfEMP1), the protein product of *var* genes, characterized by its large molecular mass and modular domain architecture, is one of the most sophisticated virulence factors identified in *P. falciparum* and one of the most diverse adhesion molecules in nature[57,58]. PfEMP1 variants differ in the number and structural configuration of tandem cysteine-rich Duffy-binding-like domains (DBLs) and cysteine-rich interdomain regions (CIDRs)[57]. The distinct adhesion domains of surface-expressed PfEMP1 molecules dictate the specific endothelial receptor interactions of parasitized erythrocytes within the microvasculature of different tissue, including the brain. The vast number of possible adhesion domain combinations resulting from homologous recombination between *var* gene loci generates heterogeneity in receptor-binding phenotypes among parasitized erythrocytes. The frequency of recombination in *var* genes is particularly high, leading to new *var* forms through gene conversion[59–61]. This recombination is facilitated by the nuclear architecture, where telomeric and internal *var* genes localize to the nuclear periphery[60,62]. Importantly, this genomic organization facilitates both antigenic variation and differential gene expression patterns. The subtelomeric location of variant antigen genes enables them to switch between active and silent states through epigenetic mechanisms[30,62,63], while their clustered arrangement promotes recombination and genetic diversity.

Although most of the *var* gene repertoire is highly variable, the *var2csa* variant is both remarkably well-conserved and the primary variant expressed in parasitized erythrocytes of the placenta. Exon 1 of *var2csa* contains a unique CSA binding site that interacts specifically with chondroitin sulfate A (CSA) on syncytiotrophoblasts within the placenta resulting in their sequestration within the placental intervillous space **(Fig. I.2)**[31,64,65]. A 5' regulatory region of *var2csa* transcripts acts as a repressor of mRNA translation, a feature not found in other *var* genes[66]. Translation of *var2csa* can be stimulated by *P. falciparum* translation enhancing factor (PTEF) binding of this upstream open reading frame (uORF)[67]. Pregnant women are highly susceptible to developing severe symptoms and pregnancy-associated malaria (PAM) adversely affects the fetus, resulting in low birth weight and preterm birth, and increased risk of fetal mortality[68–70]. Evolutionary conservation of elements necessary for its specialized function in placental adhesion make *var2csa* a promising target for vaccine development.

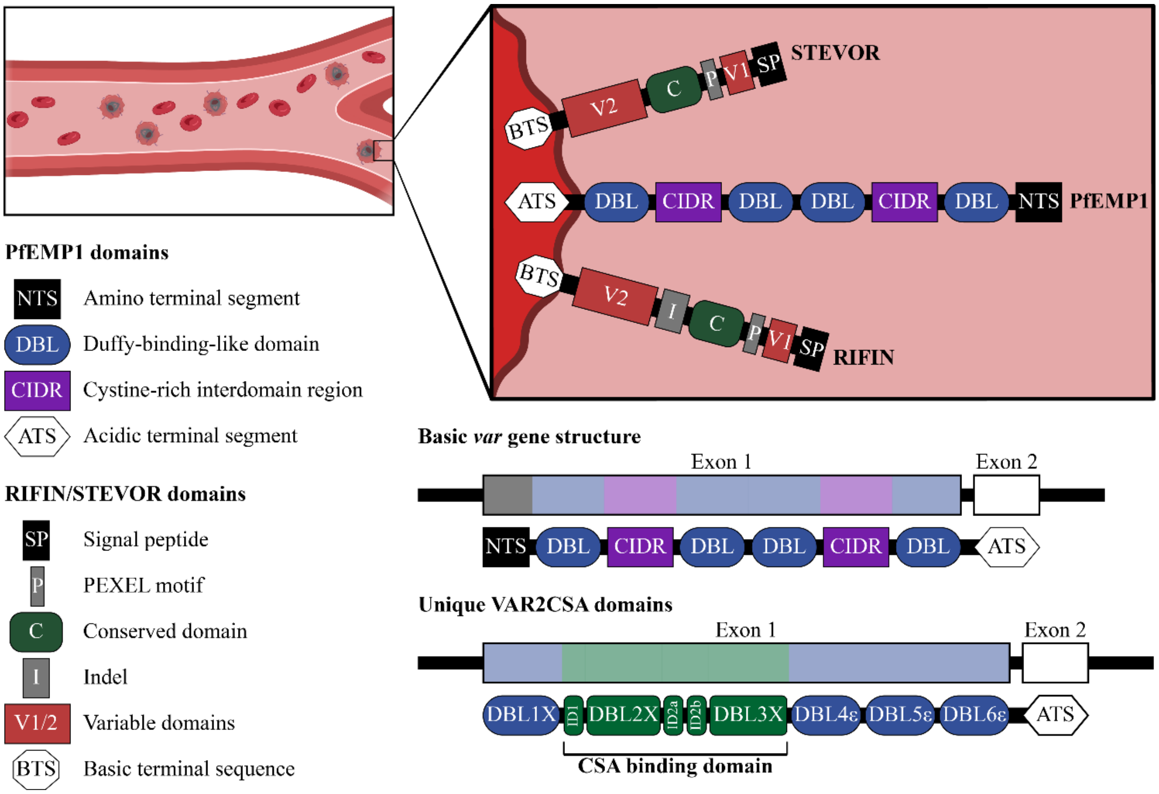

Fig. I.2. *Plasmodium* variant surface antigens. Domain architecture of major *P. falciparum* variant surface antigens (VSAs). Typical PfEMP1 structure showing the N-terminal segment (NTS) followed by alternating DBL and CIDR domains, and the conserved acidic terminal segment (ATS). RIFIN protein showing the A-RIFIN variant with the signal peptide (SP), first variable region (V1), PEXEL motif (P), semi-conserved domain (C) with the 25aa insertion found in A-RIFINs (I), second variable region (V2), and C-terminal transmembrane domain (BTS). STEVOR protein displaying similar domain structure to A-RIFINs but with a shorter variable region and no insertion sequence. Also shown is the exon structure of *var* genes from which PfEMP1 is transcribed and VAR2CSA's unique arrangement of six DBL domains (DBL1X, DBL2X-pam, DBL3X, DBL4ε, DBL5ε, DBL6ε) and lacking CIDR domains. The CSA binding domain is responsible for adhesion to syncytiotrophoblasts of the placenta.

Additional multigene families contributing to P. falciparum antigenic variation

The *rifin* family is substantially larger with ~150 members, and demonstrates more diverse organization and expression patterns[71,72]. These genes are categorized into five defined clusters (A1, A2, AB, B, and C), with 70% belonging to subgroup A[31,73]. The family shows stage-specific expression patterns, with A2 *rifin* variants preferentially expressed in trophozoites and early gametocytes[73,74]. Each *rifin* gene contains two exons encoding proteins with distinct conserved and variable domains, and a PEXEL motif that mediates export to the erythrocyte surface **(Fig. I.2)**[31,75]. A-RIFINs mediate adhesion to blood group A antigens and sialic acid on glycophorin A[71]. These interactions have significant clinical implications, as non-O blood groups are overrepresented in severe malaria cases. Beyond their role in infected erythrocytes, RIFINs are also expressed in merozoites and gametocytes, suggesting broader functions in parasite biology[31,74].

The *stevor* family, comprising ~30 genes, shares structural similarities with *rifin* but demonstrates distinct expression patterns and functional characteristics[50,72]. Like *rifin*, *stevor* genes contain PEXEL motifs and display stage-specific expression, however they are expressed later in the parasite life cycle compared to *var* and *rifin* genes **(Fig. I.2)**[32,76]. The *stevor* family resembles B-RIFINs in terms of gene number and structure, notably lacking the insertion sequence characteristic of A-RIFINs[31,73]. Like RIFINs, STEVOR proteins appear in multiple life cycle stages including merozoites and gametocytes, suggesting that they may be involved in host cell invasion[77,78]. While *var* genes typically demonstrate mutually exclusive expression, *rifin* and *stevor* families can express multiple members simultaneously[30,74,79].

These variant surface antigens work cooperatively as PfEMP1 is typically expressed alongside RIFINs and STEVORs, with precise timing of expression orchestrated throughout the parasite life cycle[71,77]. PfEMP1 and RIFINs appear on the infected erythrocyte surface simultaneously, while STEVORs emerge later during schizont development, and together they coordinate complex adhesion phenotypes through multiple receptor-ligand interactions. Complex adhesion phenotypes involving multiple variant surface antigens are strongly associated with severe malaria, making it essential to elucidate the molecular mechanisms controlling their expression and function for developing effective therapeutic interventions[80].

Exported effector proteins of *Toxoplasma gondii*

In North America and Europe, *Toxoplasma gondii* populations are predominantly composed of three major clonal lineages (types I, II, and III) that exhibit differences in virulence and interaction with host immunity[81,82]. Type I strains are more virulent and are associated with severe acute infections but are rarely isolated from human hosts, while Type II strains are significantly more common in humans and efficiently establish chronic infections through tissue cyst formation. Strains from South America, Asia, and Africa exhibit extensive genetic diversity, comprising distinct regional lineages that likely emerged through sexual recombination[83–85]. The molecular basis for geographical variation in *T. gondii* virulence lies in the varying composition of their secreted effector proteins. Genetic recombination and selective pressure resulted a unique combinations of rhoptry (ROP) and dense granule (GRA) proteins determining host-pathogen interactions.

Rhoptry Proteins

Rhoptries are specialized secretory organelles that define the apical end of *T. gondii* and play essential roles in host cell invasion and manipulation. These organelles are divided into two distinct regions: the rhoptry neck, which releases RON proteins, and the rhoptry bulb, which secretes ROP proteins[86]. This compartmentalization ensures precise temporal control over protein secretion during the invasion process. Most rhoptry proteins are injected into host cells during the initial stages of invasion before the parasitophorous vacuole is fully formed and include active kinases and pseudokinases that serve crucial roles in host cell manipulation and parasite survival.

ROP16 is a key effector in human infection through its direct manipulation of the JAK/STAT cell signaling pathway. This serine-threonine kinase directly phosphorylates STAT3 and STAT6 transcription factors, leading to downregulation of IL-12 production and modulation of inflammatory responses[86,87]. ROP38 regulates host cell signaling by specifically targeting the MAP kinase pathway, modulating apoptosis and cell proliferation signals[88]. Unlike other ROPs, ROP38 appears to function primarily through downstream modification of host transcriptional responses rather than direct kinase activity[88]. ROP54, a recently characterized pseudokinase, localizes to the cytoplasmic face of the parasitophorous vacuole membrane and restricts immune loading of GBP2[89]. Among the most well-characterized ROP proteins is ROP18, a serine/threonine kinase that complexes with ROP5 and ROP17 to target immunity-related GTPases (IRGs)[90,91]. However, this complex plays a minimal role in human cells due to the limited IRG family compared to mice. These effectors demonstrate strain-specific variations in their activity against human cells, suggesting ongoing evolutionary adaptation to human immunity.

Dense Granules

Dense granule proteins (GRA) are a diverse family of effectors that orchestrate modifications to host cells following invasion. Unlike rhoptries that are secreted in a rapid burst during invasion, GRA proteins exhibit sustained secretion throughout intracellular infection, enabling long-term manipulation of host cell functions. They utilize complex molecular machinery for export beyond the parasitophorous vacuole[92].

Several GRA proteins exhibit sophisticated nuclear targeting capabilities. GRA16 forms a complex with host phosphatase PP2A-B55 and HAUSP, modulating p53 levels and cell cycle progression[93]. This interaction promotes host cell survival under stress conditions through coordinated regulation of metabolic and cell cycle pathways.

GRA24 functions through direct p38α MAPK activation. It contains two kinase-interacting motifs that bind and scaffold two p38 molecules, enabling sustained kinase activation while preventing regulatory phosphatase binding[94,95]. This leads to prolonged inflammatory responses through enhanced cytokine production. While this may seem counterintuitive, by bypassing the classical MAPK phosphorylation cascade and maintaining sustained p38 autophosphorylation, GRA24 enables the parasite to fine-tune the inflammatory response in a way that balances parasite survival with host viability.

T. gondii inhibitor of STAT1 transcriptional activity (TgIST) inhibits STAT1-dependent transcription by forming an inhibitory complex with STAT1 and the Mi-2 nucleosome remodeling and deacetylase complex (Mi-2/NuRD)[96–99]. TgIST has a unique ability to modify the host cell epigenetic landscape by maintaining the activating H3K4me3 mark, while enforcing transcriptional repression of interferon-stimulated genes[98]. Critically, activity is most effective early in infection as cells have not been primed with IFN-γ, revealing a temporal aspect to its function that aligns with the parasite's need to ensure survival without completely compromising host immunity[99].

T. gondii NCoR/SMRT modulator (TgNSM) employs similar epigenetic regulatory functions by targeting a nuclear corepressor complex (NCoR/SMRT) and recruiting histone deacetylase enzymes such as HDAC3 to promoters of necroptosis-related genes[96]. TgNSM colocalizes with TgIST within the nucleus to block interferon-driven expression of PKR and MLKL. This coordinated function ensures survival of the host cell, therefore enabling the parasite to establish chronic infection.

Toxoplasma E2F4-associated EZH2-inducing gene regulator (TEEGR) manipulates transcriptional regulation by selectively binding E2F/DP transcription factors leading to activation of EZH2, a catalytic subunit of the Polycomb repressive complex 2 (PRC2) that inhibits expression to NF-κB regulated genes through methylation of histone H3 lysine 27 (H3K27me3)[100]. TEEGR exhibits selective suppression of NF-κB-regulated cytokines and chemokines, while maintaining expression of others, suggesting intricate control over the host immune response rather than broad immunosuppression.

Toxoplasma gondii secretes an array of effector proteins that fundamentally alter host cell function and enable successful parasitism. These proteins execute precise modifications to cell signaling pathways, gene expression patterns, and cell survival mechanisms. Understanding their molecular mechanisms illuminates fundamental aspects of cell signaling and immune regulation, while identifying potential therapeutic targets for toxoplasmosis treatment.

Epigenetic mechanisms that regulate gene expression and pathogen virulence.

Nuclear Architecture and Spatial Organization

Nuclear architecture in apicomplexan parasites undergoes structural modifications throughout developmental progression, directly influencing transcriptional regulation[50,51]. During intraerythrocytic development, *P. falciparum* exhibits stage-specific nuclear reorganization coordinated with its 48-hour replication cycle[38]. The ring stage maintains a compact genome structure with restricted transcriptional machinery access[52]. This organization transforms significantly during the trophozoite stage, where the nucleus shows substantial volume increase and chromatin expansion[53,54]. Hi-C experiments have revealed sophisticated 3D genome organization in *Plasmodium* species[27,50,55,56]. While lacking classical topologically associating domains found in other eukaryotes[50,57,58], *P. falciparum* shows distinct organizational features including centromeres clustering at one nuclear pole while telomere clusters forming at the opposite pole[27,56], chromosome arms folding into parallel structures anchored at centromeres[56,59], formation of distinct heterochromatin clusters, and stage-specific chromosomal interactions **(Fig. I.2)**[27,50,55,56].

Sexual stage development reveals distinct nuclear characteristics between male and female gametocytes. Female gametocytes develop smaller, more compact nuclei containing stored, translationally repressed transcripts, while male gametocytes exhibit larger nuclear volumes to accommodate mitotic preparation[55]. These features vary between species; *P. berghei* shows nuclear pores distant from DNA, while *P. falciparum* displays morphologically atypical gametocyte nuclear organization[55]. During gametocytogenesis, *P. falciparum* maintains telomeric and centromeric clusters while undergoing specific intrachromosomal reorganization[56].

The organization of ribosomal DNA (rDNA) reveals sophisticated nuclear compartmentalization. In *P. falciparum*, rDNA units dispersed on different chromosomes localize to a single nucleolus[60]. Hi-C data confirms enriched contacts between chromosomes 5 and 7, placing A-type rDNA units (active during IDC) near the nuclear periphery[27]. This organization suggests the existence of perinuclear transcriptionally active compartments that may help segregate active genes from compact heterochromatin regions[27]. Nuclear architecture plays a crucial role in virulence gene regulation through spatial organization. FISH studies have located *var* gene families in subtelomeric regions at the nuclear periphery[61], with Hi-C experiments confirming that internal *var* genes and other repressed genes physically associate with subtelomeric regions through chromatin looping[50]. This organization creates repressive centers enriched in H3K9me3 and HP1 protein **(Fig. I.2)**[62,63]. Similar clustering patterns appear in *P. knowlesi SICAvar* genes, suggesting a conserved mechanism for mutually exclusive expression control[64].

The transition between hosts to complete the parasite's life cycle requires specific nuclear reorganization. During sporozoite stages, Hi-C experiments reveal strong associations between repressive clusters and genes involved in invasion and virulence in both *P. falciparum* and *P. vivax*[55,56]. Additional long-range interactions develop for genes involved in liver stage development, particularly affecting *csp*, *trap*, and *gest* genes required for hepatocyte invasion[65,66].

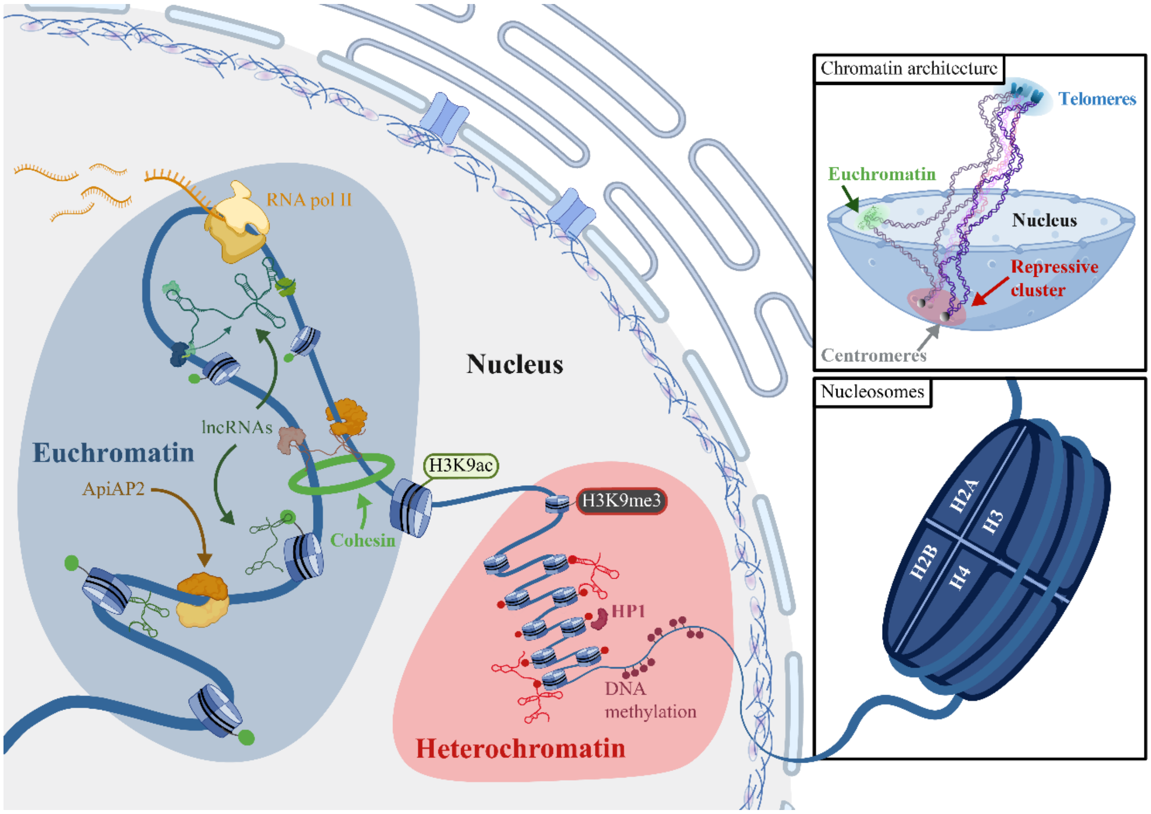

Figure I.3. Apicomplexan epigenetic landscape. The epigenetic landscape of Apicomplexa combines conserved eukaryotic features with unique regulatory mechanisms, including ApiAP2 transcription factors, activating (H3K9ac) and silencing (H3K9me3) histone modifications, and chromatin regulators like HP1. Apicomplexan nuclear architecture is characterized by distinctive clustering of centromeres and telomeres, with actively transcribed regions occupying distinct nuclear territories, all built upon the fundamental nucleosome structure composed of canonical histones H2A, H2B, H3, and H4.

Chromatin Structure and Nucleosome Dynamics

Within the nucleus of eukaryotic cells, including apicomplexan parasites, genomic DNA exists as a complex with proteins forming elaborate chromatin structures. This organization enables extensive DNA compaction by wrapping chromatin around histone octamers to create nucleosomes, which further aggregate into chromatin fibers. When chromatin adopts a highly condensed state, transcriptional machinery access becomes restricted, limiting the activation of gene expression and other protein-mediated processes. The structure of chromatin responds to various conditions including cellular environment, cell cycle stage, and cell type, enabling modulation of gene expression patterns[22,67–69]. Modification of chromatin structure and nucleosome positioning proves essential for transcriptional activation of target genes. Changes to the histone core occur through incorporation of histone variants or through post-translational modifications of histone tails, both mechanisms playing central roles in chromatin organization. The standard histone octamer comprises duplicate copies of H2A, H2B, H3, and H4, though this composition can vary significantly **(Fig. I.2)**. Several histone variants have emerged as crucial factors in chromatin reorganization events[70–72].

The incorporation of histone variants provides an important regulatory mechanism in apicomplexan parasites. The histone variant H2A.Z appears at transcription start sites of both active and inactive genes[70]. In *P. falciparum*, the positioning of H2A.Z exhibits unique characteristics, showing enrichment in intergenic AT-rich regions regardless of their transcriptional state[70,71]. This distribution pattern differs notably for clonally variant genes expressing PfEMP1, a protein essential for immune evasion and successful proliferation. Additional variants in *P. falciparum* include H2B.Z, CenH3, and H3.3, each conferring specific functional properties[73–75].

The nucleosome architecture in *Plasmodium* reveals distinctive features that facilitate its complex developmental program. Examination of active and highly transcribed genes shows consistent patterns of nucleosome-depleted regions near transcription start sites, accompanied by positioned nucleosomes within coding sequences. This arrangement permits transcriptional machinery access while maintaining structural organization[76–79]. In contrast, transcriptionally silent genes exhibit elevated nucleosome occupancy in promoter regions despite retaining nucleosome-depleted regions near their transcription start sites, establishing a barrier to transcription initiation [77–79]. This organizational pattern becomes particularly apparent within the IDC, where genes associated with sexual and transmission stages maintain dense nucleosome coverage in their regulatory regions[78,80]. Throughout the parasite cell cycle, nucleosome architecture undergoes substantial reorganization[73,78,80–82].

The transition between developmental stages requires precise coordination of chromatin structure modifications. During the *P. falciparum* asexual cycle, genes known to be upregulated in sexual and transmission stages maintain distinctive nucleosome profiles[78,80]. At the ring and schizont stages, the genome adopts a more condense chromatin state. Notable exceptions include genes involved in virulence factors, particularly those mediating red blood cell remodeling and host cell invasion[31,83,84]. This pattern undergoes significant alteration during the trophozoite stage, where substantial chromatin decompaction facilitates the extensive transcriptional activity required for growth and replication[85].

Gametocyte development introduces additional complexity to chromatin organization. The chromatin state reflects the distinct transcriptional requirements of male and female gametocytes, with female gametocytes having more compact chromatin organization associated with stored, transcriptionally repressed genes[55]. These changes coincide with stage-specific histone modifications, including H3K36me2/me3 in early gametocytes and H3K27me1 and H3K36me1 in late gametocytes[61,86].

The arrangement of chromatin structure carries significant implications for gene regulation and cellular development. In *P. falciparum*, genes that physically associate tend to show coordinated activation or repression patterns, demonstrating the importance of three-dimensional chromatin organization in transcriptional regulation[55,85]. This coordination becomes particularly evident in the expression of virulence genes, where chromatin structure helps maintain the mutually exclusive expression patterns vital for immune evasion[32,50,87–89]. The resolution of current 3C techniques has provided insights into chromatin remodeling events mediated by these proteins, though improved experimental approaches may reveal additional regulatory mechanisms. These species-specific variations in chromatin organization likely reflect evolutionary adaptations to different host environments and life cycle requirements.

Epigenetic Modifications

The dynamic rearrangement of histone cores and variants represents only one aspect of chromatin regulation in apicomplexan parasites. Post-translational modifications (PTMs) of histone tails, including methylation, acetylation, and phosphorylation, constitute a 'histone code' that significantly influences nucleosome-DNA interactions. These modifications play crucial roles in chromatin remodeling and gene expression by modulating the recruitment or inhibition of transcriptional machinery in promoter regions[72,90,91]. Quantitative mass spectrometry studies have identified extensive histone PTMs in *Plasmodium*, with over 230 modifications detected during the IDC of *P. falciparum*[92].

Certain epigenetic modifications represent universal markers of transcriptional states across eukaryotes. In *Plasmodium*, H3K9me3 interacts with heterochromatin protein 1 (HP1) and, along with H3K36me3, associates with silenced promoters and heterochromatin regions[55,72,90,93]. These marks appear mutually exclusive from H3K9ac and H3K4me3, which typically indicate active transcription[73,81,94]. Advanced machine learning analysis of *P. falciparum* histone modifications suggests that H3K9me3, H4K20me3, and H3K9ac serve specific regulatory functions during the schizont stage of erythrocytic development[95].

The regulation of sexual differentiation and gametocytogenesis involves distinct histone PTM patterns that respond to environmental stimuli. Mass spectrometry has identified specific gametocyte-associated modifications, including H3K36me2, H3K36me3, and H3K79me3 in early or late gametocytes, while H3K9me2 and H3K18me1 appear exclusively during the IDC[91].

Recent studies have expanded our understanding of epigenetic modifications across different *Plasmodium* species. In *P. vivax* sporozoites, H3K9me3 shows significant enrichment in subtelomeric regions, likely due to the concentration of virulence and invasion genes in these areas[96]. H3K9ac and H3K4me3 marks appear within 1000 kb upstream and downstream of highly expressed genes rather than within coding regions **(Fig. I.2)**[96]. The rodent malaria parasite *P. berghei* displays similar patterns to other *Plasmodium* species, with H3K9me3 and HP1 restricted to telomeric and subtelomeric regions[97]. H3K9ac shows broad distribution across chromosomes in all stages, though notably, H3K9ac enrichment in the 5′-UTR does not correlate with transcript upregulation in female gametocytes, possibly because these transcripts are produced and stored earlier[97].

The reversible nature of histone PTMs requires precise control by specialized proteins categorized as writers, erasers, and readers. The histone acetylome undergoes regulation by histone acetyltransferases, including GNAT and MYST family proteins, counterbalanced by histone deacetylases (HDACs) and sirtuin proteins[98]. These enzymes maintain temporal precision of epigenetic marks, as demonstrated by PfGCN5 and HDACs, whose inhibition disrupts *P. falciparum* gene expression patterns[94,99]. Histone lysine methylation relies on methyltransferases (HKMTs) and demethylases, though their roles remain less thoroughly characterized in these parasites[100].

Various regulatory proteins have demonstrated roles in virulence gene control. Disruption of Sir2A and Sir2B[101,102], along with the HKMT enzyme PfSETvs[103,104], increases transcription of multiple *var* genes, indicating their involvement in virulence gene silencing. The critical nature of these enzymes has prompted investigation into inhibitors as potential therapeutic agents, with some compounds showing high specificity in disrupting parasite life cycle progression[86].

Beyond writers and erasers, histone marks require recognition by reader proteins that facilitate recruitment or stabilization of protein complexes mediating diverse biological processes. Proteomic studies have identified multiple readers in *P. falciparum*[105], including HP1 and *Pf*GCN5-PfADA2, components of the SAGA complex involved in transcription initiation and elongation[106]. The bromodomain protein *Pf*BDP1, working with *Pf*BDP2 and *Pf*AP2-I, binds acetylated histones in invasion gene promoters, ensuring proper erythrocyte invasion[107,108].

Numerous non-histone proteins associate with chromatin to regulate DNA maintenance, replication, and transcription. Recent chromatin enrichment proteomics identified several such factors in *P. falciparum*, including SMC and CRWN-like proteins[109]. The latter share functional similarity with plant lamina proteins[110] and associate with heterochromatin regions in *Plasmodium*[109]. SMC proteins, conserved across eukaryotes, participate in chromosome condensation, segregation, and DNA repair. ChIP-seq experiments have shown *Pf*SMC3 localizing to centromere regions on all chromosomes[109], suggesting roles in sister chromatid separation and cell cycle regulation similar to those observed in yeast[111].

Epigenetic modifications play crucial roles during developmental transitions across different parasite life stages. In *P. cynomolgi*, studies using histone methyltransferase inhibitors demonstrated acceleration of hypnozoite activation in primary hepatocyte cultures[112]. While the precise mechanisms remains unclear—whether through direct effects on parasite histones or host methylation status—these findings confirm the importance of epigenetic regulation across all major life cycle stages[112].

Epigenetic modifications are particularly vital in regulating virulence gene families. In *P. falciparum*, *var* gene regulation involves a complex interplay between various modifications[73,81,94]. The active *var* gene exhibits H3K9ac and H3K4me2/me3 marks[98,99], while silent *var* genes show H3K9me3 enrichment in the nuclear periphery[63,100,113,114]. These processes involve stage-specific chromatin modifications and long-range chromosomal interactions mediated by novel architectural proteins.

The transition to sexual stages involves specific epigenetic signatures. Recent mass spectrometry analysis has identified stage-specific modifications throughout gametocyte development[94], including distinct patterns of H3K36me2/me3 in early gametocytes and H3K27me1 and H3K36me1 in late gametocytes[61,86]. In *P. berghei*, detailed epigenomic mapping across asexual stages, male and female gametocytes, and ookinetes has confirmed the restricted presence of H3K9me3 and HP1 in telomeric and subtelomeric regions[97]. The distribution of H3K9ac marks across chromosomes showed distinctive patterns in female gametocytes, where transcript enrichment patterns suggested earlier production and storage of specific mRNAs[97].

The essential role of epigenetic regulation in parasite development presents promising therapeutic opportunities. Inhibitors targeting histone-modifying enzymes have shown efficacy in disrupting parasite life cycle progression[86].

Transcriptional Regulation

Transcriptional activity depends heavily on nucleosome occupancy and epigenetic states at promoter regions, which modulate transcriptional machinery access to chromatin. RNA polymerase II (RNA Pol II) and its associated subunits carry out protein-coding gene transcription, with recruitment to promoter regions requiring general transcription factors including TFII subunits[115]. While eukaryotes typically employ TATA-binding protein (TBP) and TBP-associated factors (TAFs) in the TFIID complex, *Plasmodium* TAFs lack the histone fold domain responsible for TAF heterodimerization[116]. This structural difference necessitates alternative regulatory mechanisms, exemplified by *Pf*TBP and *Pf*TFIIE binding to promoters independently of transcriptional activity and acetylated histone presence[117].

Analysis of RNA Pol II occupancy reveals sophisticated regulatory mechanisms during the IDC. Two distinct binding phases occur that do not directly correlate with detected mRNA levels, suggesting regulatory processes beyond transcription initiation[118]. Recent comparative analysis of genome-wide nascent RNA with steady-state mRNA profiles demonstrated elevated RNA Pol II pausing during early ring stages compared to other blood stages, confirming the importance of polymerase pausing in gene expression regulation [119]. These patterns align with the synchronous pattern of expression observed during development, involving coordinated action between multiple regulatory components, including transcription factors, nucleosome positioning mechanisms, histone modifications, and chromatin structural regulators[120,121].

The discovery of transcription factors in *Plasmodium* revealed a unique regulatory landscape. Beyond the limited presence of common transcription factors[122], the identification of the apicomplexan-specific ApiAP2 family marked a significant breakthrough[123]. These factors, comprising approximately 27 members in *Plasmodium*, contain one to four AP2 DNA-binding domains analogous to plant Apetala-2 proteins[124]. Despite their relatively small number compared to the approximately 5,500 genes in the parasite genome[122], these factors serve as master regulators of stage-specific transitions.

Several ApiAP2 transcription factors function as master regulators of stage-specific transitions in *Plasmodium* development. Factors regulating specific stages include AP2-G and AP2-G2 control of gametocyte development[37,39], AP2-O for ookinete development[125], AP2-SP for sporozoite stages[125], and AP2-L for liver stage progression in both *P. falciparum* and *P. berghei*[126]. Recent knockout screening of ApiAP2 family members in *P. berghei* and *P. yoelii* has revealed complex regulatory patterns, with these factors acting as both repressors and activators during stage transitions[127,128]. Analysis of AP2exp mutants showed transcriptional changes affecting exported proteins associated with clonally variant gene families, including RIFIN and STEVOR[129].

Despite the presence of AP2 motifs upstream of genes showing variable expression across the IDC, machine learning algorithms suggest these motifs may not be critical determinants of gene expression during the *Plasmodium* erythrocytic cycle[95]. This observation suggests that ApiAP2 DNA binding likely requires both specific motif recognition and appropriate epigenetic states at target loci. The relationship between DNA accessibility, epigenetic state, and transcription factor binding parallels regulatory mechanisms observed in other eukaryotes[130]. The limited number of transcription factors (1.46% of protein-coding genes compared to humans at >7.5%) implies the existence of regulatory mechanisms beyond conventional transcriptional control[121,129,131].

The synchronous pattern of expression during development necessitates coordinated action between multiple regulatory components, including transcription factors, nucleosome positioning mechanisms, histone modifications and their regulatory enzymes, DNA methylation-associated enzymes, non-coding RNAs, chromatin structural regulators, and RNA-binding proteins[120,121,124]. ChIP-sequencing experiments targeting RNA polymerase II have revealed sophisticated transcriptional regulation mechanisms, showing that RNA polymerase II binds to targeted promoters during early asexual stages and remains stationary until specific molecular signaling thresholds trigger transcriptional activation during the trophozoite stage[132].

The regulation of virulence gene families highlights the complexity of transcriptional control mechanisms. The transcriptional control of *var* genes involves interplay between various regulators, including ApiAP2 heterochromatin-associated factors (PfAP2-HF). Recent studies identified eight new PfAP2-HFs showing preferential enrichment in heterochromatic regions, with their knockout resulting in significant changes in *var* gene expression[133,134]. AP2-exp and AP2-exp2 mutations affect expression of antigenic variant genes (*rifin*, *stevor*, and *Pfmc*-2TM) and cell remodeling genes[134–136].

The limited transcription factor repertoire in *Plasmodium* necessitates integration with other regulatory mechanisms. AP2 factors associate with various chromatin-associated proteins, including AP2-HC[135] and PfSIP2, which show enrichment in subtelomeric regions[133,134]. The SAGA complex members PfGCN5-PfADA2 participate in multiple transcriptional steps including initiation and elongation[106], while bromodomain proteins like PfBDP1 and PfBDP2 work with PfAP2-I to regulate invasion gene expression[107,108]. These interactions demonstrate how transcription factors coordinate with chromatin modifiers to achieve precise gene regulation.

Transcriptional control mechanisms respond to environmental signals and stress conditions, particularly during host cell invasion and immune evasion. ApiAP2 factors in

Plasmodium coordinate responses to changing host environments during transmission between mosquito and human hosts. This environmental responsiveness requires precise coordination between transcription factors and chromatin remodeling machinery to ensure appropriate stage-specific gene expression.

Non-coding RNA Regulation

Non-coding RNAs (ncRNAs) play crucial regulatory roles in cellular processes, though their initial classification as transcriptional noise has evolved with advanced RNA-seq platforms revealing their importance as regulatory elements[137]. These transcripts conventionally divide into small non-coding RNAs and long non-coding RNAs (lncRNAs) based on a 200-nucleotide length threshold[138]. Unlike other eukaryotes, *Plasmodium* lacks conventional RNA interference mechanisms (miRNAs, siRNAs, piRNAs)[137], instead evolving an expanded repertoire of RNA-binding proteins and lncRNAs serving diverse regulatory functions[132].

LncRNAs undergo processing similar to messenger RNAs, including capping, splicing, and polyadenylation, despite not encoding proteins. In higher eukaryotes, lncRNAs participate in various regulatory processes including chromosome dosage compensation, imprinting, transcription, translation, splicing, epigenetic regulation, and cell differentiation[138]. These molecules often function as molecular signals responding to various stimuli, acting as signals, decoys, guides, and scaffolds for key regulatory proteins. In *P. falciparum*, hundreds of ncRNAs show stage-specific expression patterns[139], with specific subsets transcribed from telomere-associated repetitive elements

(TAREs) contributing to regional stability and heterochromatin maintenance[140–142] These TARE-DNA sequences extend from telomeres to the first subtelomeric *var* genes, with TARE-lncRNAs transcribed throughout the IDC[140,141]. They potentially regulate telomeric heterochromatin through recruitment of chromatin modifiers, contributing to the broader regulatory network controlling variant surface antigen expression[140–142].

The regulation of sexual commitment involves complex interactions between lncRNAs and protein factors. Gametocytogenesis, controlled by the master regulator PfAP2-G, involves intricate regulation through the GDV1 protein[139,143]. During asexual stages, a five-exon antisense lncRNA transcribed downstream from the *Pfgdv1* locus represses *Pfgdv1* expression. Recent studies have demonstrated that GDV1 overexpression can dramatically increase sexual conversion rates from 8% to 75%[143], highlighting the essential nature of this gene-lncRNA regulatory mechanism. At the post-transcriptional level, lncRNAs influence multiple aspects of RNA metabolism including mRNA export, turnover, decay, and translation[138,144]. In the cytoplasm, they function as scaffolds for protein complexes involved in gene expression, directly interacting with translation factors to modulate translation efficiency.

The interaction between ncRNAs and chromatin structure represents a crucial regulatory mechanism. In *P. falciparum*, the involvement of lncRNAs and GC-rich ncRNAs in *var* gene activation demonstrates the complex interplay between RNA-based regulation and chromatin modification. TARE-lncRNAs contribute to heterochromatin formation and maintenance, suggesting a direct role in chromatin organization[140–142]. While lncRNAs serve essential functions across apicomplexan species, their specific roles and mechanisms show considerable variation. The diverse functions of lncRNAs in higher eukaryotes[138], including their roles in chromatin structure modification and transcriptional control through protein complex recruitment, suggest that additional regulatory mechanisms likely remain to be discovered in apicomplexan parasites[145]. The essential roles of ncRNAs in parasite development and host-pathogen interactions present potential therapeutic opportunities. Stage-specific expression of certain ncRNAs and their involvement in developmental transitions suggest they could serve as therapeutic targets.

Virulence Gene Regulation

The strongest evidence for the importance of genome architecture in transcriptional regulation emerges from the study of virulence genes, particularly the VRSM genes (*var*, *rifin*, *stevor*, and *Pfmc-2tm* genes) in *P. falciparum*[27]. These virulence proteins are encoded by large polymorphic gene families expressed in a mutually exclusive manner, essential for the parasite's immune evasion capabilities. This process parallels the "allelic exclusion" mechanism described for mammalian immunoglobulin and odorant receptor genes[146,147]. Understanding the principles governing mutually exclusive expression of VRSM genes represents a key question in malaria research[148,149].

Initial FISH experiments revealed that VRSM genes colocalize at the nuclear periphery[27,55]. Subsequent Hi-C experiments confirmed that chromatin looping brings all VRSM genes, including internal *var* genes and repressed genes, into physical proximity near subtelomeric chromosomal clusters[27,55,85,88]. Similar clustering patterns appear in *P. knowlesi SICAvar* genes, suggesting that nuclear architecture-mediated control of mutually exclusive expression represents a conserved mechanism across *Plasmodium* species[56]. The arrangement of virulence genes within chromosome territories significantly influences their expression. In *P. falciparum*, telomeric clustering creates repressive centers enriched in H3K9me3 and HP1, facilitating interactions between *var* gene families crucial for mutually exclusive expression[62,63]. These perinuclear repressive clusters incorporate additional subtelomeric multigene families associated with reinvasion, including *stevor* and *Pfmc-2TM*, which undergo similar mutually exclusive expression during the parasite cell cycle[156]. However, these findings from Hi-C and FISH experiments remain correlative, requiring additional research to establish direct mechanistic links between nuclear organization and virulence gene control.

Histone modifications play crucial roles in *var* gene regulation as well. Silent *var* genes display clear H3K9me3 enrichment[32,63,150], while a single active *var* gene exhibits H3K9ac and H3K4me2/me3 marks[100,150]. The importance of these modifications becomes evident through studying regulatory enzymes Sir2A and Sir2B[101,102] and the HKMT enzyme PfSETvs[103,104], whose disruption increases transcription of multiple *var* genes.

LncRNAs also regulate expression of variant surface antigens. In *P. falciparum*, they modulate antigenic switching of PfEMP1 variants through both sense and antisense transcription from *var* gene introns[61,151–155]. The RUF6 family of ncRNAs, located near internal *var* genes, demonstrates clonally variant expression patterns that influence active *var* gene expression through interactions with ApiAP2 transcription factors. Additionally, TARE-lncRNAs contribute to heterochromatin formation and maintenance near immune evasion genes, extending from telomeres to the first subtelomeric *var* genes[140–142].

Virulence gene regulation demonstrates the integration of multiple regulatory layers, including nuclear architecture, epigenetic modifications, transcription factors, and non-coding RNAs. The orchestration of these mechanisms ensures appropriate timing and level of virulence gene expression while maintaining mutually exclusive expression patterns where required. This complexity reflects the critical nature of virulence gene regulation for parasite survival and highlights the evolutionary pressure to maintain sophisticated control mechanisms. The complex regulation of virulence genes presents multiple potential therapeutic targets. Inhibitors targeting histone-modifying enzymes involved in *var* gene regulation have shown promise in disrupting parasite life cycle progression[86]. The distinct mechanisms employed by different species suggest opportunities for species-specific therapeutic interventions. Targeting *P. falciparum var* gene regulatory complexes may provide selective approaches to combat these infections.

Additional features of chromatin organization in higher eukaryotes

The hierarchical organization of eukaryotic chromatin enables precise spatial and temporal regulation of genomic functions within the confined nuclear space. While chromatin dynamics mediated by histone modifications and interactions with chromatin remodeling proteins and transcription factors are consistent features throughout most eukaryotic organisms, Apicomplexa do not exhibit the higher-order chromatin structure of more complex eukaryotes. This divergent organization likely reflects the specialized requirements of parasitic life cycles and the need for rapid, coordinated gene expression changes during host cell invasion and adaptation.

At the intermediate scale chromatin forms loop structures typically ranging from 50-500kb, which are stabilized by architectural proteins, primarily CTCF and cohesin complexes **(Fig. I.3)**[157]. The primary model for loop formation involves active loop extrusion by cohesin complexes, which function as molecular motors to progressively expand chromatin loops until encountering boundary elements, typically CTCF binding sites[158-160]. High-resolution Hi-C studies demonstrated that over 90% of chromatin loops occur between CTCF sites arranged in a convergent orientation[157]. This directional preference provides a structural basis for proper genome folding and regulation, with functional implications for gene expression and developmental processes. When CTCF is knocked out or its binding sites are inverted, it disrupts the normal looping patterns and alters gene regulation due to reduced interactions within topological domains but increased inappropriate interactions between neighboring domains[159,161,162]. Although cohesin is vital to the establishment and maintenance of chromatin loops, depletion of cohesin has revealed that while widespread changes in three-dimensional chromatin organization occur, the impact on gene expression is relatively modest, suggesting that additional regulatory mechanisms help maintain transcriptional stability even when chromatin architecture is perturbed[161,163,164]. However, genes whose promoters participate in loops show significantly higher expression compared to those without loops, demonstrating the importance of these structures in transcriptional regulation[157].

These loops facilitate specific enhancer-promoter interactions while preventing inappropriate regulatory contacts, creating insulated neighborhoods called Topologically Associating Domains (TADs), which represent self-interacting genomic regions typically spanning 200kb-1Mb **(Fig. I.3)**[57,165]. TADs demonstrate remarkable evolutionary conservation across cell types and species, suggesting their essential role in maintaining proper gene regulation and genomic stability[166]. While TAD boundaries demonstrate considerable conservation, internal domain organization exhibits cell-type-specific variations that correlate with transcriptional states and developmental potential[167-169]. During cellular differentiation, TAD architecture undergoes remodeling at key developmental loci, facilitating lineage-specific transcriptional programs through modulation of enhancer accessibility and promoter targeting. The disruption of TAD boundaries, whether through genetic alterations or epigenetic modifications, can lead to inappropriate enhancer-promoter interactions and subsequent dysregulation of gene expression, highlighting their importance in development and disease pathogenesis[170,171].

Moving up the organizational hierarchy, chromatin segregates into A/B compartments—megabase-scale domains that correspond to active and inactive chromatin states respectively **(Fig. I.3)**[157,172]. A compartments are characterized by open chromatin, high gene density, and active transcription, and are enriched in H3K36me3 and H3K4me3 modifications in higher eukaryotes. Conversely, B compartments exhibit dense chromatin packing, lower gene density, and transcriptional repression, and are enriched in H3K9me3 and H3K27me3 histone marks. The molecular basis for compartmentalization appears to be driven by factors such as nucleosome density, histone modifications, and associated protein complexes[172-174]. The cumulative effect of these molecular interactions leads to distinct nuclear microenvironments that facilitate or restrict transcriptional activity. High-resolution Hi-C studies have revealed that compartments can be further subdivided into subcompartments characterized by differing patterns of histone modifications and transcriptional activity[157,175].

At the highest level, individual chromosomes occupy discrete nuclear regions termed chromosome territories, where physical constraints and regulatory requirements arrange chromosomes to best facilitate efficient gene regulation **(Fig. I.3)**[172,176]. This level of organizational hierarchy is not unlike the antipodal organization of the *P. falciparum* genome[27,56]. Spatial constraints within chromosome territories establish probabilistic interactions where DNA within the same chromosome (cis) naturally interact at a higher frequency than segments from different chromosomes (trans)[172,177]. Territories are not rigid structures but rather dynamic entities that can undergo repositioning during cellular differentiation and in response to physiological stimuli[178,179].

Understanding the higher-order chromatin organization in complex eukaryotes provides critical insights into Apicomplexa biology, as these parasites have evolved distinct modifications to canonical genome architecture. While apicomplexan parasites maintain fundamental nucleosome organization with core histones, they notably lack key architectural features present in higher eukaryotes, such as clearly defined TADs and canonical A/B compartmentalization. Instead of CTCF-mediated boundary elements, these organisms employ alternative mechanisms for chromatin organization, particularly around virulence genes. The comparative analysis between higher eukaryotic and apicomplexan chromatin organization thus illuminates both the conserved foundational elements of genome architecture and the specialized regulatory innovations that enable parasitic survival strategies.

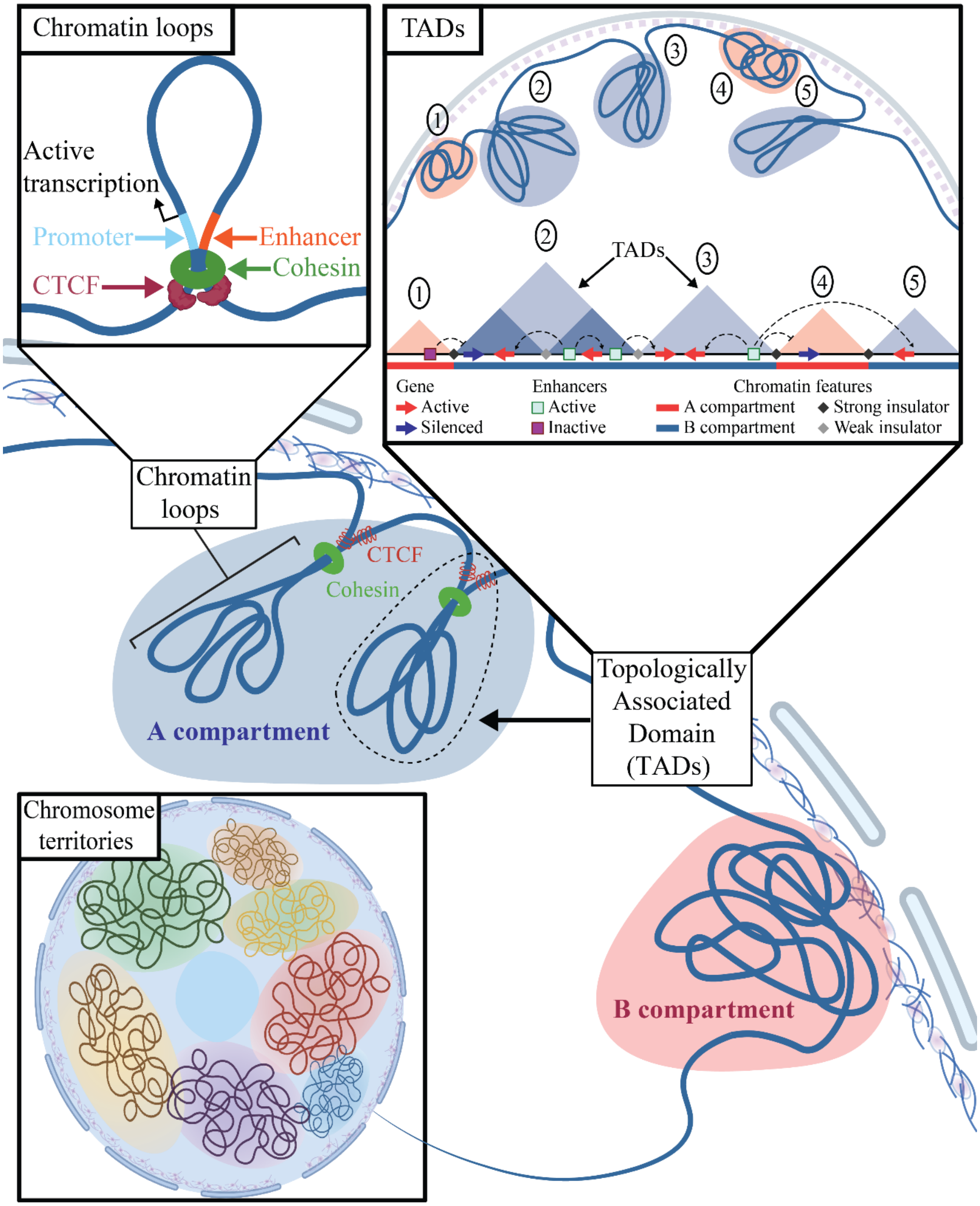

Figure I.4. The complex chromatin organization of higher eukaryotes. Chromosome organization in higher eukaryotes occurs at multiple hierarchical scales, from small chromatin loops mediated by cohesin and CTCF (50-500kb), to topologically associating domains (TADs) grouping related regulatory elements (200kb-1Mb), to larger A/B compartments that segregate active and inactive chromatin (1-10Mb), and finally to discrete chromosome territories within the nucleus.

REFERENCES

1. By-Nc-Sa, C. World malaria report 2023.
2. Dubey, J. P. Advances in the life cycle of Toxoplasma gondii. *International Journal for Parasitology* **28**, 1019–1024 (1998).
3. Weiss, L. M. & Dubey, Jitender. P. Toxoplasmosis: A history of clinical observations. *International Journal for Parasitology* **39**, 895–901 (2009).
4. Vannier, E. G., Diuk-Wasser, M. A., Ben Mamoun, C. & Krause, P. J. Babesiosis. *Infectious Disease Clinics of North America* **29**, 357–370 (2015).
5. Chartier, C. & Paraud, C. Coccidiosis due to Eimeria in sheep and goats, a review. *Small Ruminant Research* **103**, 84–92 (2012).
6. Morrissette, N. S. & Sibley, L. D. Cytoskeleton of Apicomplexan Parasites. *Microbiol Mol Biol Rev* **66**, 21–38 (2002).
7. McFadden, G. I. The apicoplast. *Protoplasma* **248**, 641–650 (2011).
8. Okada, M. *et al.* Critical role for isoprenoids in apicoplast biogenesis by malaria parasites. *eLife* **11**, e73208 (2022).
9. Ralph, S. A., D'Ombrain, M. C. & McFadden, G. I. The apicoplast as an antimalarial drug target. *Drug Resistance Updates* **4**, 145–151 (2001).
10. LeRoux, M., Lakshmanan, V. & Daily, J. P. Plasmodium falciparum biology: analysis of in vitro versus in vivo growth conditions. *Trends in Parasitology* **25**, 474–481 (2009).
11. Tantular, I. S. *et al.* Preservation of Wild Isolates of Human Malaria Parasites in Wet Ice and Adaptation Efficacy to In Vitro Culture. *Trop.Med.Health* **40**, 37–45 (2012).
12. Roobsoong, W. *et al.* Improvement of culture conditions for long-term in vitro culture of Plasmodium vivax. *Malar J* **14**, 297 (2015).
13. Valenciano, A. L., Gomez-Lorenzo, M. G., Vega-Rodríguez, J., Adams, J. H. & Roth, A. In vitro models for human malaria: targeting the liver stage. *Trends in Parasitology* **38**, 758–774 (2022).
14. Gural, N. *et al.* In Vitro Culture, Drug Sensitivity, and Transcriptome of Plasmodium Vivax Hypnozoites. *Cell Host & Microbe* **23**, 395-406.e4 (2018).

15. Müller, J. & Hemphill, A. In vitro culture systems for the study of apicomplexan parasites in farm animals. *International Journal for Parasitology* **43**, 115–124 (2013).

16. Feix, A. S., Cruz-Bustos, T., Ruttkowski, B. & Joachim, A. In vitro cultivation methods for coccidian parasite research. *International Journal for Parasitology* **53**, 477–489 (2023).

17. Wongsrichanalai, C. & Sibley, C. H. Fighting drug-resistant Plasmodium falciparum: the challenge of artemisinin resistance. *Clinical Microbiology and Infection* **19**, 908–916 (2013).

18. Suresh, N. & Haldar, K. Mechanisms of artemisinin resistance in Plasmodium falciparum malaria. *Current Opinion in Pharmacology* **42**, 46–54 (2018).

19. Vaughan, A. M. & Kappe, S. H. I. Malaria Parasite Liver Infection and Exoerythrocytic Biology. *Cold Spring Harb Perspect Med* **7**, a025486 (2017).

20. Ejigiri, I. & Sinnis, P. Plasmodium sporozoite–host interactions from the dermis to the hepatocyte. *Current Opinion in Microbiology* **12**, 401–407 (2009).

21. Milner, D. A. Malaria Pathogenesis. *Cold Spring Harb Perspect Med* **8**, a025569 (2018).

22. Mueller, I., Zimmerman, P. A. & Reeder, J. C. Plasmodium malariae and Plasmodium ovale – the 'bashful' malaria parasites. *Trends in Parasitology* **23**, 278–283 (2007).

23. Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S. & Mitchell, G. H. A Brief Illustrated Guide to the Ultrastructure of Plasmodium falciparum Asexual Blood Stages. *Parasitology Today* **16**, 427–433 (2000).

24. Warncke, J. D. & Beck, H.-P. Host Cytoskeleton Remodeling throughout the Blood Stages of Plasmodium falciparum. *Microbiol Mol Biol Rev* **83**, e00013-19 (2019).

25. *Pathogen-Host Interactions: Antigenic Variation v. Somatic Adaptations*. vol. 57 (Springer International Publishing, Cham, 2015).

26. Anders, R. F., McColl, D. J. & Coppel, R. L. Molecular variation in Plasmodium falciparum: Polymorphic antigens of asexual erythrocytic stages. *Acta Tropica* **53**, 239–253 (1993).

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 Research* **24**, 974–988 (2014).

28. Bozdech, Z. *et al.* The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum. *PLoS Biol* **1**, e5 (2003).

29. Daily, J. P. *et al.* In Vivo Transcriptome of *Plasmodium falciparum* Reveals Overexpression of Transcripts That Encode Surface Proteins. *J INFECT DIS* **191**, 1196–1203 (2005).

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

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

32. Scherf, A., Lopez-Rubio, J. J. & Riviere, L. Antigenic Variation in *Plasmodium falciparum*. *Annu. Rev. Microbiol.* **62**, 445–470 (2008).

33. Lee, W.-C., Russell, B. & Rénia, L. Sticking for a Cause: The Falciparum Malaria Parasites Cytoadherence Paradigm. *Front. Immunol.* **10**, 1444 (2019).

34. Turrini, F., Giribaldi, G., Carta, F., Mannu, F. & Arese, P. Mechanisms of band 3 oxidation and clustering in the phagocytosis of *Plasmodium falciparum* -infected erythrocytes. *Redox Report* **8**, 300–303 (2003).

35. Venugopal, K., Hentzschel, F., Valkiūnas, G. & Marti, M. Plasmodium asexual growth and sexual development in the haematopoietic niche of the host. *Nat Rev Microbiol* **18**, 177–189 (2020).

36. Eksi, S. *et al.* Plasmodium falciparum Gametocyte Development 1 (Pfgdv1) and Gametocytogenesis Early Gene Identification and Commitment to Sexual Development. *PLoS Pathog* **8**, e1002964 (2012).

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

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

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

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

41. Sologub, L. *et al.* Malaria proteases mediate inside-out egress of gametocytes from red blood cells following parasite transmission to the mosquito: Inside-out egress of malaria gametocytes. *Cellular Microbiology* **13**, 897–912 (2011).

42. Montoya, J. & Liesenfeld, O. Toxoplasmosis. *The Lancet* **363**, 1965–1976 (2004).

43. Robert-Gangneux, F. & Dardé, M.-L. Epidemiology of and Diagnostic Strategies for Toxoplasmosis. *Clin Microbiol Rev* **25**, 264–296 (2012).

44. Meissner, M., Schlüter, D. & Soldati, D. Role of *Toxoplasma gondii* Myosin A in Powering Parasite Gliding and Host Cell Invasion. *Science* **298**, 837–840 (2002).

45. Sheffield, H. G. & Melton, M. L. The Fine Structure and Reproduction of Toxoplasma gondii. (2024).

46. Lyons, R. E., McLeod, R. & Roberts, C. W. Toxoplasma gondii tachyzoite–bradyzoite interconversion. *Trends in Parasitology* **18**, 198–201 (2002).

47. Francia, M. E., Dubremetz, J.-F. & Morrissette, N. S. Basal body structure and composition in the apicomplexans Toxoplasma and Plasmodium. *Cilia* **5**, 3 (2015).

48. Tomasina, R. & Francia, M. E. The Structural and Molecular Underpinnings of Gametogenesis in Toxoplasma gondii. *Front. Cell. Infect. Microbiol.* **10**, 608291 (2020).

49. Dubey, J. P., Lindsay, D. S. & Speer, C. A. Structures of *Toxoplasma gondii* Tachyzoites, Bradyzoites, and Sporozoites and Biology and Development of Tissue Cysts. *Clin Microbiol Rev* **11**, 267–299 (1998).

50. Gardner, M. J. *et al.* Genome sequence of the human malaria parasite Plasmodium falciparum. **419**, (2002).

51. Mok, B. W., Ribacke, U., Sherwood, E. & Wahlgren, M. A highly conserved segmental duplication in the subtelomeres of Plasmodium falciparum chromosomes varies in copy number. *Malar J* **7**, 46 (2008).

52. Calderwood, M. S., Gannoun-Zaki, L., Wellems, T. E. & Deitsch, K. W. Plasmodium falciparum var Genes Are Regulated by Two Regions with Separate Promoters, One Upstream of the Coding Region and a Second within the Intron. *Journal of Biological Chemistry* **278**, 34125–34132 (2003).

53. Dzikowski, R. *et al.* Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Reports* **8**, 959–965 (2007).

54. Smith, J. D. *et al.* Switches in expression of plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101–110 (1995).

55. Roberts, D. J. *et al.* Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* **357**, 689–692 (1992).

56. Recker, M. *et al.* Antigenic Variation in Plasmodium falciparum Malaria Involves a Highly Structured Switching Pattern. *PLoS Pathog* **7**, e1001306 (2011).

57. Smith, J. D. The role of PfEMP1 adhesion domain classification in Plasmodium falciparum pathogenesis research. *Molecular and Biochemical Parasitology* **195**, 82–87 (2014).

58. Baruch, D. I. *et al.* Cloning the P. falciparum Gene Encoding PfEMPl, a Malarial Variant Antigen and Adherence Receptor on the Surface of Parasitized Human Erythrocytes.

59. Claessens, A. *et al.* Generation of Antigenic Diversity in Plasmodium falciparum by Structured Rearrangement of Var Genes During Mitosis. *PLoS Genet* **10**, e1004812 (2014).

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

61. Taylor, H. M., Kyes, S. A. & Newbold, C. I. Var gene diversity in Plasmodium falciparum is generated by frequent recombination events. *Molecular and Biochemical Parasitology* **110**, 391–397 (2000).

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

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

64. Fried, M. & Duffy, P. E. Adherence of *Plasmodium falciparum* to Chondroitin Sulfate A in the Human Placenta. *Science* **272**, 1502–1504 (1996).

65. Ma, R. *et al.* Structural basis for placental malaria mediated by Plasmodium falciparum VAR2CSA. *Nat Microbiol* **6**, 380–391 (2021).

66. Amulic, B., Salanti, A., Lavstsen, T., Nielsen, M. A. & Deitsch, K. W. An Upstream Open Reading Frame Controls Translation of var2csa, a Gene Implicated in Placental Malaria. *PLoS Pathog* **5**, e1000256 (2009).

67. Chan, S. *et al.* Regulation of PfEMP1–VAR2CSA translation by a Plasmodium translation-enhancing factor. *Nat Microbiol* **2**, 17068 (2017).

68. Steketee, R., Nahlen, B., Parise, M. & Menendez, C. The burden of malaria in pregnancy in malaria-endemic areas. *The American Journal of Tropical Medicine and Hygiene* **64**, 28–35 (2001).

69. Chua, C. L. L., Hasang, W., Rogerson, S. J. & Teo, A. Poor Birth Outcomes in Malaria in Pregnancy: Recent Insights Into Mechanisms and Prevention Approaches. *Front. Immunol.* **12**, 621382 (2021).

70. Chua, C. L. L. *et al.* Malaria in Pregnancy: From Placental Infection to Its Abnormal Development and Damage. *Front. Microbiol.* **12**, 777343 (2021).

71. Goel, S. *et al.* RIFINs are adhesins implicated in severe Plasmodium falciparum malaria. *Nat Med* **21**, 314–317 (2015).

72. Cheng, Q. *et al.* stevor and rif are Plasmodium falciparum multicopy gene families which potentially encode variant antigens_. *Molecular and Biochemical Parasitology* (1998).

73. Joannin, N., Abhiman, S., Sonnhammer, E. L. & Wahlgren, M. Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. *BMC Genomics* **9**, 19 (2008).

74. Petter, M. *et al.* Variant proteins of the Plasmodium falciparum RIFIN family show distinct subcellular localization and developmental expression patterns. *Molecular and Biochemical Parasitology* **156**, 51–61 (2007).

75. De Koning-Ward, T. F., Dixon, M. W. A., Tilley, L. & Gilson, P. R. Plasmodium species: master renovators of their host cells. *Nat Rev Microbiol* **14**, 494–507 (2016).

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

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

78. Niang, M. *et al.* STEVOR Is a Plasmodium falciparum Erythrocyte Binding Protein that Mediates Merozoite Invasion and Rosetting. *Cell Host & Microbe* **16**, 81–93 (2014).

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

80. Heddini, A. *et al.* Fresh Isolates from Children with Severe *Plasmodium falciparum* Malaria Bind to Multiple Receptors. *Infect Immun* **69**, 5849–5856 (2001).

81. Sibley, L. D. & Boothroyd, J. C. Virulent strains of Toxoplasma gondii comprise a single clonal lineage. *Nature* **359**, 82–85 (1992).

82. Howe, D. K. & Sibley, L. D. Toxoplasma gondii Comprises Three Clonal Lineages: Correlation of Parasite Genotype with Human Disease. *Journal of Infectious Diseases* **172**, 1561–1566 (1995).

83. Minot, S. *et al.* Admixture and recombination among *Toxoplasma gondii* lineages explain global genome diversity. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13458–13463 (2012).

84. Su, C. *et al.* Globally diverse *Toxoplasma gondii* isolates comprise six major clades originating from a small number of distinct ancestral lineages. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5844–5849 (2012).

85. Galal, L., Hamidović, A., Dardé, M. L. & Mercier, M. Diversity of Toxoplasma gondii strains at the global level and its determinants. *Food and Waterborne Parasitology* **15**, e00052 (2019).

86. Besteiro, S., Dubremetz, J.-F. & Lebrun, M. The moving junction of apicomplexan parasites: a key structure for invasion: The moving junction of apicomplexan parasites. *Cellular Microbiology* **13**, 797–805 (2011).

87. Butcher, B. A. *et al.* Toxoplasma gondii Rhoptry Kinase ROP16 Activates STAT3 and STAT6 Resulting in Cytokine Inhibition and Arginase-1-Dependent Growth Control. *PLoS Pathog* **7**, e1002236 (2011).

88. Ong, Y.-C., Reese, M. L. & Boothroyd, J. C. Toxoplasma Rhoptry Protein 16 (ROP16) Subverts Host Function by Direct Tyrosine Phosphorylation of STAT6. *Journal of Biological Chemistry* **285**, 28731–28740 (2010).

89. Peixoto, L. *et al.* Integrative Genomic Approaches Highlight a Family of Parasite-Specific Kinases that Regulate Host Responses. *Cell Host & Microbe* **8**, 208–218 (2010).

90. Kim, E. W. *et al.* The Rhoptry Pseudokinase ROP54 Modulates Toxoplasma gondii Virulence and Host GBP2 Loading. *mSphere* **1**, e00045-16 (2016).

91. Fentress, S. J. *et al.* Phosphorylation of Immunity-Related GTPases by a Toxoplasma gondii-Secreted Kinase Promotes Macrophage Survival and Virulence. *Cell Host & Microbe* **8**, 484–495 (2010).

92. Niedelman, W. *et al.* The Rhoptry Proteins ROP18 and ROP5 Mediate Toxoplasma gondii Evasion of the Murine, But Not the Human, Interferon-Gamma Response. *PLoS Pathog* **8**, e1002784 (2012).

93. Curt-Varesano, A., Braun, L., Ranquet, C., Hakimi, M.-A. & Bougdour, A. The aspartyl protease TgASP5 mediates the export of the *Toxoplasma* GRA16 and GRA24 effectors into host cells: TgASP5 is essential for GRA16 and GRA24 export. *Cellular Microbiology* **18**, 151–167 (2016).

94. Bougdour, A. *et al.* Host Cell Subversion by Toxoplasma GRA16, an Exported Dense Granule Protein that Targets the Host Cell Nucleus and Alters Gene Expression. *Cell Host & Microbe* **13**, 489–500 (2013).

95. Pellegrini, E. *et al.* Structural Basis for the Subversion of MAP Kinase Signaling by an Intrinsically Disordered Parasite Secreted Agonist. *Structure* **25**, 16–26 (2017).

96. Braun, L. *et al.* A *Toxoplasma* dense granule protein, GRA24, modulates the early immune response to infection by promoting a direct and sustained host p38 MAPK activation. *Journal of Experimental Medicine* **210**, 2071–2086 (2013).

97. Rosenberg, A. & Sibley, L. D. Toxoplasma gondii secreted effectors co-opt host repressor complexes to inhibit necroptosis. *Cell Host & Microbe* **29**, 1186-1198.e8 (2021).

98. Huang, Z. *et al.* The intrinsically disordered protein TgIST from Toxoplasma gondii inhibits STAT1 signaling by blocking cofactor recruitment. *Nat Commun* **13**, 4047 (2022).

99. Gay, G. *et al.* Toxoplasma gondii TgIST co-opts host chromatin repressors dampening STAT1-dependent gene regulation and IFN-γ–mediated host defenses. *Journal of Experimental Medicine* **213**, 1779–1798 (2016).

100. Olias, P., Etheridge, R. D., Zhang, Y., Holtzman, M. J. & Sibley, L. D. Toxoplasma Effector Recruits the Mi-2/NuRD Complex to Repress STAT1 Transcription and Block IFN-γ-Dependent Gene Expression. *Cell Host & Microbe* **20**, 72–82 (2016).

101. Braun, L. *et al.* The Toxoplasma effector TEEGR promotes parasite persistence by modulating NF-κB signalling via EZH2. *Nat Microbiol* **4**, 1208–1220 (2019).

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

103. Watzlowik, M. T., Das, S., Meissner, M. & Längst, G. Peculiarities of Plasmodium falciparum Gene Regulation and Chromatin Structure. *IJMS* **22**, 5168 (2021).

104. Weiner, A. *et al.* 3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite Plasmodium falciparum: Dynamics of nuclear architecture in P. falciparum. *Cellular Microbiology* **13**, 967–977 (2011).

105. Dahan-Pasternak, N. *et al.* PfSec13 is an unusual chromatin associated nucleoporin of *Plasmodium falciparum* , which is essential for parasite proliferation in human erythrocytes. *Journal of Cell Science* jcs.122119 (2013) doi:10.1242/jcs.122119.

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

107. Bunnik, E. M. *et al.* Comparative 3D genome organization in apicomplexan parasites. *Proc Natl Acad Sci USA* **116**, 3183–3192 (2019).

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

109. Sexton, T. *et al.* Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell* **148**, 458–472 (2012).

110. Reers, A. B. *et al.* Histone modification analysis reveals common regulators of gene expression in liver and blood stage merozoites of Plasmodium parasites. *Epigenetics & Chromatin* **16**, 25 (2023).

111. Mancio-Silva, L., Zhang, Q., Scheidig-Benatar, C. & Scherf, A. Clustering of dispersed ribosomal DNA and its role in gene regulation and chromosome-end associations in malaria parasites. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15117–15122 (2010).

112. Epp, C., Li, F., Howitt, C. A., Chookajorn, T. & Deitsch, K. W. 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).

113. Freitas-Junior, L. H. *et al.* Telomeric Heterochromatin Propagation and Histone Acetylation Control Mutually Exclusive Expression of Antigenic Variation Genes in Malaria Parasites. *Cell* **121**, 25–36 (2005).

114. Lopez-Rubio, J.-J., 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* **5**, 179–190 (2009).

115. Galinski, M. R. *et al. Plasmodium knowlesi:* a superb *in vivo* nonhuman primate model of antigenic variation in malaria. *Parasitology* **145**, 85–100 (2018).

116. Kumar, H. & Tolia, N. H. Getting in: The structural biology of malaria invasion. *PLoS Pathog* **15**, e1007943 (2019).

117. Talman, A. M. *et al.* PbGEST mediates malaria transmission to both mosquito and vertebrate host: PbGEST mediates malaria transmission. *Molecular Microbiology* **82**, 462–474 (2011).

118. Poespoprodjo, J. R., Douglas, N. M., Ansong, D., Kho, S. & Anstey, N. M. Malaria. *The Lancet* **402**, 2328–2345 (2023).

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

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

121. Hoeijmakers, W. A. M. *et al.* H 2 A . Z / H 2 B . Z double-variant nucleosomes inhabit the AT -rich promoter regions of the *P lasmodium falciparum* genome. *Molecular Microbiology* **87**, 1061–1073 (2013).

122. 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 *P lasmodium falciparum* . *Molecular Microbiology* **87**, 1167–1182 (2013).

123. Salcedo-Amaya, A. M. *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).

124. Bártfai, R. *et al.* H2A.Z Demarcates Intergenic Regions of the Plasmodium falciparum Epigenome That Are Dynamically Marked by H3K9ac and H3K4me3. *PLoS Pathog* **6**, e1001223 (2010).

125. Sullivan, W. J. Histone H3 and H3.3 Variants in the Protozoan Pathogens *Plasmodium falciparum* and *Toxoplasma gondii*. *DNA Sequence* **14**, 227–231 (2003).

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

127. Westenberger, S. J., Cui, L., Dharia, N., Winzeler, E. & Cui, L. Genome-wide nucleosome mapping of Plasmodium falciparum reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes. *BMC Genomics* **10**, 610 (2009).

128. Ponts, N., Harris, E. Y., Lonardi, S. & Le Roch, K. G. Nucleosome occupancy at transcription start sites in the human malaria parasite: A hard-wired evolution of virulence? *Infection, Genetics and Evolution* **11**, 716–724 (2011).

129. Kensche, P. R. *et al.* The nucleosome landscape of *Plasmodium falciparum* reveals chromatin architecture and dynamics of regulatory sequences. *Nucleic Acids Res* **44**, 2110–2124 (2016).

130. Ponts, N. *et al.* Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Research* **20**, 228–238 (2010).

131. Ruiz, J. L. *et al.* Characterization of the accessible genome in the human malaria parasite Plasmodium falciparum. *Nucleic Acids Research* **46**, 9414–9431 (2018).

132. Toenhake, C. G. *et al.* Chromatin Accessibility-Based Characterization of the Gene Regulatory Network Underlying Plasmodium falciparum Blood-Stage Development. *Cell Host & Microbe* **23**, 557-569.e9 (2018).

133. Dzikowski, R. & Deitsch, K. W. Genetics of antigenic variation in Plasmodium falciparum. *Curr Genet* **55**, 103–110 (2009).

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

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

136. Dzikowski, R., Frank, M. & Deitsch, K. Mutually Exclusive Expression of Virulence Genes by Malaria Parasites Is Regulated Independently of Antigen Production. *PLoS Pathog* **2**, e22 (2006).

137. Lenz, T. *et al.* Epigenetics and chromatin structure regulate var2csa expression and the placental binding phenotype in Plasmodium falciparum. Preprint at https://doi.org/10.7554/eLife.93632.1 (2024).

138. Trelle, M. B., Salcedo-Amaya, A. M., Cohen, A. M., Stunnenberg, H. G. & Jensen, O. N. 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).

139. Coetzee, N. *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).

140. Saraf, A. *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).

141. Fraschka, S. A. *et al.* Comparative Heterochromatin Profiling Reveals Conserved and Unique Epigenome Signatures Linked to Adaptation and Development of Malaria Parasites. *Cell Host & Microbe* **23**, 407-420.e8 (2018).

142. Cui, L. *et al.* PfGCN5-Mediated Histone H3 Acetylation Plays a Key Role in Gene Expression in *Plasmodium falciparum*. *Eukaryot Cell* **6**, 1219–1227 (2007).

143. Read, D. F., Cook, K., Lu, Y. Y., Le Roch, K. G. & Noble, W. S. Predicting gene expression in the human malaria parasite Plasmodium falciparum using histone modification, nucleosome positioning, and 3D localization features. *PLoS Comput Biol* **15**, e1007329 (2019).

144. Muller, I. *et al.* Transcriptome and histone epigenome of Plasmodium vivax salivary-gland sporozoites point to tight regulatory control and mechanisms for liver-stage differentiation in relapsing malaria. *International Journal for Parasitology* **49**, 501–513 (2019).

145. Witmer, K., Fraschka, S. A., Vlachou, D., Bártfai, R. & Christophides, G. K. An epigenetic map of malaria parasite development from host to vector. *Sci Rep* **10**, 6354 (2020).

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

147. Chaal, B. K., Gupta, A. P., Wastuwidyaningtyas, B. D., Luah, Y.-H. & Bozdech, Z. Histone Deacetylases Play a Major Role in the Transcriptional Regulation of the Plasmodium falciparum Life Cycle. *PLoS Pathog* **6**, e1000737 (2010).

148. Cui, L., Fan, Q., Cui, L. & Miao, J. Histone lysine methyltransferases and demethylases in Plasmodium falciparum. *International Journal for Parasitology* **38**, 1083–1097 (2008).

149. Duraisingh, M. T. *et al.* Heterochromatin Silencing and Locus Repositioning Linked to Regulation of Virulence Genes in Plasmodium falciparum. *Cell* **121**, 13–24 (2005).

150. Tonkin, C. J. *et al.* Sir2 Paralogues Cooperate to Regulate Virulence Genes and Antigenic Variation in Plasmodium falciparum. *PLoS Biol* **7**, e1000084 (2009).

151. Ukaegbu, U. E. *et al.* Recruitment of PfSET2 by RNA Polymerase II to Variant Antigen Encoding Loci Contributes to Antigenic Variation in P. falciparum. *PLoS Pathog* **10**, e1003854 (2014).

152. Jiang, L. *et al.* PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. *Nature* **499**, 223–227 (2013).

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

154. Cheon, Y., Kim, H., Park, K., Kim, M. & Lee, D. Dynamic modules of the coactivator SAGA in eukaryotic transcription. *Exp Mol Med* **52**, 991–1003 (2020).

155. Josling, G. A. *et al.* A Plasmodium Falciparum Bromodomain Protein Regulates Invasion Gene Expression. *Cell Host & Microbe* **17**, 741–751 (2015).

156. Santos, J. M. *et al.* Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor. *Cell Host & Microbe* **21**, 731-741.e10 (2017).

157. Batugedara, G. *et al.* The chromatin bound proteome of the human malaria parasite. *Microbial Genomics* **6**, (2020).

158. Wang, H., Dittmer, T. A. & Richards, E. J. Arabidopsis CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. *BMC Plant Biol* **13**, 200 (2013).

159. Glynn, E. F. *et al.* Genome-Wide Mapping of the Cohesin Complex in the Yeast Saccharomyces cerevisiae. *PLoS Biol* **2**, e259 (2004).

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

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

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

163. Coulson, R. M. R., Hall, N. & Ouzounis, C. A. Comparative Genomics of Transcriptional Control in the Human Malaria Parasite *Plasmodium falciparum*. *Genome Res.* **14**, 1548–1554 (2004).

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

165. Gopalakrishnan, A. M., Nyindodo, L. A., Ross Fergus, M. & López-Estraño, C. Plasmodium falciparum: Preinitiation complex occupancy of active and inactive promoters during erythrocytic stage. *Experimental Parasitology* **121**, 46–54 (2009).

166. Rai, R. *et al.* Genome-wide analysis in Plasmodium falciparum reveals early and late phases of RNA polymerase II occupancy during the infectious cycle. *BMC Genomics* **15**, 959 (2014).

167. Lu, X. M. *et al.* Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite Plasmodium falciparum. *Nucleic Acids Research* **45**, 7825–7840 (2017).

168. Shang, X. *et al.* PfAP2-EXP2, an Essential Transcription Factor for the Intraerythrocytic Development of Plasmodium falciparum. *Front. Cell Dev. Biol.* **9**, 782293 (2022).

169. 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 Research* **50**, 3413–3431 (2022).

170. Templeton, T. J. *et al.* Comparative Analysis of Apicomplexa and Genomic Diversity in Eukaryotes. *Genome Res.* **14**, 1686–1695 (2004).

171. Balaji, S. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Research* **33**, 3994–4006 (2005).

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

173. Yuda, M. *et al.* Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. *Molecular Microbiology* **71**, 1402–1414 (2009).

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

175. Modrzynska, K. *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).

176. Zhang, C. *et al.* Systematic CRISPR-Cas9-Mediated Modifications of *Plasmodium yoelii* ApiAP2 Genes Reveal Functional Insights into Parasite Development. *mBio* **8**, e01986-17 (2017).

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

178. MacQuarrie, K. L., Fong, A. P., Morse, R. H. & Tapscott, S. J. Genome-wide transcription factor binding: beyond direct target regulation. *Trends in Genetics* **27**, 141–148 (2011).

179. Ponts, N. *et al.* Genome-wide Mapping of DNA Methylation in the Human Malaria Parasite Plasmodium falciparum. *Cell Host & Microbe* **14**, 696–706 (2013).

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

181. Flueck, C. *et al.* A Major Role for the Plasmodium falciparum ApiAP2 Protein PfSIP2 in Chromosome End Biology. *PLoS Pathog* **6**, e1000784 (2010).

182. Sierra-Miranda, M. *et al.* *Pf*AP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds to *Plasmodium falciparum* telomeres. *Cellular Microbiology* **19**, e12742 (2017).

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

184. Chahine, Z. & Le Roch, K. G. Decrypting the complexity of the human malaria parasite biology through systems biology approaches. *Front. Syst. Biol.* **2**, 940321 (2022).

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

186. Rinn, J. L. & Chang, H. Y. Long Noncoding RNAs: Molecular Modalities to Organismal Functions. *Annu. Rev. Biochem.* **89**, 283–308 (2020).

187. Broadbent, K. M. *et al.* Strand-specific RNA sequencing in Plasmodium falciparum malaria identifies developmentally regulated long non-coding RNA and circular RNA. *BMC Genomics* **16**, 454 (2015).

188. Sierra-Miranda, M. *et al.* Two long non-coding RNAs generated from subtelomeric regions accumulate in a novel perinuclear compartment in Plasmodium falciparum. *Molecular and Biochemical Parasitology* **185**, 36–47 (2012).

189. Broadbent, K. M. *et al.* A global transcriptional analysis of Plasmodium falciparum malaria reveals a novel family of telomere-associated lncRNAs. *Genome Biol* **12**, R56 (2011).

190. Raabe, C. A. *et al.* A global view of the nonprotein-coding transcriptome in Plasmodium falciparum. *Nucleic Acids Research* **38**, 608–617 (2010).

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

192. Batugedara, G. *et al.* Novel insights into the role of long non-coding RNA in the human malaria parasite, Plasmodium falciparum. *Nat Commun* **14**, 5086 (2023).

193. Li, Y., Baptista, R. P. & Kissinger, J. C. Noncoding RNAs in Apicomplexan Parasites: An Update. *Trends in Parasitology* **36**, 835–849 (2020).

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

195. Giorgetti, L. *et al.* Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).

196. Zhang, X. *et al.* A coordinated transcriptional switching network mediates antigenic variation of human malaria parasites. *eLife* **11**, e83840 (2022).

197. Chookajorn, T. *et al.* Epigenetic memory at malaria virulence genes. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 899–902 (2007).

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

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

200. Guizetti, J., Barcons-Simon, A. & Scherf, A. Trans-acting GC-rich non-coding RNA at *var* expression site modulates gene counting in malaria parasite. *Nucleic Acids Res* gkw664 (2016) doi:10.1093/nar/gkw664.

201. Wei, G., Zhao, Y., Zhang, Q. & Pan, W. Dual regulatory effects of non-coding GC-rich elements on the expression of virulence genes in malaria parasites. *Infection, Genetics and Evolution* **36**, 490–499 (2015).

202. 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**, e03054-19 (2020).

203. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, 1665–1680 (2014).

204. Nichols, M. H. & Corces, V. G. A CTCF Code for 3D Genome Architecture. *Cell* **162**, 703–705 (2015).

205. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, (2015).

206. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports* **15**, 2038–2049 (2016).

207. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 996–1001 (2014).

208. Guo, Y. *et al.* CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **162**, 900–910 (2015).

209. Seitan, V. C. *et al.* Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res.* **23**, 2066–2077 (2013).

210. Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320.e24 (2017).

211. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).

212. Vietri Rudan, M. *et al.* Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Reports* **10**, 1297–1309 (2015).

213. Phillips-Cremins, J. E. *et al.* Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. *Cell* **153**, 1281–1295 (2013).

214. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).

215. Bonev, B. *et al.* Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572.e24 (2017).

216. Lupiáñez, D. G. *et al.* Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell* **161**, 1012–1025 (2015).

217. Rajderkar, S. *et al.* Topologically associating domain boundaries are required for normal genome function. *Commun Biol* **6**, 435 (2023).

218. Lieberman-Aiden, E. *et al.* Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**, 289–293 (2009).

219. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. U.S.A.* **115**, (2018).

220. Falk, M. *et al.* Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature* **570**, 395–399 (2019).

221. Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).

222. Cremer, T. & Cremer, M. Chromosome Territories. *Cold Spring Harbor Perspectives in Biology* **2**, a003889–a003889 (2010).

223. Zhang, Y. *et al.* Spatial Organization of the Mouse Genome and Its Role in Recurrent Chromosomal Translocations. *Cell* **148**, 908–921 (2012).

224. Dekker, J. & Mirny, L. The 3D Genome as Moderator of Chromosomal Communication. *Cell* **164**, 1110–1121 (2016).

225. Branco, M. R. & Pombo, A. Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. *PLoS Biol* **4**, e138 (2006).

CHAPTER 1

Epigenetics and chromatin structure regulate *var2csa* expression and the placental binding phenotype in *Plasmodium falciparum*

Todd Lenz[1,†], Madle Sirel[2,†], Hannes Hoppe[2], Sulman Shafeeq[2], Karine Le Roch[1,*] and Ulf Ribacke[2,3,*].

[1] Department of Molecular, Cell and Systems Biology, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA.
[2] Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institutet, Tomtebodavägen 16, SE-17165, Solna, Sweden.
[3] Department of Cell and Molecular Biology (ICM), Uppsala University, Husargatan 3, SE-75237, Uppsala, Sweden.
[†] These authors contributed equally to this work.
* Correspondence: Ulf Ribacke – Tel: +46 704975196; Email: ulf.ribacke@ki.se.
* Correspondence: Karine Le Roch – Tel: +1 951 8275422; Email: karinel@ucr.edu.

A version of this chapter is available in ELife. 07, March 2024. doi.org/10.7554/eLife.93632.1. I performed all computational analyses and generated the majority of the figures presented, with the exception of Figures 1.1A-D. I was also actively involved in the manuscript preparation and editing process.

ABSTRACT

Plasmodium falciparum is responsible for what appears to be a never-ending public health issue in the developing world. With repeated infections, a gradual semi-immunity to severe malaria can be acquired but this is disrupted when women become pregnant as the parasite cytoadheres in the placenta to prevent splenic clearance. This change in tissue tropism is due to specific expression of the antigenically variable adhesin VAR2CSA. To better understand the molecular mechanisms activating *var2csa* and antigenic variation overall, we used a combination of phenotypic and systems biology assays. We first established phenotypically homogenous populations of VAR2CSA expressing and placenta binding parasites that were shown to exclusively transcribe *var2csa* while all other *var* genes remained silenced. We also confirmed that the transcriptional activation was strongly associated with distinct depletion of repressive H3K9me3 marks. Further, we used chromatin conformation capture as a high-resolution approach to determine interchromosomal interactions and established that transcriptional activation is linked to a small yet significant repositioning of *var2csa* relative to heterochromatic telomeric clusters. Lastly, we demonstrated that occupancy of 5-methylcytosine was present in all *var* genes but independent of transcriptional activation and switching. Altogether, these findings provide insights at high resolution into the potential role of 5-methylcytosine in *P. falciparum* and increase our understanding of the mechanisms regulating antigenic variation at the epigenetics and chromatin structure level.

INTRODUCTION

The morbidity and mortality associated with *Plasmodium falciparum* malaria are negatively correlated to the number of infections contracted[1]. As a consequence, young children and other malaria naïve individuals are at greatest risk of developing severe illness and succumb to the disease. With repeated exposure, a gradual immunity to clinical manifestations is acquired and symptomatic episodes in adults are relatively rare in endemic areas[1]. This holds true until women become pregnant, when the built-up immunity is disrupted, and they again become highly susceptible to develop severe disease[2]. Besides affecting the mother, the pregnancy-associated malaria (PAM) also results in adverse outcomes for the fetus, including miscarriages, stillbirths, preterm births and low birth weights, of which the latter two often result in predisposition for morbidities later in life[3-5]. Thus, PAM represents a devastating public health problem that targets the most vulnerable populations in the resource scarce malaria endemic regions.

The onset of PAM is tightly linked to the appearance of the placenta, which presents a new niche for parasitized red blood cells (pRBCs) to sequester and thrive. To cytoadhere in the microvasculature, the parasite employs antigenically variable adhesins from the *var* gene encoded *P. falciparum* Erythrocyte Membrane Protein 1 (PfEMP1) protein family, a key virulence feature of the parasite directly linked to the pathogenesis of severe disease[6, 7]. While harmful for the human host, cytoadhesion provides the parasite immune evasive opportunities, such as the shielding of parasite antigens expressed on pRBC surfaces by other cells and evasion of splenic clearance[8]. The PfEMP1 repertoire is large, with approximately 60 *var* genes per parasite genome from which only one is assumed to be expressed at a time, a process referred to as mutually exclusive expression. Transcriptional switching between different *var* loci creates antigenic variation, which aids in immune evasion and binding of a plethora of human receptors and variable tissue tropism for the parasite[8-10]. Between parasites, genetic variation among the majority of *var* genes is vast, with one important exception, the remarkably well-conserved *var2csa.* This gene encodes the PfEMP1 VAR2CSA that mediates cytoadherence to chondroitin sulfate A (CSA) on syncytiotrophoblasts in the placenta and is considered the culprit behind PAM[11, 12].

The variable expression of PfEMP1 is thought to be accomplished through several layers of regulatory activities to ensure that only one, or very few, *var* genes are translated to functional adhesins and exported to the pRBC surface[9, 10, 13-15]. Antigenic variation is mainly achieved through transcriptional regulation, where altered DNA accessibility is orchestrated through variations in nucleosome composition, histone modifications and organization of chromatin into active or repressive clusters. In addition, non-coding RNAs and specific transcription factors have been proposed to compose additional layers of regulation that govern repression, activation and switching of expression[15]. Lastly, the peculiar VAR2CSA appears to also be regulated on a translational level[16-18].

While major advances have been made recently to better understand gene expression of these important virulence factors, mechanisms regulating these events at the molecular level are complex and many questions remain. An improved understanding of the detailed molecular mechanisms that mediate parasite adhesion is vital if we want to grasp the molecular factors that control malaria pathology and to identify new therapeutic strategies. Here, using a combination of advanced phenotypic assays and systems biology approaches including transcriptomics, epigenetics and chromatin structure features, we provide a detailed view in high-resolution on how expression of the PAM mediating *var2csa* is regulated at transcriptional initiation level. Intriguingly, we also show that occupancy of the epigenetic mark 5-methylcytosine (5mC) is significant in antigenically variable genes but appears to be disconnected to transcriptional activation and switching.

RESULTS

Repeated *in vitro* selection on CSA results in a highly homogenous population of VAR2CSA expressing and placenta binding parasites

P. falciparum pRBCs have different receptor preferences for cytoadherence to diverse host cells. This has been previously established *in vitro* by panning on antibody-coated beads or either cell-bound or soluble human receptors[13, 35, 36]. While several parasite molecules have been identified as ligands for cytoadherence, the antigenically variant PfEMP1 proteins, encoded by the *var* genes, have been demonstrated as the genuine adherence molecules. In PAM, VAR2CSA encoded by a subfamily of the *var* genes, plays a vital role for the cytoadherence of pRBCs to the CSA expressed on the surface of placental syncytiotrophoblasts. To better understand the molecular mechanisms regulating *var2csa* expression, we first employed an approach that entailed several rounds of repeated panning of NF54 pRBCs on chondroitin sulfate A (CSA) coated plates[13] to select for parasites with homogenous expression of VAR2CSA and a placenta binding phenotype. Prior to the panning procedure, we confirmed by scanning electron microscopy (SEM) the presence of knob-structures, which are protrusions where PfEMP1s are surface exposed on pRBCs in patient isolates but are easily lost upon long-term *in vitro* cultivation **(Fig. 1.1A)**. Several phenotypic analyses were used to confirm the accuracy of our selection process. We analyzed pRBC surface expression of VAR2CSA by flow cytometry-based antibody recognition throughout the trajectory of repeated panning and revealed a gradual transition from a phenotypically heterogeneous to homogeneous parasite population **(Fig. 1.1B)**. In addition, we interrogated the cytoadhesive potential using placental sections from malaria non-immune donors[21]. For the repeatedly panned parasites, substantial numbers of pRBCs were observed binding to the placental syncytiotrophoblasts **(Fig. 1.1C)**. We were able to quantify bound pRBCs per mm^2 using a VAR2CSA surface negative PTEF knockout parasite[18] as control for background binding and revealed an approximate 6-fold higher binding of the CSA-selected parasite line **(Fig. 1.1D)**. Thus, the repeatedly panned NF54CSA (from here on referred to as NF54CSAh) was deemed phenotypically highly homogenous with a VAR2CSA/PAM relevant cytoadherence profile and considered suitable for down-stream analyses of *var2csa* expression and regulation.

The *P. falciparum* placental binding phenotype is strictly linked to expression of *var2csa*

Next, we sought to decipher any transcriptomic differences responsible for the phenotypic disparity observed between the unselected NF54 and NF54CSAh. As PfEMP1 encoding *var* genes are actively transcribed during the first third of the parasite's 48h life cycle (also referred to as the ring-stage)[37, 38], we harvested NF54 and NF54CSAh from three independent selections at 16 ± 2h post invasion (p.i.) of RBCs for RNA-seq. We also collected parallel samples for ChIP-seq, MeDIP-seq and Hi-C.

The RNA-seq analysis revealed differential expression in only 64 of the 5285 protein coding genes of the parasite **(Supp. Data 1.1)**. Hierarchical clustering and principal component analysis performed on the complete set of data revealed little deviation ($\rho > 0.99$) between both replicates and samples when accounting for genome-wide transcription **(Supp. Fig. 1.1A-B)**. We however observed a very strong correlation between replicates of NF54 ($\rho > 0.97$) and between replicates of NF54CSAh ($\rho > 0.63$) when only taking *var* genes into account **(Fig. 1.1E)**, confirming a high degree of phenotypic maintenance among collected samples. In line with this, we observed a near exclusive representation of *var* genes among differentially expressed genes **(Fig. 1.1F)**, with *var2csa* being the only upregulated and almost all other of the 60 *var* genes downregulated in NF54CSAh compared to NF54 **(Fig. 1.1F-G)**. Besides the confirmation of *var2csa* expression being closely associated to the placental binding phenotype of malaria parasites, the high degree of phenotypic homogeneity of NF54CSAh and the extensive heterogeneity of NF54 provided an excellent opportunity for down-stream analyses of gene regulation in antigenic variation via epigenetic mechanisms and chromatin structure.

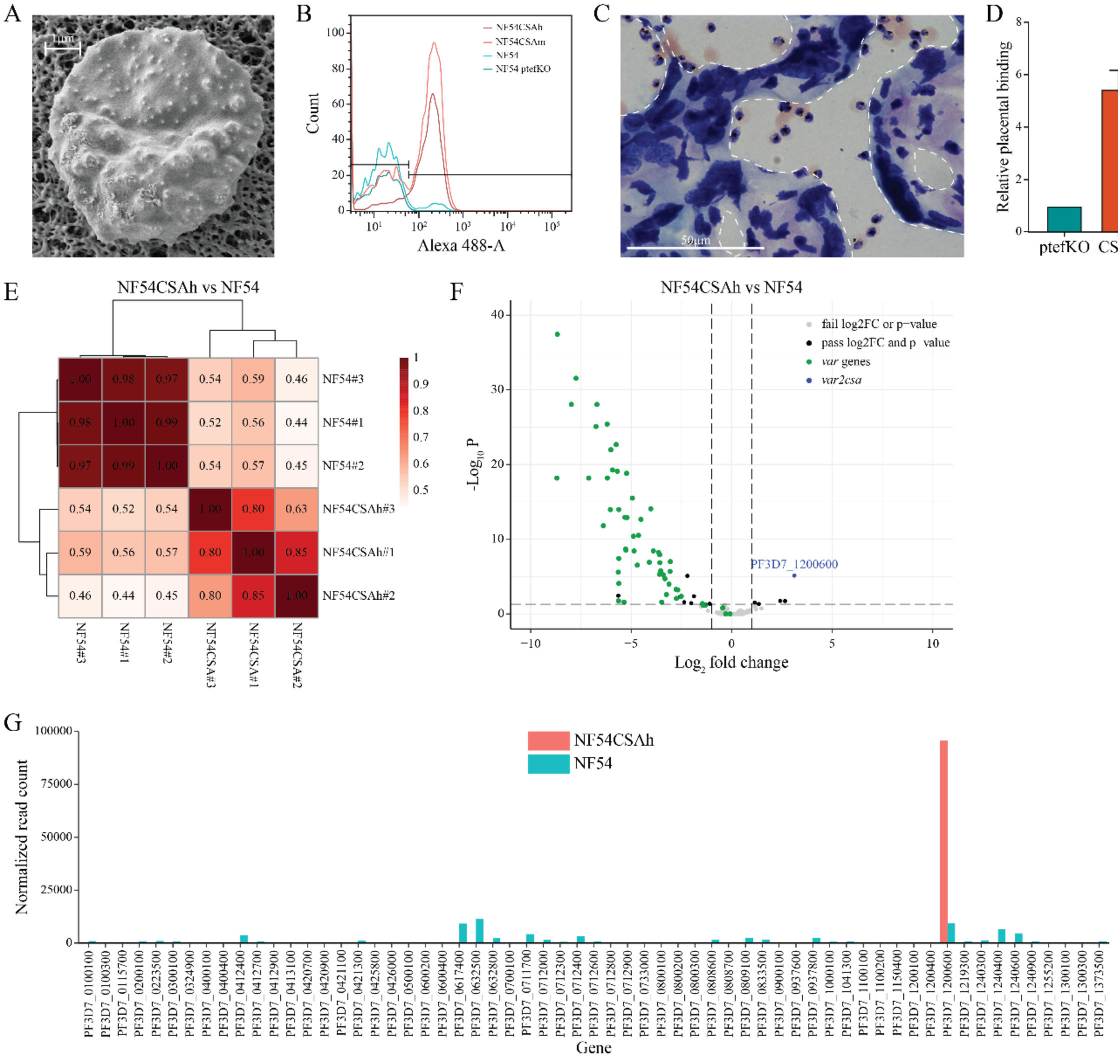

Figure 1.1. Generation of parasites with a homogenous placental binding phenotype. (A) The presence of *in vivo* knob-structures on pRBCs was confirmed by SEM prior to phenotypic selection. Shown is an image of a pRBC with knobs appearing as white protrusions. **(B)** Flow cytometry-based antibody recognition of VAR2CSA surface expression on repeatedly selected pRBCs. NF54 ptefKO served as negative control, NF54 was the unselected parasite line, NF54CSAm was an intermediately selected line and NF54CSAh was panned to achieve a homogenous population of VAR2CSA expressing pRBCs. **(C)** Giemsa-stained placental section with bound NF54CSAh pRBCs. Tissue boundaries are indicated by white dashed lines. **(D)** Quantification of relative placental binding of NF54CSAh from three biological replicates. pRBCs bound per mm^2 were normalized to the VAR2CSA negative NF54 ptefKO parasite line. **(E)** Spearman correlation of *var* gene expression profiles determined by RNA-seq for triplicates of NF54CSAh and the unselected NF54. **(F)** Differential gene expression between NF54CSAh and NF54 identified a limited number of significant genes (FDR <0.05, Log_2 fold change >1) of which the majority were *var* genes (*var2csa* in blue, all other *var* genes in green). **(G)** Normalized read counts for all *var* genes revealed NF54 to be phenotypically highly heterogeneous whereas *var2csa* was the only *var* gene expressed by NF54CSAh.

Transcriptional activity of *var2csa* is associated to a near complete lack of H3K9me3 occupancy

Besides its known role in repressing repetitive elements and gene depleted regions in eukaryotes, the histone H3 lysine 9 trimethylation mark (H3K9me3) has been demonstrated to be an important player in heterochromatin formation, to silence lineage-inappropriate genes and control cell fate. This appears particularly true for *Plasmodium spp.*, where H3K9me3-marks have been identified as potentially critical in regulating parasite specific genes involved in pathogenicity and sexual commitment[39, 40]. We therefore utilized chromatin ChIP-seq as previously described[41] to assess the distribution of H3K9me3 and examine whether NF54CSAh displays differentially bound heterochromatin/euchromatin marks due to elevated expression of *var2csa*. We performed ChIP-seq experiments in triplicates using NF54CSAh and NF54 samples collected at the same time as the samples used for the RNA-seq experiments described above. ChIP-seq libraries were sequenced, processed and mapped to the genome. Following input normalization, we observed a clear targeted enrichment of H3K9me3 within both subtelomeric and internal *var* gene clusters as previously described[42] **(Fig. 1.2A)**. H3K9me3 occupancy within the gene body of *var2csa* of NF54CSAh was almost entirely eliminated, differentiating it from the NF54 control which shows little variation from other *var* genes, which goes in line with its highly heterogenous *var* phenotype and consequently small contribution to H3K9me3 for individual genes on a population level **(Fig. 1.2A-B)**. Due to highly homologous sequences preventing unique mapping, several *var* genes display no H3K9me3 coverage **(Fig. 1.2B)**.

This observation was confirmed by performing MACS3 broad peak calling of input normalized H3K9me3 enrichment with results showing an average of 2234 significant consensus peaks ($q < 0.05$) per sample and approximately 79% of peaks within the coding sequence (CDS) of protein coding genes **(Supp. Data 1.2)**. Differential binding analysis via DiffBind showed strong correlation ($r > 0.75$) between replicates and samples, with only 4 differentially bound sites ($FDR < 0.05$) **(Supp. Fig. 1.2C-D, Supp. Fig. 1.3)**, all within the CDS or just upstream of the transcription start site (TSS) of *var2csa* **(Supp. Data 1.2)**. Given that *var2csa* is significantly downregulated in NF54 these results demonstrate the correlation between loss of *var2csa* expression and H3K9me3 mediated transcriptional silencing.

P. falciparum variant antigen gene families such as *var, rif, stevor and pfmc-2tm* cluster within subtelomeric regions of most chromosomes and enrichment of H3K9me3 within the gene body is similarly linked to their repression[43-45]. NF54CSAh and NF54 both display this same pattern of elevated H3K9me3 within the exons of these multicopy gene families, and after normalizing each region (5' UTR, exons, introns and intergenic) the mean H3K9me3 within each is higher than genes outside these families as well as all intergenic regions **(Fig. 1.2C-D)**.

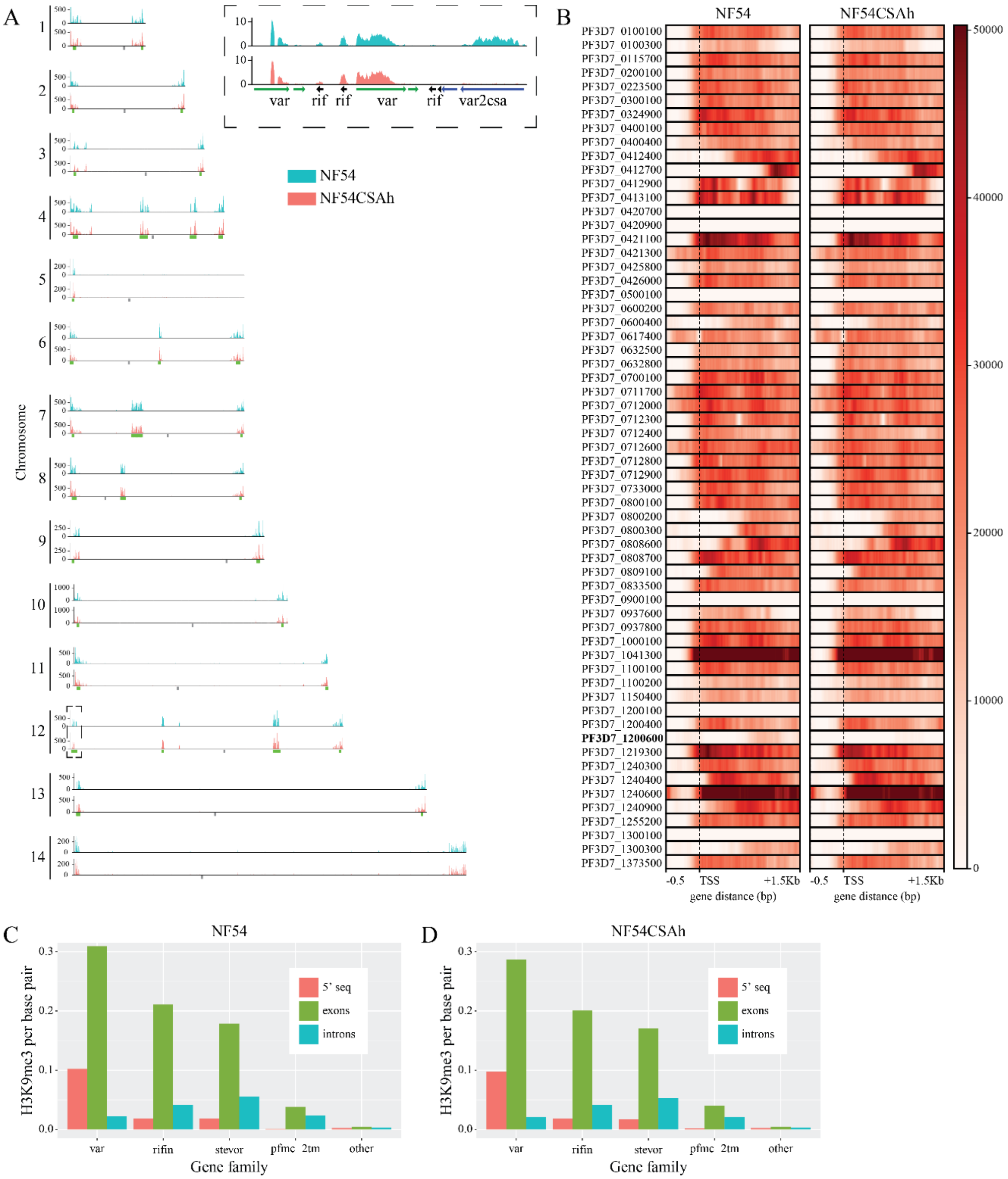

Figure 1.2. The *var2csa* locus is devoid of repressive H3K9me3 in placental binding parasites. (A) Chromosomal distribution of repressive H3K9me3 as determined by ChIP-seq from merged biological triplicates of NF54 and NF54CSAh. *var* gene clusters (green) and centromeres (gray) are annotated. The only notable difference between the parasite lines was at the *var2csa* locus (insert). **(B)** Heatmaps of H3K9me3 occupancy in NF54 and NF54CSAh for all *var* genes (0.5 kb 5' of the TSS and 1.5kb into the gene body) with *var2csa* displayed with bolded gene ID. **(C and D)** H3K9me3 enrichment in regions of multicopy gene families *var*, *rif*, *stevor* and *pfmc-2tm*, and the rest of the genome, in NF54 and NF54CSAh.

Increased interchromosomal interaction and perinuclear repositioning is involved in silencing of *var* genes and activation of *var2csa* respectively

To further investigate the effect chromatin organization and accessibility has on transcription of highly variant multi-copy gene families such as *var* and *rif*, we performed Hi-C experiments on tightly synchronized trophozoites to ensure peak expression of the target genes. Hi-C libraries of three biological replicates for each sample (NF54 and NF54CSAh) were prepared as previously described[41, 46] and sequenced to a mean depth of ~155 million reads per replicate. The libraries were processed (aligning, pairing, mapping, and quality filtering) using HiC-Pro[30] and resulted in a mean of ~44 million valid interaction pairs per replicate. Due to the *P. falciparum* genome size and frequency of MboI cut sites, we elected to bin our reads at a 10-kb resolution to identify intrachromosomal and interchromosomal interactions.

A high stratum adjusted correlation coefficient (SCC ≈ 0.87 to 0.94) suggests that chromatin structures were highly similar between replicates and samples **(Supp. Fig. 1.4A)**. We therefore combined the biological replicates for downstream analyses. Due to variation in sequencing depth between the merged NF54 and NF54CSAh samples, random sampling was performed on NF54CSAh to obtain ~100 million consensus reads for comparative and differential analyses. Heatmaps generated from the ICED normalized matrices show patterns similar to previous studies[41, 47, 48], with a negative log-linear relationship between contact probability and genomic distance demonstrating that our experiment worked as expected **(Supp. Fig. 1.4B)**. Most intrachromosomal interactions occur at a distance less than 10% the total length of each chromosome; with heterochromatin clustering occurring in telomeric regions and internal *var* gene clusters at a higher frequency than the other distant regions **(Fig. 1.3A-C, Supp. Fig. 1.5)**.

To evaluate the correlation between gene expression and chromatin architecture, we used Selfish[32] to identify differential intrachromosomal and interchromosomal interactions. Although there were slight variations in the number of intrachromosomal interactions across most regions, the consistent pattern that emerged was the increased number of interactions between subtelomeric regions on most chromosomes ($p < 0.05$, $\log_2FC > 1$) in NF54CSAh over NF54, indicating tighter heterochromatin control of *var* gene regions **(Fig. 1.3A-C, Supp. Fig. 1.7)**. The subtelomeric regions on chromosomes 2, 3, 4, and 10 showed the largest increase in interaction frequency and all of the *var* genes within those regions were significantly down-regulated in the RNA-seq data generated in this study. Most interesting is perhaps the mix of increased and decreased interactions between subtelomeric regions on chromosome 12, due to the proximity of *var2csa* to other nearby *var* genes within that region confirming a potential role of the chromatin 3D structure in gene regulation.

Additional analysis of these small-scale structural changes was performed by 3D chromatin modeling using the Poisson-based algorithm, PASTIS[33]. Modeling shows co-localization of centromeres and telomeric clustering in distinct regions within the nucleus **(Fig. 1.3D-E)**. The overall 3D structure of the genome was similar between NF54 and NF54CSAh. Using the coordinates output from PASTIS we computed the distance between various bins/regions and found that the mean distance between telomeres containing *var* genes decreased by 6% in NF54CSAh, indicating overall telomeric compaction. What is more significant is that although there is a 52% increase in distance of *var2csa* to the nearest bin in spatial proximity and 19% greater distance from the telomeric cluster, the *var* gene nearest to the end of chromosome 12 only shows a 3% increase in distance from the telomeric cluster. This stark difference in distances for two regions only separated by 20 kb shows that small and localized changes in the chromatin structure is enough to allow transcriptional activation of *var2csa*.

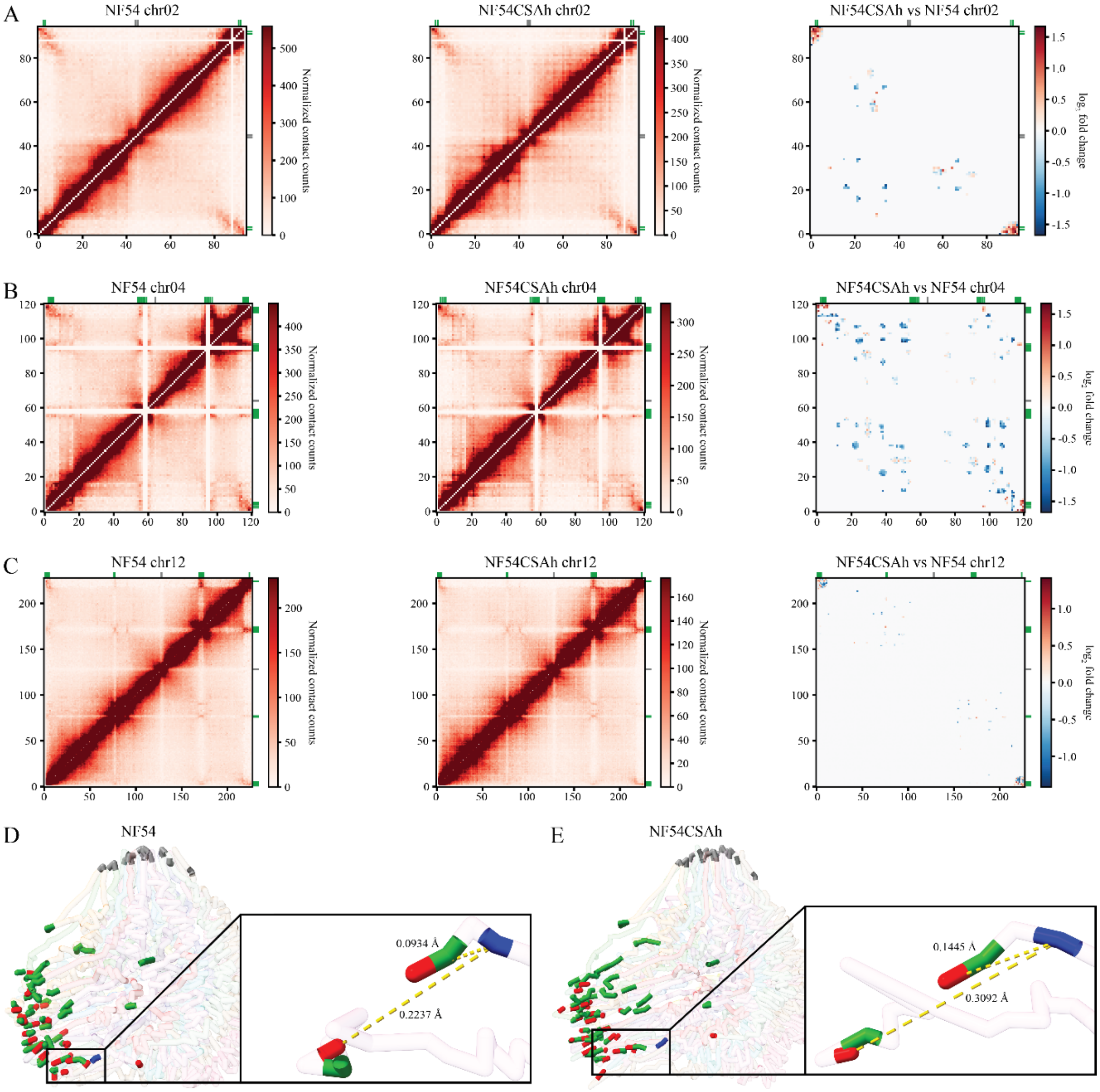

Figure 1.3. Changes in chromatin organization and perinuclear repositioning upon *var2csa* activation and silencing. (A-C) Hi-C normalized contact count heatmaps representing intrachromosomal interactions for chromosomes 2, 4 and 12 in NF54 (left) and NF54CSAh (middle). Differential intrachromosomal interactions between the two parasite lines (right) displayed a general increase in interactions between subtelomeric and internal *var* gene clusters for NF54CSAh compared to NF54. Chromosome 12 stood out as the only exception with decreased interactions in NF54CSAh for the subtelomere containing the *var2csa* locus. **(D)** 3D chromatin modeling for NF54 displays a polarized nucleus with clustering of centromeres (gray) and telomeres (red) in distinct regions. The majority of *var* genes (green), including *var2csa* (blue) and internal *var* genes, are located in close proximity to the telomeric cluster. **(E)** The 3D chromatin model for NF54CSAh displayed a similar topology to NF54 with the exception of decreased distance between *var* genes (green) and telomere ends (red) overall, and an increased distance of *var2csa* (blue) from the telomeric cluster.

Distribution of DNA methylation influences gene expression overall but does not mediate transcriptional activation and switching in antigenic variation

DNA methylation of cytosine residues serves as yet another epigenetic transcriptional antagonist and is found in many model eukaryotic organisms; however, the context and level of methylation differs among plants and animals[49-52]. In higher eukaryotes, while DNA methylation of promoter sequences has been shown to be a repressive epigenetic mark that down-regulates gene expression, DNA methylation is more prevalent within gene-bodies and positively correlated with gene expression levels[53]. Recent studies into *P. falciparum* erythrocytic stages have identified the possible presence of low levels of 5-methylcytosine (5mC) and under-characterized 5-hydroxymethylcytosine (5hmC)-like marks throughout the genome[54-56]. Presence of low levels of 5mC has also been detected in sporozoites and liver stage forms in *P. vivax*[57]. Due to the GC-poor nature of the *P. falciparum* genome, and the consequential low levels of 5-methylcytosine (5mC) coverage, there remains uncertainty as to what level of transcriptional control is conferred through DNA methylation[54]. Recent findings suggest perturbation of levels of the DNA methyltransferase PfDNMT2 to have significant impact on gene expression and cell proliferation[58] but any potential role in antigenic variation and cytoadhesive phenotypes remains to be elucidated.

To deduce the potential gene regulatory relevance of different DNA methylation marks, we first analyzed 5mC and 5hmC levels in NF54CSAh using MeDIP-seq, which allows for differentiation of 5mC from 5hmC due to antibody specificity, compared to prior efforts that have used bisulfite sequencing as the method of choice[55]. We also investigated the presence of 6-methyladenine (6mA), a common DNA modification in prokaryotes that has only recently been studied in humans and other eukaryotic organisms[59-63].

Mapping and peak calling of genome-wide 5mC coverage shows ~94% of a mean 6876 significant consensus peaks mapping to genomic coordinates within the CDS of protein coding genes **(Fig. 1.4A, Supp. Data 1.3)**. 5hmC and 6mA coverages were detected at considerably lower levels with only 269 and 0 significant peaks, respectively **(Supp. Fig. 1.8, Supp. Data 1.3)**. We therefore considered that 6mA modifications are most likely absent from the *P. falciparum* genome. 5hmC distribution across the genome does not preferentially map to specific gene families or regions within each gene **(Supp. Fig. 1.8B)**. It is therefore likely that the weak 5hmC signal detected across the genome could be attributed to background noise or 5hmC coverage is non-specific and serves a more generalized function rather than being key to transcriptional regulation.

Due to the more abundant and distinct distribution of 5mC, we then examined the potential correlation between 5mC coverage and gene expression. Because no significant difference in exon coverage was detected between NF54CSAh and NF54 in *var2csa* or any other *var* gene **(Fig. 1.4A)**, we investigated the difference in distribution of 5mC marks within high versus lowly expressed genes. We detected a negative correlation for 5mC marks in the promoter of highly expressed genes, whereas 5mC across the rest of the gene body displays the opposite correlation **(Fig. 1.4B)**. Functional differentiation of methylated cytosines within these two distinct regions has already been observed in *P. falciparum* and in other eukaryotic organisms[49, 52, 54] but not in the context of antigenic variation. Over the years, genome wide DNA methylation studies have confirmed this "DNA methylation paradox" and demonstrated that methylation in exons is found in constitutively expressed genes and has an important role in maintaining transcriptional fidelity by preventing spurious transcription initiation and can affect histone modification and alternative splicing[64]. We therefore investigated the DNA methylation pattern across the exon/intron junction **(Fig. 1.4C)**. Altogether these results confirm the presence of 5mC in the parasite genome with a higher level within gene bodies and in genes that are often marked by H3K9me3. While additional experiments will be needed to further validate the role of DNA methylation in *P. falciparum*, this epigenetic mark does not seem to directly control transcriptional activation and switching of *var2csa* expression but is likely essential in maintaining transcriptional elongation and fidelity, guiding histone modification enzymes and controlling splicing.

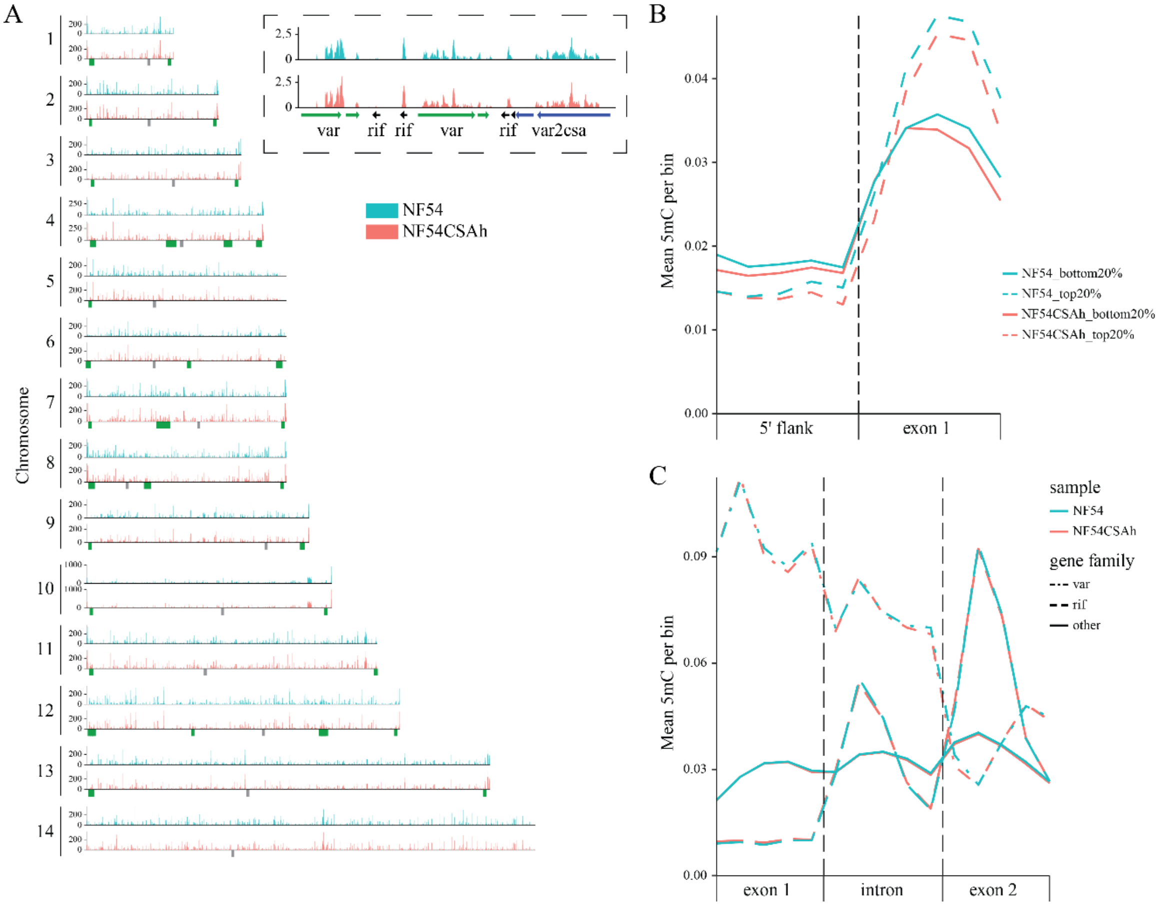

Figure 1.4. Regional differences in 5mC occupancy is dissociated from *var2csa* regulation. (A) Genome-wide distribution of input normalized MeDIP-seq counts for 5mC in NF54 and NF54CSAh reveal a high level of concordance between the analyzed parasite lines. Regional clusters of antigenically variable genes are denoted in green below each plot with an area of chromosome 12 marked by a dashed box and a zoomed-in insert displaying minor differences in 5mC occupancy in the *var2csa* locus between NF54 and NF54CSAh (*var2csa* in blue, other *var* genes in green and *rif* in black). **(B)** Levels of 5mC marks binned in 5' flanks and first exon for the top 20% highest expressed genes and 20% lowest expressed genes in NF54 and NF54CSAh. **(C)** Distribution of 5mC across exon-intron boundaries for the 5mC enriched *var*, *rif* and other two-exonic genes.

DISCUSSION

The complex and highly variable repertoire of *var* genes constitutes an essential tool for *P. falciparum* pRBCs to express different PfEMP1 proteins in order to cytoadhere to vascular endothelium and avoid the host immune system. The ability to cytoadhere and thereby withdraw from the circulation is crucial for the highly rearranged, rigid and immunogenic pRBCs to avoid splenic clearance and for the intracellular parasites to survive[65, 66]. Although the prevention of splenic clearance helps the parasite to escape the immune defense of the host, the exposure of the cytoadhesive proteins presents another issue, namely the immune recognition of the adhesins themselves. Therefore, varied expression of these surface proteins through antigenic variation is a necessity, both for the sake of immune escape but also for altered receptor preference and tissue tropism.

Excessive cytoadherence of *P. falciparum* in the microvasculature of various organs has been strongly associated with malaria disease severity and death. Perhaps the most striking example is the VAR2CSA-mediated sequestration of pRBCs to syncytiotrophoblasts in the placental intervillous space that render previously semi-immune women at risk of severe malaria while pregnant[12]. The causative *var2csa* gene is unique from other *var* genes in that it is relatively conserved between parasite genomes and is regulated at both a transcriptional and translational level. Parasites cultured continuously *in vitro* often modify their patterns of *var* gene expression by transcribing *var2csa* without translating the mRNA, as this additional layer of regulation presents the parasites with energy saving and competitive advantages in the absence of the need to cytoadhere for its survival[16, 17]. Similarly, the complex regulation of the gene allows for the parasites to not express VAR2CSA until infecting a pregnant woman and being presented with the new placental niche.

In the work presented here, we were adamant about analyzing the regulation of *var2csa* in parasites with an appropriate *in vivo* phenotype instead of parasites with *in vitro* artifactual and non-productive transcription of the gene. Therefore, several downstream analyses were performed post phenotypic selection, where we were able to confirm the placental binding parasites with abundant surface expression of VAR2CSA. In concordance with the adhesive phenotype observed, the selected parasite NF54CSAh showed exclusive and abundant transcription of *var2csa*, which was in stark contrast to the original line which displayed transcript evidence from a vast number of *var* genes. These parasite lines therefore presented an excellent opportunity to scrutinize the molecular requirements for *var2csa* selection and activation on a population level, similar to what would occur for isogenic parasites in the *in vivo* context of PAM.

Several studies have previously and conclusively linked gene activation in *P. falciparum* to alterations in DNA accessibility through the change from heterochromatic to euchromatic state[40-44]. In line with our transcriptomic findings, which revealed very limited and *var* exclusive differences, we observed a remarkable similarity in global H3K9me3 occupancy between the two analyzed parasite lines. We found H3K9me3 primarily in antigenic variable gene contexts with over-representation in *var*, *rif* and *stevor*, which highlights the high degree of heterochromatic gene silencing required for these genes. Differential binding analysis revealed only four significant differences between the analyzed lines all centered within or around *var2csa*, with a near absence in H3K9me3 occupancy for NF54CSAh. This clearly confirms the pivotal role of epigenetic regulation through histone modifications in the establishment of the placental binding phenotype in *P. falciparum*.

Besides histone modifications, other molecular events are known to be involved in the transition of a gene from a heterochromatic to euchromatic state in eukaryotes. One such factor is the macromolecular events that involve physical repositioning of genes within the nuclear context to allow for their activation or repression[67]. In *P. falciparum*, nuclear repositioning has previously been suggested to be involved in regulating the activation of the gametocyte-specific transcription factor locus, pfap2-g, in the transition from the asexual to the early sexual stages in the parasite life cycle progression[41]. Nuclear repositioning has also been suggested in antigenic variation[68-71], but its importance has remained somewhat unclear. This is partially due to how well the nuclear context of the parasite is resolved upon interrogation. Previous studies have most often relied on fluorescent *in situ* hybridizations (FISH) to determine the proximity of genes of interest to a single marker of heterochromatin, such as a subtelomeric repetitive DNA sequence. Considering the small size of the *P. falciparum* nucleus, the naturally close proximity of many *var* genes to these subtelomeric repeats and the limited resolution of fluorescent probe approaches, this is a non-trivial task even in the case of robustly selected parasites with discrete *var* phenotypes. This indeed applies to the subtelomerically located *var2csa* for which FISH based endeavors have been inconclusive[13, 68]. Our current Hi-C experiments, performed on parasites with vastly different *var* phenotypes proved able to provide the resolution needed to resolve this matter. We observed a significant decrease in differential interchromosomal interactions for NF54CSAh in the subtelomere of chromosome 12 where *var2csa* is located, whereas interactions around other silent *var* genes were increased. Upon 3D structural modeling it became evident that the genome of NF54CSAh displays an overall compaction of telomere ends and *var* gene-containing intrachromosomal regions, which is in line with the observed heterochromatic silencing of *var* genes. On the contrary, *var2csa* was distanced from the nearest telomeric cluster to a transcription permissive area of the nucleus. This conclusively suggests that, besides alteration of epigenetic marks, nuclear repositioning and small and localized changes in the chromatin structure are indeed required for activation of *var2csa* transcription.

Lastly, we investigated the differences in DNA methylation between NF54CSAh and NF54. DNA methylation exists in most eukaryotic organisms, including plants, and engages in various physiological processes and has a complex correlation with gene expression. While the role of promoter hypermethylation in repressing gene transcription has been well documented, emerging evidence indicates that gene body methylation can promote transcription in up to 60% of genes in some species. While the presence of low levels of 5mC and 5hmC marks has been identified throughout the *P. falciparum*[54, 55] and *P. vivax*[57] genomes, the function and the molecular mechanism underlying its regulation has remained obscure. However, recent studies where perturbations of the methylation process by either small molecules[56, 57] or reverse genetics[58] suggest essentiality and a role in gene regulation. Here, using MeDIP-seq, we show that DNA methylation is mostly absent from the promoter regions of most *var* genes but is high throughout the CDS.

Although there is no clear indication that 5mC regulates the transcriptional activation of *var2csa*, mechanistically, it may be a required modification to guide and stabilize RNA polymerases as well as histone modifying enzymes, and ultimately control transcription at the initiation or elongation level. While the presence and function of gene body methylation in some species remains controversial, a growing body of evidence suggests that its effects, while minor, may be shaped by natural selection[72] and correlated with fitness[73]. A high level of DNA methylation in genes known to be involved in antigenic variation in malaria parasites may be crucial for parasite adaptation and survival *in vivo* and deserve additional investigation. Collectively our data confirm a strong association between epigenetics, genome organization and gene expression in regulating parasite transcripts involved in malaria pathogenicity but also what appears as a fundamentally different role for DNA methylation through 5mC.

MATERIALS AND METHODS

Parasites and in vitro culturing

The *P. falciparum* strains NF54 and NF54CSA-ptefKO[18] were cultivated according to standard methods[19]. All parasite cultures were maintained in blood group O RBCs at 4% hematocrit in RPMI-1640 medium (Gibco) containing 27 nM $NaHCO_3$ (Sigma), 2mM L-glutamine (Hyclone) and 2.5 µg/mL gentamicin (Gibco) and supplemented with 10% human A+ serum. The microaerophilic environment was kept constant by gassing of culture flasks with a mixture of 90% N_2, 5% O_2 and 5% CO_2 and cultures were kept in suspension by orbital shaking at 37°C. Parasite cultures were regularly synchronized by treatments with 5% sorbitol as described previously[20].

Scanning electron microscopy

To confirm proper pRBC surface morphology prior to the establishment of cytoadhesive phenotype, scanning electron microscopy was used. Magnetically enriched pRBCs were fixed in 2.5% glutaraldehyde (Polysciences), 1% paraformaldehyde (Santa Cruz) in PBS for 1 h, rotating at 4ºC. The samples were washed twice with cold PBS. Following the fixation, specimens were adhered onto alcian blue pre-treated 0.45mm Supor® PES membrane filter (Pall Corp.) and washed with 0.1M phosphate buffer pH7.4 followed by MilliQ water. The membranes were then subjected to stepwise ethanol dehydration, transferred to acetone and critical-point-dried using carbon dioxide (Leica EM CPD030). The membranes were finally mounted on specimen stubs using carbon adhesive tabs and sputter coated with a 10nm layer of platinum (Quorum Q150T ES). SEM images were acquired using an Ultra 55 field emission scanning electron microscope (Zeiss) at 3kV and the SE2 detector.

Establishment and preservation of parasite phenotypes

The chondroitin sulphate A (CSA) and placental binding phenotype of NF54CSA was maintained by bi-weekly panning on CSA-coated plastic plates with minor modifications from Brolin et al[13]. Briefly, 100 µg/mL CSA from bovine trachea (Sigma) in phosphate-buffered saline (PBS) was coated on plates overnight leaving a fraction of the plate without CSA-coating. Thereafter, to prevent non-specific binding, plates were blocked with 2% bovine serum albumin (BSA, fraction V, HyClone) in PBS for 1 h. Mature trophozoites were purified using a MACS magnetic cell sorter (Miltenyi BioTec) and resuspended in RPMI-1640 with 10% human serum. Parasites were added to CSA-coated plates and incubated for 1 h at 37°C in microaerophilic conditions with occasional gentle swirling. Plates were thereafter washed with RPMI-1640 until background binding was low, as confirmed by comparison with the non-CSA coated area. Thereafter, the bound pRBCs were recovered from the plates and returned to culture. It is well known that parasites selected to bind to CSA (and other cytoadherence receptors) revert back to a non-CSA binding state over time. With the selected panning interval, minor reversion was noted. Prior to any down-stream analyses, phenotypes were nevertheless confirmed by flow cytometry (see below)

Flow cytometric analysis of pRBC surface levels of VAR2CSA

For pRBC surface staining of VAR2CSA, cultures with late-stage pRBCs (24-40 hpi) were blocked in PBS with 2% bovine serum albumin (BSA, fraction V, HyClone) for 30 min at room temperature. Primary antibody against the VAR2CSA (goat anti-DBL1-6) was added to blocked cultures at a concentration of 100 µg/mL in PBS with 2% BSA and incubated for 1 h at room temperature. Non-immune goat IgG (Jackson ImmunoResearch) was used for non-specific binding control. Cells were washed twice with 90 µL of PBS before incubated with a secondary rabbit anti-goat antibody coupled to Alexa647 (Invitrogen) at 1:100 dilution, 5 µg/mL Dihydroethidium (DHE) (Invitrogen) and 10 µg/mL Hoechst 33342 (Invitrogen) in PBS for 1 h at room temperature. Unbound secondary antibody was washed away by 2 x 90 µL washes with PBS before cells were resuspended in PBS to a final hematocrit of 0.3%. 3000 pRBCs per sample were analyzed using a BD FACSVerse (BD Bioscience) and surface positivity was analyzed in FlowJo version 10 by first gating on intact cells on SSC-A and FSC-A, followed by FSC-H and FSC-A for single cells and from there, the Hoechst-DHE double positive pRBC population was selected. This population was analyzed for percentage of Alexa647 positive cells based on a cutoff from the non-immune IgG background binding. The parasite NF54CSA-ptefKO[18] was used as negative control as the knock-out of the gene encoding the *Plasmodium* Translation Enhancing Factor (PTEF) has previously been shown to be essential for the translation of VAR2CSA. Thus, NF54CSA-ptefKo pRBCs are devoid of VAR2CSA and therefore non-adherent to CSA.

Placental binding

A placental binding assay was adapted from Flick et al[21]. Cryo-sections (8 µm in thickness and approximately 6 mm in diameter) from placentas donated by healthy Swedish pregnant women were prepared in a Microm HM 560 cryotome (ThermoScientific). Sections were mounted onto 10-well microscopy slides (Novakemi AB) and stored at -80 ºC. Prior to use, the placental sections were submerged in ice cold PBS. Late-stage pRBCs were enriched using magnetic columns (Miltenyi Biotec) and washed 3 times before resuspended in binding medium (RPMI-1640 (Gibco) with HEPES adjusted to 20 mM, pH 6.8, and 10% human serum at a density of 2.4×10^5 pRBCs / µL. 25 µL (0.6×10^7 cells) of pRBC suspension was added to each placental section and were incubated at 37°C for 1 h in a humid chamber. Unbound cells were washed away by dipping the slides 3 x 5 min in RPMI-1640 with HEPES. Preparations were thereafter fixed with 100% methanol for 15 s and stained with 5% Giemsa (Merck) for 15 minutes, washed with distilled water and air-dried. Samples were visualized using an Eclipse 80i (Nikon) microscope at 1000 x magnification and blinded for counting. Samples were imaged using an Infinity 3 color camera (Teledyne Lumenera). Number of bound cells per experimental condition was counted for 30-45 fields containing on average 29 pRBCs per field and relative binding was determined compared to the one of VAR2CSA non-expressing pRBCs.

RNA purification, library preparation and sequencing

For RNA sequencing, NF54 and NF54CSAh were harvested from three independent selections at 16 ± 2h post invasion (p.i.) of RBCs. Cell pellets were lysed using five volumes of TRIzol Reagent (Life Technologies) before RNA was purified using the NucleoSpin miRNA Mini kit for miRNA and RNA purification (Macherey-Nagel) without size fractionation, according to the instructions of the manufacturer. The high retention of RNA on the columns of the kit allows for sequential purification without any apparent loss of RNA (not shown). Therefore, all RNA samples were passed over columns in total three times, with the first two elutions being treated with Turbo DNAse (Invitrogen) in solution to ensure a complete lack of contaminating gDNA in the preparations. Strand specific RNA sequencing (RNA-seq) libraries were generated using the KAPA Stranded RNA-Seq kit (Roche) and were amplified for 5 PCR cycles (45 s at

98°C for initial denaturing and cycles of 15 s at 98°C, 30 s at 60°C, 30 s at 68°C before final extension for 60 s at 68°C) using the KAPA HiFi HotStart ready mix (KAPA Biosystems) before purification twice over 1X KAPA Pure beads (KAPA Biosystems). Prior to sequencing, quality of libraries was evaluated using Agilent 2100 Bioanalyzer (Agilent) and a High sensitivity DNA kit (Agilent). Libraries were quantified using Collibri Library Quantification kit (Invitrogen) and thereafter sequenced in 75bp paired-end reads on an Illumina NextSeq550 instrument. PhiX library (Illumina) was used for sequencing control and to diversify libraries.

RNA-seq data processing and differential expression analysis

Sequenced RNA-seq libraries were assessed for quality using FastQC (v0.11.9). In addition to the index adapters, 12 bp were trimmed from the ends of reads based on sequence quality using Trimmomatic (v0.39)[22]. Trimmed reads were then aligned to the *P. falciparum* 3D7 genome assembly (PlasmoDB v58) using HISAT2 (v2.1.0)[23]. Samtools (v1.10)[24] was used to filter and sort aligned reads using a quality score of 30 and retrieve properly mapped and paired reads. The output high quality, properly paired reads were mapped to protein coding genes to retrieve total read count using HTseq-count (v1.99.2)[25]. Differential gene expression analysis between NF54CSAh and NF54 was performed using DESeq2 (v1.32.0)[26].

Chromatin immunoprecipitation (ChIP) and ChIP-seq library preparation

For the Chip-seq experiment, we harvested three Biological replicates of NF54 and NF54CSAh (16 ± 2h p.i) in parallel with the samples collected for RNA-seq. Parasites were freed from RBCs using 0.1% saponin (from quillaja bark, Sigma) and washed thrice with ice-cold PBS. Parasites were cross-linked in 1% methanol-free formaldehyde (Polysciences Inc) at 37°C for 10 min while kept in suspension on an orbital shaker (50 rpm). The cross-linking was thereafter quenched using 125 mM glycine (Sigma) for 5 min at 37°C before cell pellets were washed two times with PBS at 4°C. Nuclei were extracted using a 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 Halt protease inhibitor cocktail from ThermoFischer) and incubated for 30 min on ice before addition of Nonidet P 40 substitute (Roche) to a final concentration of 0.5%. Samples were homogenized by passage through 23G needles three times and 25G needles seven times. Parasite nuclei were thereafter collected and resuspended in shearing buffer (1% SDS, 10 mM EDTA, 50 mM Tris HCl pH 8.1 and 1 x Halt protease inhibitor cocktail, ThermoFisher). Chromatin was fragmented using the Covaris ultra sonicator (ME220) for 5 min with the following settings: 25% duty cycle, 75W intensity peak incident power, 1000 cycles per burst. Insoluble material was removed by centrifugation for 10 min at 15700 rcf at 4°C before samples were diluted tenfold in ChIP dilution buffer (16.7 mM Tris-HCl pH 8, 1.2 mM EDTA, 0.01% SDS, 150 mM NaCl. 1.1% Triton and 1x Halt protease inhibitor cocktail from ThermoFisher). To reduce non-specific background, samples were precleared with ChIP-grade protein A/G magnetic beads (Thermo

Scientific) / 0.2mg/ml Salmon Sperm DNA (Invitrogen) for 2 h at 4°C with end-to-end rotation. To retrieve H3K9 trimethylated histones, samples were incubated with 2 µg anti-H3K9me3 antibodies (ab8898, Abcam) or a non-immune rabbit IgG as negative control (12-370, Upstate) overnight at 4°C. An input sample was saved at 4°C until the de-crosslinking step. Antibody–protein complexes were recovered during 2 h incubation at 4°C with protein A/G magnetic beads (Thermo Scientific), followed by duplicate washes of 15 min with low-salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), high-salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl), LiCl wash buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.1) and TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA). Chromatin immunocomplexes were thereafter eluted from the beads twice with elution buffer (1% SDS, 0.1 M NaHCO3) for 15 min at room temperature and elutes were combined. RNA was removed by incubation with 0.48 µg/µL RNase A (Invitrogen) for 30 min at 37°C, followed by a 2 h incubation at 45°C with proteinase K (final concentration 0.24 µg/µL, Ambion). Samples were de-crosslinked overnight at 65°C by adding NaCl to a final concentration of 200 mM. DNA was thereafter extracted using phenol:chloroform:isoamyl alcohol (25:24:1, Sigma) and ethanol precipitation and was further purified with KAPA Pure beads (KAPA Biosystems).

ChIP-seq libraries were prepared using the KAPA LTP library preparation kit (KAPA Biosystems). Immunoprecipitated and input libraries were amplified for 16 and 5 PCR cycles respectively (45 s at 98°C for initial denaturing and cycles of 15 s at 98°C, 30 s at 60°C, 30 s at 68°C before final extension for 60 s at 68°C) using the KAPA HiFi HotStart ready mix (KAPA Biosystems) before purification twice over 1X KAPA Pure beads (KAPA Biosystems). Prior to sequencing, quality of libraries was evaluated using Agilent 2100 Bioanalyzer (Agilent) and a High sensitivity DNA kit (Agilent). Libraries were quantified using Colibri Library Quantification kit (Invitrogen) and thereafter sequenced in 75bp paired-end reads on an Illumina NextSeq550 instrument. PhiX library (Illumina) was used for sequencing control and to diversify libraries.

Methylated DNA immunoprecipitation (MeDIP) and MeDIP-seq library preparation

Biological triplicates of NF54 and NF54CSAh (16 ± 2h p.i) were harvested for MeDIP-seq in parallel with samples collected for RNA-seq and ChIP-seq. Parasites were initially released from host RBCs by 0.1% saponin (from quillaja bark, Sigma) and washed thrice with ice-cold PBS. Thereafter genomic DNA was extracted using QIAamp DNA Blood Kit (Qiagen) following the manufacturer instructions, including the optional RNase A treatment. Extracted DNA was eluted in NE buffer (Macherey-Nagel) and quantified with dsDNA HS assay kit (Invitrogen) on a Qubit 3 fluorometer (Invitrogen). gDNA was fragmented in TE buffer using the Covaris ultra sonicator (ME220) for 130 s with the following settings: 20% duty cycle, 70W intensity peak incident power, 1000 cycles per burst, resulting in average fragment size of approximately 300 bp. Thereafter gDNA was precipitated and concentrated by addition of 0.1 volumes of sodium acetate (Alfa Aesar) and 3 volumes of 100% ethanol. gDNA was quantified with dsDNA HS assay kit (Invitrogen) for Qubit 3 fluorometer (Invitrogen) and 100 / 30 ng (immunoprecipitation/input) used for library preparation using KAPA LTP library preparation kit (KAPA Biosystems). Library preparation was stopped prior to PCR amplification. The quality of libraries was assessed with Agilent 2100 Bioanalyzer (Agilent) and High sensitivity DNA kit (Agilent) and quantity with the Collibri Library Quantification kit (Invitrogen). Thereafter equal amounts of libraries from NF54 and NF54CSAh were pooled for the same bio-replicate for the same antibody. Libraries were denatured at 95°C for 10 min and then incubated on ice for 10 min. The Methylated DNA Immunoprecipitation (MeDIP) Kit (Active Motif) was used according to the instructions by the manufacturer with some modifications. Three µg of polyclonal rabbit anti-5mC (#61255, Active Motif) and rabbit IgG (#103524, Active Motif) were added to the denatured libraries. For the elucidation of potential presence of 5hmC and 6mA in NF54CSAh only, three µg of polyclonal rabbit anti-5mhC (#39069, Active Motif) and polyclonal rabbit anti-6mA (#202003, Synaptic Systems) was used, respectively. Immune complexes were collected using 20 µL of protein G magnetic beads (Active Motif). After elution from the beads, samples were purified using 1X KAPA Pure beads (KAPA Biosystems) and PCR amplified using the KAPA HiFi HotStart ready mix (KAPA Biosystems) for 8 PCR cycles (45 s 98°C of initial denaturing and amplification at 15 s at 98°C, 30 s at 60°C, 30 s at 68°C and final extension 60 s at 68°C). Libraries were analyzed for quality using an Agilent 2100 Bioanalyzer (Agilent) and the High sensitivity DNA kit (Agilent), quantified using Collibri Library Quantification kit (Invitrogen).and thereafter sequenced in 75bp paired-end reads on an Illumina NextSeq550 instrument. PhiX library (Illumina) was used for sequencing control and to diversify libraries.

ChIP-seq and MeDIP-seq data processing and peak calling

ChIP-seq read quality was assessed with FastQC (v0.11.9). Adapters and low-quality bases were trimmed using Trimmomatic (v0.39)[22]. Bowtie2 (v2.3.5.1)[27] was used to align reads to the *P. falciparum* 3D7 genome (PlasmoDB v58). PCR duplicates were tagged using Picardtools (v2.26.11, http://broadinstitute.github.io/picard). Quality filtering with mapping score of 30, sorting and indexing were performed using Samtools (v1.10)[24] keeping only properly paired and mapped reads. Bedtools (v2.27.1)[28] was used to retrieve per-base genome-wide coverage. All replicates for each sample (H3K9me3, IgG, and input) were per-million normalized then IgG and input were subtracted from H3K9me3 to remove background. Genome-wide coverage was binned at 10 bp resolution. Broad peaks were called using MACS3 (v3.0.0a7, https://github.com/macs3--project/MACS) with FDR < 0.05. TSS coverage profiles were mapped with deeptools2 (v3.5.1)[29] bamCompare to subtract input reads and regions from 500 bp 5’ and 1500 bp 3’ of the TSS were scored. DiffBind (v3.2.7) was utilized for differential peak calling.

MeDIP-seq data processing was performed the same as above. Per-million normalization was performed prior to subtracting IgG and input reads from 5mC, 5hmC and 6mA for genome-wide visualization at 10 bp resolution. MACS3 peak calling was performed using the default narrow peak calling for 5mC, 5hmC and 6mA, but minimum fold coverage for model building was reduced to 2. Default DiffBind parameters were used for differential peak calling between NF54CSAh and NF54 for 5mC, and between NF54CSAh 5mC, 5hmC, and 6mA.

Chromosome conformation capture sequencing (Hi-C) library preparation

NF54 and NF54CSAh (20 ± 2 hpi) were fixed for 25 min at 37°C with 1.25% methanol-free formaldehyde (Polysciences Inc) and quenched with 150 mM glycine for 15 min at 37°C. Samples were incubated at 4°C for 15 min, then washed with ice-cold PBS. Supernatant was removed then pellet resuspended in lysis buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 2 mM AEBSF, 0.10% Igepal CA-360 (v/v), 1X protease inhibitor cocktail) and incubated for 30 min on ice. Samples were homogenized by 26.5G needle, then washed in lysis buffer. Pellets were resuspended in 0.5% SDS and incubated at 62°C to solubilize the chromatin. DNA was digested overnight with 100U MboI (NEB) restriction enzyme then ends filled using dTTP, dGTP, dATP, biotinylated dCTP. Blunt ends were ligated using 4000U T4 DNA ligase then de-crosslinked in decrosslinking buffer (50 mM Tris-HCl, pH 8.0, 1% SDS, and 500 mM NaCl) and RNase A. Ligated DNA was sheared to 300-500 bp using a Covaris S220 sonicator (10% duty factor, 140 peak incident power, 200 cycles per burst, 65 seconds). Fragments pulled down with streptavidin T1 beads (Invitrogen) and libraries prepped using a KAPA library preparation kit (KAPA biosystems). Libraries PCR amplified with NEB multiplex oligos (45 sec at 98°C, 12 cycles of 15 sec at 98°C, 30 sec at 55°C, 30 sec at 62°C and a final extension step of 5 min at 62°C) and sequenced using the NOVASeq 6000 (Illumina).

Hi-C data processing, differential interaction analysis and 3D modeling

Paired-end Hi-C libraries were processed (mapping, filtering, and normalizing) using HiC-Pro (v3.1.0)[30] with mapping quality cutoff of 30 at 10 kb resolution while aligning to the *P. falciparum* 3D7 genome (PlasmoDB v58). Interaction heatmaps were generated using ICED-normalized interaction matrices after per-million normalization to allow for direct comparison between NF54CSAh and NF54. Correlation among biological replicates was evaluated with HiCRep[31], then replicates were merged to generate a single representative sample. To enhance visualization, the first two diagonals of the matrices were set to 0 and data was scaled to the 90th percentile. Differential interactions between NF54CSAh and NF54 were identified using Selfish[32] for all intrachromosomal and interchromosomal matrices. Coordinate matrices were generated by PASTIS[33] and visualized as 3D chromatin models in ChimeraX[34], highlighting bins containing *var* genes, telomeres and centromeres.

ACKNOWLEDGEMENTS

The authors thank Lars Haag at the Electron microscopy unit (Emil), Karolinska Institutet, for excellent technical assistance related to scanning electron microscopy.

FUNDING

This work was funded by The Knut and Alice Wallenberg Foundation (KAW 2017.0055 to U.R.); The Swedish Research Council (VR 2018-05814, VR 2016-02917 to U.R.); The National Institutes of Allergy and Infectious Diseases and the National Institutes of Health (R01 AI136511, R21 AI142506 to K.L.R.); and The University of California, Riverside (NIFA-Hatch-225935 to K.L.R.). Funding for open access charge: The Swedish Research Council.

AUTHOR CONTRIBUTIONS

Todd Lenz: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Visualization, Writing – Original draft. **Madle Sirel**: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Visualization, Writing – Review and editing. **Hannes Hoppe**: Methodology, Writing – Review and editing. **Sulman Shafeeq**: Methodology, Writing – Review and editing. **Karine Le Roch**: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Supervision, Writing – Original draft, Project administration. **Ulf Ribacke**: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Supervision, Writing – Original draft, Project administration

CONFLICT OF INTEREST

All contributing authors declare no competing interests nor conflicts of interest.

DATA AVAILABILITY

Data supporting the findings of this manuscript are available from corresponding authors upon reasonable request. Sequencing datasets generated and analyzed for the study are available in the SRA repository under the accession number PRJNA947327.

SUPPLEMENTARY DATA

Supplementary data and figures for this chapter are available at doi.org/10.1101/2023.01.31.526483.

REFERENCES

1. Doolan, D. L., Dobano, C. & Baird, J. K. Acquired immunity to malaria. *Clin. Microbiol. Rev.* **22**, 13-36 (2009).

2. Steketee, R. W., Nahlen, B. L., Parise, M. E. & Menendez, C. The burden of malaria in pregnancy in malaria-endemic areas. *Am. J. Trop. Med. Hyg.* **64**, 28-35 (2001).

3. Rogerson, S. J. & Beeson, J. G. The placenta in malaria: mechanisms of infection, disease and foetal morbidity. *Ann. Trop. Med. Parasitol.* **93**, 35-42 (1999).

4. Nyirjesy, P., Kavasya, T., Axelrod, P. & Fischer, P. R. Malaria during pregnancy: neonatal morbidity and mortality and the efficacy of chloroquine prophylaxis. *Clin. Infect. Dis.* **16**, 127-32 (1999).

5. Chua, C. L. L., Hasang, W., Rogerson, S. J. & Teo, A. Poor birth outcomes in malaria in pregnancy: recent insights into mechanisms and prevention approaches. *Front. Immunol.* **12**, 621382 (2021)

6. Scherf, A., Lopez-Rubio, J. J. & Riviere, L. Antigenic variation in Plasmodium falciparum. *Annu. Rev. Microbiol.* **62**, 445-470 (2008).

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

8. Lee, W. C., Russel, B. & Rénia, L. Sticking for a cause: the falciparum malaria parasites cytoadherence paradigm. *Front. Immunol.* **10**, 1444 (2019).

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

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

11. Fried, M. & Duffy, P. E. Adherence of Plasmodium falciparum to chondroitin sulfate A in the human placenta. *Science* **272**, 1502-1504 (1996).

12. Salanti, A. et al. Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. *J. Exp. Med.* **200**, 1197-1203 (2004).

13. Brolin, K. J. M. et al. Simultaneous transcription of duplicated var2csa gene copies in individual Plasmodium falciparum parasites. *Genome Biol.* **10**, R117 (2009).

14. Joergensen, L. et al. Surface co-expression of two different PfEMP1 antigens on single Plasmodium falciparum-infected erythrocytes facilitates binding to ICAM1 and PECAM1. *PLoS Pathog.* **6**, e1001083 (2010).

15. Hollin, T. & Le Roch, K. G. From genes to transcripts, a tightly regulated journey in Plasmodium. *Front. Cell. Infect. Microbiol.* **10**, 618454 (2020)

16. Mok, B. W. et al. Default pathway of var2csa switching and translational repression in Plasmodium falciparum. *PLoS ONE* **3**, e1982 (2008).

17. Amulic, B. et al. An upstream open reading frame controls translation of var2csa, a gene implicated in placental malaria. *PLoS Pathog.* **5**, e1000256 (2009).

18. Chan, S. et al. Regulation of PfEMP1-VAR2CSA translation by a Plasmodium translation-enhancing factor. *Nat. Microbiol.* **2**, 17068 (2017).

19. Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science* **193**, 673-675 (1976).

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

21. Flick, K. et al. Role of nonimmune IgG bound to PfEMP1 in placental malaria. *Science* **293**, 2098-2100 (2001).

22. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

23. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357-360 (2015).

24. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

25. Anders, S., Pyl, P. T. & Huber, W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).

26. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

27. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-359 (2012).

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

29. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160-W165 (2016).

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

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

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

33. Varoquaux, N., Ay, F., Noble, W. S. & Vert, J. P. A statistical approach for inferring the 3D structure of the genome. *Bioinformatics* **30**, i26-i33 (2014).

34. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14-25 (2018).

35. Roberts, D. J. et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* **357**, 689-692 (1992).

36. Staalsoe, T. et al. In vitro selection of Plasmodium falciparum 3D7 for expression of variant surface antigens associated with severe malaria in African children. *Parasite Immunol.* **25**, 421-427 (2003).

37. Bozdech, Z. et al. The transcriptome of the developmental cycle of Plasmodium falciparum. *PLoS Biol.* **1**, e5 (2003).

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

39. Ninova, M., Tóth, K. F. & Aravin, A. A. The control of gene expression and cell identity by H3K9 trimethylation. *Development* **146**, dev181180 (2019).

40. Lopez-Rubio, J. J. et al. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* **5**, 179-190 (2009).

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

42. Michel-Todó, L. et al. Patterns of heterochromatin transitions linked to changes in the expression of Plasmodium falciparum clonally variant genes. *Microbiol. Spectr.* **11**, e0304922 (2023).

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

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

45. Howitt, C. A. et al. Clonally variant gene families in Plasmodium falciparum share a common activation factor. *Mol. Microbiol.* **73**, 1171-1185 (2009).

46. Gupta, M. K., Lenz, T. & Le Roch, K. G. Chromosomes conformation capture coupled with next-generation sequencing (Hi-C) in Plasmodium falciparum. *Methods Mol. Biol.* **2369**, 15-25 (2021).

47. 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.* **24**, 974-988 (2014).

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

49. Zhang, X. et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* **126**, 1189-1201 (2006).

50. Lister, R. et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315-322 (2009).

51. Feng, S. et al. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. USA* **107**, 8689-8694 (2010).

52. Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916-919 (2010).

53. Jjingo, D., Conley, A. B., Yi, S. V., Lunyak, V. V. & Jordan, I. K. On the presence and role of human gene-body DNA methylation. *Oncotarget* **3**, 462-474 (2012).

54. Ponts, N. et al. Genome-wide mapping of DNA methylation in the human malaria parasite Plasmodium falciparum. *Cell Host Microbe* **14**, 696-796 (2013).

55. Hammam, E. 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).

56. Nardella, F. et al. DNA methylation bisubstrate inhibitors are fast-acting drugs active against artemisinin-resistant Plasmodium falciparum parasites. *ACS Cent. Sci.* **6**, 16-21 (2020).

57. Maher, S. P. et al. A drug repurposing approach reveals targetable epigenetic pathways in Plasmodium vivax hypnozoites. *bioRxiv* https://doi.org/10.1101/2023.01.31.526483 (2023).

58. Lucky, A. B. et al. Characterization of the dual role of Plasmodium falciparum DNA methyltransferase in regulating transcription and translation. *Nucleic Acids Res.* gkad248 (2023).

59. Wion, D. & Casadesús, J. N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat. Rev. Microbiol.* **4**, 183-192 (2006).

60. Zhang, G. et al. N6-methyladenine DNA modification in Drosophila. *Cell* **161**, 893-906 (2015).

61. Luo, G. Z., Blanco, M. A., Greer, E. L., He, C. & Shi, Y. DNA N(6)-methyladenine: a new epigenetic mark in eukaryotes? *Nat. Rev. Mol. Cell Biol.* **16**, 705-710 (2015).

62. Liang, Z. et al. DNA N6-adenine methylation in Arabidopsis thaliana. *Dev. Cell* **45**, 406-416 (2018).

63. Xiao, C. L. et al. N6-methyladenine DNA modification in the human genome. *Mol. Cell* **71**, 306-318 (2018).

64. Zilberman, D. An evolutionary case for functional gene body methylation in plants and animals. *Genome Biol.* **18**, 87 (2017).

65. Del Portillo, H. A. et al. The role of the spleen in malaria. *Cell Microbiol.* **14**, 343-355 (2012).

66. Maier, A. G. et al. Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. *Cell* **134**, 48-61 (2008).

67. Ferrai, C., de Castro, I. J., Lavitas, L., Chotalia, M. & Pombo, A. Gene positioning. *Cold Spring Harb. Perspect. Biol.* **2**, a000588 (2010).

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

69. Duraisingh, M. T. et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. *Cell* **121**, 13-24 (2005).

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

71. Coleman, B. I. et al. Nuclear repositioning precedes promoter accessibility and is linked to the switching frequency of a Plasmodium falciparum invasion gene. *Cell Host Microbe* **12**, 739-750 (2012).

72. Dixon, G., Liao, Y., Bay, L. K. & Matz, M. V. Role of gene body methylation in acclimatization and adaptation in a basal metazoan. *Proc. Natl. Acad. Sci. USA* **115**, 13342-13346 (2018).

73. Muyle, A. M., Seymour, D. K., Lv, Y., Huettel, B. & Gaut, B. S. Gene body methylation in plants: mechanisms, functions, and important implications for understanding evolutionary processes. *Genome Biol. Evol.* **14**, evac038 (2022)

CHAPTER 2

Chromatin Structure and *var2csa* – A Tango In Regulation Of *var* Gene Expression In The Human Malaria Parasite, *Plasmodium falciparum*?

Todd Lenz[1], Xu Zhang[2], Abhijit Chakraborty[3], Abbas Roayaei Ardakany[3], Jacques Prudhomme[1], Ferhat Ay[3], Kirk Deitsch[2] and Karine G Le Roch[1,*].

[1] Department of Molecular, Cell and Systems Biology, University of California Riverside, 900 University Avenue, Riverside, CA 92521, USA.
[2] Department of Microbiology and Immunology, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10065, USA.
[3] La Jolla Institute for Immunology, 9420 Athena Circle, La Jolla, CA 92037, USA.
* Correspondence: Karine Le Roch – Email: karinel@ucr.edu.

A version of this chapter is available in BioRxiv. 13, Feb. 2024. doi.org/10.1101/2024.02.13.580059. The Deitch lab provided the *var2csa* knockout line used for this project, but I conducted all experiments, performed all data analyses, generated all figures, and took the lead role in manuscript preparation and editing.

ABSTRACT

Over the last few decades, novel methods have been developed to study how chromosome positioning within the nucleus may play a role in gene regulation. Adaptation of these methods in the human malaria parasite, *Plasmodium falciparum*, has recently led to the discovery that the three-dimensional structure of chromatin within the nucleus may be critical in controlling expression of virulence genes (*var* genes). Recent work has implicated an unusual, highly conserved *var* gene called *var2csa* in contributing to coordinated transcriptional switching, however how this gene functions in this capacity is unknown. To further understand how *var2csa* influences *var* gene switching, we deleted the gene using CRISPR-Cas9 targeted DNA double-strand breaks within the sub-

telomeric region of chromosome. To characterize changes in chromatin architecture stemming from this deletion and their effect on *var* gene expression, we used a combination of RNA-seq, Chip-seq and Hi-C to pinpoint epigenetic and chromatin structural modifications in regions of differential gene expression. We observed a net gain of interactions in sub-telomeric regions and internal *var* gene regions following *var2csa* knockout, indicating an increase of tightly controlled heterochromatin structures. Our results suggest that disruption of *var2csa* results not only in changes in *var* gene transcriptional regulation but also a significant tightening of heterochromatin clusters thereby disrupting coordinated activation of *var* genes throughout the genome.

INTRODUCTION

Although the recent coronavirus pandemic has dominated the news for the last few years, it is important to note that human malaria parasites still infect 249 million people worldwide annually. *Plasmodium*, the causative agent of malaria, remains one of the deadliest parasite-borne diseases, with an estimated 608,000 deaths in 2022[1] affecting mostly children and immunocompromised individuals in sub-Saharan Africa and Southeast Asia. There are five members of the *Plasmodium* genus known to infect humans, *Plasmodium falciparum* being most deadly. *P. falciparum* can not only evade the host immune system but can sequester and cyto-adhere to blood vessels leading to circulatory obstruction resulting in dysfunction of multiple organs including the brain. Understanding the mechanisms by which *P. falciparum* evades the immune response, and thus sustains further replication within the host, is key to identifying novel drug targets and abating the impact of malaria on global health.

To evade the host immune responses, the parasite switches the expression of specific exported proteins or antigens on the surface of infected red blood cells. The molecular mechanisms underlying 'antigenic variation' have been extensively studied over the years with particular focus on the *var* gene family encoding *Plasmodium falciparum* erythrocyte membrane protein-1 (PfEMP1)[2, 3]. Expression of this gene family is thought to be regulated, at least partially, by epigenetic mechanisms. In eukaryotes, epigenetics involves various covalent modifications of nucleic acids and histone proteins which lead to changes in chromatin structure and gene expression. Similar mechanisms of gene regulation have been detected in *P. falciparum* with tight control of virulence genes organized in heterochromatin clusters with repressive histone mark H3K9me3[4-6] localized at the nuclear periphery[4, 7, 8].

PfEMP1 is exported to the plasma membrane of an infected red blood cell and mediates the adhesion of the infected erythrocytes to receptors on endothelial cells of the post capillary endothelium. Although there are approximately 60 members of the *var* gene family, only one *var* gene is expressed at any given time through a process known as mutually exclusive gene expression[9, 10]. To circumvent the host immune response, the parasite will switch the *var* gene expressed at a rate of 2-18% per generation preventing the recognition of parasitized cells by host antibodies[3, 11-13]. During this switching process, designated the “single-many-single” (SMS) model, it has been hypothesized that an intermediate state exists in which there is no dominantly expressed *var* gene and instead transcription is dispersed across several *var* genes before settling on a single dominant transcript[13]. This was originally referred to as the “many” state and data supporting this model were recently obtained through transcriptomic analysis[14]. Specific *var* genes serve as ‘sink’ nodes, notably *var2csa*, and are upregulated enabling the transition from the “many” state back to expression of a single *var* transcript, thus explaining how a large population of parasites can express a small subset of *var* genes[13, 14]. However, those results were obtained using clonal populations of parasites and future single cell RNA-seq studies would aid in fully understanding the “many” state. It is also not fully understood what epigenetic mechanisms allow parasites to transition between these transcriptional states; however, this model implies a mechanism that unites the entire *var* gene family into a single network, potentially providing the basis for coordination of *var* gene expression within a large population of infected RBCs. Transcriptional dominance is complex and recent evidence in other models, including transcriptional regulation of odorant receptor genes, has identified possible non-coding mechanisms within mRNA that mediate the exclusive expression of a single gene within a large variable gene family[15].

Of the 60 *var* genes comprising the PfEMP1 family, *var2csa* is one of the most conserved between parasite strains and is the only variant whose transcription is upregulated in *P. falciparum*-parasitized erythrocytes in the placenta, causing clinical symptoms in the expectant mother and serious harm to the fetus[16-18]. Pregnancy associated malaria (PAM) results from sequestration of infected erythrocytes by binding chondroitin sulfate A (CSA) on the surface of syncytiotrophoblasts in the placental tissue[16]. Although expression in CSA-bound infected RBCs are exclusive to pregnant women, *var2csa* can be translationally repressed in infected erythrocytes while still being transcriptionally active at very low levels in non-pregnant populations, providing evidence that *var2csa* serves an additional, possible regulatory function[19-22].

Although considerable progress has been made in elucidating the genetic and epigenetic basis for coordinated *var* gene switching, the hierarchy of these mechanisms controlling regulation remains unknown. Switching between different *var* genes is highly coordinated and imperative to parasite survival, but evidence suggests that some *var* genes may be more important than others and thus the rate of switching is variable depending on its genomic location—subtelomeric versus internal regions of the chromosomes[13,23-26]. Internal *var* genes are remarkably stable and transcriptional switching is uncommon compared to subtelomeric *var* genes which readily switch in the presence of environmental pressure[25-27]. Manipulation of H3K9me3 and H3K36me3 epigenetic marks results in upregulation of *var2csa* expression within a mixed population, regardless of previous *var* gene expression, suggesting that var2csa inhabits a prominent position within the *var* gene activation hierarchy[25]. Furthermore, when *var2csa* was deleted, *var* gene switching was disabled or greatly reduced, resulting in much more stable *var* gene expression over time[14]. This coincided with an overall reduction in the level of *var* transcripts, in particular the more dispersed transcriptional pattern observed in parasites transitioning through the hypothetical switching intermediate state. This implied that *var* gene expression had somehow become more tightly regulated, with fewer genes displaying even low-level transcriptional activity.

However, how the loss of the *var2csa* locus led to this broader transcriptional phenotype is not known.

Using a combination of transcriptomic, epigenetic and chromatin structure analyses we determined that *var2csa* deletion leads to a significant tightening of heterochromatin clusters throughout the genome, thus providing a likely explanation for reduced *var* transcription and disrupted transcriptional switching. These results confirm the importance of chromatin architecture during the process of *var* gene switching and provide a molecular basis for the phenotype resulting from deletion of the *var2csa* locus.

RESULTS

Deletion of *var2csa* alters *var* gene expression

To evaluate the effect of *var2csa* disruption on changes in overall *var* gene expression, we used Clustered Regularly Interspaced Short Palindromic Repeat/CRISPR-associated protein-9 (CRISPR/Cas9) targeted DNA double-strand breaks (DSBs) within the sub-telomeric region of chromosome 12 near the *var2csa* promoter **(Fig. 2.1A)**. The successfully edited cell line described in this manuscript as *var2csa*-deleted line (ΔV2) was previously shown to have undergone repair of the DSB through telomere healing by the addition of telomeric heptad repeat elements near the site of the break[32]. This resulted in the loss of the *var2csa* locus and surrounding region of the chromosome. Furthermore, although *var* gene transcription was not lost, deletion of *var2csa* was shown to greatly reduce *var* gene switching and resulted in a significant reduction in the *var* transcripts that make up a semi-conserved pattern of what were referred to as *var* "minor" transcripts[14]. These are low-level transcripts that come from a subset of *var* genes dispersed across the genome. How loss of *var2csa* led to a reduction in expression of transcripts coming from *var* genes located on other chromosomes was not clear, and any effects on members of other variant gene families was not determined.

To investigate the phenotype resulting from deletion of *var2csa*, we used RNA-seq to compare *var* gene expression of the ΔV2 line to an NF54 wild-type line **(Fig. 2.1B).** Because lines with varying sites of telomere healing all underwent identical recombination events and most displayed the 'many' phenotype rather than trend toward expression of a single or small subset of *var* genes, we performed RNA-seq on only a single *var2csa*-mutant line[14,31]. We previously observed that parasite populations expressing different levels of *var* transcripts progress through the ring-stage of the replicative cycle at different rates[14], however all cultures converged to the same point in the cycle prior to the initiation of DNA replication. Therefore, to avoid complications due to differences in cell cycle progression, total RNA was collected from cultures at ~24 hpi, a time point when *var* gene mRNA is still prevalent, but all parasites have returned to synchronicity. RNA was obtained from two independent cultures[14]. Expression profiles presented a strong correlation ($s > 0.98$) between both replicates and samples **(Supp. Fig. 2.1A)**; however, the correlation for *var* gene expression was much weaker ($s \approx 0.82$ to 0.84) between cell lines indicating that much of the transcriptional difference observed was due to changes in *var* gene expression **(Fig. 2.1C and Supp. Fig. 2.1A-C)**. Differential expression analysis revealed 51 significantly differentially expressed genes ($FDR < 0.05$ and $\log_2FC > 1$), with 10 upregulated and 41 downregulated **(Fig. 2.1D and**

Supp. Data 2.3). Of the 51 significantly differentially expressed genes, 2 *var* genes display increased transcription ($\log_2$FC ≈ 4.61 and 0.77) and 10 *var* genes show decreased transcription ($\log_2$FC ≈ -0.74 to -7.66), including *var2csa* ($\log_2$FC ≈ -7.21). Furthermore, 11 of 14 *rifin* genes–an antigenic gene family found in close proximity to *var* gene clusters–are also downregulated. Our gene ontology analysis shows significant enrichment of genes involved in antigenic variation for the downregulated genes **(Supp. Fig. 2.1C)**. These results confirm the total loss of *var2csa* expression following the introduction of the DSB, as well as the subsequent downregulation of several other *var* genes. Interestingly, the only highly upregulated *var* gene is located within an internal *var* gene cluster on chromosome 4, whereas all of the downregulated *var* genes are located in subtelomeric *var* gene clusters. This suggests that the knockout of *var2csa* results in epigenetic and chromatin structural changes that cause the downregulation of these subtelomerically located genes. Furthermore, there is not a major *var* gene expressed in ΔV2 but rather widespread low-level transcription when compared to the WT line which shows one major *var* gene with elevated PF3D7_1200100 transcripts. This result is consistent with previous work which showed that knockout of *var2csa* coincides with the 'many' *var* gene expression phenotype in most clonal populations[32].

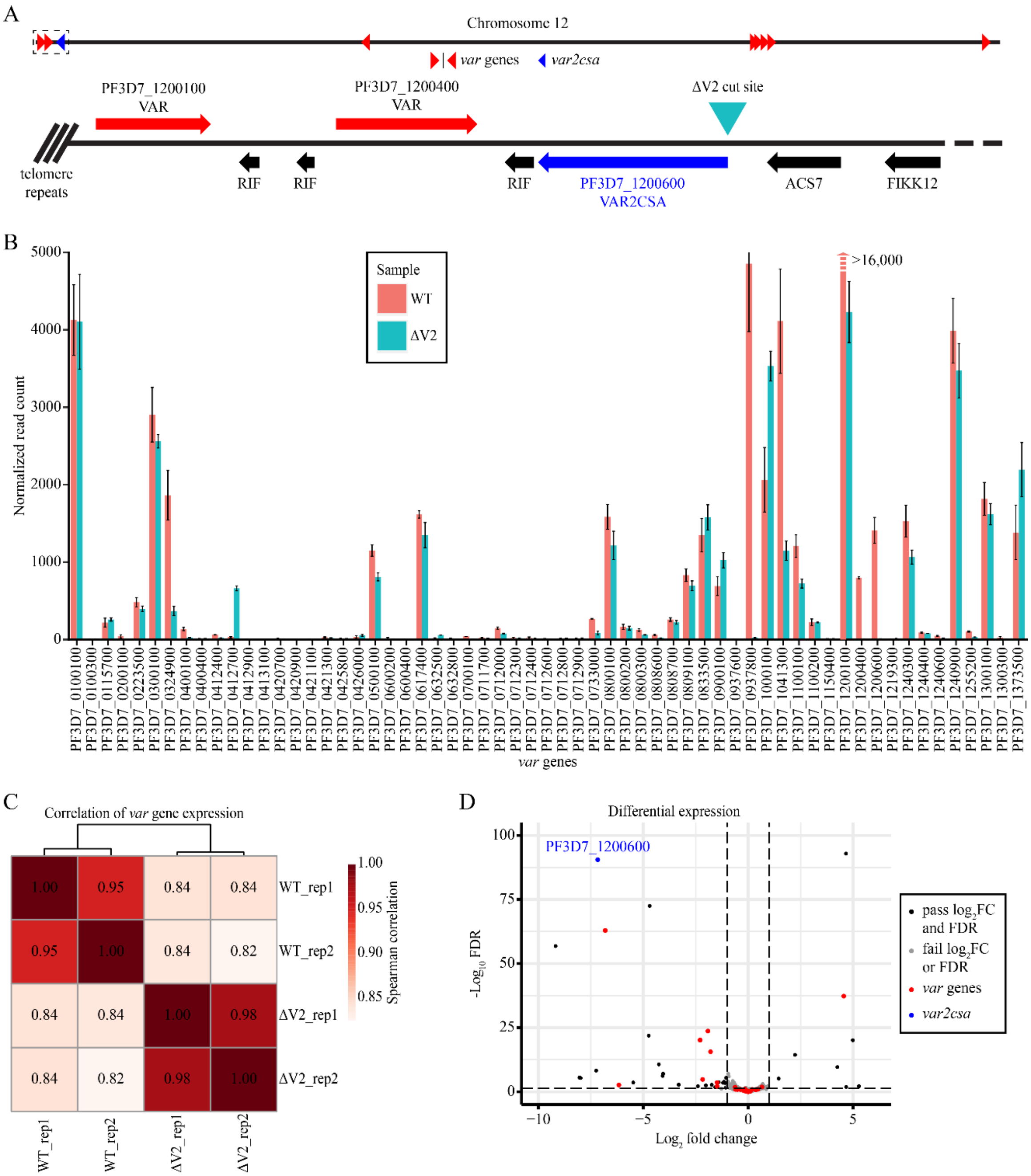

Figure 2.1. Overall *var* gene transcription decreased following a targeted DSB within the *var2csa* promoter. (A) Location on chromosome 12 of the CRISPR/Cas9 targeted DSB within the *var2csa* promoter region. **(B)** Read count normalized *var* gene expression comparing the WT and ΔV2 cell lines with error bars indicating the maximum and minimum values of replicates. The figure was truncated to enhance visibility of other *var* genes due to significantly higher expression of PF3D7_1200100 in the WT cell line. **(C)** Spearman correlation for *var* gene expression in both replicates and cell lines. **(D)** Significant differentially expressed genes ($\log_2$FC > 1 and FDR < 0.05) with most *var* genes highlighted in red and *var2csa* in blue.

Tighter *var* gene expression control by repressive histone marks

To better understand the role of epigenetics in *var* gene expression in our ΔV2 line, we assessed the distribution of repressive histone mark H3K9me3 and examined whether ΔV2 displays differentially bound heterochromatin due to *var2csa* deletion. We performed ChIP-seq experiments in duplicate on ΔV2 and NF54 WT trophozoite stage parasites (~24 hpi) using anti-H3K9me3 antibodies to identify whether the downregulated *var* genes strongly correlate with repressive histone marks. Our analysis confirmed an abundance of H3K9me3 on both subtelomeric and internal *var* gene regions in each chromosome. This result correlates well with previous evidence suggesting *var* gene-containing regions are tightly controlled in heterochromatin cluster(s) **(Fig. 2.2A)**[4,7]. H3K9me3 is characteristically absent from other regions of the genome, confirming that a sizable percentage of the genome is available for openly permissive gene expression with the exception of members of the variably expressed gene families, including *var* gene loci[33].

Peak calling analysis found 198 and 227 significant consensus peaks for the WT and ΔV2, respectively, covering less than 10% of the entire genome. Of the significantly called peaks located within 1 kb of a gene coding region, ~25% of peaks are near *var* genes and an additional ~36% are located near *rifin* genes **(Supp. Data 2.4 and 2.5)**. While analyzing read coverage within different regions of *var* genes, we found that ΔV2 contains significantly more reads ($p < 0.05$) in the 500 bp 5’ region, exon 1 and exon 2 after normalizing for library size and region length and subtracting input reads, as determined by a Mann-Whitney U test **(Fig. 2.2B)**. These data support our transcriptomic analysis and indicate that the downregulation of *var* genes following *var2csa* knockout is associated with an increase in heterochromatin-associated histone marks. Furthermore, we do not see H3K9me3 expansion from upstream and downstream regions surrounding differentially expressed *var* genes, but rather there is an overall increase in H3K9me3 coverage of downregulated genes **(Fig. 2.2C)**[33]. Differential peak calling analysis found significant overlap between WT and ΔV2 peaks and only 12 differential binding sites **(Supp. Fig. 2.2)**. Of the 12 differentially called peaks, 4 reside within the telomeric or subtelomeric region of various chromosomes outside any coding or promoter sequences, and 4 others are found in a subtelomeric region of chromosome 2 that is frequently spontaneously lost in cultured parasites **(Supp. Table 2.6)**[34, 35]. This leaves only 4 broad coding regions that are differentially bound by H3K9me3 in ΔV2, suggesting that the loss of *var* gene expression following *var2csa* knockout is not solely due to an increase in repressive histone marks.

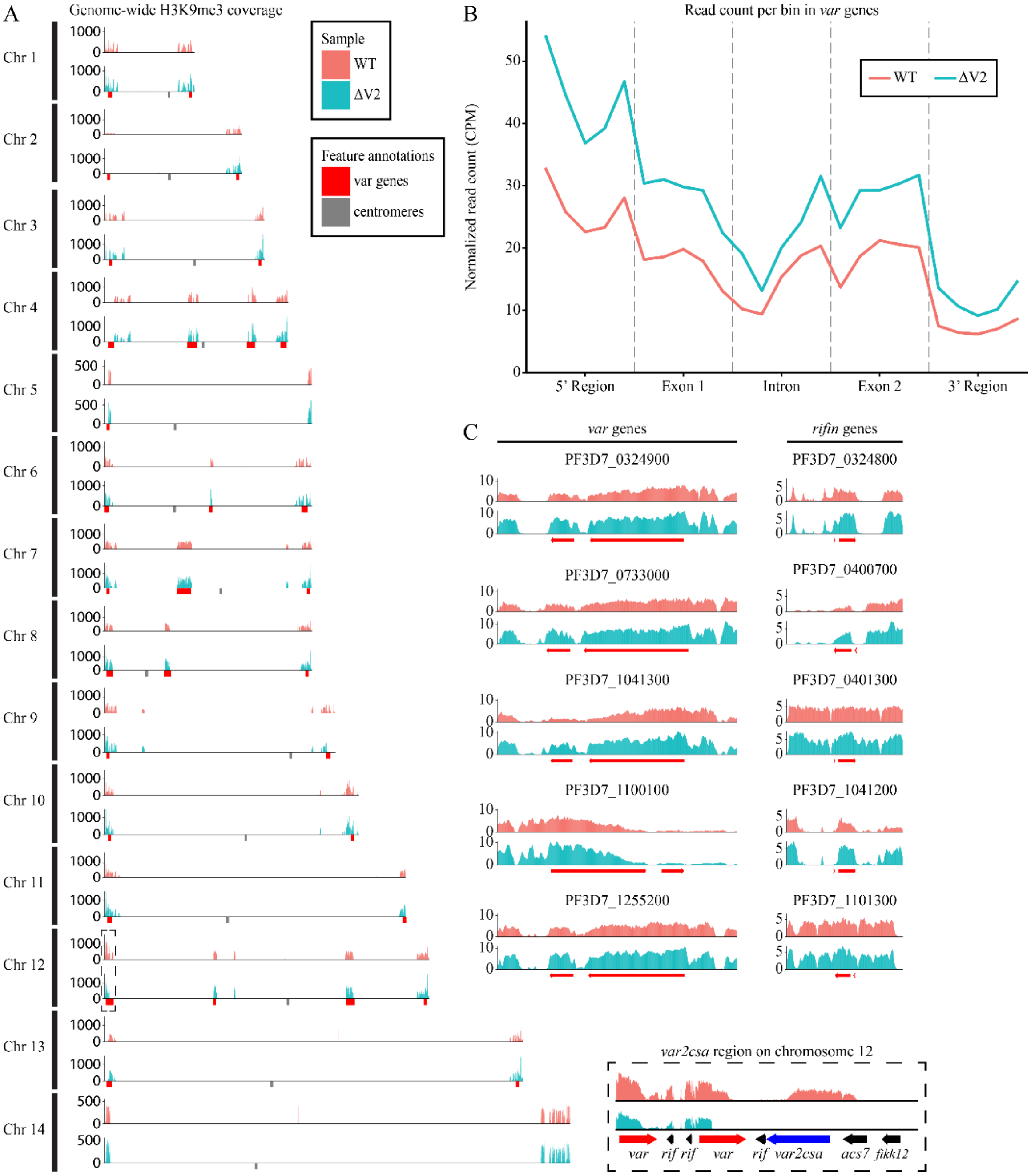

Figure 2.2. H3K9me3 is concentrated within *var* gene clusters and occupancy increased following *var2csa* disruption. (A) Genome-wide H3K9me3 coverage per-million normalized and input subtracted. The insert shows the *var2csa* region within the subtelomeric region of chromosome 12. *Var* gene clusters (red) and centromeres (gray) are highlighted along the x-axes. **(B)** Binned coverage of *var* genes from 500 bp 5' of the TSS to 500 bp 3' from the end of the CDS. The 5' and 3' regions consist of five 100 bp bins. Exons and the intron for all *var* genes are divided into five equal sized bins and normalized by length. **(C)** H3K9me3 coverage of five downregulated *var* and *rifin* genes from 2 kb up- and downstream of the CDS.

Structural changes to chromatin architecture linked to *var* expression

To confirm the link between epigenetics and chromatin structure, we then performed Hi-C experiments on tightly synchronized trophozoite stage parasites (~24 hpi)[36]. Hi-C libraries of two biological replicates for each sample (WT and ΔV2) were sequenced to a mean depth ~56 million reads per replicate. After processing the libraries (aligning, pairing, mapping and quality filtering), we obtained ~15-25 million valid interaction pairs per sample and binned the data at 10 kb resolution. Biological replicates display a strong stratum adjusted correlation (SCC > 0.94) with stronger similarity among ΔV2 samples compared to the WT control **(Fig S2.4A)**. However, correlation between WT and ΔV2 samples was also strong (SCC > 0.80) suggesting concentrated structural differences in chromatin architecture. Due to strong correlation, biological replicates for each condition/genotype were merged to increase the data resolution for downstream analysis. Random sampling was then performed on the WT due to differences in sequencing depth and number of total valid interaction pairs. Interactions between telomeres and *var* gene-containing regions across all chromosomes, including internal *var* genes, are significantly higher than the background for both intrachromosomal and interchromosomal interactions in both WT and ΔV2 **(Fig 2.3A and Supp. Fig 2.5, 2.6 and 2.8A)**.

The difference in interaction counts between the background and subtelomeric regions in the wild type and ΔV2 is made even more apparent by the differential interaction analysis ($q < 0.05$), which revealed an increase in interchromosomal interactions between bins containing *var* genes and a general trend of decreased interactions between subtelomeric and internal *var* gene regions and other bins throughout the genome **(Fig 2.3A and Supp. Fig 2.7)**. This indicates a chromatin restructuring where telomeres are moving away from the rest of the genome and condensing into even tighter heterochromatin structures. Decreases in chromatin interactions are also detected near the centromeres, resulting in a loss of centromeric A compartments in ΔV2 **(Fig. 2.3B)**. The chromatin structure of the centromere is so distinct that principal component 1 (PC1) of the compartment analysis highlighted the B compartment surrounding the centromere in the WT rather than the dense heterochromatin *var* gene regions **(Supp. Fig. 2.8B)**. The interactions near *P. falciparum* centromeres resemble the jet-like projections found in mammalian chromatin[37]. These chromatin jets are small regions of open chromatin that permeate nearby dense B compartments and extend perpendicularly, similar to the X-shaped structures near annotated centromeres and internal *var* gene clusters in the WT heatmaps **(Fig. 2.3A and Supp. Fig. 2.5)**. Hi-C analysis of chromatin jets in cohesin-depleted human lymphocytes shows substantial weakening of cohesin-dependent loop-extrusion and loss of chromatin jets[37]. Thus far no studies have been conducted on jet-like structures in *Plasmodium spp.*; therefore, additional experiments will be required to identify the breadth of molecular factors responsible for centromere and chromatin maintenance which result in these features. However, we provide compelling evidence to suggest that expression of *var2csa* is a key component necessary for the maintenance of the overall chromatin structure. Increased interaction in telomeric regions may lead to a loss of the jet-like structures observed near centromeres and internal *var* genes in *P. falciparum*.

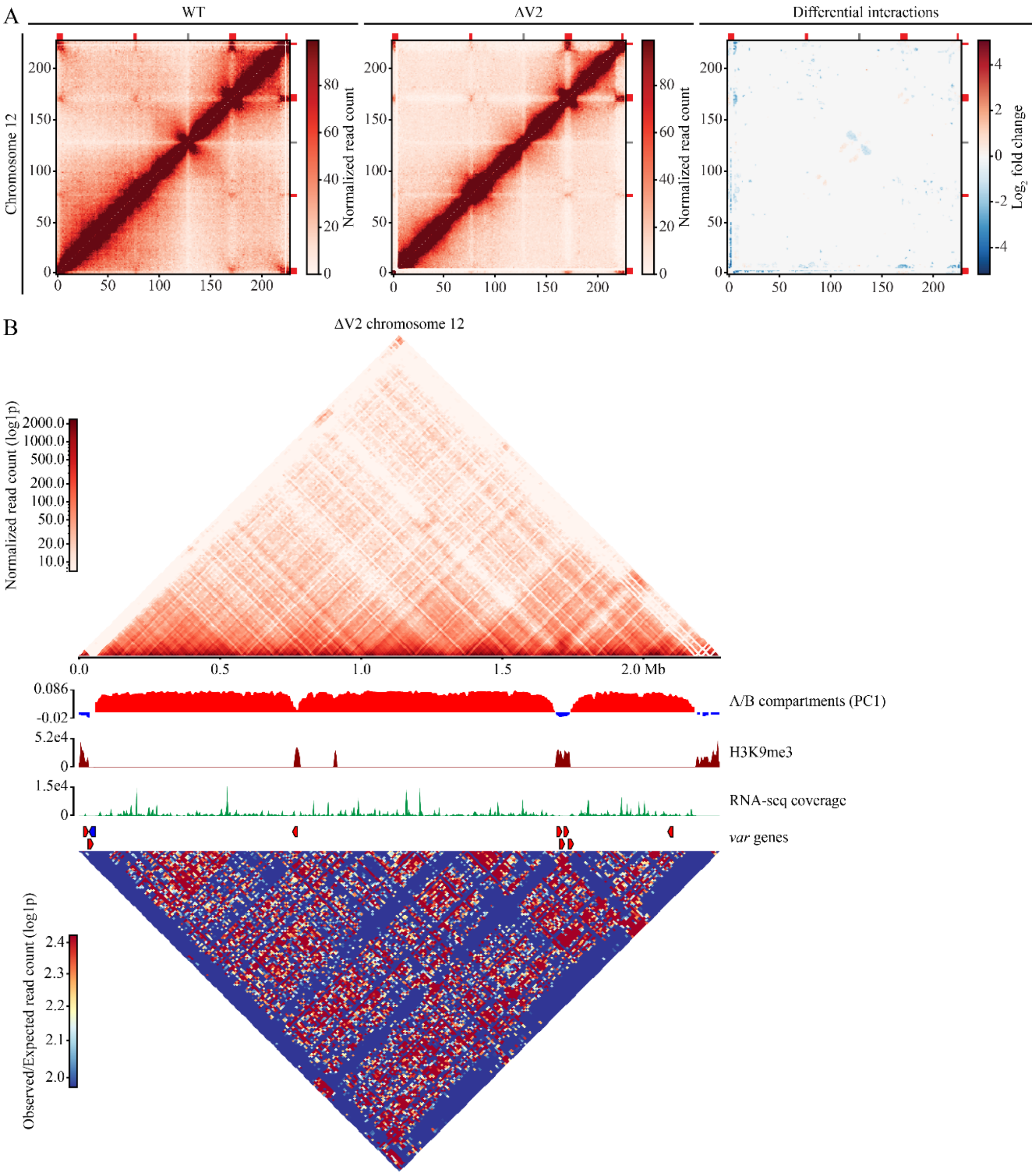

Figure 2.3. The targeted DSB resulted in a loss of chromatin interactions throughout most of the genome and compaction near *var* genes. (A) Hi-C contact count and differential interaction heatmaps for chromosome 12 binned at 10 kb resolution. The WT and ΔV2 interaction heatmaps are ICED normalized and the WT has been subsampled to the same number of reads as ΔV2. The color is scaled to the highest value in both datasets. **(B)** A series of plots for chromosome 12 of ΔV2, including log1p transformed ICED normalized read coverage, A/B compartment analysis, H3K9me3 coverage, RNA-seq transcript coverage, location of the *var* genes and plot of observed/expected interaction counts.

Chromatin 3D modeling reveals global structural differences

To gain a more comprehensive view of the chromatin architecture and identify structural changes resulting from deletion of *var2csa*, we generated 3D models using Pastis[38]. The goal was to identify regions that show clear structural changes that would indicate chromatin remodeling resulting from deletion of *var2csa*, and perhaps the resulting recombination events, which would likely result in more compacted chromatin structure within *var* gene regions. Our models are consistent with previous work that utilized 3D chromatin modeling in that our models show centromeres and telomeres clustered into two separate distinct regions **(Fig. 2.4)**[7, 30, 31]. We see that compared to wildtype parasites, ΔV2 shows a more compacted chromatin structure involving *var* regions, with telomeres concentrated into closed structures with a decreased distance between telomeric *var* gene bins. Compaction of the telomeres in closed structures prevents access to transcriptional machinery and thus suppresses their expression, which is supported by the RNA-seq data showing downregulation of several subtelomeric *var* genes.

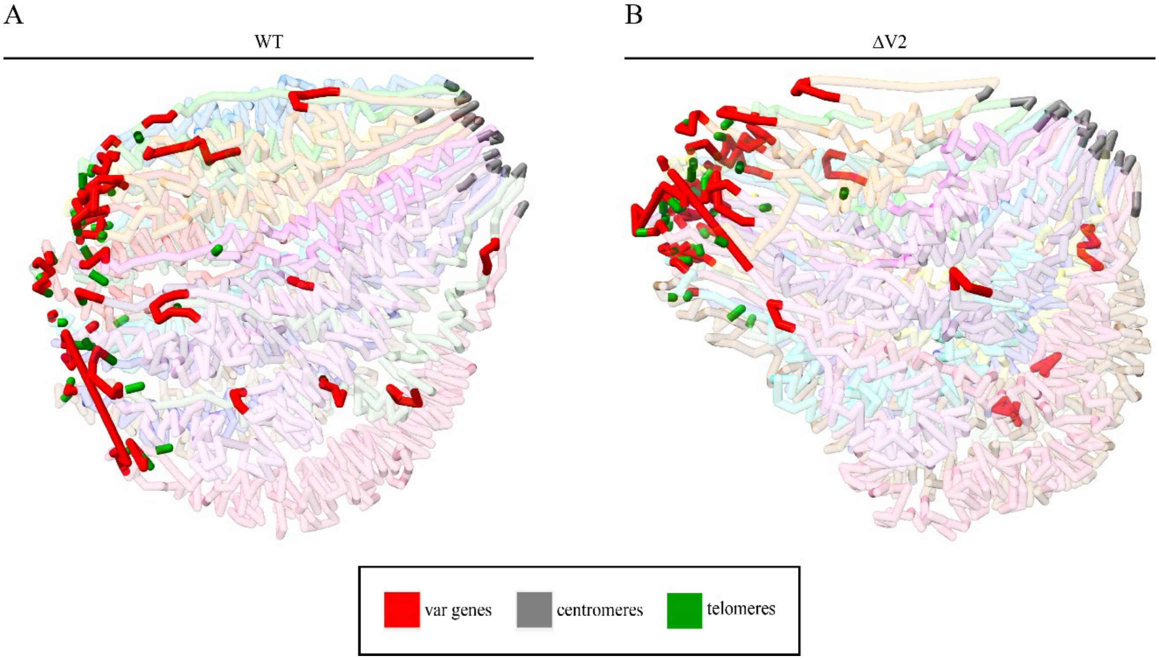

Figure 2.4. Genome-wide 3D models. (A, B) 3D chromatin model of **(A)** WT and **(B)** ΔV2 with *var* genes (red), centromeres (gray), and telomeres (green) highlighted.

Structural changes are consistent between *var2csa* knockout clones

To confirm that the results obtained were not specific to our selected *var2csa* knockout and subsequent recombination observed at the genomic level, we evaluated additional clones of *var2csa* knockout[14, 32]. Although some variation was observed between lines, a similar pattern was observed for all clones analyzed with overall increased interactions between telomeres and *var* genes indicating that a compacted heterochromatin structure is consistent between all *var2csa* knockouts analyzed compared to WT lines **(Supp. Fig. 2.9 and 2.10)**. This result suggests that knockout of *var2csa* is linked to a more condensed heterochromatin center and likely the reason the parasite is unable to significantly upregulate a single *var* gene.

DISCUSSION

Conservation of the highly variable *var* gene family within subtelomeric regions throughout the *P. falciparum* genome is imperative for adhesion of infected erythrocytes to receptors on endothelial cells. Expression of PfEMP1 and excessive cytoadherence of iRBCs within the microvasculature of various organs increases the severity of symptoms in malaria infections. Erythrocytes infected by VAR2CSA-expressing parasites bind to syncytiotrophoblasts within the placenta of pregnant women and can induce symptoms in previously immunocompetent women and cause significant damage to the fetus. The *var2csa* gene is unique among the *var* gene repertoire due to its specialized receptor preference and highly conserved sequence between parasite genomes, as well as being both transcriptionally and translationally regulated. Low-level transcription of *var2csa* in non-pregnant populations allows the parasite to be primed for proliferation within the placental tissue when infecting a pregnant woman while also not wasting energy on translation.

Maintenance of functional *var* genes aids the parasite in surviving the host immune response and avoiding splenic clearance through intermittent *var* gene switching, preventing immune recognition of the PfEMP1 surface proteins on iRBCs. Due to the importance of *var2csa* and maintenance of functional expression of PfEMP1, it is important to fully understand the underlying mechanisms controlling *var* gene expression and antigenic variation. The CRISPR-Cas9 targeted DSB within the promoter region of *var2csa* resulted in a complete loss of the locus. DNA repair via a cascade of non-homologous recombination events with the telomeres of other chromosomes produces chimeric *var* genes and ultimately supports parasite survival[30]. Previous work indicated that loss of *var2csa* reduced *var* transcriptional switching and led to a reduction in detectable *var* transcripts[14]. Our analysis confirmed a moderate downregulation of several *var* genes and other nearby subtelomeric genes. Differential expression analysis indicated that *var2csa* deletion affected transcription of subtelomeric genes almost exclusively, while transcription within the rest of the genome remained steady.

A model of promoter competition is frequently described as a possible mechanism underlying mutually exclusive expression within multicopy gene families in model organisms[39, 40]. We have similarly proposed that competition between *var* promoters could play an important role in selecting which *var* gene is active as well as contributing to *var* gene switching[41, 14]. The *var2csa* promoter appears to be unique in its competitiveness compared to other *var* promoters[21, 42], and thus its presence in the genome could serve to maintain expression levels of other *var* promoters within a range that is conducive to periodic switching. Its loss therefore could upset this balance, leading to reduced switching and lower levels of the *var* "minor" transcripts associated with *var* promoter activity. Consistent with this model, our Hi-C experiments showed that compared to the wildtype line, the *var2csa* deleted line displayed greater compaction of heterochromatin with increased interchromosomal interactions across subtelomeric *var* gene regions. Compartment analysis also highlighted chromatin modifications near the centromere and loss of jet-like structures. The 3D chromatin models further emphasized the clear compaction of subtelomeric regions, which is in accordance with increased H3K9me3 silencing and downregulation of subtelomeric *var* gene expression.

Our data confirm the importance of chromatin maintenance and *var2csa* expression in *var* gene regulation, building on previous research highlighting several key molecular mechanisms. Changes in histone acetylation, regulated in part by the histone deacetylase PfSir2, have been implicated in *var* gene regulation[43-45], as has methylation of H3K36[5, 46] and association of long noncoding RNAs with *var* loci[5, 47, 48]. We observed increased compaction of subtelomeric regions in response to *var2csa* deletion, an effect possibly related to telomeric position, something not studied extensively in *Plasmodium*[45]. Importantly, increased availability of the universal methyl donor S-adenosyl methionine in response to environmental changes can lead to increased H3K9me3 in genomic regions found in heterochromatin[49, 42]. This in turn leads to activation of *var2csa*[42]. Thus, these molecular pathways must intersect in ways that lead to mutually exclusive expression and coordinated switching. Detection of additional molecular factors, proteins or RNAs, interacting with this locus are needed to fully comprehend at the mechanistic level sequential coding and non-coding components required for *var* gene expression.

MATERIALS AND METHODS

Samples and culturing

Parasite lines of NF54 isolates, one WT control line contributed by Megan G. Dowler and obtained through BEI Resources, NIAID, NIH: *Plasmodium falciparum* strain NF54 (Patient Line E), Catalog No. MRA-1000; and one line containing CRISPR/Cas9 targeted DSBs in a subtelomeric region of chromosome 12 (ΔV2). The parasites were maintained at ~5-10% parasitemia in human erythrocytes at 5% hematocrit. Cultures were synchronized twice at the ring stage ~8 hr apart using 5% sorbitol and then samples collected during the trophozoite stage at ~6% parasitemia.

RNA library preparation

Infected erythrocytes were treated with 0.15% saponin then flash frozen. Total RNA was extracted using TRIzol LS Reagent (Invitrogen) and chloroform and incubated overnight at -20°C in isopropanol. Samples were then treated with 4 units of DNase I (NEB) for 1 hr at 37°C. mRNA was then purified using NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, E7490) according to manufacturer's instructions. Libraries were prepared using NEBNext Ultra Directional RNA Library Prep Kit (NEB, E7420) and PCR amplified (15 min at 37°C, 12 cycles of 30 sec at 98°C, 10 sec at 55°C, and 12 sec at 62°C, then 5 min at 62°C) with KAPA HiFi HotStart Ready Mix (Roche). Libraries were quantified by Bioanalyzer (Agilent) and sequenced using the NovaSeq 6000 (Illumina) and S4 300 flow cell for paired-end reads.

RNA-seq data processing and differential expression analysis

Sample quality of the RNA-seq libraries was assessed via FastQC (v.0.11.9) and per-base sequence quality and content was used to determine read trimming length. The index adapters were trimmed using Trimmomatic (v0.39)[50]. Paired and trimmed reads were then aligned to the *P. falciparum* 3D7 genome (PlasmoDB v58) using HISAT2 (v2.1.0)[51]. Non-properly mapped or paired reads and reads with a quality score less than

30 were filtered using Samtools[52]. Reads were mapped to protein coding genes using HTseq (v1.99.2)[53]. DESeq2 (v1.32.0)[54] was used for differential expression analysis.

ChIP-seq library preparation

Infected erythrocytes were collected in 1X PBS and lysed using 0.15% saponin. Chromatin was crosslinked using 1% formaldehyde then extracted with nuclear extraction buffer (10 mM HEPES, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 2 mM AEBSF, 1X protease inhibitor and 1X phosphatase inhibitor) and homogenized via 26.5-gauge needle. DNA was sheared to ~200-700 bp using a Covaris S220 sonicator for 8 min (5% duty cycle, 140 PIP, 200 cycles per burst). Immunoprecipitation was performed with IGG and anti-H3K9me3 antibodies after aliquoting samples for non-immunoprecipitated inputs. Immunoprecipitated DNA was extracted with phenol/chloroform, isoamyl (25:24:1). Libraries prepared using the KAPA Library Preparation Kit (KAPA Biosystems) and amplified using NEB index adapters with 12 PCR cycles (15 sec at 98°C, 30 sec at 55°C, and 30 sec at 62°C). Libraries were quantified by Bioanalyzer (Agilent), then sequenced using the NovaSeq 6000 (Illumina) and S4 300 flow cell for paired-end reads.

ChIP-seq data processing and peak calling

Paired-end ChIP-seq libraries were analyzed by FastQC (v0.11.9) for quality assessment. Adapters and low-quality bases were trimmed using Trimmomatic (v0.39)[50]. Bowtie2[55] was used to align paired reads to the P. falciparum 3D7 genome (PlasmoDB v58). PCR duplicates were tagged using Picardtools (v2.26.11)[56]. Samtools[52] was utilized to filter, sort and index the aligned, de-duplicated reads with a mapping quality cutoff of 40, keeping only mapped and properly paired reads. Reads were mapped to the genome at 1-bp resolution using Samtools. Per-million normalized inputs were subtracted from H3K9me3 samples. Broad peaks were called using MACS3 (v3.0.0a7)[57] (FDR < 0.05). Differential peak calling was performed using DiffBind (v3.2.7)[58]. Reads were mapped to binned regions within *var* genes (Fig. 2.2B). The 5' and 3' regions were split into five 100 bp bins and exons and introns were split into five bins of equal length, then each region was length normalized and samples per-million normalized before merging replicates. The merged input was then subtracted from the merged H3K9me3.

Hi-C library preparation

Parasite chromatin was crosslinked with 1.25% formaldehyde in warm 1X PBS. Cells were incubated in lysis buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 2 mM AEBSF, 0.10% Igepal CA-360 (v/v), and 1X protease inhibitor cocktail) then homogenized with a 26-gauge needle. Crosslinked DNA was digested using MboI (NEB) restriction enzyme then 5' overhangs were filled in by dNTPs with Biotin-14-dCTP (Invitrogen) using DNA Polymerase I and Large (Klenow) Fragment (NEB). Blunt-ends were ligated with 4000 units T4 DNA Ligase and chromatin de-crosslinked using de-crosslinking buffer (50 mM Tris-HCl at pH 8.0, 1% SDS, 1 mM EDTA, and 500 mM NaCl). Biotinylated DNA was sheared to 300-500 bp using a Covaris S220 (settings: 10% duty factor, 140 peak incident power, and 200 cycles per burst for 65 seconds) and biotinylated DNA fragments were pulled down using MyOne Streptavidin T1 beads (Invitrogen). End repair, A-tailing, and adapter ligation were all performed in lo-bind tubes using the KAPA library preparation kit (KAPA biosystems). Library was amplified using the HiFi HotStart ReadyMix (KAPA Biosystems) with a universal forward primer and barcoded reverse primer and incubated following the PCR program: 45 sec at 98°C, 12 cycles of 15 sec at 98°C, 30 sec at 55°C, 30 sec at 62°C and a final extension of 5 min at 62°C. The library was then purified using double-SPRI size selection, with 0.5Å~ right-side selection (25 µl AMPure XP beads) and 1.0Å~ left-side selection (25 µl AMPure XP beads). Libraries were quantified by NanoDrop (Thermo Scientific) and Bioanalyzer (Agilent), prior to multiplexing and sequencing in a 150-bp paired-end run on a NovaSeq 6000 (Illumina).

Hi-C Data Analysis

Paired-end HiC library reads were processed (i.e., mapping, filtering, and normalization) using HiC-Pro[59] and the *P. falciparum* 3D7 genome (PlasmoDB v58) with mapping quality cutoff of 30 and binning at 10 kb resolution. Intra-chromosomal and inter-chromosomal ICED-normalized matrices were per-million normalized before generating interaction heatmaps, with 2 bins from the diagonal set to 0 to enhance visualization. Differential interaction analysis was conducted at 10 kb resolution using Selfish while removing sparse regions[60]. HiCExplorer was used to identify and plot A/B compartments and O/E plots[61]. 3-dimensional chromatin remodeling was performed by generating 3-D coordinate matrices using Pastis[38] and then visualizing using ChimeraX[62].

FUNDING

This work was funded by the National Institutes of Allergy and Infectious Diseases and the National Institutes of Health (R21 AI142506-01, R01 AI136511).

AUTHOR CONTRIBUTIONS

Todd Lenz: Methodology, Investigation, Data curation, Formal analysis, Software, Validation, Visualization, Writing – Original draft, Writing – Review & editing. **Xu Zhang**: Methodology. **Abhijit Chakraborty**: Software. **Abbas Roayaei Ardakany**: Software. **Jacques Prudhomme**: Resources. **Ferhat Ay**: Software. **Kirk Deitsch**: Methodology, Writing – Original draft, Writing – Review & editing. **Karine Le Roch**: Project administration, Validation, Supervision, Writing – Original draft, Writing – Review & editing.

CONFLICT OF INTEREST

All contributing authors declare no conflict of interests nor competing interests.

DATA AVAILABILITY

All sequence files (fastq) used in analyses within in this study are available in the NCBI sequence read archive (SRA) under the project accession number PRJNA1069851.

SUPPLEMENTARY DATA

Supplementary data and figures for this chapter are available at doi.org/10.1101/2024.02.13.580059.

REFERENCES

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

2. Baruch, D. I. et al. Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell* **82**, 77-87 (1995).

3. Smith, J. D. et al. Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101-110 (1995).

4. Lopez-Rubio, J. J., 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* **5**, 179-190 (2009).

5. Jiang, L. et al. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. *Nature* **499**, 223-227 (2013).

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

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

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

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

10. Dzikowski, R., Frank, M. & Deitsch, K. Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog.* **2**, e22 (2006).

11. Roberts, D. J. et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* **357**, 689-692 (1992).

12. Gatton, M. L., Peters, J. M., Fowler, E. V. & Cheng, Q. Switching rates of Plasmodium falciparum var genes: faster than we thought? *Trends Parasitol.* **19**, 202-208 (2003).

13. Recker, M. et al. Antigenic variation in Plasmodium falciparum malaria involves a highly structured switching pattern. *PLoS Pathog.* **7**, e1001306 (2011).

14. Zhang, X. et al. A coordinated transcriptional switching network mediates antigenic variation of human malaria parasites. *eLife* **11**, e83840 (2022).

15. Pourmorady, A. D. et al. RNA-mediated symmetry breaking enables singular olfactory receptor choice. *Nature* **625**, 181-188 (2024).

16. Fried, M. & Duffy, P. E. Adherence of Plasmodium falciparum to chondroitin sulfate A in the human placenta. *Science* **272**, 1502-1504 (1996).

17. Salanti, A. et al. Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. *J. Exp. Med.* **200**, 1197-1203 (2004).

18. Trimnell, A. R. et al. Global genetic diversity and evolution of var genes associated with placental and severe childhood malaria. *Mol. Biochem. Parasitol.* **148**, 169-180 (2006).

19. Duffy, M. F. et al. Transcribed var genes associated with placental malaria in Malawian women. *Infect. Immun.* **74**, 4875-4883 (2006).

20. Lavstsen, T. et al. Expression of Plasmodium falciparum erythrocyte membrane protein 1 in experimentally infected humans. *Malar. J.* **4**, 21 (2005)

21. Mok, B. W. et al. Default Pathway of var2csa switching and translational repression in Plasmodium falciparum. *PLoS ONE* **3**, e1982 (2008).

22. Amulic, B., Salanti, A., Lavstsen, T., Nielsen, M. A. & Deitsch, K. W. An upstream open reading frame controls translation of var2csa, a gene implicated in placental malaria. *PLoS Pathog.* **5**, e1000256 (2009).

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

24. Noble, R. et al. The antigenic switching network of Plasmodium falciparum and its implications for the immuno-epidemiology of malaria. *eLife* **2**, e01074 (2013).

25. Ukaegbu, U. E. et al. A Unique Virulence Gene Occupies a Principal Position in Immune Evasion by the Malaria Parasite Plasmodium falciparum. *PLoS Genet.* **11**, e1005234 (2015).

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

27. Bachmann, A. et al. Highly co-ordinated var gene expression and switching in clinical Plasmodium falciparum isolates from non-immune malaria patients. *Cell Microbiol.* **13**, 1397-1409 (2011).

28. Kyes, S. A. et al. A well-conserved Plasmodium falciparum var gene shows an unusual stage-specific transcript pattern. *Mol. Microbiol.* **48**, 1339-1348 (2003).

29. Salanti, A. et al. Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Mol. Microbiol.* **49**, 179-191 (2003).

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

31. 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.* **24**, 974-988 (2014).

32. Zhang, X. et al. Rapid antigen diversification through mitotic recombination in the human malaria parasite Plasmodium falciparum. *PLoS Biol.* **17**, e3000271 (2019).

33. Michel-Todó, L. et al. Patterns of Heterochromatin Transitions Linked to Changes in the Expression of Plasmodium falciparum Clonally Variant Genes. *Microbiol. Spectr.* **11**, e0304922 (2023).

34. Horrocks, P., Kyes, S., Pinches, R., Christodoulou, Z. & Newbold, C. Transcription of subtelomerically located var gene variant in Plasmodium falciparum appears to require the truncation of an adjacent var gene. *Mol. Biochem. Parasitol.* **134**, 193-199 (2004).

35. Figueiredo, L. M., Freitas-Junior, L. H., Bottius, E., Olivo-Marin, J. C. & Scherf, A. A central role for Plasmodium falciparum subtelomeric regions in spatial positioning and telomere length regulation. *EMBO J.* **21**, 815-824 (2002).

36. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).

37. Guo, Y. et al. Chromatin jets define the properties of cohesin-driven in vivo loop extrusion. *Mol. Cell* **82**, 3769-3780 (2022).

38. Varoquaux, N., Ay, F., Noble, W. S. & Vert, J. P. A statistical approach for inferring the 3D structure of the genome. *Bioinformatics* **30**, i26-i33 (2014).

39. Hutchinson, S. et al. The establishment of variant surface glycoprotein monoallelic expression revealed by single-cell RNA-seq of Trypanosoma brucei in the tsetse fly salivary glands. *PLoS Pathog.* **17**, e1009904 (2021).

40. Constância, M., Pickard, B., Kelsey, G. & Reik, W. Imprinting mechanisms. *Genome Res.* **8**, 881-900 (1998).

41. Dzikowski, R. & Deitsch, K. W. Active transcription is required for maintenance of epigenetic memory in the malaria parasite Plasmodium falciparum. *J. Mol. Biol.* **382**, 288-297 (2008).

42. Schneider, V. M. et al. The human malaria parasite Plasmodium falciparum can sense environmental changes and respond by antigenic switching. *Proc. Natl Acad. Sci. USA* **120**, e2302152120 (2023).

43. Tonkin, C. J. et al. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. *PLoS Biol.* **7**, e84 (2009).

44. Duraisingh, M. T. et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. *Cell* **121**, 13-24 (2005).

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

46. Ukaegbu, U. E. et al. Recruitment of PfSET2 by RNA polymerase II to variant antigen encoding loci contributes to antigenic variation in P. falciparum. *PLoS Pathog.* **10**, e1003854 (2014).

47. Epp, C. et al. 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).

48. Amit-Avraham, I. et al. Antisense long noncoding RNAs regulate var gene activation in the malaria parasite Plasmodium falciparum. *Proc. Natl Acad. Sci. USA* **112**, E982-E991 (2015).

49. Harris, C. T. et al. Sexual differentiation in human malaria parasites is regulated by competition between phospholipid metabolism and histone methylation. *Nat. Microbiol.* **8**, 1280-1292 (2023).

50. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

51. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357-360 (2015).

52. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

53. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).

54. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

55. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-359 (2012).

56. Broad Institute. Picard toolkit (Broad Institute, GitHub repository, 2019); https://broadinstitute.github.io/picard/

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

58. Stark, R. & Brown, G. DiffBind: differential binding analysis of ChIP-Seq peak data (2011); https://bioconductor.org/packages/release/bioc/html/DiffBind.html

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

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

61. Ramírez, F. et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).

62. Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14-25 (2018).

CHAPTER 3

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

Zeinab Chahine[1,#], Mohit Gupta[1,†,#], **Todd Lenz**[1,#], Thomas Hollin[1,#], Steve Abel[1], Charles AS Banks[2], Anita Saraf[2,$], Jacques Prudhomme[1] , Suhani Bhanvadia[1] , Laurence A Florens[2] and Karine G Le Roch[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 – Email: karine.leroch@ucr.edu.

A version of this chapter has been published in Elife. 28, Aug. 2024. doi.org/10.1101/2023.09.11.557253. My role in this collaborative project focused specifically on the analysis of ChIP-seq and Hi-C datasets, generating figures from this data (Fig. 3.2, 3.5, and 3.6), and writing their respective results and methods sections.

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 with complementary molecular and genomics approaches, we demonstrate that *Pf*MORC 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 *Pf*MORC 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 *Pf*MORC impact histone marks and induces the collapse of parasite heterochromatin structure leading to its death. These findings confirm that *Pf*MORC 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 small 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)[1, 6]. The *Plasmodium* genome, however, encompasses over 5,000 protein-coding genes suggesting there are 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]. Together these components work in concert 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 investigate important these 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 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 and 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 specific loci to regulate chromatin accessibility, restricting sexual commitment[33, 35]. *Tg*MORC-depleted cells also resulted in changes 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 in *P. falciparum* identified *Pf*MORC as one of the highest enriched chromatin-bound proteins at various stages of the parasite intraerythrocytic development cycle (IDC)[28]. *Pf*MORC was also detected at relatively high levels throughout the parasite life cycle, including sporozoite and liver stages[33, 41], and has been identified in protein pull down experiments as targeting 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 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 IDC. Immunoprecipitation of HA-tagged parasites validated the role of *Pf*MORC in heterochromatin maintenance. Using the TetR-DOZI system we downregulated *Pf*MORC demonstrated the functional significance of *Pf*MORC in heterochromatin stability and gene repression. Immunofluorescence assays and ChIP-seq experiments showed that *Pf*MORC localizes to heterochromatin clusters near *var* genes with significant overlap of well-known heterochromatin signatures of H3K9-trimethylation (H3K9me3) marks. When *Pf*MORC was downregulated, H3K9me3 was detected at a lower level, demonstrating a possible role for the protein in epigenetic regulation and gene repression. Finally, Hi-C analyses demonstrated that downregulation of *Pf*MORC results in significant dysregulation of chromatin architecture stability resulting in parasite death. Altogether, our work provides significant insight into the role of *Pf*MORC for maintenance of epigenetic signatures and chromosome 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 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, Supp. Data 3.1)**. 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, Supp. 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 of 599. Mutations in *gdv1* have been detected in the past and seem to emerge relatively often indicating a fitness benefit for *in vitro* culture systems[8]. The involvement of GDV1 as essential in sexual commitment is well established, but the precise role of *Pf*MORC in sexual commitment has yet to be fully addressed[7,8,14,46-48] **(Supp. Data 3.2)**. Despite this, 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 **(Supp. Fig. 3.1)**. Regardless of potential protein degradation, results showed an expected band size at ~290 kDa that was absent in the NF54 control compared to our tagged line. Immunofluorescence assay (IFA) of intracellular parasites revealed that *Pf*MORC 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 more diffuse foci could be detected as the number of DNA copies increased at the schizont stage. Moreover, punctuates had strong colocalization signals with histone H3K9me3 marks present at different parasite developmental stages. This was most evident during the early ring stage when chromatin is more compact with a single DNA copy. Altogether, our findings indicated that *Pf*MORC is a true nuclear protein likely associated with 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 clustering, we performed immunoprecipitation (IP) followed by mass-spectrometry (MS) analysis using mature stages of *P*fMORC-HA and parental control line. 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 peptide (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 large size of the MORC protein (295kDa) and 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. To define a set of *Pf*MORC associated proteins, we used the QPROT statistical framework[49] to compare proteins detected in *Pf*MORC samples versus controls. We identified 73 *Pf*MORC associated proteins ($\log_2$ FC ≥ 2 and Z statistic > 5) **(Fig. 3.1D, Supp. Data 3.3)**. 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 N-acetyltransferase (PF3D7_1020700), HDAC1 (PF3D7_0925700), and SET3 (PF3D7_0827800). Several associations 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 found that many 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, leading to similar enrichment levels. Two of these three proteins have been reported to interact with MORC in several studies, further supporting a strong interaction. 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 detected SIP2, which engages in heterochromatin formation and telomere regulation[51], and 8 AP2 transcription factors, including AP2-G2, AP2-O2, AP2-G5. Among other AP2s detected, 5 were previously identified by ChIP-seq as *P. falciparum* AP2 Heterochromatin-associated Factors (*Pf*AP2-HFs)[52]. Similar to previous work, our findings indicate that *Pf*MORC interacts with multiple ApiAP2 TFs, chromatin remodelers, and epigenetic players associated with heterochromatin regions.

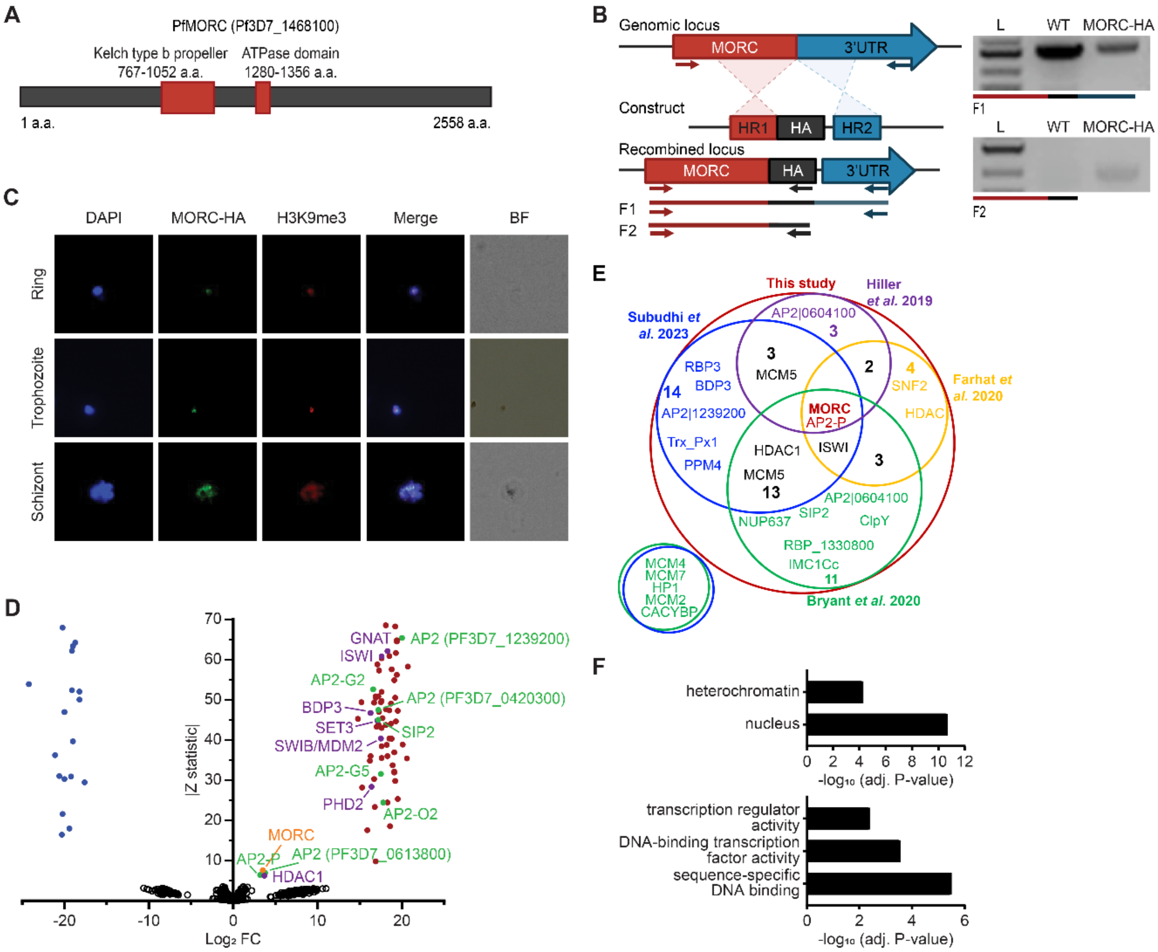

Figure 3.1. *Pf*MORC-HA is associated with heterochromatin. (A) Illustration of *Pf*MORC containing domains including Kelch type b propeller and ATPase domains using InterProScan. **(B)** Design strategy applied for *Pf*MORC C-terminal HA tagging. PCR amplification of the genomic C-terminus end of *Pfmorc* region extending towards the 3'UTR (F1) as well as extension from the C-terminus toward the HA flanking sequence (F2) verifies the correct insertion site. NF54 genomic DNA was used as a negative control. **(C)** IFA of *Pf*MORC foci 22 expressing colocalization with H3K9me3 (red). Cell nuclei are stained with DAPI (blue). BF: brightfield **(D)** Significance plot representing *Pf*MORC interactome recovered through protein immunoprecipitation followed by mass spectrometry (IP-MS). Graph shows MORC (orange) binding 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 ≥ 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, Supp. Data 3.4)**. 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, Supp. Fig. 3.2B)**.

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, Supp. Fig. 3.2C-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 **(Supp. Fig. 3.2D)**. *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.

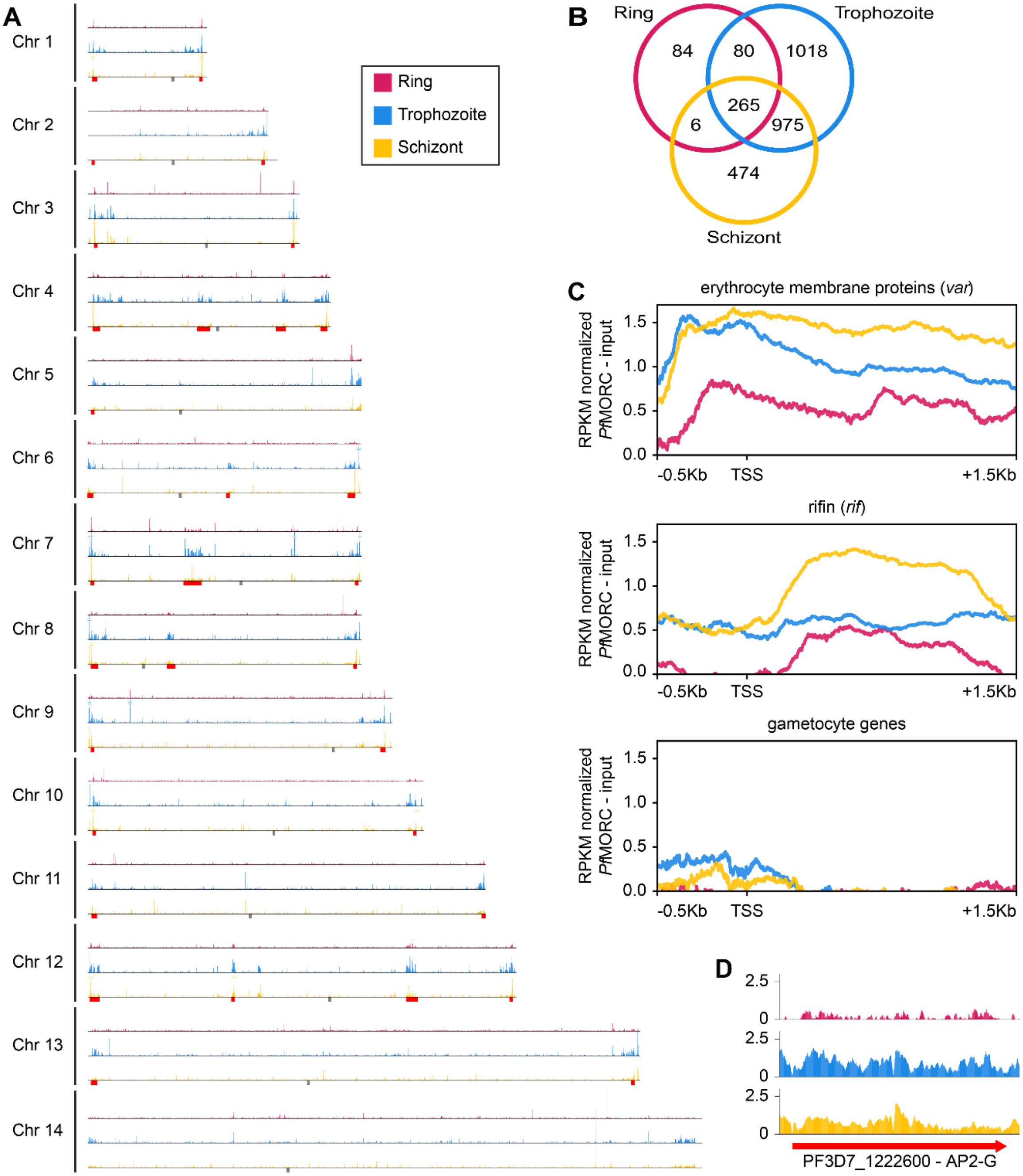

Figure 3.2. Genome wide distribution of *Pf*MORC proteins. (A) Chromosome distribution of *Pf*MORC binding shows a predisposition for subtelomeric and internal *var* gene regions (red). Each track is input subtracted and per-million normalized. Grey boxes indicate the position of centromeres. **(B)** Overlap of called ChIP peaks between time points. **(C)** Profile of *Pf*MORC coverage from 0.5 kb 5' of the TSS to 1.5 kb 3' of the TSS in the ring, trophozoite, and schizont stages. Each plot includes per-million 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 **(Supp. Data 3.2)** 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 **(Supp. Fig. 3.3)**. 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, Supp. data 3.5)**. 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, Supp. Data 3.5)**, confirming the importance of the protein at the trophozoite and early schizont stages.

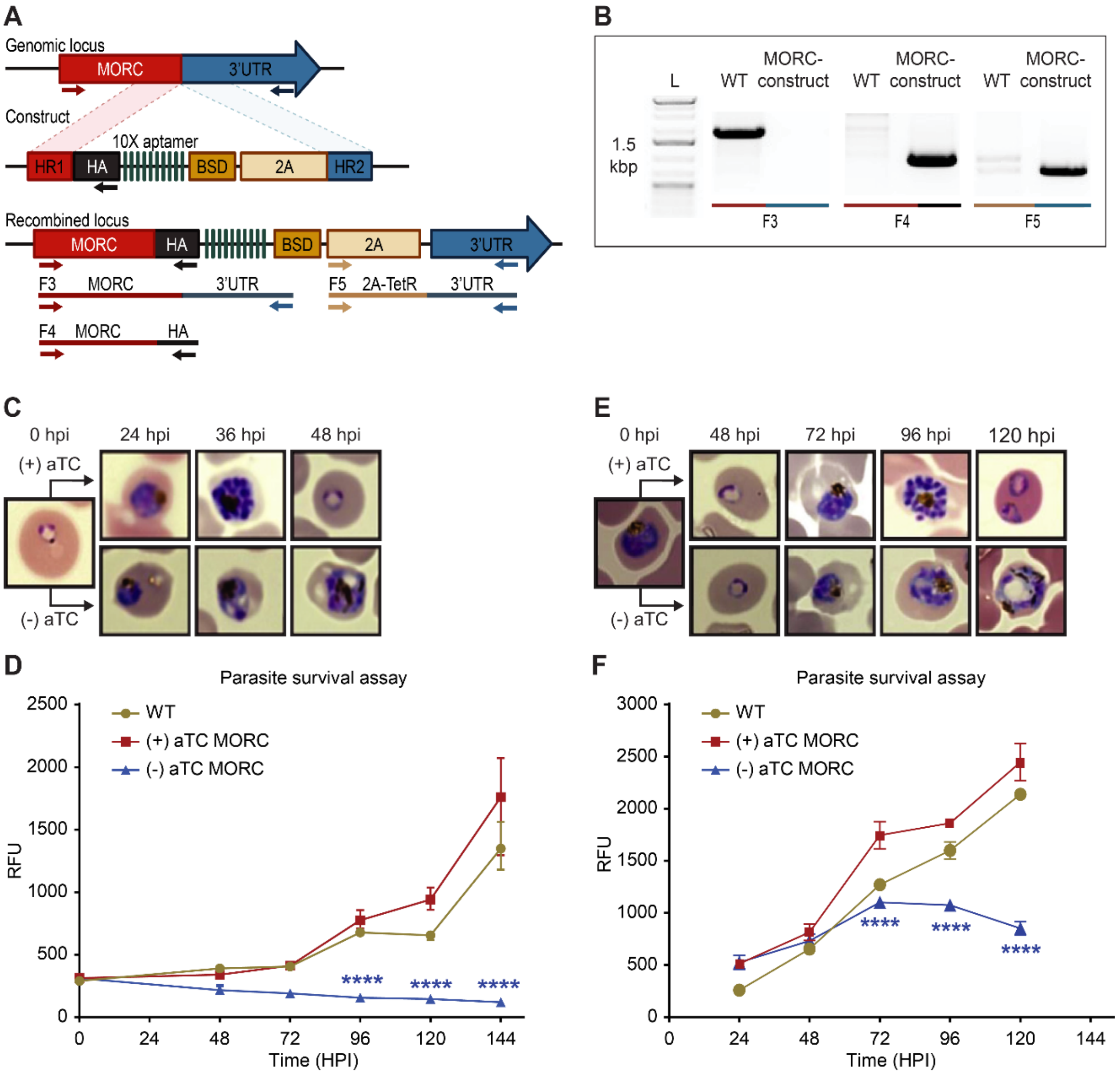

Figure 3.3. *Pf*MORC is essential for cell survival. (A) Diagram representation of *Pf*MORC-HA-TetRDOZI plasmid. **(B)** PCR amplification is used to verify genomic insertion using primer sets targeting 1.5 kb of WT *Pfmorc* genome locus absent in transgenic line (MORC construct) (F3) as well as verification of HA insertion (F4) and TetR-DOZI system extending along 3' UTR of the construct (F5). **(C)** Phenotypic and **(D)** quantitative analysis of parasite cell cycle progression after aTC withdrawal at the ring stage (0-6 hpi) (2-way ANOVA, n=3, $p \leq 0.0001$). **(E)** Phenotypic and **(F)** quantitative analysis of parasite cell cycle progression after aTC removal at the trophozoite stage (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 schizont (36 hpi) stages to allow for detection of early and late changes in gene expression after aTC removal. Pairwise correlations analysis **(Supp. Fig. 3.4A)** 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, Supp. Data 3.6)**. This data agrees 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 **(Supp. Fig. 3.3)**. It is highly possible that the observed down-regulated genes are due to an indirect effect of cell cycle arrest. At 36 hpi the number of genes exhibiting changes in gene expression were significantly higher with 1319 upregulated and 1150 downregulated (FDR < 0.05, $\log_2$FC > 0.5 or < -0.5) in (-) aTC *Pf*MORC, 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 to the surface of red blood cells that could be linked to significantly decreased *Pf*MORC binding of these gene families and heterochromatin cluster(s) disorganization at the trophozoite and schizont stages. GO enrichment analysis for downregulated genes 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 arrest in cell cycle progression at the trophozoite and schizont stages **(Supp. data 3.5)**[59].

Although these results are compelling, the small overlap 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)**.

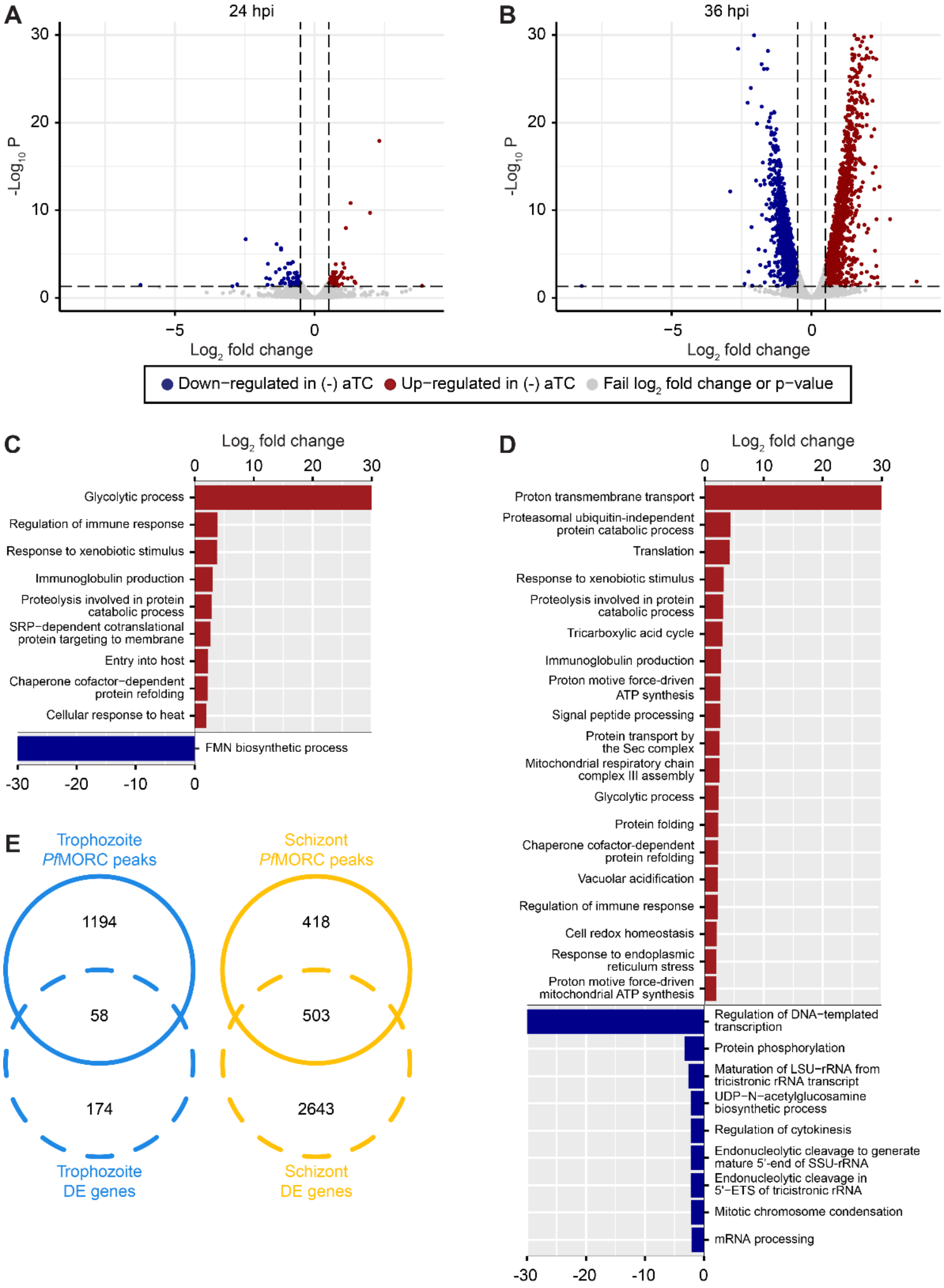

Figure 3.4. *Pf*MORC KD on parasite transcriptome. (A, B) Volcano plots denoting upregulated (red), and downregulated (blue) genes discovered through differential expression analysis following *Pf*MORC knockdown at **(A)** 24 hpi **(B)** 36 hpi. **(C, D)** Gene ontology enrichment analysis for upregulated (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 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, Supp. Data 3.6)**. 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.

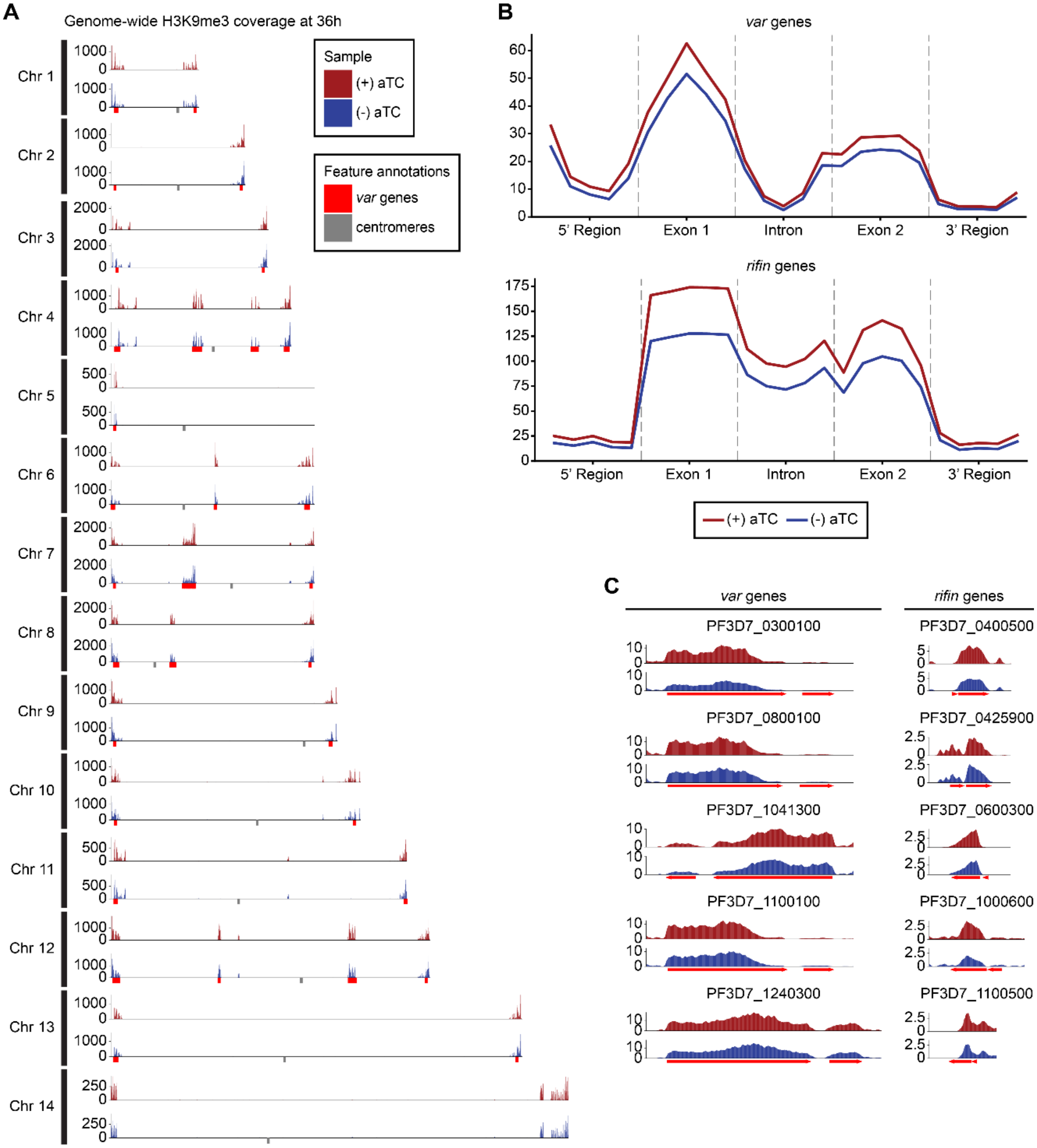

Figure 3.5. Impact of *Pf*MORC KD on heterochromatin markers. (A) H3K9me3 coverage of IGG subtracted and per-million normalized (+/-) aTC show similar distribution and concentration within antigenic gene clusters (red). Replicates (n = 2) are merged using the mean read coverage per bp. **(B)** Binned coverage of *var* and *rifin* genes from 1 kb 5' of the TSS to 1 kb 3' of the end. Exons and introns are split into five equal length bins and 5' and 3' regions are binned in five 200 bp bins. Read count within each bin is per-million and length normalized. **(C)** Coverage of the five most upregulated *var* and *rifin* genes as determined by 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 (σ = ~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 **(Supp. Fig. 3.4)**, 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, Supp. Fig. 3.6-3.9)** 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, Supp. Fig. 3.10-3.11)**. 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 (**Supp. 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.

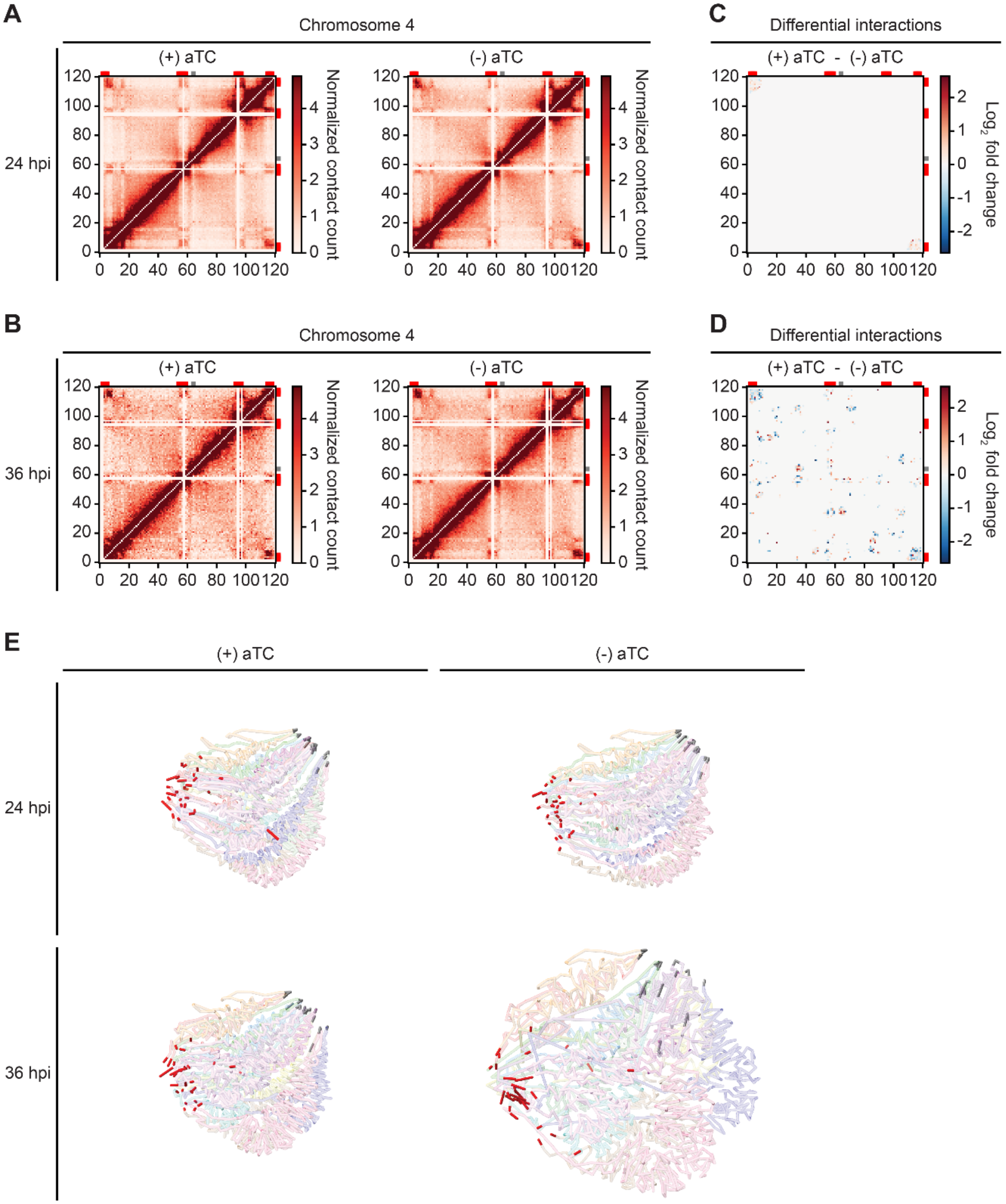

Figure 3.6. Loss of *Pf*MORC expression correlates with heterochromatin expansion. Intrachromosomal interaction heatmaps of (+/-) aTC *Pf*MORC for chromosome 4 at **(A)** 24 hpi and **(B)** 36 hpi displaying heterochromatin clustering within antigenic (*var, rifin* and *stevor*) gene-dense regions (red). 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 function 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 **(Supp. Data 3.1)** 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 **(Supp. Data 3.1)**. A 3x HA-tag was fused to the C-terminal 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 *Pf*MORC were amplified with respective primers **(Supp. Data 3.1)** 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 RBCs infected with 3-5% early ring stage parasites. After one asexual cycle, transfected cultures were supplemented with 1.5 µM WR99210 (2.6nM) (provided by 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, transfected parasites were serially diluted to approximately 0.5% parasite/well in 96 well plates.

Molecular analysis of the transgenic lines

Genomic DNA (gDNA) was extracted and purified using a DNeasy Blood & Tissue kit (Qiagen) following manufacturer instructions. Diagnostic PCR analysis was used to genotype transfected lines using the primers listed in Supplementary data 3.1. PCR amplification was conducted using KAPA HiFi HotStart ReadyMix (Roche) and amplicons analyzed by gel electrophoresis then sequenced. For whole genome sequencing, genomic DNA was fragmented using a Covaris S220 ultrasonicator and libraries generated using KAPA LTP Library Preparation Kit (Roche, KK8230). To verify that the insert was present and at the correct location of the genome 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 insert sequence at 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 (v0.7.17) 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 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. Samples were then 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 cultures were divided 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 lysed parasites after 30 min 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 then blocked and 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 resuspended in 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)). 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. Lysates were precleared with Dynabeads™ Protein A (Invitrogen) for 1h at 4°C. Anti-HA tag antibody (1:2500, Abcam, ab9110) was added to the control and HA tagged *Pf*MORC precleared protein extracts for 1hr at 4°C under constant rotation followed by addition of fresh Dynabeads™ Protein A beads 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[73].

Proteomics data processing and analysis

Tandem mass (MS/MS) spectra were interpreted using ProluCID v.1.3.3[74] against a database 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 using Contrast v1.9 and in-house developed sandmartin (v.0.0.1, https://github.com/tzw-wen/kite/tree/master/kitelinux)[75]. Our in-house developed software NSAF7 v.0.0.1 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 replicate *Pf*MORC 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 adding NP-40/Igepal to a final of 0.25%. After lysis, the parasites are homogenized 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, 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. The diluted chromatin was 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 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. 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, 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. Then 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 hrs 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 the DNA using KAPA LTP Library Preparation Kit (Roche, KK8230) and sequenced on a NovaSeq 6000.

ChIP-seq analysis

FastQC (v0.11.8) was used to analyze read quality. Adapter sequences were removed using Trimmomatic (v0.39). Bases with Phred scores below 20 were trimmed using Sickle (v1.33). Reads were mapped against the *P. falciparum* genome (v48) using Bowtie2 (v2.4.4). Using Samtools (v1.11), only properly paired reads with mapping quality 40 or higher were retained and PCR duplicates were removed by PicardTools MarkDuplicates (v2.18.0). Genome-wide read counts per nucleotide were normalized by dividing millions of mapped reads for each sample, then IgG read counts were subtracted from the anti-HA IP counts. Peak calling was performed using MACS3 (v3.0.0a7)[78] with the upper and lower limits for fold enrichment set to 2 and 50, respectively.

Phenotypic analyses

P. falciparum NF54 and transgenic *Pf*MORC-HA-TetR-DOZI lines were synchronized, grown for one cycle post-synchronization and subjected to 3 washes with 1X PBS before supplementing with or without (aTC) for 48 hrs. or 72 hrs. Culture media were replaced daily with or without aTC supplementation.

Quantitative Growth assay

Synchronous cultures of ring or trophozoite stage at 0.5 % parasitemia were grown with or without aTC (500 nM) in a 96 well plate in triplicate for each time point. Parasites were collected every 24 hrs. 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 cell 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 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 libraries were assessed by Bioanalyzer (Agilent Technology Inc). Libraries were sequenced using a NovaSeq 6000 (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, https://www.htslib.org/). 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 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. 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 using HiC-Pro[60] using a mapping quality cutoff of 30 while aligning to the *P. falciparum* genome (PlasmoDB, v58). Reads were binned at 10 kb resolution then ICED-normalized. Stratum-adjusted correlation coefficients were calculated using HiCRep[80], then replicates merged to generate a single consensus sample for each time point and condition. Normalized matrices were per-million normalized and heatmaps scaled on each chromosome to the maximum value for all samples to allow for direct visual comparisons between each condition and time point. Due to overrepresentation of short-range interactions, maximum values were set to the 90th percentile of each chromosome to compress highly interacting regions and enhance visualization. Significant intrachromosomal interactions were identified with FitHiC[81] to identify the log-linear relationship between contact probability and genomic distance. Differential interactions between (+) aTC and (-) aTC were identified using Selfish[62] with default parameters (FDR < 0.05). The maximum and minimum values in the differential heatmaps were set to the absolute value of the highest $\log_2$ fold change within each chromosome. PASTIS[63] was used to generate coordinate matrices and then visualized as 3D chromatin models in ChimeraX[82], highlighting *var* genes, telomeres, and centromeres.

Statistical analyses

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

ACKNOWLEDGEMENTS

This work was supported by NIH grants to KGLR (R01 AI136511 and R21 AI142506-01) and by the University of California, Riverside to KGLR (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).

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, MSV000092353).

AUTHOR CONTRIBUTIONS

Zeinab Chahine: Methodology, Investigation, Writing – Original draft, Writing – Review and editing. **Mohit Gupta**: Methodology, Investigation, Writing – Review and editing. **Todd Lenz**: Methodology, Software, Formal analysis, Visualization, Writing – Original draft, Writing – Review and editing. **Thomas Hollin**: Methodology, Formal analysis, Writing – Original draft, Writing – Review and editing. **Steve Abel**: Methodology, Software, Formal analysis, Writing – Review and editing. **Charles AS**

Banks: Methodology, Software, Formal analysis, Writing – Review and editing. **Anita Saraf**: Methodology, Writing – Review and editing. **Jacques Prudhomme**: Resources, Writing – Review and editing. **Suhani Bhanvadia**: Formal analysis. **Laurence A Florens**: Formal analysis, Supervision, Writing – Review and editing. **Karine G Le Roch**: Conceptualization, Formal analysis, Supervision, Writing – Original draft, Writing – Review and editing.

CONFLICT OF INTEREST

All contributing authors declare no competing interests.

SUPPLEMENTARY DATA

Supplementary data and figures for this chapter are available at doi.org/10.7554/eLife.92499.1.

REFERENCES

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

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

3. Balaji, S., Babu, M. M., Iyer, L. M. & 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.* **33**, 3994-4006 (2005).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

28. Batugedara, G. et al. The chromatin bound proteome of the human malaria parasite. *Microb. Genom.* **6**, 20 (2020).

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

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

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

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

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

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

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

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

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

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

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

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

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

44. Singh, M. K. et al. A Plasmodium falciparum MORC protein complex modulates epigenetic control of gene expression through interaction with heterochromatin. *eLife* **12**, e83840 (2023).

45. Ganesan, S. M. et al. Synthetic RNA-protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nat. Commun.* **7**, 10727 (2016).

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

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

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

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

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

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

52. 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.* **50**, 3413-3431 (2022).

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

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

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

56. Zhang, M. et al. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, eaap7847 (2018).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

72. Rajaram, K., Liu, H. B. & Prigge, S. T. Redesigned TetR-Aptamer System To Control Gene Expression in Plasmodium falciparum. *mSphere* **5**, e00457-20 (2020).

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

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

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

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

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

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

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

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

81. Ay, F., Bailey, T. L. & Noble, W. S. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res.* **24**, 999-1011 (2014).

82. Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14-25 (2018).

CHAPTER 4

The *Toxoplasma* rhoptry protein ROP55 is a major virulence factor that prevents lytic host cell death

Margarida T Grilo Ruivo[1], Ji-hun Shin[2ψ], **Todd Lenz**[3ψ], Stephanie Y Matsuno[2], Katherine Olivia Yanes[2], Arnault Graindorge[1], Maguy Hamie[4], Laurence Berry-Sterkers[1], Mathieu Gissot[5], Hiba El Hajj[4], Karine G Le Roch[3], Melissa B Lodoen[2], Maryse Lebrun[1*§], Diana Marcela Penarete-Vargas[1*§]

[1] Laboratory of Pathogens and Host Immunity, UMR 5294 CNRS, UA15 INSERM, Université de Montpellier, Montpellier 34095, France.
[2] Department of Molecular Biology and Biochemistry and the Institute for Immunology, University of California, Irvine, CA 92697, USA.
[3] Department of Molecular, Cell and Systems Biology, University of California Riverside, Riverside, CA 92521 USA.
[4] Department of Experimental Pathology, Immunology and Microbiology, American University of Beirut, Beirut, Lebanon.
[5] U1019 - UMR 9017, Univ. Lille, CNRS, INSERM, CHU Lille, Institut Pasteur de Lille, Center for Infection and Immunity of Lille, Lille F-59000, France.
[ψ §] These authors contributed equally to this work.
[*] Correspondence: Maryse Lebrun – maryse.lebrun@umontpellier.fr
[*] Correspondence:Diana Marcela Penarete-Vargas–diana.penaretevargas@umontpellier.fr

The manuscript for this chapter is currently under review. For this project, I performed the transcriptomic analysis and generated the related figures (Fig 4.4, Supp. Fig. 4.4).

ABSTRACT

Programmed-cell death is an antimicrobial defense mechanism that promotes clearance of intracellular pathogens. *Toxoplasma* counteracts host immune defenses by secreting effector proteins into host cells; however, how the parasite evades lytic cell death and the effectors involved remain poorly characterized. We identified ROP55, a rhoptry protein that promotes parasite survival by preventing lytic cell death in absence of IFN-γ stimulation. RNA-Seq analysis revealed that ROP55 acts as a repressor of host pro-inflammatory responses. In THP-1 monocytes ΔROP55 infection increased NF-κB p65 nuclear translocation, IL-1β production, and GSDMD cleavage compared to wild type or complemented parasites. ΔROP55 infection also induced RIPK3-dependent necroptosis in human and mouse primary macrophages. Moreover, ΔROP55 parasites were dramatically impaired in virulence in female mice and prevented NF-κB activation and parasite clearance in mBMDM. These findings place ROP55 as a major virulence factor, dampening lytic cell death and enabling *Toxoplasma* to evade clearance from infected cells.

INTRODUCTION

Toxoplasma gondii is an obligate intracellular parasite that is estimated to infect a third of the world human population, making it one of the most successful apicomplexan parasites[1, 2]. Although most infections are generally asymptomatic, toxoplasmosis poses a serious health threat to immunocompromised individuals and to the developing foetus[3]. The recent emergence of highly virulent strains causing diseases in immunocompetent individuals also reinforces its global importance[4, 5]. *Toxoplasma* can establish a chronic infection that persists for the lifetime of the host. To do so, the parasite has evolved several mechanisms to subvert the host's immune response. These strategies depend on an arsenal of parasite effector proteins that are secreted into the host cell from two distinct intracellular compartments: the rhoptries and the dense granules[6, 7]. These parasite effectors counteract host signaling pathways and dictate the variations in the virulence between *Toxoplasma* strains, the most common of which in Europe and North America are referred to as Type I, II and III[6-8].

Programmed cell death is a critical aspect of the innate immune response and controls infections by killing and clearing infected cells, thereby depriving intracellular pathogens of their replicative niche[9]. Cell death pathways can either induce cell lysis and inflammation, as observed in both pyroptosis and necroptosis, or engage a nonlytic death program that does not induce inflammation, such as apoptosis[10]. Pyroptosis occurs through the formation of multiprotein complexes called inflammasomes, in which inflammatory caspases play a central role for the maturation of IL-1 family cytokines and the initiation of cell death[11]. The activation of NLRP3, which is the most well-studied inflammasome, is a tightly regulated process requiring two distinct signals[11]. The first signal arises from PAMPs and DAMPs that activate pattern recognition receptors and NF-κB signaling, leading to a priming stage in which the inactive forms of NLRP3 and IL-1β are synthesized[12]. A second signal is then necessary for the maturation of IL-1β by caspase-1. This signal is triggered by disturbances to cellular homeostasis caused by multiple agonists; among them, ionic flux, reactive oxygen species, and mitochondrial dysfunction, inducing NLRP3 inflammasome assembly to promote caspase-1 activation[13]. Caspase-1 activity is also required for cleaving Gasdermin D (GSDMD), which, upon insertion in the plasma membrane, forms pores that induces cell swelling and osmotic rupture, resulting in pyroptotic cell death. Necroptotic cell death, on the other hand, is a caspase-independent program that can be initiated by diverse triggers including ligation of death receptors, Toll-like receptors (TLRs) or cytosolic nucleic acid sensors[10, 14]. These diverse stimuli converge to activate the serine/threonine protein kinase RIPK3, which, in turn, phosphorylates the pseudokinase MLKL. Active MLKL anchors in the plasma membrane and forms pores, leading to the release of damage-associated molecular patterns (DAMPs), cytokines and chemokines, and the lytic death of the cell[10, 14]. Depending on the trigger that initiates necroptosis, RIPK3 undergoes oligomerization with MLKL with or without the involvement of RIPK1. In either scenario, necroptosis can lead to NF-kB signaling and transcriptional activation[15-18] and can only proceed when caspase-8 is inhibited or insufficiently activated, since caspase-8 inhibits necroptosis by cleaving RIPK1 and RIPK3[10, 19]. *Toxoplasma* repression of lytic cell death pathways have emerged as an important parasite strategy to favor survival, assuring long term persistence. For example, *Toxoplasma* bradyzoites block interferon-induced necroptosis by the combined function of two dense granule proteins (TgNSM and TgIST)[20], contributing to survival of intracellular cysts. These findings suggest that necroptosis might be regulated in *Toxoplasma*-infected cells, but the extent to which these effectors or others play a role in vivo or during acute infection, remain unknown. In contrast, pyroptosis has been identified as the pathway associated with the resistance to chronic *Toxoplasma* infection in Lewis rats[21]. This resistance mechanism is linked to the activation of the NLRP1 sensor, resulting in host cell lysis and parasite clearance in vivo[22]. However, across different host species, *Toxoplasma* infection does not consistently lead to pyroptotic cell death. Indeed, inflammasome activation relies on the specific combination of host cell type and *Toxoplasma* strain. For instance, whereas infection of human THP-1 cells with type II parasites induces the nuclear translocation of

NF-κB, production of IL-1β and activation of NLRP3 inflammasome[23-25], infection with type I parasites does not[23, 26]. This difference due in part to allelic differences in the gene encoding GRA15, between the type I RH strain and type II strains[23, 26]. Additionally, type I RH parasites actively employ counter-measures to block the NF-κB/IL-1β signalling axis, as type I *Toxoplasma* infection inhibits LPS-induced NF-κB nuclear translocation and IL-1β production, dampening host inflammatory responses[27-29]. In neutrophils, this strategy of immune evasion was shown to be dependent on actively invading parasites[27, 30]. However, the identity of the parasite effectors involved in this suppression remains poorly characterized. Here we identify a new *Toxoplasma* rhoptry bulb protein, ROP55, that prevents the induction of host lytic cell death through the repression of NF-κB signaling. The absence of ROP55 triggers both pyroptosis and RIPK3-dependent necroptosis of infected cells. Furthermore, we show that deletion of the ROP55 gene leads to a significant decrease in parasite virulence in vivo. Our findings reveal the role of ROP55 as a major virulence factor in *Toxoplasma*, preventing lytic host cell death as a conserved mechanism of evading parasite clearance in human and mouse cells.

RESULTS

TGGT1_300220 is a new rhoptry protein, named ROP55.

Rhoptry genes display a characteristic expression pattern and can be identified through bioinformatic analysis[31]. Using this approach, we identified TGGT1_300220, encoding a hypothetical protein of 1823aa, with a predicted signal sequence and a mRNA periodicity typical of rhoptry proteins **(Fig. 4.1A)**. TGGT1_300220 is not predicted to be essential for growth in fibroblasts (fitness score of -0.15)[32], has no characterized domains using PROSITE[33], HHpred[34] or Phyre-2[35] searches, and appears to be restricted to coccidian parasites **(Supp. Fig. 4.1)**.

To determine its localization, we endogenously tagged TGGT1_300220 with a triple HA tag at its C-terminus in the type I background strain RH Δ*ku80* (Tg300220-HA3 line) **(Supp. Fig. 4.2A-B)** confirming its expected size of 206 kDa **(Fig. 4.1B)**. Immunofluorescence analysis (IFA) revealed a clear apical dot, suggestive of rhoptry localization that partially colocalized with the rhoptry neck protein RON2, but to a greater degree with ROP5, a rhoptry bulb protein **(Fig. 4.1C)**. This rhoptry bulb localization fits with the hyperLOPIT (spatial proteomics method hyperplexed Localization of Organelle Proteins by Isotopic Tagging) analysis[36] which predicted TGGT1_300220 to be associated with rhoptry fraction 1, which contains luminal proteins that are secreted and injected into the host cell upon invasion. These data show that TGGT1_300220 is a new rhoptry bulb protein, named hereafter ROP55, for rhoptry protein 55.

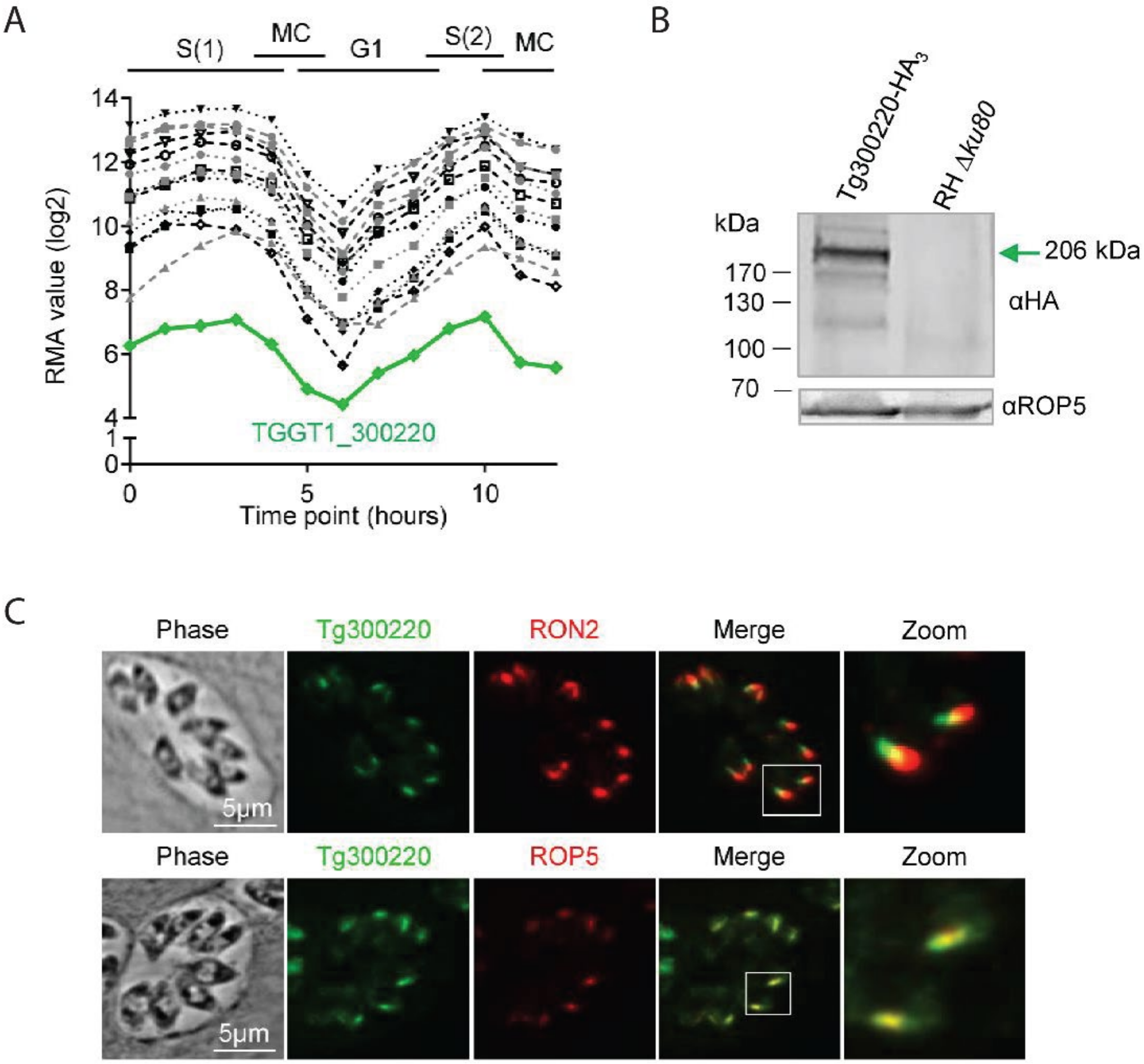

Figure 4.1. ROP55 is a new rhoptry bulb protein. (A) Expression pattern of known rhoptry proteins and TGGT1_300220 during the tachyzoite cell cycle. Data obtained from ToxoDB (Behnke et al. 2010). **(B)** Immunoblot of lysates from parental (RH Δku80) and Tg300220-HA3 lines using anti-HA antibodies. ROP55 migrates at the expected size of 206 kDa. Loading control using anti-ROP5 antibodies. **(C)** Colocalization of TGGT1_300220-HA3 (Tg300220) (green) with rhoptry neck RON2 or rhoptry bulb ROP5 (red). Inset: higher magnification.

ROP55 is a major virulence factor.

To study the function of ROP55, we generated a knock-out ROP55-HA3 line (ΔROP55) **(Supp. Fig. 4.2C-D**) and a complemented line (compl.) by reintroducing the HA-tagged *rop55* gene under its own promotor in the *UPRT* locus of the ΔROP55 line **(Supp. Fig. 4.2E-F)**. Western blot and IFA confirmed the deletion and re-expression of ROP55 in the complemented cell line and its targeting to the correct locus **(Fig. 4.2A-B)**.

To assess the ability of ΔROP55 to complete the lytic cycle in human foreskin fibroblasts (HFFs) we performed plaque assays over 6 days. We observed a reduction in both size and number of plaques **(Fig. 4.2C-D)** compared to the wild-type line (WT), and this phenotype was fully rescued in the complemented strain **(Fig. 4.2C-D)**. We next evaluated the role of ROP55 *in vivo* during acute infection of BALB/c mice injected intraperitoneally (i.p.) with increasing doses of ΔROP55 tachyzoites. All mice died by day 8 following injection of 100 WT or complemented tachyzoites **(Fig. 4.2E)**. Remarkably, all mice infected with 10^4 and 10^5 ΔROP55 tachyzoites survived, and only the injection of 1 million tachyzoites resulted in the death of 50% of mice **(Fig. 4.2E)**. Moreover, when we assessed parasite burden at 4 days post infection (dpi), by quantitative RT-PCR (qPCR), we observed substantially decreased Tg*SAG1* expression in the spleens of mice infected with 1000 ΔROP55 tachyzoites, suggesting that the survival of mice infected with ΔROP55 is associated with a reduction in parasite burden **(Fig. 4.2F, Supp. Fig. 4.2I-J)**. Altogether, these results show that ROP55, although not essential for *Toxoplasma* growth *in vitro*, has a prominent role *in vivo* and contributes to the virulence of the RH strain in female mice.

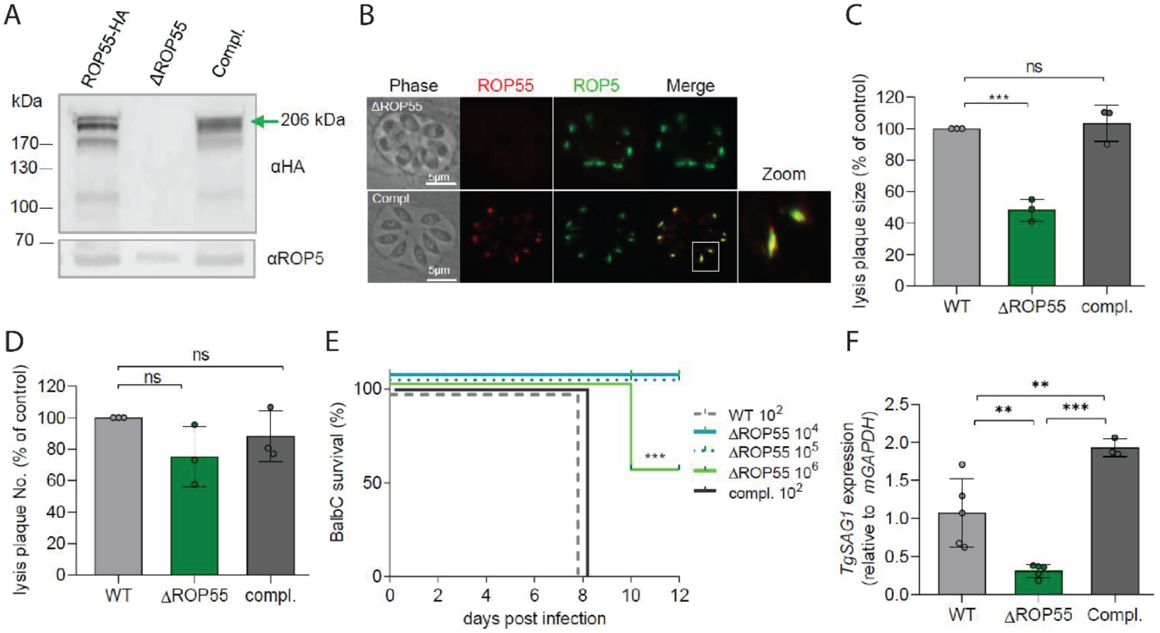

Figure 4.2. ROP55 is dispensable for the completion of the lytic cycle in vitro and is essential for virulence in mice. (A) Western blot (WB) showing absence of ROP55 in the ΔROP55 line and its re-expression in the complemented strain (Compl). ROP55 was detected using an αHA antibody. **(B)** Absence of ROP55 in ΔROP55 strain and correct localization of ROP55 (red) at the bulb of the rhoptries in the complemented strain. ROP55 (green). **(C, D)** Plaque assay of RH Δku80 parental (WT), ΔROP55 and complemented (compl.) lines on HFF monolayers after 6 days. Values are normalized relative to the percentage of lysis in the control strain. Bars represent mean ± SD of 3 biological replicates. **(C)** Lysis plaque areas were measured in 3 independent experiments. **(D)** Number of lysis plaques showing 3 independent experiments. **(E)** BALB/c mice were intraperitoneally infected with 100 tachyzoites of the WT and complemented strains (5 mice per condition, n=3), or with increasing doses of 104 tachyzoites (5 mice, n=3), 105 tachyzoites (5 mice, n=2) or 106 tachyzoites (10 mice, n=1) of ΔROP55 knock-out parasites. **(F)** Expression of TgSAG1 measured by qRT-PCR in spleens of BALB/c mice 4 days post-infection. Representative graph of 3 independent experiments. Bars represent mean ± SD of three to five mice per group. Significance was determined using one-way ANOVA (**C**, **D** and **F**). **=p<0.01; ***=p<0.001

ROP55 controls parasite and host cell survival.

To determine which stage of the lytic cycle was impaired in ΔROP55 parasites, we analyzed the individual steps involved in tachyzoites propagation *in vitro*. We observed a modest increase (25%) in the percentage of intracellular ΔROP55 parasites relative to the WT strain after 5 minutes of contact with HFF cells **(Fig. 4.3A)**. When we counted the number of parasites per vacuole at 24 hours-post infection (hpi), ΔROP55 parasites were not affected in their capacity to replicate **(Fig. 4.3B)**. By triggering egress from infected host cells with the Ca^{++} ionophore A23187, ΔROP55 parasites exhibited similar egress to the WT line **(Fig. 4.3C)**. However, when we measured parasite burden by qRT-PCR at 24 hpi, we detected a significant reduction in parasite load in cells infected with ΔROP55 parasites compared to WT or complemented parasites **(Fig. 4.3D, Supp. Fig. 4.3A-B)**. To evaluate whether this reduction was due to elimination of parasite vacuoles, we counted the vacuoles at 4 and 24 hpi. We observed a marked reduction in the number of vacuoles containing ΔROP55 parasites at 4 hpi, which persisted at 24 hpi, suggesting that elimination occurred shortly after invasion **(Fig. 4.3E, Supp. Fig. 4.3C)**.

Intriguingly, when culturing ΔROP55 parasites, we noticed the presence of abnormal host cells with elongated and condensed nuclei **(Supp. Fig. 4.3D)**. In addition, the cells appeared to be detaching and dying. The elongated nucleus phenotype was detectable within 2 hpi and increased to 6 hpi **(Supp. Fig. 4.3E)**. To assess host cell membrane integrity and cell viability, infected fibroblasts were stained with propidium iodide (PI), a non-permeant fluorescent dye commonly used to detect dead cells. At 6 dpi, the number of cells staining with PI was significantly increased when infected with ΔROP55 parasites compared to WT and complemented lines **(Fig. 4.3F)**. This was concomitant with an increased level of lactate dehydrogenase (LDH) release into the culture medium of ΔROP55 infected cells **(Fig. 4.3G, Supp. Fig. 4.3F-G)**, which is indicative of lytic cell death[37]. These data show that infection with RH tachyzoites lacking ROP55 leads to an increase in host cell mortality, accompanied by the elimination of parasite vacuoles and a decrease in parasite load. Importantly, the cell death was observed in the absence of any stimulation, such as IFN-γ or LPS. However, ROP55 deficiency did not induce the death of all infected HFFs, and when the parasites were found within a viable host cell, they grew normally, as shown in Fig. 4.3B. Accordingly, electron microscopy images did not reveal any ultra-structural defects of the host cell, the parasites, or the PV at 24 hpi **(Supp. Fig. 4.3H-I)**. These images also showed normal rhoptry shape. Altogether, these results suggest a role for ROP55 in controlling host cell viability, which may contribute to maintaining the parasite's replicative niche.

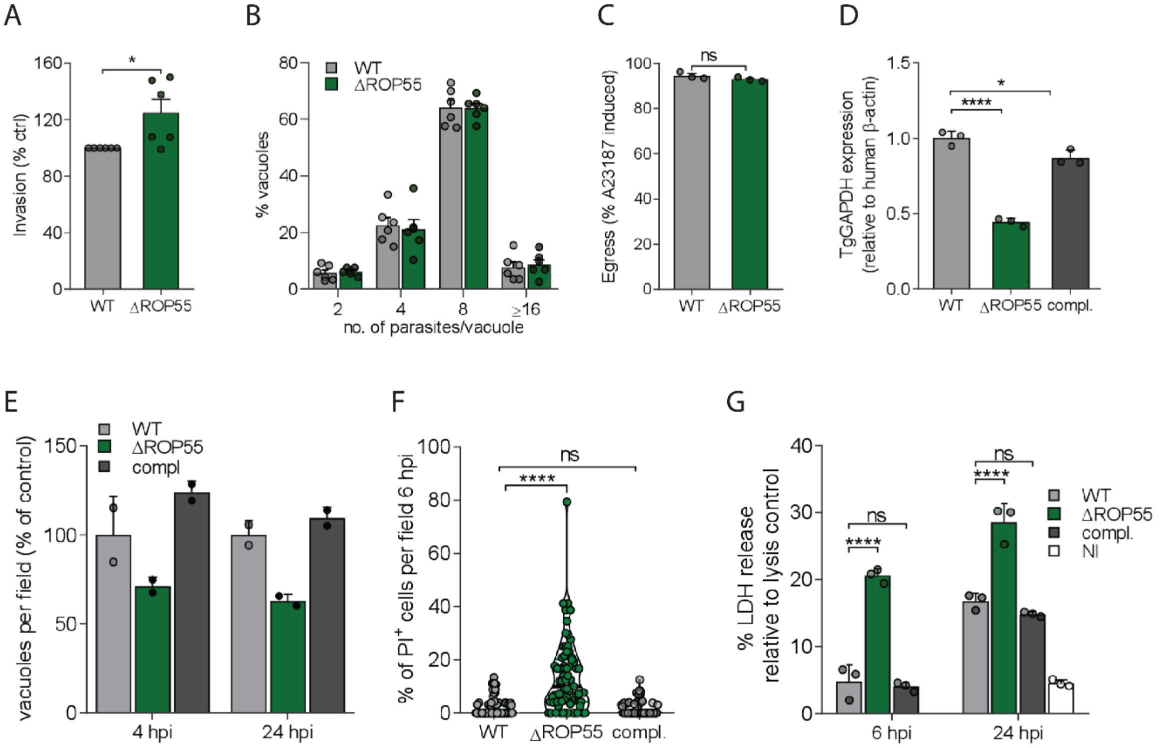

Figure 4.3. Absence of ROP55 induces cell death, resulting in decreased parasite burden. (A) Percent of invaded HFFs quantified at 5 minutes post-invasion. Values are normalized relative to the percentage of cells invaded by the control strain. Bars represent mean ± SEM of 6 biological replicates, each with 3 technical replicates. **(B)** Number of parasites per vacuole at 24 hpi. Bars represent mean ± SEM of 6 biological replicates. **(C)** Percent of egress events following A23187 treatment at 30 hpi. 200 vacuoles counted per coverslip. Bars represent mean ± SD of 3 biological replicates, each with 3 technical replicates. **(D)** Parasite burden measured by qRT-PCR in HFFs at 24 hpi. Bars represent mean ± SD of 3 technical replicates. **(E)** Number of vacuoles per field in HFFs at 4 or 24 hpi. 200 vacuoles counted per coverslip. Values expressed as percentage relative to WT strain. Bars represent mean ± SD of 2 technical replicates. **(F)** Percent of HFF nuclei staining positive for propidium iodide (PI) per field at 6 hpi. At least 15 fields were counted per coverslip using 2 coverslips per condition. **(G)** Quantification of LDH released into medium of HFFs at 6 or 24 hpi. Values normalized to 100% of lysis control (with addition of lysis buffer). Bars represent mean ± SD of 3 technical replicates. Statistical significance was determined using a two-tailed paired t test (**A** and **C**), one-way ANOVA (**D** and **F**), or two-way ANOVA for (**B** and **G**). *=$p<0.05$; ****=$p<0.0001$.

ROP55 is a repressor of inflammatory host cell death.

The ability of ROP55 to affect host cell fate prompted us to investigate the host transcriptional response during infection with ΔROP55 parasites. At 6 hpi, a substantially different gene expression profile was displayed in HFFs infected with ΔROP55 relative to WT **(Fig. 4.4A)**. We found 388 differentially expressed genes (DEGs) ($p\text{-adj} < 0.05$), of which 283 were up-regulated and 105 were down-regulated. Interestingly, many of these DEGs were associated with cell death and the inflammatory response of the host cell **(Fig. 4.4A, Supp. Table 4.1)**. This profile was similar at 24 hpi, with 125 differentially expressed genes, of which 44 were up-regulated and 81 were down-regulated **(Supp. Fig. 4.4A-B, Supp. Table 4.2)**. When focusing on the set of genes showing the greatest change ($\text{Log}_2\text{FC} > 2$) at 6 hpi, we found a group of 62 genes specifically overexpressed after infection with ΔROP55 **(Fig. 4.4B)**. Among them was a set of inflammatory mediators known to be upregulated during inflammation in fibroblasts[38]. These genes included IL6, LIF, IL23 and CSF3 and their related transcription factors ELF3 and STAT4, as well as IL1A, IL1B and IL1RN (encoding the IL1 receptor antagonist), three important members of the IL1 family (IL1F). The complete list of DEGs found at 6 hpi is presented in Supplementary Table 4.1 and are accessible through NCBI SRA (PRJNA881613). Importantly, the complementation of ΔROP55 parasites broadly reverted the expression profile to that of the WT strain **(Fig. 4.4A, Supp. Fig. 4.4C)**, strongly supporting the ROP55-dependent regulation of host gene expression.

A Subsequent Gene Ontology (GO) enrichment analysis on the 283 specifically up regulated genes at 6 hpi showed that the biological processes associated with ΔROP55 infection involved a range of stress-related responses, such as regulation of cell proliferation and the induction of immune responses **(Fig. 4.4C, Supp. Table 4.3)**. Notably, an Ingenuity Pathway Analysis (IPA) performed on the overexpressed genes revealed proinflammatory stimuli, such as pathogen/damage-associated molecular patterns (PAMPs/DAMPs) and TLR ligands as potential up-stream regulators of these cellular responses (TNF, LPS, IL1B, poly rI:rC-RNA) **(Fig. 4.4D)**. All of these stimuli converge on NF-κB and MAPK signaling activation[39], which are critical hubs of inflammatory responses and cell death programming. Accordingly, RELA (NF-κB p65 subunit), p38 MAPK, and ERK1/2 were also found in the IPA analysis **(Fig. 4.4D)**. Altogether, these findings indicate a potential role for ROP55 as a suppressor of proinflammatory pathways triggered during infection with the RH *Toxoplasma* strain.

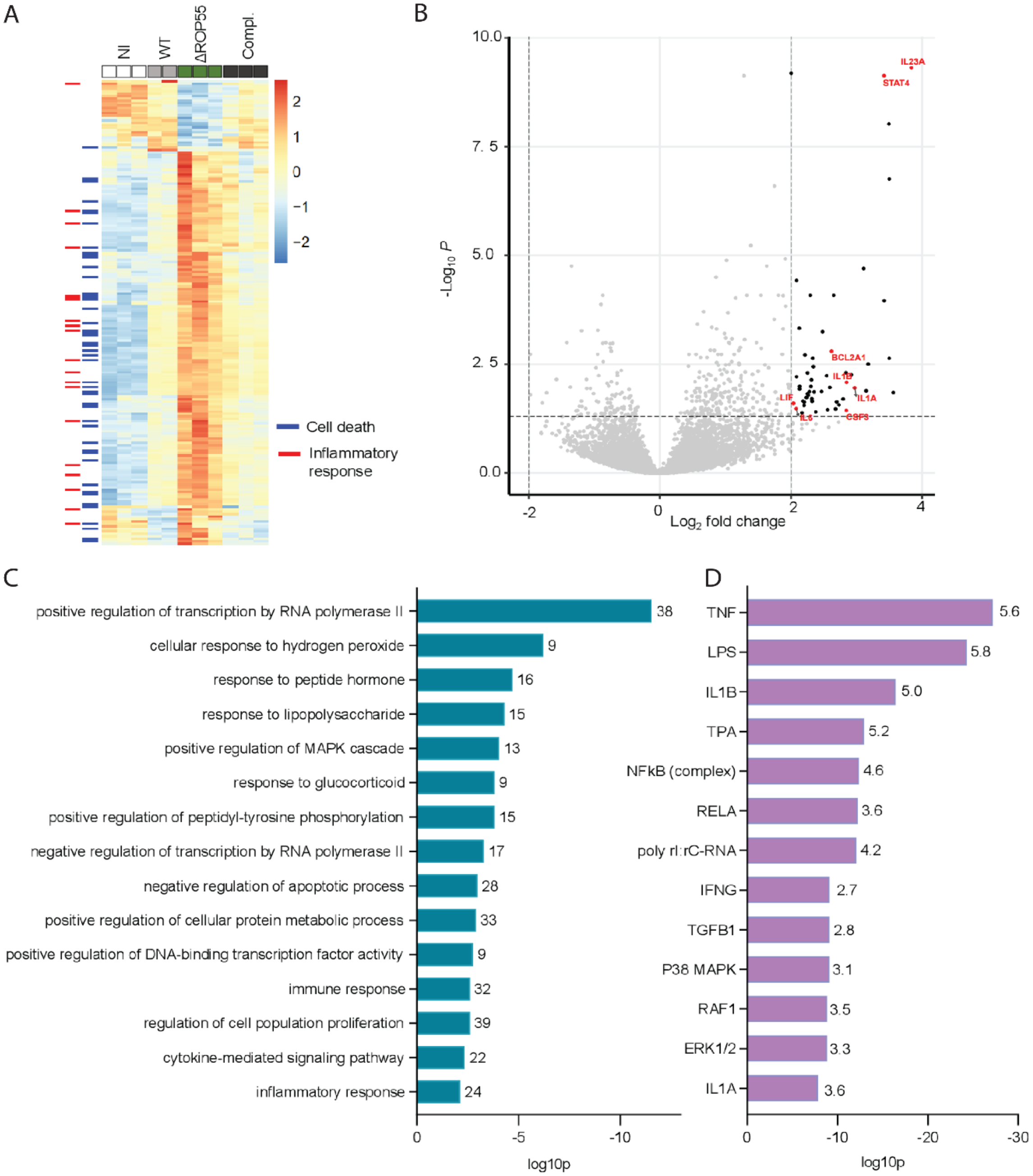

Figure 4.4. ΔROP55 infected fibroblasts display a pro-inflammatory transcriptomic profile at 6 hpi. (A) Heatmap of HFFs infected with RH Δku80 parental strain (WT), ΔROP55, or complemented (compl.) strains for 6 h, and non-infected cells (NI). **(B)** Volcano plot highlighting differentially expressed genes (padj < 0.05, $\log_2$ fold change > 2) in black. Genes labelled in red are among the top 25 DEGs. **(C)** Bar plot representing selected list of enriched biological processes from gene ontology (GO) analysis for the upregulated genes (padj < 0.05). The number of genes in each category is indicated and the scale represents the weighted Fisher p-value. **(D)** IPA analysis representing top upstream regulators of overexpressed DEGs. Numbers indicate the activation z-score, and the scale represents the BH corrected p-value.

ROP55 suppresses activation of the NF-κB/IL-1β signaling axis in fibroblasts.

The inhibition of NF-κB activation is responsible for the repression of cellular inflammatory responses during infection by the RH *Toxoplasma* strain[27, 28, 30, 40]. As transcriptomic analysis of ΔROP55- infected cells revealed the upregulation of several mediators of the NF-κB/IL-1β axis **(Fig. 4.4)**, we hypothesized that ROP55 might control this signaling pathway. To validate the observed increase in *IL1B* expression **(Fig. 4.4B)**, we conducted qPCR in cells infected for 6 and 24 hours. A significant increase in *IL1B* mRNA was observed at 6 hpi that further increased at 24 hpi in ΔROP55-infected HFFs **(Fig. 4.5A, Supp. Fig. 4.5A-D).** This resulted in increased levels of IL-1β detected in the cell supernatant by ELISA **(Fig. 4.5B)**.

Since NF-κB is the main transcriptional activator of IL-1β, we tested whether ROP55 might inhibit NF-κB activation. We infected HFFs and measured the nuclear translocation of the NF-κB p65 subunit by IFA at 2 and 6 hpi **(Fig. 4.5C)**. The type II Pru strain, known to induce p65 NF-κB nuclear translocation in a $GRA15^{II}$ -dependent manner was used as a positive control[26]. A significant increase in p65 nuclear translocation was observed in cells infected with ΔROP55 relative to cells infected with the WT and complemented parasites **(Fig. 4.5D-E, Supp. Fig. 4.5E-F)**, supporting a model in which ROP55 inhibits NF-κB p65 activation shortly after invasion.

To assess whether the NF-κB translocation was coupled with a downstream transcriptional activity, we infected the human embryonic kidney HEK 293 reporter cell line that expresses GFP under the control of NF-κB responsive elements[26] and quantified the GFP expression at 6 and 24 hpi **(Fig. 4.5F)**. Negligible levels of GFP fluorescence were observed following infection with the WT or complemented strains at both timepoints **(Fig. 4.5G-H, Supp. Fig. 4.5G)**. In contrast, reporter cells infected with ΔROP55 displayed a robust and sustained GFP expression, corroborating the p65 nuclear translocation observed in HFFs. Overall, these findings suggest a regulatory role of ROP55 in controlling NF-κB activation and the subsequent proinflammatory response in human fibroblasts.

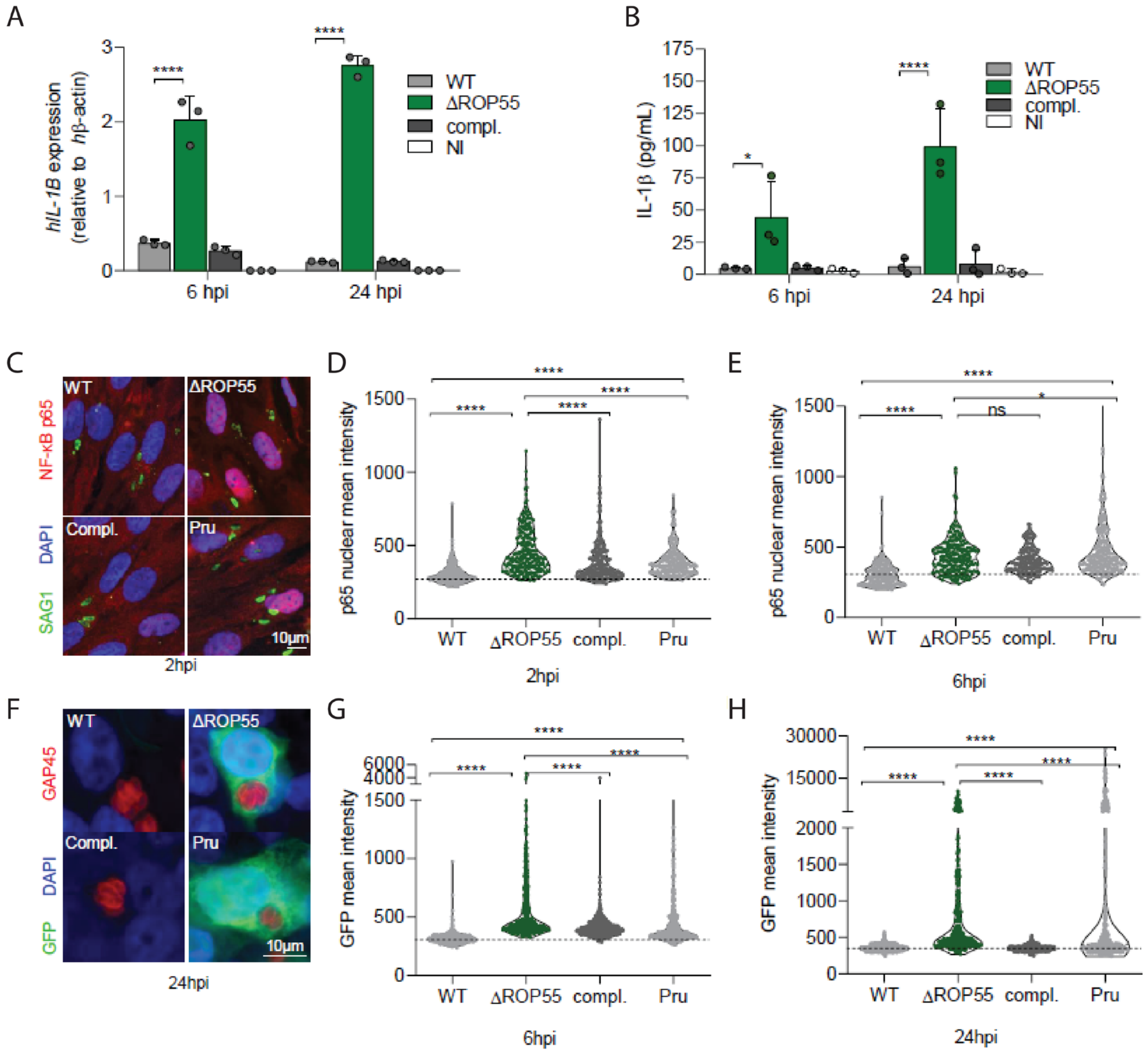

Figure 4.5. ROP55 acts as a negative regulator of the NF-κB/IL-1β pathway. (A) Expression of IL1B measured by qRT-PCR in HFFs at 6 and 24 hpi or non-infected cells (NI). Bars represent mean ± SD of 3 biological replicates, each with 3 technical replicates. **(B)** IL-1β secretion from HFFs at 6 or 24 hpi measured by ELISA. Bars represent mean ± SD of 3 biological replicates. **(C)** IFA of NF-κB p65 subunit nuclear translocation (red) in HFFs at 2 hpi. Intracellular parasites are stained with anti-SAG1 antibody (green), and nuclei are stained with DAPI (blue). **(D, E)** Quantification of p65 nuclear signal intensity in HFFs at 2 hpi **(D)** and 6 hpi **(E)**. Plots from 2 independent experiments quantifying at least 200 nuclei per condition. Horizontal line represents the mean fluorescence of p65 staining in cells infected with the WT strain. **(F)** IFA of GFP expression in HEK 293 NF-κB reporter cells at 24 hpi. Intracellular parasites are stained with anti-GAP45 antibody (red), and nuclei are stained with DAPI (blue). Scale represents 10 µm. **(G, H)** Quantification of GFP intensity in HEK 293 NF-κB reporter cells at 6 hpi **(G)** or 24 hpi **(H)**. Plots from 2 independent experiments quantifying at least 300 cells. Horizontal line represents the mean GFP fluorescence of cells infected with the WT strain. Statistical significance was determined using a two-way ANOVA. Adjusted p-values of Tukey's multiple comparison test *=p<0.05; ***=p<0.001 ****=p<0.0001.

ROP55 function is conserved in human and murine immune cells.

We next sought to determine if ROP55 also plays a role in controlling host cell death and inflammation in immune cells. Consistent with the findings in human fibroblasts, the absence of ROP55 induced significant lytic cell death of primary murine bone-marrow-derived macrophages (mBMDMs) **(Fig. 4.6A, Supp. Fig. 4.6A)**. This was accompanied by a decrease in parasite load **(Fig. 4.6B, Supp. Fig. 4.6B-D)**, that was not due to a defect of invasion, since ΔROP55 invaded mBMDM at the same level as WT parasites **(Fig. 4.6C, Supp. Fig. 4.6E)**. Furthermore, a significant increase in the nuclear translocation of p65 NF-κB subunit was observed upon infection of mBMDMs with ΔROP55 parasites **(Fig. 4.6D-E, Supp. Fig. 4.6F-G)**.

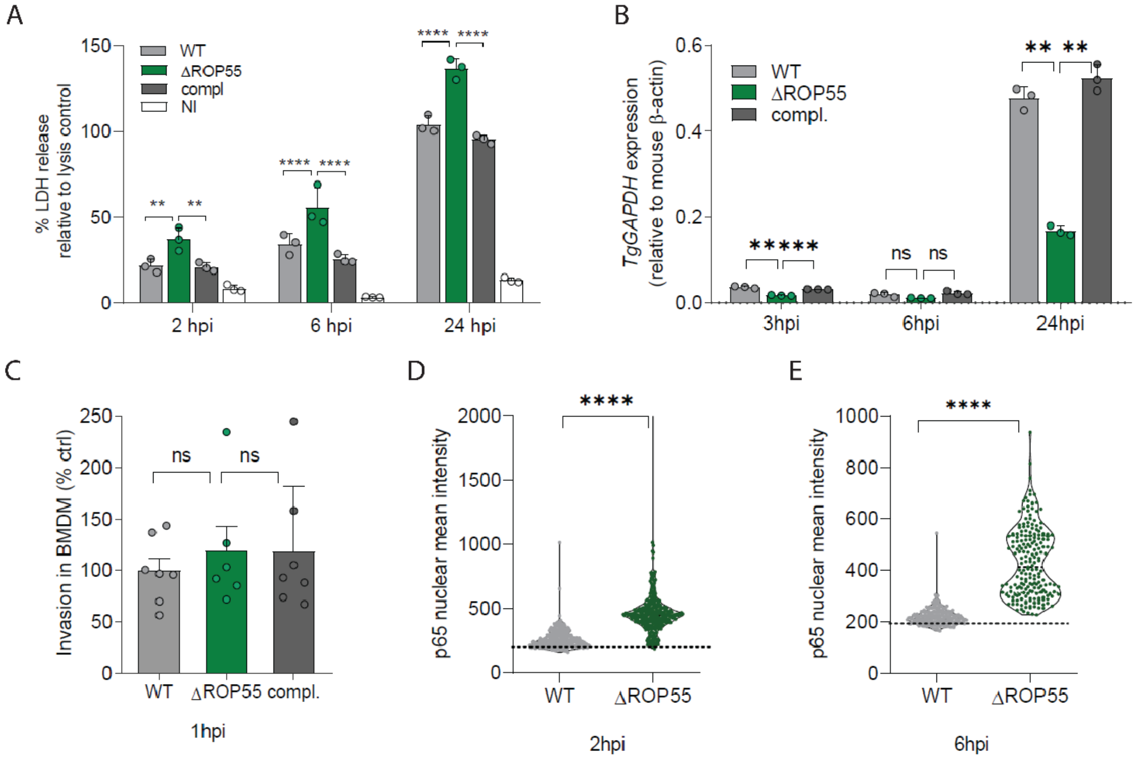

Figure 4.6. ROP55 prevents the lytic death of primary murine phagocytes. (A) Quantification of LDH released into medium of mBMDM at 2, 6 or 24 hpi. Representative of 2 independent experiments. Values normalized to 100% of lysis control (with addition of supplier lysis buffer). Bars represent mean ± SD of 3 technical replicates. **(B)** Expression of TgGADPH in mBMDM at 3, 6 or 24 hpi measured by qRT-PCR. Representative of 4 independent experiments. Bars represent mean ± SD of 3 technical replicates. **(C)** Percent of invaded mBMDM quantified at 1 hpi. Values normalized relative to the percent of cells invaded by the control strain. Bars represent mean ± SEM of 3 technical replicates. Representative of 2 independent experiments. **(D, E)** Quantification of p65 nuclear signal intensity in mBMDMs at 2 hpi **(D)** and 6hpi. **(E)**. Representative violin plots from 2 independent experiments quantifying at least 200 nuclei for each condition. Horizontal line represents the mean fluorescence of p65 staining in cells infected with the WT strain. Statistical significance determined using one-way ANOVA (**C**, **D** and **E**) or two-way ANOVA (**A** and **B**). **=p<0.01; ****=p<0.0001.

As the effectors shaping immune responses against *Toxoplasma* might differ between murine and human hosts, we also evaluate ΔROP55 infection in human immune cells. Monocytes play a central role in initiating innate defenses against *Toxoplasma*[41, 42] and the THP-1 cell line recapitulates many features of primary human monocytes. ΔROP55-infected THP-1 monocytes also displayed increased mortality, as assessed by LDH release **(Fig. 4.7A, Supp. Fig. 4.7A-B)** and PI incorporation **(Supp. Fig. 4.7C)**. Whereas THP-1 were equally infected by WT and ΔROP55 parasites at 2 hpi, significantly reduced parasite load was observed at 24 hpi in ΔROP55-infected cells **(Fig. 4.7B, Supp. Fig. 4.7D-E)**. Accordingly, at 16 hpi, we observed a higher proportion of PI positive cells infected with CFSE-labelled ΔROP55 parasites versus CFSE-labelled WT or complemented parasites **(Supp. Fig. 4.7C)**. Since CFSE parasites enables to discern infected from uninfected cells, we observed that a high proportion of non-infected cells were also dying in ΔROP55 THP-1 cultures **(Supp. Fig. 4.7C)**.

As previously described, nuclear translocation of the NF-κB p65 subunit was not detected in THP-1 cells infected with WT RH parasites **(Fig. 4.7C-D)**[23, 26]. In contrast, and consistent with our findings in human fibroblasts and murine mBMDMs, unstimulated THP-1 monocytes infected with ΔROP55 parasites displayed a significant increase in nuclear intensity of p65 NF-κB **(Fig. 4.7C-D, Supp. Fig. 4.7F)**. In the canonical NF-κB signaling pathway, p65 is maintained in the cytoplasm by IκBα, an inhibitor of NF-κB, and activation of the pathway results in ubiquitination and proteasomal degradation of IκBα, releasing p65 to enter the nucleus. Consistent with the increase in p65 nuclear translocation, we observed a significant reduction in IκBα protein levels in ΔROP55-infected THP-1 **(Fig. 4.7E-F)**. Consequently, the induction of IL-1β transcripts and the release of IL-1 protein in the culture media were observed in ΔROP55-infected THP-1 **(Fig. 4.7G-H, Supp. Fig. 4.7G-H)**. Furthermore, we detected both immature and cleaved forms of the cytokine in the cellular lysates and supernatants of ΔROP55- infected THP-1 **(Fig. 4.7I)**.

Collectively, these data indicate infection of THP-1 monocytes with ΔROP55 parasites induces early NF-κB signaling that activates IL-1β transcription, maturation and secretion, coupled with lytic cell death. These observations suggest a function for ROP55 control of upstream NF-κB signaling to escape clearance by inflammatory cell death.

Perplexingly, we noticed that ΔROP55 parasites cultured continuously for several months showed a consistent decline in their ability to induce IL-1β secretion in THP-1 cells, eventually reaching almost undetectable levels of the cytokine **(Supp. Fig. 4.7I)**. However, the increased lytic cell death and NF-κB translocation in cells infected with ΔROP55 parasites were sustained over the time **(Supp. Fig. 4.7J-K)**, indicating that the mutant line was able to adapt *in vitro* to avoid specifically the induction of IL-1β. Given this adaptation, for all the following experiments described here, unless otherwise stated, freshly thawed ΔROP55 parasites capable of inducing IL-1β secretion were utilized.

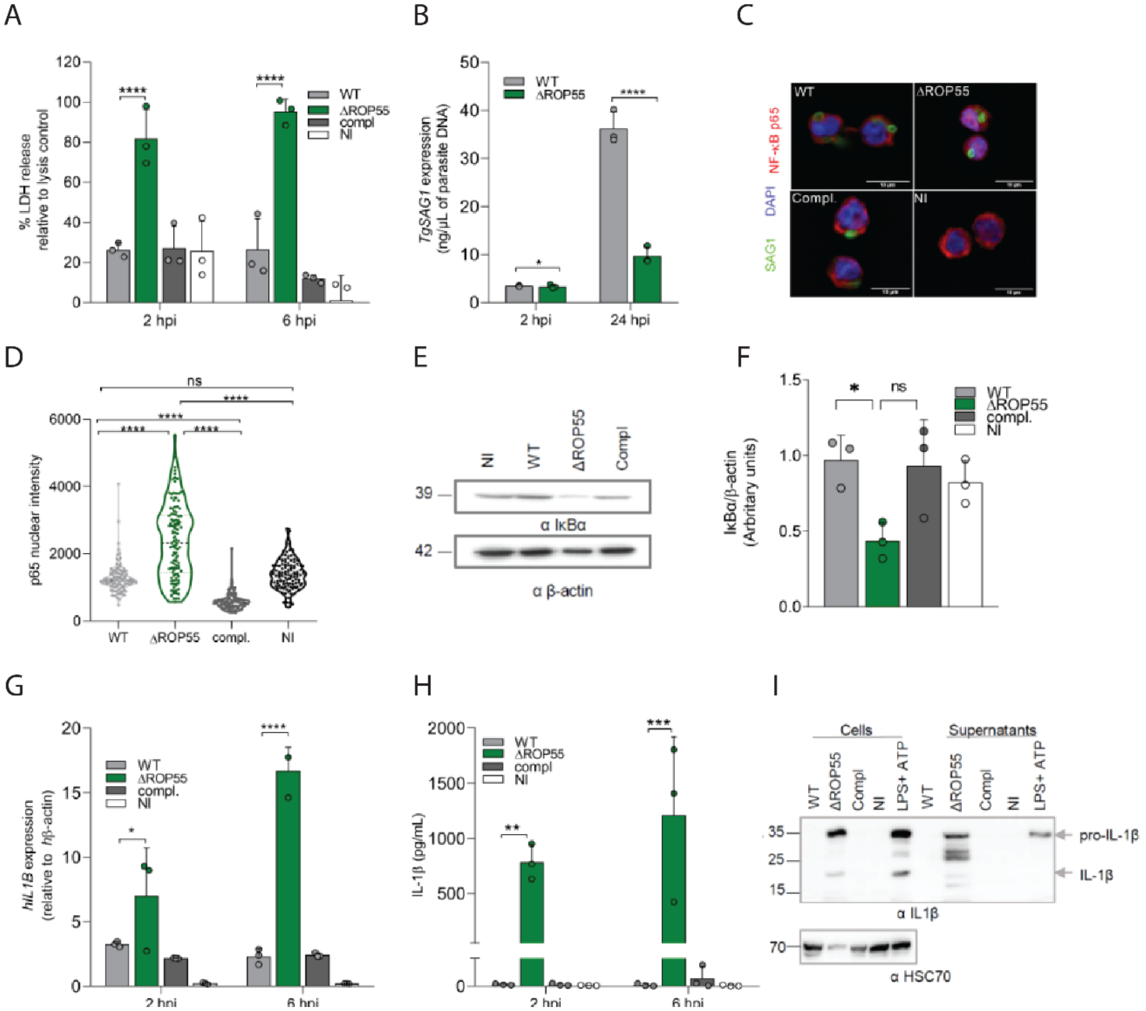

Figure 4.7. In THP-1 monocytes, ROP55 prevents NF-κB signaling and downstream inflammatory cell death. (A) LDH released into medium of THP-1 cells 2 or 6 hpi or non-infected cells (NI). Pool of 3 independent experiments. Values normalized against 100% of lysis control showing mean ± SD of 3 technical replicates. **(B)** TgSAG1 levels measured by qPCR in THP-1 at 2 and 24 hpi. **(C)** IFA showing NF-κB p65 nuclear translocation (red) in THP-1 cells 2 hpi with WT, ΔROP55 and complemented parasites stained with anti-SAG1 antibody (green). Nuclei are labelled with DAPI (blue). **(D)** Quantification of p65 nuclear intensity in THP-1 from **C**. >120 nuclei quantified per condition. Representative of 2 independent experiments. **(E)** WB of IκBα in THP-1 lysates 16 hpi or NI. **(F)** Densitometry quantification of IκBα bands observed in **E**. Arbitrary units obtained after normalization to the β-actin signal of corresponding lines. Pool of 3 independent experiments. **(G)** Expression of hIL1B in THP-1 at 2 and 6 hpi measured by qRT-PCR. Bars represent mean ± SD of 3 biological replicates, each with 3 technical replicates. **(H)** ELISA measuring IL-1β secretion from supernatant of THP-1 after 2 and 6 hpi. Bars represent mean ± SD of 3 biological replicates. **(I)** WB of mature (IL-1β) and immature (pro-IL-1β) forms of IL-1β present in cell lysate or supernatant of THP-1 cells 6 hpi or NI. Control THP-1 stimulated with LPS for 6h and ATP for the last 30 minutes before collection. Loading control: human HSC70. Representative of 3 independent experiments. Statistical significance determined using one-way ANOVA (**D** and **F**) or two-way ANOVA (**A**, **B**, **G** and **H**). *=$p<0.05$; **=$p<0.01$; ***=$p<0.001$ and ****=$p<0.0001$.

Activation of the NLRP3 inflammasome and pyroptotic cell death in THP-1 cells infected by ΔROP55 parasites.

IL-1β production is regulated by the NLRP3 inflammasome in human monocytes infected by *Prugniaud* strain *Toxoplasma*[24, 25]. To assess NLRP3 inflammasome involvement during ΔROP55 infection, we tested inflammasome priming by evaluating the transcriptional upregulation of NLRP3 in THP-1 cells. While transcript levels of *NLRP3* remained low in non-infected cells and those infected with control parasites, ΔROP55-infected THP-1 showed increased *NLRP3* transcripts **(Fig. 4.8A, Supp. Fig. 4.8A-B)**. To test the requirement of NLRP3 inflammasome activation for IL-1β secretion we pre-treated THP-1 with MCC950, an NLRP3 inhibitor that prevents inflammasome oligomerization and activation[43]. In THP-1 infected with ΔROP55 parasites, MCC950 treatment resulted in a substantial decrease in IL-1β secretion **(Fig. 4.8B)**. As a control, THP-1 were treated with canonical NLRP3 inflammasome activators LPS and ATP, after which MCC950 still reduced IL-1β release. Since the classical NLRP3 inflammasome recruits caspase-1 to process pro-IL-1β into mature IL-1β[11], we evaluated the levels of IL-1β secretion in caspase-1 knockout THP-1 cells (ΔCASP1)[44] and observed significantly reduced IL-1β secretion **(Fig. 4.8C)**. These data show the involvement of the canonical NLRP3 inflammasome pathway for IL-1β production in response to infection with ΔROP55 parasites. These findings also provide evidence for the involvement of ROP55 in hindering the initial cascade that leads to NLRP3 priming in THP-1 cells infected with *Toxoplasma*.

To investigate whether NLRP3 inflammasome activation induced pyroptotic cell death upon infection with ΔROP55 parasites, we analyzed Gasdermin D (GSDMD) cleavage, a hallmark of pyroptosis. We detected the 30 kDa cleaved N-terminal fragment of GSDMD only in cells infected with ΔROP55 **(Fig. 4.8D, Supp. Fig. 4.8C)**. We then compared the mortality of ΔCASP1 and ΔGSDMD[25] THP-1 against wild-type THP-1 cells upon infection with ΔROP55. Interestingly, a significant but only partial reduction in mortality was observed in these pyroptosis-deficient cell lines **(Fig. 4.8E-F, Supp. Fig. 4.8D)**. These findings indicate that ΔROP55-infected THP-1 are dying, in part, through GSDMD-dependent pyroptosis. These results also highlight the triggering of a pyroptosis independent cell death mechanism in THP-1 cells infected with parasites lacking ROP55.

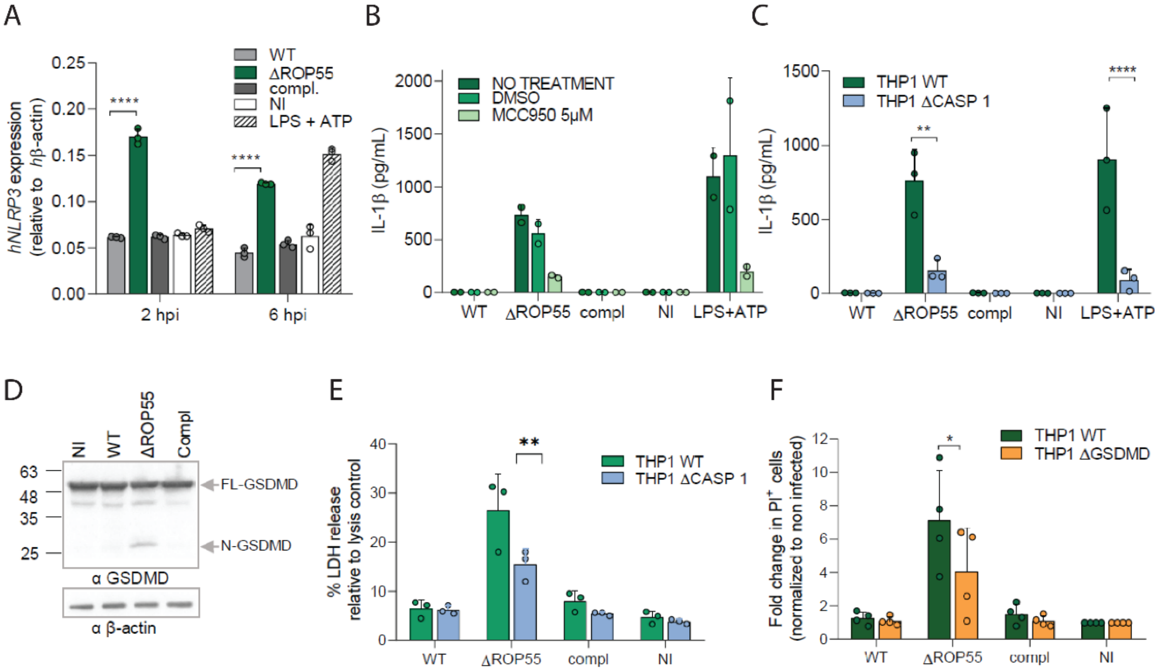

Figure 4.8. Activation of NLRP3 inflammasome and pyroptotic cell death in the absence of ROP55. (A) Expression of NLRP3 measured by qRT-PCR in THP-1 cells 2 and 6 hpi or non-infected cells (NI). Bars represent mean ± SD of 3 technical replicates. **(B)** Quantification by ELISA of IL-1β secreted in culture supernatant of THP-1 cells pretreated with 5 μM MCC950 or DMSO 40 min before infection with corresponding strains or NI. Bars represent mean ± SD of 2 biological replicates. **(C)** Quantification by ELISA of IL-1β secreted in media of Δcaspase1 THP-1 cells (ΔCASP1) at 2 hpi. Bars represent mean ± SD of 3 biological replicates, each with 2 technical replicates. **(D)** WB of full length-GSDMD (FL-GSDMD) and N-term-GSDMD (N-GSDMD) forms of gasdermin D in cell lysates of THP-1 cells at 14 hpi. β-actin was used as loading control. **(E)** LDH released from Δcaspase1 THP-1 cells (ΔCASP1) at 2 hpi. Bars represent mean ± SD of 3 technical replicates. **(F)** Increase of PI positive cells after infection. Values of infected THP-1 WT and ΔGSDMD THP-1 cells (ΔGSDMD) were normalized to NI. Bars represent mean ± SD of 4 biological replicates. Statistical significance was determined using two-way ANOVA (**A**, **C**, and **F**) or and one-way ANOVA (**E**). *=p<0.05; **=p<0.01; ***=p<0.001 ****=p<0.0001.

The absence of ROP55 triggers necroptosis of infected cells.

Since only partial rescue of the THP-1 cell death was observed in absence of caspase 1 or GSDMD, these findings suggested the activation of at least two different death programs in cells infected with ΔROP55. We thus explored alternative cell death pathways that could account for the residual mortality not associated with the pyroptotic cascade. The subsequent experiments were performed with adapted-ΔROP55 parasites that did not trigger IL-1β induction in THP-1 cells, but that still induced lytic host cell death. GSDME is another pore forming gasdermin involved in a secondary lytic death pathway activated by apoptotic caspases. In this process, named secondary necrosis[45], caspase-3 is responsible for GSDME cleavage, releasing its N-terminal fragment and causing pore formation and loss of membrane integrity in a similar manner to GSDMD. To assess if secondary necrosis occurs upon infection with ΔROP55, we first evaluated the activation state of caspase 3 using an intrinsically non-fluorescent probe that produces a fluorogenic response only upon caspase-3 and caspase-7 cleavage. As a positive control, treatment of THP-1 cells with staurosporine, a potent inducer of apoptosis, resulted in increased caspase-3/7 cleavage **(Fig. 4.9A)**. A slight but not significant increase in caspase-3/7 activity was observed 2 hpi with ΔROP55 **(Fig. 4.9A)**. This result prompted us to investigate the pore forming activity of GSDME. However, we did not detect any GSDME cleaved N-terminal fragment under infection of THP 1 with ΔROP55 or control parasites **(Fig. 4.9B)**. This result was consistent with the fact that GSDME knockout THP-1 cells did not display a reduction in mortality after infection with ΔROP55 **(Fig. 4.9C)**, excluding secondary necrosis as a death mechanism in ΔROP55-

infected cells.

We next investigated the occurrence of necroptosis, which along with pyroptosis, is a major mechanism of defense against intracellular pathogens. For this, we performed pharmacological inhibition of RIPK1 and RIPK3, the two kinases involved in necroptotic pathways[14], as well as MLKL, which is a critical substrate of RIPK3 for necroptosis execution. THP-1 cells were pre incubated either with Nec1, an allosteric inhibitor of RIPK1[46]; GSK872, a competitive inhibitor of RIPK3[47]; or Necrosulfonamide[48], an inhibitor of MLKL polymerization. The cells were then infected or left untreated, and the mortality was assessed by measuring LDH release after 2h and 6h. Whereas RIPK1 inhibition had no major effect on the mortality of ΔROP55-infected THP-1, RIPK3 and MLKL inhibition dramatically decreased the percentage of LDH release from ΔROP55-infected THP-1 **(Fig. 4.9D-E, Supp. Fig. 4.9A-D)**. A requirement for the triggering of the necroptotic cascade is the lack of caspase-8 activity[19]. We thus assessed the activation state of caspase-8, by evaluating its auto-proteolysis into p18 caspase-8 by western blot. As control for caspase-8 processing, THP-1 cells were treated with TNF-α and cyclohexamide, resulting in generation of the p18 subunit. In ΔROP55-infected THP-1, only the 58 kDa form of pro-caspase 8 was detected **(Fig. 4.9F)**, consistent with uncleaved caspase-8 promoting necroptosis[14]. we also found that in ΔROP55-infected ΔCASP-1 THP-1 cells, caspase-8 is not active **(Fig. 4.9F)**. ΔCASP-1 THP-1 also display a rescue of ΔROP55-induced mortality under RIPK3 and MLKL inhibition, reinforcing the involvement of necroptosis in pyroptosis-deficient cells **(Supp. Fig. 4.9E).**

In order to evaluate if ΔROP55 infection also triggers necroptosis in human primary immune cells, we isolated monocytes from healthy adult donor peripheral blood mononuclear cells (PBMCs) and differentiated them into macrophages. Importantly, we found that infection of human monocyte-derived macrophages (hMDM) with ΔROP55 also induced death of host cells at 18 hpi **(Fig. 4.9G)**. Furthermore ΔROP55-induced hMDM mortality was rescued by inhibiting the necroptotic executioner pseudokinase MLKL with necrosulfonamide **(Fig. 4.9G)**. Finally, in mBMDM, RIPK3 inhibition also suppressed ΔROP55-induced mortality at 2 hpi **(Supp. Fig. 4.9F)**, whereas RIPK1 inhibition did not induce changes at this time point. Interestingly, at 6hpi, a significant reduction in mBMDM mortality was detected with both RIPK1 and RIPK3 inhibitors **(Supp. Fig. 4.9G)**. Furthermore, we observed a dramatic change in the shape of the vast majority of nuclei of mBMDM infected with ΔROP55 after 6h. Their nuclei appeared swollen or disrupted, and importantly, this phenotype was rescued under the treatment with necroptosis inhibitors **(Supp. Fig. 4.9H)**. Overall, these results show a conserved role for ROP55 in the subversion of a RIPK3- dependent cell death pathway in mouse and human cells.

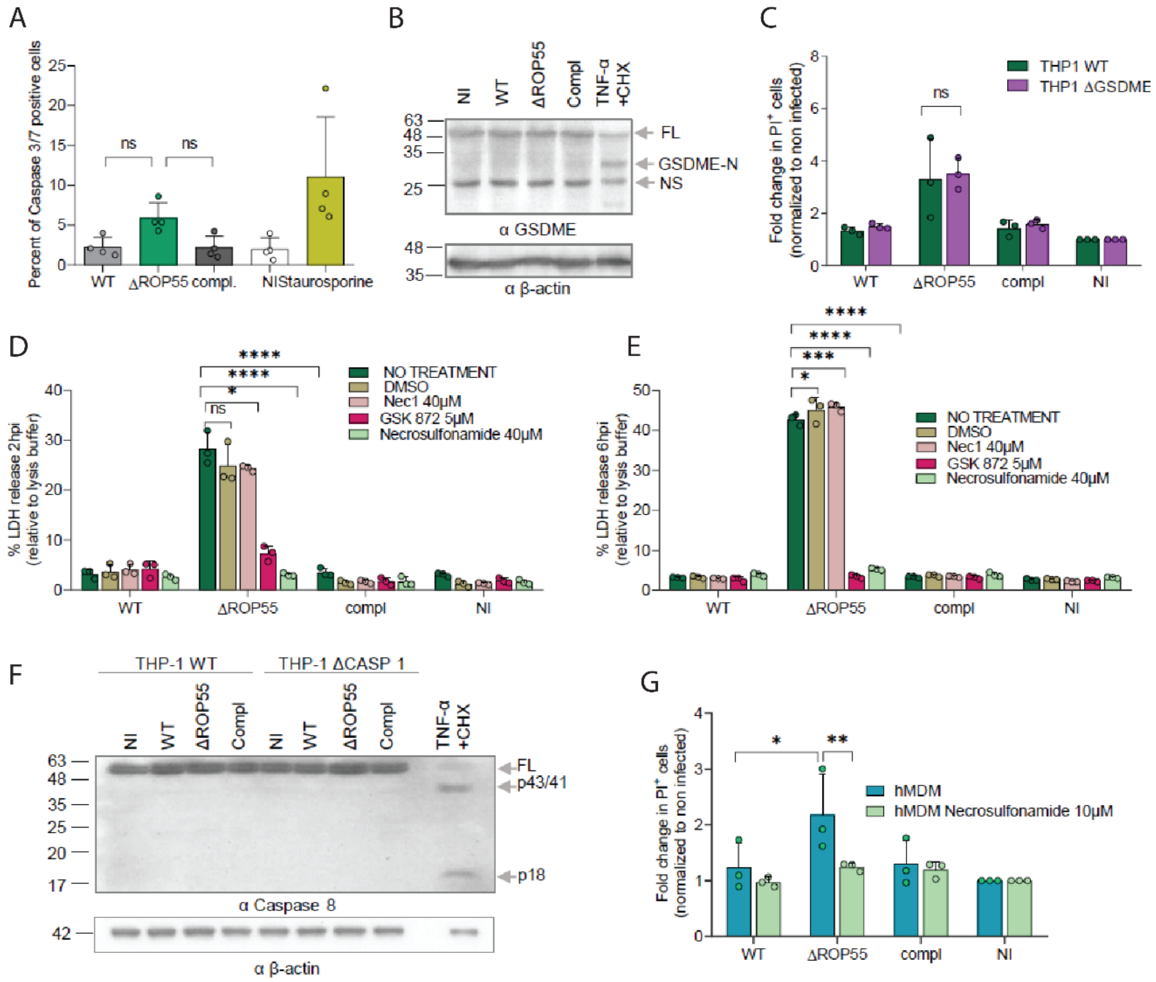

Figure 4.9. Necroptosis inhibition restores the viability of cells infected with ΔROP55 parasites. (A) Percent of Caspase 3/7 positive THP-1 cells 2 hpi or non-infected cells (NI). Bars represent mean ± SD of 4 biological replicates. **(B)** WB of full length-GSDME (FL-GSDME) 55kDa and N-term-GSDME (N-GSDME) 30kDa forms in cell lysate of THP-1 16hpi. β-actin was used as a loading control. NS: nonspecific band at 25 kDa. Representative gel of 3 independent experiments. **(C)** Fold change of PI positive cells after infection. Values of infected THP-1 WT and ΔGSDME THP-1 cells (ΔGSDME) were normalized to NI and reported as mean ± SD of 3 biological replicates. **(D, E)** LDH released of THP-1 cells pretreated with 40μM Nec1, 5 μM GSK872, 40 μM Necrosulfonamide or DMSO 45 minutes before 2 hpi **(D)** or 6 hpi **(E)**. Values normalized to 100% of lysis control and represent mean ± SD of 3 technical replicates. **(F)** WB of THP-1 WT and ΔCASP1 THP-1 cell lysates 14 hpi showing full length pro-caspase8 (FL) and p43/41 and p18 subunits after TNFα and ciclohexamide (CHX) treatment. Representative of 3 independent experiments. **(G)** Fold change in PI positive hMDM 18hpi with respective strains. Data normalized to NI cells. hMDM were pretreated or not with 10μM necrosulfonamide 40 minutes before infection. Values are reported as mean ± SD of 3 biological replicates using primary hMDM from two different donors. Statistical significance was determined using one-way ANOVA (**A**) or two-way ANOVA (**C**, **D**, **E** and **G**). *=p<0.05; **=p<0.01; ***=p<0.001; ****=p<0.0001.

DISCUSSION

During infection with intracellular pathogens, regulated cell death pathways are key components of innate immunity that sense danger signals and activate effector mechanisms critical for the subsequent establishment of adaptative immunity[9]. As a result of successful coevolution with the host, several intracellular pathogens have developed diverse strategies to suppress the activation of cell death pathways[10], and *Toxoplasma* is not an exception. Both, avirulent and highly virulent *Toxoplasma* strains have been shown able to block apoptosis[49], pyroptosis[25] and necroptosis[20]. By these means, the parasite maintains its replicative niche, which is essential for intracellular survival, host propagation and chronic persistence. However, how *Toxoplasma* circumvents inflammatory cell death, and how the virulent RH strain counteracts the activation of NF-κB, the central transcription factor involved in all these responses, remains poorly characterized. In the present study we have identified a new rhoptry protein, ROP55, that, when removed, induces significant NF-κB activation and lytic death of infected host cells. Consequently, a decrease of parasite load is observed in cells infected with ROP55-deficient parasites, supporting a role for ROP55 in the maintenance of the parasite niche. Importantly, ROP55 appears to be one of the major virulence factors required for the establishment of acute toxoplasmosis in mice.

Transcriptomic analysis of HFFs infected with ΔROP55 parasites revealed a massive inflammatory response, that we further characterized in THP-1 monocytes and found that priming of the NLRP3 inflammasome was responsible for IL-1β production and secretion. Previous studies using the RH strain have suggested the existence of parasite factors able to block NF-κB signaling and host cell inflammatory responses. These were based on the fact that: i) infection of macrophages, monocytes, neutrophils and fibroblasts by RH tachyzoites does not result in the activation of NF-κB[27-30], and ii) the RH strain is able to block NLRP3 inflammasome activation and IL-1β production during activation of human peripheral blood neutrophils with LPS[27], by suppressing the activation of NF-κB signaling. This suppression was also observed in macrophages and fibroblasts[28, 30]; however, the parasite effectors involved in this effect were unknown. Here, we show that upon infection with RH parasites deficient in ROP55, this repression is released in three human cell lines (HFF, THP-1 and HEK293) as well as in primary mouse macrophages. ROP55 acts during the first two hours after infection, which suggests an early role for ROP55 in preventing the transduction of parasite sensing through NF-κB signaling, potentially hindering NLRP3 priming. It is worth noting that ROP55 effects were observed in the absence of exogenous IFN-γ or LPS stimulation, indicating that the triggers of these responses arise from molecular patterns directly linked to *Toxoplasma* infection.

In contrast to RH, the Type II *Pru* strain constitutively activates IL-1β expression and release from infected cells, but, without inducing cell death. Two counterbalancing mechanisms for these inflammatory responses arise from dense granule effectors: i) TEEGR[50], that regulates EZH2 to antagonize NF-κB-induced cytokines; and ii) MAG1 that inhibits IL-1β induction[51]. Here, we show, in RH parasites, another level of control involving the rhoptries. Dense granule and rhoptry effectors might cooperate to modulate host inflammation and cell death. Lima *et al.*[27] used mycalolide B to block parasite invasion and dense granule release, but not rhoptry secretion in neutrophils, showing that rhoptry secretion alone does not suppress LPS-induced IL 1β production. Unlike Lima's study, we used different cell types without LPS stimulation. *Toxoplasma* PAMPs may activate different pathways than LPS for IL-1β production. Further studies will be necessary to address whether ROP55 may have additive contribution or synergize with dense granule effectors in different *Toxoplasma* strains.

ΔROP55 infection induced pyroptotic lytic death of THP-1 monocytes was dependent on caspase-1 and the executioner protein GSDMD. However, approximately 50% of cells deficient in either caspase-1 or GSDMD survived infection with ΔROP55 parasites, suggesting that an additional, pyroptosis-independent mechanism of death occurs. We then assessed apoptotic caspases and necroptosis, which are also known to play a role during *Toxoplasma* infection, particularly in IFN-γ-activated cells[20, 52]. A crosstalk between cell death programs has been documented[10,53], showing that they do not operate in a mutually exclusive manner, but rather, they are interconnected signaling processes. Although we found a slight but not significant increase in caspase-3/7 activation, our data demonstrate that secondary necrosis was not occurring in ΔROP55-infected THP-1 cells. Remarkably, using pharmacological inhibition, we demonstrated that RIPK3- and MLKL-dependent necroptosis was responsible for the death of cells infected with ΔROP55 parasites. This was demonstrated in human THP-1 monocytes as well as in primary human and mouse macrophages. The enzymatic activity of RIPK1 was not involved in ΔROP55-induced necroptosis in THP-1 cells at either time point, or in mBMDMs at 2 hpi. Interestingly, at a later time point (6 hpi), RIPK1 inhibition significantly reduced the mortality of mBMDMs. A role for RIPK1 in THP-1 cells or in mBMDMs at 2 hpi cannot be excluded, because Nec1 does not inhibit RIPK1-scaffolding function, which is important contributing to NF-κB signal transduction[14]. Altogether these results suggest that ROP55 may interfere with a necroptotic signaling cascade, which eventually intersects with the NLRP3 priming pathway through NF-κB transduction, protecting infected cells from both necroptosis and pyroptosis.

We were surprised to observe that ΔROP55 parasites cultured for several months consistently showed a gradual decline in their ability to induce IL-1β secretion in cells, eventually reaching undetectable levels of the cytokine. In this adapted-ΔROP55 mutant, NF-κB signal transduction was maintained. Interestingly, NLRP3 transcriptional induction was lost. These findings suggest that the mutant adapted by blocking the initial inflammasome priming signal to prevent inflammasome-induced pyroptosis. This also implies that the signal initiating inflammasome activation may differ from the signal triggering necroptosis, because necroptosis was maintained in the adapted-ΔROP55 mutant. It might be possible that, due to the pressure from the inflammatory environment in the culture, ΔROP55 parasites adapted by downregulating effectors that prime the inflammasome response. How the ΔROP55 parasites adapted remains an open question, which in the future might help to understand the cross-talk between necroptosis and pyroptosis and the characterization of potential new parasite cell death and inflammasome regulators.

Preventing both necroptosis and NF-κB-dependent inflammatory signaling is not unique to *Toxoplasma*. Viruses and pathogenic bacteria counteract these cellular responses as a strategy to maintain their replicative niche[54, 55]. RIPK3-dependent necroptosis has been largely characterized as an antiviral defense mechanism triggered by TL3[16] and ZBP1 sensors[56] after detection of viral genomes. In these pathways, RIPK3 activation induces both necroptotic cell death and NF-κB mediated cytokine production[15, 57]. Interestingly, RIPK1-scaffolding activity is crucial for mediating NF-κB activation in RIPK3-dependent necroptosis[16, 17, 58] whereas its kinase activity is dispensable[57]. One fascinating viral protein subverting both necroptosis[59] and inflammatory signaling is the mouse cytomegalovirus M45 protein. This protein interacts with RIP homotypic interaction motifs (RHIM) present in RIPKs and in ZBP1 and induces degradation of NF-κB essential modulator (NEMO)[60] of infected cells. As a result, M45 blocks both TLR and IL-1R signaling pathways, which are initiated by two different triggers that eventually converge to induce NF-κB translocation.

Few studies have examined necroptosis as a defense against *Toxoplasma*. The dense granule proteins TgNSM and TgIST block interferon-regulated necroptotic gene expression[20, 61], preserving intracellular cysts, but their role in tachyzoite survival is unstudied, and no clear evidence shows that necroptosis kills tachyzoites or controls acute infection. For example, RIPK3-deficient macrophages do not affect *Toxoplasma* survival. RIPK3-knockout mice resist intraperitoneal infection normally but show better survival in oral infections due to increased immune cell infiltration and edema in the intestinal lining, not to reduced parasite numbers[62]. Additionally, MLKL−/− mice are similar susceptibility to oral *Toxoplasma* infection suggesting that RIPK3 impact on host survival may be due to its pyroptotic, not necroptotic, activity[62]. Using intraperitoneal infection of RH ΔROP55 parasites, we found improved survival of mice, which is correlated with a decrease in parasite burden in the spleen at day 4 post-infection. Studies examining the infection of RIPK3−/− or MLKL−/− mice by ΔROP55 parasites are still needed to evaluate the contribution of RIPK3-dependent necroptotic cell death to the loss of ΔROP55 virulence *in vivo*.

The sensors for RIPK3 activation upon ΔROP55 infection are unknown. Interestingly, Influenza A virus is sensed by ZBP1, and this can induce RIPK3 necroptosis within the nucleus, resulting in nuclear envelope disruption[63], a feature that might be reminiscent of the shape of the nuclei in mBMDMs infected by ΔROP55 parasites. It has been proposed that ZBP1 can assist host control during *Toxoplasma* infection *in vitro* and *in vivo,* because ZBP1$^{-/-}$ mice had increased parasitemia and inflammatory cytokines compared to wild-type mice[64]. There is no direct evidence of *Toxoplasma* sensing by ZBP1; however, it has been shown that ZBP1 can also bind endogenous nucleic acids, acting as a sensor for excessive transcription states generated during infection[65, 66]. The detection of *Toxoplasma* DNA through AIM2 indicates atypical apoptosis mediated by caspase-8 in IFN-γ-primed macrophages[52]. It is worth noting that this and previously reported cell death mechanisms upon infection with *Toxoplasma*[20, 67] differ greatly from the death triggered by ΔROP55 in naive HFF, mBMDM, hMDM and undifferentiated THP-1 monocytes. The lytic cell death observed upon ΔROP55 infection is unique in the sense that it is observed in the absence of exogenous stimulation, owing to the detection of *Toxoplasma* molecular patterns or/and the sensing of danger signals coming directly from the infection.

ROP55 is crucial in orchestrating early repression mechanisms that allows escape of lytic cell death and ensures parasite survival. Notably, lytic cell death was observed in two human cell lines and in primary mouse and human macrophages, highlighting ROP55 control of conserved pathways in different cell types and species. Among the characterized *Toxoplasma* mutants for GRA and ROP effectors, ΔROP55 appears to be one of the most attenuated parasites in a mouse model of acute toxoplasmosis. Our work reveals a previously unappreciated rhoptry function in the complex mechanisms utilized by *Toxoplasma* to escape inflammatory cell death and opens new avenues to understand how the parasite avoids elimination by lytic host cell-death.

MATERIALS AND METHODS

Ethics statement

All mice protocols were approved by the Institutional Animal Care and Utilization Committee (IACUC) of the American University of Beirut (AUB). Animals were housed in a pathogen free facility with a 12h on/off light cycle. Humane endpoints were fully respected per AUB IACUC following AAALAC (Association for Assessment and Accreditation of Laboratory Animal Care International) guidelines and guide of animal care (Guide, NRC 2011). Mice were monitored daily. To verify the acute phase of infection, blood was drawn following anesthesia by inhalation of isoflurane and sera were probed against tachyzoite extracts by Western blot. Mice were sacrificed if any abnormal health features were noticed. Animals were anesthetized before cervical dislocation.

Cell line maintenance and treatments

Human Foreskin Fibroblasts (HFF) (ATCC CRL 1634) and HEK 293 NF-κB reporter cell line (System Biosciences, Cat. # TR860A-1), were cultured in Dulbecco's Modified Eagle's Medium (Gibco) supplemented with 5% or 10% (v/v) of heat inactivated Foetal Bovine Serum (FBS), 2 mM Glutamine (Gibco) and 50 µg/mL Penicillin/Streptomycin at 37°C with 5% CO2. THP-1 WT, ΔGSDMD THP-125 Δcaspase-1 THP-1, and ΔGSDME THP-1 cells[68] were cultured in RPMI 1640 (Gibco) supplemented with 100 U/mL penicillin, 100 ug/mL streptomycin, 2 mM L-glutamine and 10% (v/v) heat inactivated FBS and 1 µg/ml of puromycin. THP-1 cells were treated with a final concentration of 5 µM MCC950 (Sigma, 5381200001) as described by Gov *et al.*[24]. Mouse BMDM were isolated from BALB/c mice tibias and differentiated for 7 days following standard procedures[69]. Type I RH parasite strains lacking the *ku80* gene (RH *Δku80*) were used for all experiments unless stated otherwise and routinely passed in HFF cells. Type II *Prugniaud* (*Pru*) were cultured in the same conditions as RH parasites.

Human monocyte-derived macrophage differentiation

Human whole blood was obtained from healthy adult donors at the University of California, under the auspices of the Institute for Clinical and Translational Science, with donors providing written informed consent. The collection followed the ethical guidelines approved by the Institutional Review Board at the University of California, Irvine. Peripheral blood mononuclear cells (PBMCs) were separated from whole blood using density gradient centrifugation with lymphocyte separation media (MP Biomedicals, Santa Ana, CA). Monocytes were subsequently isolated from the PBMCs through counterflow elutriation, as previously described[70], and assessed for purity after each isolation. Typically, this method yielded highly pure monocyte cultures (>93% pure monocytes), based on CD11b+ and CD3- CD20- CD56- marker staining. After isolation, monocytes were differentiated into macrophages using 100 ng/ml recombinant human macrophage colony-stimulating factor (M CSF; Gemini Bio-Products, West Sacramento, CA) over 6 to 7 days. Fresh media with M-CSF was supplemented on day 3. On days 6 to 7, the monocyte-derived macrophages were replated and allowed 4 hours to adhere before being used in infection experiments with a multiplicity of infection (MOI) of 2.

Parasite transfections

For each transfection $20x10^6$ freshly lysed parasites were electroporated at 2.02 kV, 50Ω, 25μF in an Electro Cell Manipulator 630 (BTX) with 5 μg of PCR products obtained with the KOD polymerase in addition to 30 μg CRISPR/Cas9 plasmid or with the linearized pLIC-HA3-CAT plasmid. Transgenic parasites were selected with 1μM pyrimethamine, 20 μM Chloramphenicol or 5 μM FUDR. Clones were isolated through limiting dilution and correct integrations were verified by diagnostic PCR.

Plasmid constructs

The HA3-tagging at the endogenous locus of *ROP55* was performed by single homologous recombination using the ligation-independent-cloning strategy. The last 2 Kb of the *rop55* genomic sequence were amplified with primers ML2258 and ML2259 and inserted in frame with the sequence coding for HA3 tag in the vector pLIC-HA3-CAT. After linearization with *SfuI* and transfection, the integration of the construct at *rop55* locus was verified with primers ML2363 and ML1476.

For the generation of ΔROP55 line, two different RNA guides (gRNAs)s were cloned into the pU6-Cas9-YFP plasmid in order to target the 5' and 3' UTR of *rop55* in the ROP55-HA3-tagged line. Briefly, gRNAs were generated by annealing primers ML2258/ML2259 and primers ML2831/ML2832, respectively. The annealed DNA fragments were cloned in the *BsaI* sites of vector pU6-Cas9 to obtain plasmids p5'ROP55 and p3'ROP55. A donor DNA fragment containing the DHFR resistance cassette flanked by 30 bp homology regions to the 5' UTR and 3' UTR of *rop55* locus was amplified with primers ML2829 and ML2828 from pLIC HA3DHFR vector. The double homologous recombination events were verified by PCR using primers ML2636 and ML2454 for 5' integration and ML2601 and ML2830 for 3' integration.

To complement the ΔROP55 line, an extra copy of the *rop55* gene*,* driven by its own promotor and fused with a C-terminal sequence encoding for a triple HA tag, was inserted at the *uprt* locus. Briefly, a 7kb DNA fragment was synthetized containing the following sequences: 50bp of homology region with the 5'UTR of the *uprt* locus, the *rop55* promotor region (927 bp), the full cDNA sequence of *rop55* fused to a triple HA tag sequence, the 3'UTR of the *hxgprt* gene and 50bp of homology regions with the 3'UTR of the *uprt* locus. After cloning in pUC vector, digestion with *PstI* and *NaeI*, the 7 kb excised fragment was co-transfected with two gRNAs targeting the 5'and 3'UTR of the *uprt* locus. The plasmid guide p5'UPRT was constructed by annealing primers

ML3445 and ML3446 and the plasmid guide p3'UPRT by annealing primers ML2087 and ML2088. The integration of the exogenous copy of *rop55* at the *uprt* locus was verified with primers ML3187 and ML2081 and ML2384 and ML3190. Primer sequences are in Supplementary table 4.4.

Plaque assay

Cells were seeded on 24-well plates and grown to 100% confluence and infected with 7000 tachyzoites per well. Four-fold dilutions were then performed from the top row to the next row. The plate was incubated at 37°C, 5% CO2 and left untouched for 6 days. On day 6 post infection, cells were fixed with 4% paraformaldehyde (PFA) (w/v) for 20 minutes and stained with Diff-Quik dyes (Medion Diagnostics). Images of the lower row of the plate were acquired with an Olympus MVX10 stereo microscope. Lysis plaque size was measured using the Zen software, Blue edition (Zeiss).

Intracellular Replication and Number of vacuoles

Confluent cells on coverslips were infected with $2x10^5$ freshly released parasites for one hour, and then the remaining extracellular parasites were washed away. 24 hours later, cells were fixed in 4% PFA for 20 minutes, and immunofluorescence staining (IFA) of the inner membrane complex (using αGAP45) was used to count the number of parasites per vacuole and the number of vacuoles per field. The number of vacuoles containing either 2, 4, 6, 8 or 16 parasites were counted using a Leica DM2500, 100x oil objective NA=1.4, microscope (Leica Biosystems), and expressed as percentage. 200

vacuoles/coverslips (n=3 coverslips/experiment) were analyzed. Graphs show means of 6 independent experiments. For the number of vacuoles per field, 20 fields per coverslips were counted, in duplicates. Graphs show means of 2 independent experiments.

Invasion

Cells were infected with $3x10^6$ freshly lysed parasites in 300 µl of complete medium. Parasites incubated for 20 minutes with the plate on ice. The plate was then moved to a water bath at 38°C for 5 minutes. After 5 minutes, invasion was blocked by fixing infected cells with 4% PFA (w/v) in HBSS, (Gibco). Differential IFA was then performed to distinguish intracellular from extracellular parasites. Non-permeabilized samples were saturated with HBSS 10% FBS and then probed with anti-SAG1 antibodies to stain extracellular parasites. Samples were permeabilized with 0.1% saponin for 15 minutes with anti-ROP1 antibodies to detect PVM of intracellular parasites. Intracellular and extracellular parasites were counted in 20 fields per coverslips in triplicate. The values are expressed as the number of intracellular parasites per field and normalized to that of the control line (RH *ΔKu80*) arbitrarily fixed to 100%. Graphs show the mean of 6 independent experiments, with three technical replicates each. For mBMDM, $5x10^5$cells were seeded on coverslips and infected with $1x10^6$ tachyzoites for 1h at 37°C in triplicate for each strain. Non-permeabilized samples were probed with anti-SAG1 antibodies and then permeabilized with 0.2% Triton for 15 minutes to stain extracellular parasites with anti-GAP45 antibodies. The values were expressed as the number of intracellular parasites per field and normalized to that of the control line (RH *ΔKu80*).

Egress

HFFs grown on coverslips in a 24-well plate were infected with 1×10^5 tachyzoites and washed after 2 hours. Following ~30 hours of infection, parasite egress was stimulated with calcium ionophore A23187 (1µM) for 8 minutes and then samples were fixed with 4% PFA and processed for IFA with an antibody recognizing the dense granule protein GRA3. Egress events (GRA3 released in the PV upon egress) were quantified by analyzing 200 vacuoles per coverslip n=3 coverslips). Values were normalized to that of the control line (RH *ΔKu80*) arbitrarily fixed to 100%. Graphs show means of three independent experiments, including three technical replicates each.

Immunofluorescence analysis

To determine the localization of ROP55 in intracellular parasites, HFF cells were infected with ROP55-HA3 tachyzoites for 24 h and fixed with 100% cold methanol for 8 minutes at -20°C. After three washes, infected cells were then blocked with PBS 10% FBS with Triton X-100 0.3%, incubated with primary antibody (anti-HA, anti-ROP55, anti-ROP2 antibodies) for 1h in PBS 2% FBS plus Triton X-100 0.3%, washed three times and then incubated with secondary antibodies (Alexa Fluor 488 and 594- or 647-conjugated antibodies against mouse, rat or rabbit IgG, Molecular Probes). Coverslips were mounted in Immu-Mount (ThermoScientific). For intracellular replication and egress quantifications, cells were permeabilized with Triton X-100 0.1% in PBS for 10 minutes, staining was performed as described above.

Immunoblots

Infected HFF cells were trypsinised at the intended time-point, washed in PBS and the total number of cells was counted. The cell pellet was then resuspended in Laemmli buffer containing 10 % (v/v) of β-mercaptoethanol and protease inhibitors in order to have the same number of cells per mL. For the analysis of IL-1β in HFF cells, cells were grown in serum free media and 5 mL of supernatant were concentrated with Amicon filters (with a cut off of 10 kDa, Millipore) and then precipitated over-night with ice-cold acetone at -20 °C. For IL-1β immunoblot from THP-1 cells, samples were treated as described by Pandori et al.[25]. For GSDMD immunoblot, THP-1 cells were solubilized with 1% NP-40, and the insoluble fraction, was treated for 30 min with DNAse I in 1% NP-40 buffer and then subjected to SDS-PAGE and western blotting. A wet transfer was performed for 1 hour at room-temperature to PVDF membranes. Membranes were blocked for 1 hour at RT with 5% non-fat dry milk in TNT buffer (140 mM NaCl, 15 mM Tris, 0.05% Tween 20). The primary antibodies used: rat anti-HA (3F10, Roche, 1:1000); mouse anti-ROP5; mouse anti-IL-1β (1:5000) (3ZD clone); rabbit anti-Gasdermin D (E8G3F Cell Signalling); mouse anti-Hsc70 (1:1000) (clone 1B5 from Stressgen); rabbit anti-DFNA5/GSDME EPR19859 N-terminal (ab215191 from abcam); rabbit anti-IκBα 44D4 (#4812 from Cell Signaling); mouse anti-caspase-8 1C12 (#9746 from Cell Signaling). After washes with TNT buffer, the membranes were incubated with alkaline phosphatase conjugated secondary antibodies or with horseradish peroxidase (HRP)-conjugated goat anti-mouse or anti rabbit antibodies (Jackson ImmunoResearch).

Quantitative Real Time-PCRs

Cells were washed three times to remove extracellular parasites, centrifuged for 5 minutes at 1800 RPM, then pellets resuspended in RNA lysis buffer RA1 (NucleoSpin, Macherey-Nagel) containing 1% of β-mercaptoethanol. RNA extractions were performed according to manufacturer protocol, which includes DNase treatment to remove contaminating DNA (NucleoSpin, Macherey-Nagel). cDNA synthesis was done using a normalized quantity of RNA between all samples and performed according to manufacturer's instructions (Superscript III First-Strand Synthesis kit (Invitrogen)). For controls, samples containing water or lacking reverse transcriptase were used in cDNA synthesis. cDNA was used as a template in qRT-PCR, performed in triplicate using the SensiFAST™ SYBR No-ROX Kit (Bioline) and a LightCycler 480 Real-Time PCR System (Roche). Primers were designed using the LightCycler Probe Design Software 2.0 and are listed in Supplementary table 4.4. Data were analyzed using the LightCycler 480 program with a relative quantification measurement (ΔCT), normalized to the expression of the housekeeping gene, mouse or human β-actin. For measuring parasitemia, TgGAPDH or TgSAG1 values were normalized to the mouse or human β-actin of samples infected with the control strain. Absence of amplification was confirmed in the negative controls with water and no reverse transcriptase. For absolute quantification of TgSAG1 in DNA extracted from infected THP1 cells, a standard curve was set up with known concentrations of parasite DNA using automatic efficiency calculation from LightCycler® 480 Software, which also determined the concentration of target DNA in the sample using the second derivative maximum method.

LDH Assay

Cell viability was determined with the Pierce LDH Cytotoxicity Assay according to manufacturer's protocol (Thermo scientific). Cells were washed with PBS to remove any phenol red contamination from the culture medium, and infected with freshly lysed tachyzoites at a MOI=2 in DMEM without phenol red (Gibco). At required time-points, the supernatants were harvested, LDH detection assay was performed, and 490nm and 680nm absorbances were measured with a plate reader (TECAN). Lysis Buffer-treated cells were considered as the maximum LDH activity control. Each value was normalized to that of the control arbitrarily fixed to 100%. Graphs shows the mean of three independent experiments, including three technical replicates each.

Propidium Iodide staining for microscopy analysis

To analyze cell death, confluent HFF cells were infected with freshly lysed parasites, the plate was centrifuged for 5 minutes at 1200 RPM, and cells were washed one hour later to remove extracellular parasites. One hour before the intended time-point, propidium iodide (Molecular probes, Life Technologies) at 10 μM and DAPI (1 μg/ml) was added to the wells and incubated for 1 h at 37°C and protected from light. Live imaging of infected cells was performed using an inverted Zeiss microscope and images were acquired with the ZEN software. At least 15 fields were quantified by counting the total number of cells and the nuclei staining positive for propidium iodide. The values were expressed as the % of PI positive cells per field. Graphs shows one independent experiment out of two, including two technical replicates each.

ELISA

Tachyzoites were washed in PBS before use in infection experiments, in order to avoid transferring any IL-1β from the parasite cultures. Confluent HFFs in 24 well plates were then infected with $5x10^5$freshly lysed tachyzoites. THP-1 cells were plated in 24-well plates at $5x10^5$ cells/well and infected at a MOI=2 or stimulated at the same time with 100 ng/mL of *E. coli* LPS (List Biological Laboratories, Campbell, CA) and with 5 mM ATP (Sigma-Aldrich, St. Louis, MO) added in the 30 minutes prior to harvest. Harvested supernatants were centrifuged for 5 minutes at 2000 RPM. ELISA was then performed according to manufacturer's protocol (IL-1β ELISA Kit, Thermofisher). Absorbance at 450 nm was measured with the Magellan software in a plate reader (TECAN). A standard curve of known concentrations of human IL1β was prepared in each experiment to convert the absorbance to IL-1β concentrations in pg/ml. Graphs shows the mean of three independent experiments, with three technical replicates each.

NF-κB activation measurements

For GFP reporter quantifications, glass coverslips were incubated in 24-well plates at 37°C overnight with 20 μg/mL Poly-D lysine prepared in PBS. Reporter cells were then plated at a concentration of $5x10^5$ into each well and 24 hours later cells were infected with 1 million parasites/well. After 6 or 24 hpi, coverslips were fixed with 4% PFA and probed whit anti-GAP45 antibody followed by Alexa Fluor 594-conjugated anti-rabbit IgG (Molecular Probes) to visualize infected cells. For quantification of NF-κB activation in HFFs, a confluent monolayer of cells grown on coverslips in 24-well plates were infected with 1 million parasites/well and then fixed and probed with anti NF-κB p65 (D14E12) Rabbit mAb (1:800) (Cell Signalling) and anti-SAG1 mouse antibodies (1:100) followed by Alexa Fluor 594-conjugated anti-rabbit IgG and Alexa Fluor 488 anti-mouse IgG (Molecular Probes). For quantification of NF-κB activation in THP-1 cells, cells were plated at a concentration of $5x10^5$ in 24-well plates and infected with 1 million parasites/well. After 2 or 6 hours of infection, the cells were fixed with 4% PFA and probed with antibodies as HFF cells in 96 round-bottomed well plates. Image acquisitions were performed with a Leica Thunder microscope and the mean intensity of the fluorescence of the red channel was measured for each nucleus using the quantification tool in the Leica Application Suite X program. The mean value of the fluorescence from at least 100 nuclei for each condition was plotted in GraphPad.

Flow Cytometry

THP-1 cells were plated at $5x10^5$ cells/well and infected with the indicated strain at a MOI=2. Cells were harvested at 2 or 16 hours after infection and resuspended in binding buffer (10 mM HEPES, 140 mM NaCl, and 2.5 mM CaCl2, pH 7.4). Propidium Iodide (Life Technologies, Eugene, OR) was used at 1 μg/mL per 100 μL of resuspended cells. Cells were stained for 15 minutes at room temperature, protected from light. 400 μL of buffer was then added to the cells. The cells were immediately analyzed using a NovoCyte flow cytometer with NovoExpress software (Agilent) or a FACSCalibur flow cytometer with CellQuest Software (BD Biosciences). FCS files were analyzed using FlowJo software (TreeStar, Ashland, OR).

PI values were normalized to the non-infected controls considered as 1 for each cell line. For Caspase 3/7 activity assay, cells were labelled with the CellEvent™ probe C10423 from Invitrogen for 30 minutes according with the manufacturer instructions.

Total RNA isolation for RNA sequencing analysis

Total RNA from cell pellets was resuspended in RNA lysis buffer RA1 (NucleoSpin, Macherey-Nagel) containing 1% of β-mercaptoethanol and extracted according to the manufacturer's protocol (NucleoSpin, Macherey-Nagel). To remove potential DNA contamination, RNA samples were treated twice with 1 U DNase I (Life Technologies) per 10 µg of RNA for 30 minutes at 37°C, followed by inactivation of the DNase I enzyme.

mRNA isolation and cDNA synthesis

mRNA was isolated using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7420L) per the manufacturer instructions. Briefly, 8 µl NEBNext first strand synthesis reaction buffer, 2 µl NEBNext primers, and 10 µl nuclease-free buffer, were mixed and placed in ice. 20 µl of NEBNext oligo d(T)-25 beads were washed with 100 µl NEBNext RNA binding buffer (2X) in a 1.5 ml microcentrifuge tube and placed on a magnetic rack at room temperature for 2 minutes. The supernatant was removed, and beads were washed with another 100 µl of RNA binding buffer (2X). The beads were resuspended in a volume of 50 µl RNA binding buffer (2X), transferred to 0.2 ml PCR tubes, and 50 µl of total RNA was added. To denature the RNA and facilitate binding to the beads, tubes were heated in a thermocycler at 65°C for 5 minutes with the lid set to 75°C. To assist further binding of mRNA, beads were resuspended, and samples let sit at room temperature for 5 minutes for two cycles. To remove contaminants, we placed the samples on the magnetic rack for 2 minutes and removed the supernatant and washed in 200 µl of RNA wash buffer twice. 50 µl of NEBNext Tris buffer was added and tubes were placed in a thermocycler at 80°C for 2 minutes with the lid set to 90°C, then cooled to 25°C. The beads were mixed with 50 µl of RNA binding buffer (2X) at room temperature for 5 minutes. The supernatant was removed and 200 µl of NEBNext wash buffer was added to each tube. Samples were mixed, placed back on the rack for 2 minutes at room temperature, and the supernatant removed. After spinning down and removing any additional wash buffer, the RNA was eluted by adding 15.5 µl of the prepared first strand reaction buffer and incubating at 94°C for 15 minutes. Using the magnet, 13.5 µl of supernatant from each sample was transferred to clean PCR tubes and placed on ice.

First strand cDNA was synthetized also using the NEBNext Ultra II Directional RNA Library Prep per the manufacturer instructions. RNA polymerase inhibitor, Actinomycin D, was dissolved in nuclease-free water to 0.1 µg/µl and 5 µl was added to each sample along with 0.5 µl murine RNAse inhibitor and 1 µl ProtoScript II reverse transcriptase. Samples were then incubated in a thermocycler at 25°C for 10 minutes, 42°C for 15 minutes, and 70°C for 15 minutes. The Second strand was synthesized by adding 48 µl of nuclease-free water, 8 µl NEBNext second strand reaction buffer, and 4 µl of second strand synthesis enzyme to each sample and incubating in a thermocycler for 1 hour at 16°C with the lid set to 40°C. cDNA was then purified using 1.8 volumes (144 µl) of AMPure XP magnetic beads (ref# A63881) to cDNA samples and incubated at room temperature for 5 minutes. cDNA was separated using a magnet rack and the supernatant removed. The beads were washed twice with 500 µl of 80% ethanol, the supernatant removed, and beads were spun down one more time to remove any excess ethanol. Beads were dried for 5 minutes on the magnet and resuspended in 60 µl of 0.1X TE buffer and incubated at room temperature for 2 minutes. The supernatant was collected and transferred to a new PCR tube. Ends were repaired using 6.5 µl of NEBNext end repair reaction buffer (10X) and 3 µl of NEBnext end prep enzyme mix and incubated in a thermocycler at 20°C for 30 minutes and at 65°C for 30 minutes with the lid set to 75°C.

Adapter ligation for next generation sequencing and PCR enrichment

Adapters were ligated using 15 µl of NEB blunt/TA ligase master mix, 1 µl of 1.5 µM NEBNext multiplexing Illumina adaptors, and 2.5 µl nuclease-free water to each sample. Samples were mixed and incubated at 20°C for 15 minutes and an additional 16.5 µl of nuclease-free water added. The ligation reactions were purified on DNA beads using 100 µl of AMPure beads and eluted in 52 µl of 0.1X TE buffer. 50 µl of supernatant was transferred to a new tube and purified again, this time using 50 µl of AMPure beads and eluted in 19 µl of 0.1X TE buffer. 17 µl of the supernatant was mixed with 10 µl of universal barcode/primer, 25 µl KAPA HiFi HotStart Ready Mix

(2X), and 3 µl of NEBNext USER enzyme. cDNA was amplified in a thermocycler at 37°C for 15 minutes, 12 cycles of 98°C for 30 seconds, 55°C for 10 seconds, and 62°C for 75 seconds, then 62°C for 5 minutes. cDNA was purified using 45 µl of AMPure beads and eluted with 23 µl of 0.1X TE buffer. Libraries were sequenced using the Novaseq 6000 sequencing system from Illumina in order to obtain at least 50 million reads per sample.

RNAseq pipeline

Quality of the fastq sequences was checked using FastQC and per base sequence quality and content were used to determine the number of nucleotides to trim from the ends using Trimmomatic[71]. Sequences were aligned to the most recent version of the human genome (GRCh38.p13) using Hisat2[72]. Samtools[73] was used for filtering, sorting and indexing. We utilized HTSeq[74] and the human genome feature file (gff) to count the number of uniquely mapped reads for each gene, then performed differential expression analysis using DESeq2[75]. Using the differential expression data from DESeq2, we performed GO enrichment using topGO[76].

IPA analysis

Using the Ingenuity Pathway Analysis (IPA)[77] software (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis), the lists of differentially expressed genes output from DESeq2 were analyzed more deeply. Tab-delimited matrices of the differentially expressed genes, including -log10 p-value and log2 fold change, were loaded into IPA and an initial core analysis was performed by the software. The core analysis provided summarized data about listed genes regarding canonical pathways involved, upstream regulators, diseases and disorders, molecular and cellular functions, physiological and system development functions, functions in toxicity, regulator effect networks, and associated network functions. Additionally, figures were generated from regulatory networks and associated molecular and cellular functions to visualize the data.

In vivo studies

Eight to ten weeks old female BALB/c mice were intraperitoneally injected with wild type (WT), knock-out ROP55 (ΔROP55) or complemented parasites. To assess the ΔROP55 virulence *in vivo*, freshly harvested tachyzoites (10^2, 10^4, 10^5) of WT or ΔROP55 parasites, or lethal dose (10^6) ΔROP55 parasites were intraperitoneally injected into 5-10 mice per condition and monitored for survival. To study parasite load, mice were infected with 10^3 tachyzoites and 4 dpi the animals were sacrificed, and spleens were collected in order to perform RT-qPCR.

Statistical analysis

GraphPad was used for all statistical analyses other than RNA-seq. A normality test was performed to analyze sample distribution. When data was normally distributed and had equal variance, a two-tailed unpaired t-test was used. When data was normally distributed, but sample size and variance were unequal, a Welch t-test was used. For non-normally distributed data, a two-tailed Mann-Whitney U test was used. Group analyses were performed with Tukey's or Sidaks's multiple comparison tests. Statistical significance was considered for p-values below 0.05. Bar graphs show the mean ± SD.

ACKNOWLEDGEMENTS

We thank Sebastian Lourido for the pU6-Universal plasmid and Dominique Soldati-Favre for anti-GAP45 antibodies. We also thank J. Saeij for kindly gifting HEK 293 NF-κB reporter cell line (System Biosciences, Cat. # TR860A-1). We are grateful to Derek Abbott for sharing the ΔGSDMD THP-1 line. We also thank Elodie Jublanc and Vicky Diakou at the imaging facility MRI at the University of Montpellier, part of the national infrastructure France-BioImaging supported by the French National Research Agency (ANR-10-INBS-04). Dr Maryse Lebrun is an INSERM researcher.

FUNDING

This work was supported by the Laboratoire d'Excellence (LabEx) (ParaFrap ANR-11-LABX-0024) and Fondation pour la Recherche Médicale (Equipe FRM DEQ20130326508) to M.L., the National Institutes of Allergy and Infectious Diseases and the National Institutes of Health (R01 AI136511 and R21 AI142506) and the University of California, Riverside (NIFA-Hatch 225935) to K.L.R, NIH (R01AI120846 and R21AI156452) to M.B.L. M.T.G.R. was funded by the French Parasitology consortium ParaFrap (ANR-11-LABX-0024) and was a recipient of FRM (award no. FDT20190400862).

AUTHOR CONTRIBUTIONS

Margarida T Grilo Ruivo: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – Original draft, Project administration. **Ji-hun Shin**: Methodology, Investigation, Resources, Writing – Review and editing. **Todd Lenz**: Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Writing – Review and editing. **Stephanie Y Matsuno**: Investigation. **Katherine Olivia Yanes**: Investigation. **Arnault Graindorge**: Investigation. **Maguy Hamie**: Methodology, Investigation. **Mathieu Gissot**: Investigation. **Hiba El Hajj**: Investigation. **Karine G Le Roch**: Methodology, Formal analysis, Investigation. **Melissa B Lodoen**: Methodology, Investigation. **Maryse Lebrun**: Conceptualization, Writing – Original draft, Supervision, Project administration. **Diana Marcela Penarete-Vargas**: Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original draft, Supervision.

CONFLICTS OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY

All data are present in the main text and the supplementary materials. All biological materials and data are available from the authors upon request. Source data are provided with this paper. RNA-seq data is available in NCBI SRA https://www.ncbi.nlm.nih.gov/sra (PRJNA881613).

REFERENCES

1. Flegr, J., Prandota, J., Sovičková, M. & Israili, Z. H. Toxoplasmosis--a global threat. Correlation of latent toxoplasmosis with specific disease burden in a set of 88 countries. *PLoS ONE* **9**, e90203 (2014).

2. Pappas, G., Roussos, N. & Falagas, M. E. Toxoplasmosis snapshots: global status of Toxoplasma gondii seroprevalence and implications for pregnancy and congenital toxoplasmosis. *Int. J. Parasitol.* **39**, 1385-1394 (2009).

3. Montoya, J. G. & Liesenfeld, O. Toxoplasmosis. *Lancet* **363**, 1965-1976 (2004).

4. Demar, M. et al. Fatal outbreak of human toxoplasmosis along the Maroni River: epidemiological, clinical, and parasitological aspects. *Clin. Infect. Dis.* **45**, e88-95 (2007).

5. Blaizot, R. et al. Outbreak of Amazonian Toxoplasmosis: A One Health Investigation in a Remote Amerindian Community. *Front. Cell. Infect. Microbiol.* **10**, 401 (2020).

6. Hunter, C. A. & Sibley, L. D. Modulation of innate immunity by Toxoplasma gondii virulence effectors. *Nat. Rev. Microbiol.* **10**, 766-778 (2012).

7. Boothroyd, J. C. & Hakimi, M.-A. in *Toxoplasma gondii* (eds. Weiss, L. M. & Kim, K.) 789-806 (Academic Press, 2020).

8. Barragan, A. & Sibley, L. D. Transepithelial migration of Toxoplasma gondii is linked to parasite motility and virulence. *J. Exp. Med.* **195**, 1625-1633 (2002).

9. Jorgensen, I., Rayamajhi, M. & Miao, E. A. Programmed cell death as a defence against infection. *Nat. Rev. Immunol.* **17**, 151-164 (2017).

10. Tummers, B. & Green, D. R. The evolution of regulated cell death pathways in animals and their evasion by pathogens. *Physiol. Rev.* **102**, 411-454 (2022).

11. Broz, P. & Dixit, V. M. Inflammasomes: mechanism of assembly, regulation and signalling. *Nat. Rev. Immunol.* **16**, 407-420 (2016).

12. Bauernfeind, F. G. et al. Cutting edge: NF-kappaB activating pattern recognition and cytokine receptors license NLRP3 inflammasome activation by regulating NLRP3 expression. *J. Immunol.* **183**, 787-791 (2009).

13. Groslambert, M. & Py, B. F. Spotlight on the NLRP3 inflammasome pathway. *J. Inflamm. Res.* **11**, 359-374 (2018).

14. Grootjans, S., Vanden Berghe, T. & Vandenabeele, P. Initiation and execution mechanisms of necroptosis: an overview. *Cell Death Differ.* **24**, 1184-1195 (2017).

15. Orozco, S. L. et al. RIPK3 Activation Leads to Cytokine Synthesis that Continues after Loss of Cell Membrane Integrity. *Cell Rep.* **28**, 2275-2287 (2019).

16. Meylan, E. et al. RIP1 is an essential mediator of Toll-like receptor 3-induced NF-kappa B activation. *Nat. Immunol.* **5**, 503-507 (2004).

17. Kaiser, W. J., Upton, J. W. & Mocarski, E. S. Receptor-Interacting Protein Homotypic Interaction Motif-Dependent Control of NF-κB Activation via the DNA-Dependent Activator of IFN Regulatory Factors. *J. Immunol.* **181**, 6427-6434 (2008).

18. Moriwaki, K. et al. The Necroptosis Adaptor RIPK3 Promotes Injury-Induced Cytokine Expression and Tissue Repair. *Immunity* **41**, 567-578 (2014).

19. Mocarski, E. S., Guo, H. & Kaiser, W. J. Necroptosis: The Trojan horse in cell autonomous antiviral host defense. *Virology* **479-480**, 160-166 (2015).

20. Rosenberg, A. & Sibley, L. D. Toxoplasma gondii secreted effectors co-opt host repressor complexes to inhibit necroptosis. *Cell Host Microbe* **29**, 1186-1198 (2021).

21. Cavaillès, P. et al. The rat Toxo1 locus directs toxoplasmosis outcome and controls parasite proliferation and spreading by macrophage-dependent mechanisms. *Proc. Natl Acad. Sci. USA* **103**, 744-749 (2006).

22. Cirelli, K. M. et al. Inflammasome sensor NLRP1 controls rat macrophage susceptibility to Toxoplasma gondii. *PLoS Pathog.* **10**, e1003927 (2014).

23. Gov, L., Karimzadeh, A., Ueno, N. & Lodoen, M. B. Innate immunity to Toxoplasma gondii infection. *mBio* **4**, e00255-13 (2013).

24. Gov, L., Schneider, C. A., Lima, T. S., Pandori, W. & Lodoen, M. B. NLRP3 and Potassium Efflux Drive Rapid IL-1β Release from Primary Human Monocytes during Toxoplasma gondii Infection. *J. Immunol.* **199**, 2855-2864 (2017).

25. Pandori, W. J. et al. Toxoplasma gondii activates a Syk-CARD9-NF-κB signaling axis and gasdermin D-independent release of IL-1β during infection of primary human monocytes. *PLoS Pathog.* **15**, e1007923 (2019).

26. Rosowski, E. E. et al. Strain-specific activation of the NF-κB pathway by GRA15, a novel Toxoplasma gondii dense granule protein. *J. Exp. Med.* **208**, 195-212 (2011).

27. Lima, T. S., Gov, L. & Lodoen, M. B. Evasion of Human Neutrophil-Mediated Host Defense during Toxoplasma gondii Infection. *mBio* **9**, e00973-17 (2018).

28. Butcher, B. A., Kim, L., Johnson, P. F. & Denkers, E. Y. Toxoplasma gondii tachyzoites inhibit proinflammatory cytokine induction in infected macrophages by preventing nuclear translocation of the transcription factor NF-kappa B. *J. Immunol.* **167**, 2193-2201 (2001).

29. Shapira, S. et al. Initiation and termination of NF-kappaB signaling by the intracellular protozoan parasite Toxoplasma gondii. *J. Cell Sci.* **118**, 3501-3508 (2005).

30. Butcher, B. A. & Denkers, E. Y. Mechanism of entry determines the ability of Toxoplasma gondii to inhibit macrophage proinflammatory cytokine production. *Infect. Immun.* **70**, 5216-5224 (2002).

31. Camejo, A. et al. Identification of three novel Toxoplasma gondii rhoptry proteins. *Int. J. Parasitol.* **44**, 147-160 (2014).

32. Sidik, S. M. et al. A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes. *Cell* **166**, 1423-1435 (2016).

33. Sigrist, C. J. et al. New and continuing developments at PROSITE. *Nucleic Acids Res.* **41**, D344-347 (2013).

34. Zimmermann, L. et al. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J. Mol. Biol.* **430**, 2237-2243 (2018).

35. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845-858 (2015).

36. Barylyuk, K. et al. A Comprehensive Subcellular Atlas of the Toxoplasma Proteome via hyperLOPIT Provides Spatial Context for Protein Functions. *Cell Host Microbe* **28**, 752-766 (2020).

37. Kovacs, S. B. & Miao, E. A. Gasdermins: Effectors of Pyroptosis. *Trends Cell Biol.* **27**, 673-684 (2017).

38. Nguyen, H. N. et al. Autocrine Loop Involving IL-6 Family Member LIF, LIF Receptor, and STAT4 Drives Sustained Fibroblast Production of Inflammatory Mediators. *Immunity* **46**, 220-232 (2017).

39. Broz, P. & Monack, D. M. Newly described pattern recognition receptors team up against intracellular pathogens. *Nat. Rev. Immunol.* **13**, 551-565 (2013).

40. Robben, P. M. et al. Production of IL-12 by macrophages infected with Toxoplasma gondii depends on the parasite genotype. *J. Immunol.* **172**, 3686-3694 (2004).

41. Dunay, I. R. et al. Gr1(+) inflammatory monocytes are required for mucosal resistance to the pathogen Toxoplasma gondii. *Immunity* **29**, 306-317 (2008).

42. Robben, P. M., LaRegina, M., Kuziel, W. A. & Sibley, L. D. Recruitment of Gr-1+ monocytes is essential for control of acute toxoplasmosis. *J. Exp. Med.* **201**, 1761-1769 (2005).

43. Coll, R. C. et al. MCC950 directly targets the NLRP3 ATP-hydrolysis motif for inflammasome inhibition. *Nat. Chem. Biol.* **15**, 556-559 (2019).

44. Pandori, W. J. et al. Unique role for caspase 8 in the release of IL-1β and active caspase-1 from viable human monocytes during Toxoplasma gondii infection. *bioRxiv* https://doi.org/10.1101/2022.01.01.123456 (2022).

45. Rogers, C. et al. Cleavage of DFNA5 by caspase-3 during apoptosis mediates progression to secondary necrotic/pyroptotic cell death. *Nat. Commun.* **8**, 14128 (2017).

46. Degterev, A. et al. Identification of RIP1 kinase as a specific cellular target of necrostatins. *Nat. Chem. Biol.* **4**, 313-321 (2008).

47. Zhuang, C. & Chen, F. Small-Molecule Inhibitors of Necroptosis: Current Status and Perspectives. *J. Med. Chem.* **63**, 1490-1510 (2020).

48. Sun, L. et al. Mixed Lineage Kinase Domain-like Protein Mediates Necrosis Signaling Downstream of RIP3 Kinase. *Cell* **148**, 213-227 (2012).

49. Matsuno, S. Y., Pandori, W. J. & Lodoen, M. B. Capers with caspases: Toxoplasma gondii tales of inflammation and survival. *Curr. Opin. Microbiol.* **72**, 102264 (2023).

50. Braun, L. et al. The Toxoplasma effector TEEGR promotes parasite persistence by modulating NF-κB signalling via EZH2. *Nat. Microbiol.* **4**, 1208-1220 (2019).

51. Tomita, T. et al. Toxoplasma gondii Matrix Antigen 1 Is a Secreted Immunomodulatory Effector. *mBio* **12**, e02825-20 (2021).

52. Fisch, D. et al. Human GBP1 is a microbe-specific gatekeeper of macrophage apoptosis and pyroptosis. *EMBO J.* **38**, e100926 (2019).

53. Chen, K. W. et al. Extrinsic and intrinsic apoptosis activate pannexin-1 to drive NLRP3 inflammasome assembly. *EMBO J.* **38**, e101638 (2019).

54. Koehler, H. et al. Inhibition of DAI-dependent necroptosis by the Z-DNA binding domain of the vaccinia virus innate immune evasion protein, E3. *Proc. Natl Acad. Sci. USA* **114**, 11506-11511 (2017).

55. Ashida, H., Sasakawa, C. & Suzuki, T. A unique bacterial tactic to circumvent the cell death crosstalk induced by blockade of caspase-8. *EMBO J.* **39**, e104469 (2020).

56. Thapa, R. J. et al. DAI Senses Influenza A Virus Genomic RNA and Activates RIPK3-Dependent Cell Death. *Cell Host Microbe* **20**, 674-681 (2016).

57. Peng, R. et al. Human ZBP1 induces cell death-independent inflammatory signaling via RIPK3 and RIPK1. *EMBO Rep.* **23**, e55839 (2022).

58. Rebsamen, M. et al. DAI/ZBP1 recruits RIP1 and RIP3 through RIP homotypic interaction motifs to activate NF-κB. *EMBO Rep.* **10**, 916-922 (2009).

59. Upton, J. W., Kaiser, W. J. & Mocarski, E. S. Cytomegalovirus M45 cell death suppression requires receptor-interacting protein (RIP) homotypic interaction motif (RHIM)-dependent interaction with RIP1. *J. Biol. Chem.* **283**, 16966-16970 (2008).

60. Fliss, P. M. et al. Viral mediated redirection of NEMO/IKKγ to autophagosomes curtails the inflammatory cascade. *PLoS Pathog.* **8**, e1002517 (2012).

61. Henry, B., Sibley, L. D. & Rosenberg, A. A Combination of Four Nuclear Targeted Effectors Protects Toxoplasma Against Interferon Gamma Driven Human Host Cell Death During Acute Infection. *bioRxiv* https://doi.org/10.1101/2023.01.01.123456 (2023).

62. Cervantes, P. W., Martorelli Di Genova, B., Erazo Flores, B. J. & Knoll, L. J. RIPK3 Facilitates Host Resistance to Oral Toxoplasma gondii Infection. *Infect. Immun.* **89**, e00120-21 (2021).

63. Zhang, T. et al. Influenza Virus Z-RNAs Induce ZBP1-Mediated Necroptosis. *Cell* **180**, 1115-1129 (2020).

64. Pittman, K. J., Cervantes, P. W. & Knoll, L. J. Z-DNA Binding Protein Mediates Host Control of Toxoplasma gondii Infection. *Infect. Immun.* **84**, 3063-3070 (2016).

65. Nozaki, K., Li, L. & Miao, E. A. Innate Sensors Trigger Regulated Cell Death to Combat Intracellular Infection. *Annu. Rev. Immunol.* **40**, 469-498 (2022).

66. Maelfait, J. et al. Sensing of viral and endogenous RNA by ZBP1/DAI induces necroptosis. *EMBO J.* **36**, 2529-2543 (2017).

67. Niedelman, W., Sprokholt, J. K., Clough, B., Frickel, E.-M. & Saeij, J. P. J. Cell death of gamma interferon-stimulated human fibroblasts upon Toxoplasma gondii infection induces early parasite egress and limits parasite replication. *Infect. Immun.* **81**, 4341-4349 (2013).

68. Pandori, W. J. et al. Role for Caspase-8 in the Release of IL-1β and Active Caspase-1 from Viable Human Monocytes during Toxoplasma gondii Infection. *J. Immunol.* **212**, 1161-1171 (2024).

69. Gonçalves, R. & Mosser, D. M. The Isolation and Characterization of Murine Macrophages. *Curr. Protoc. Immunol.* **111**, 14.11.11-14.11.16 (2015).

70. Bobak, D. A., Frank, M. M. & Tenner, A. J. Characterization of C1q receptor expression on human phagocytic cells: effects of PDBu and fMLP. *J. Immunol.* **136**, 4604-4610 (1986).

71. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

72. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357-360 (2015).

73. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

74. Anders, S., Pyl, P. T. & Huber, W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).

75. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

76. Alexa, A. & Rahnenfuhrer, J. topGO: Enrichment Analysis for Gene Ontology. R package version 2.58.0 (2024).

77. Krämer, A., Green, J., Pollard Jr., J. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523-530 (2014).

CHAPTER 5

Toxoplasma gondii reorganizes the host chromatin structure and epigenetic features

Todd Lenz[1], Sandeep Srivastava[2], Edward Vizcarra[2], Emma Wilson[2] and Karine G. Le Roch[1,*]

[1] Department of Molecular, Cell and Systems Biology, University of California Riverside, 900 University Avenue, Riverside, CA 92521, USA.
[2] Biomedical Sciences, University of California Riverside, 900 University Avenue, Riverside, CA 92521, USA.
*Correspondence: karinel@ucr.edu

The work presented in this chapter is ongoing and therefore any conclusions drawn from the data are preliminary. Sandeep collected most of the parasites and Edward aided in establishing a gating protocol for FACS. I performed all experiments and data analyses, generated all figures, and wrote this chapter.

ABSTRACT

Toxoplasma gondii remains one of the most pervasive pathogens globally due to its ability to employ various strategies to evade host immune responses. While there is extensive research on the transcriptional effects of secreted effector proteins, little is known about the global impact of *T. gondii* infection on host cell chromatin architecture. Here, we investigated chromatin structural changes in human foreskin fibroblasts (HFFs) during *T. gondii* infection using RNA-seq and Hi-C analyses at 6- and 24-hours post-infection. RNA-seq revealed extensive transcriptional changes, including triggering an early robust pro-inflammatory response, followed by increased anti-inflammatory signaling and suppression of immune-related genes by 24 hours. Hi-C analysis demonstrated significant alterations to host chromatin architecture, including weakening of A compartments, strengthening of B compartments, and disruption of topologically associating domain (TAD) boundaries. We observed a marked decrease in long-range chromatin interactions and increased local interactions within TAD-like structures, suggesting widespread chromatin reorganization. These findings reveal previously unknown mechanisms by which *T. gondii* manipulates host cell chromatin structure to potentially regulate gene expression and ensure parasite survival, providing new insights into host-pathogen interactions during toxoplasmosis.

INTRODUCTION

The ability to circumvent the host immune response is crucial for the survival and subsequent proliferation of eukaryotic pathogens. Understanding the underlying genetic and epigenetic mechanisms that govern these processes will enable us to develop more targeted therapeutic strategies against these infectious agents. *Toxoplasma gondii* is an intracellular protozoan parasite of the phylum Apicomplexa that infects most species of warm-blooded animals, including humans. In North American and Europe, *T. gondii* is comprised of three clonal lineages denoted Types I, II and III which differ in their virulence[1,2]. Type I strains are more virulent and are associated with severe acute infections, while Type II and to some extent Type III are significantly more common in humans and concomitant with outbreaks of Toxoplasmosis, causing long-term chronic infections by sequestering in tissue cysts. Strains from Asia, Africa and South America display more genetic diversity and instead stem from region-specific clonal lineages[3–5].

T. gondii persist within a host cell inside of parasitophorous vacuoles (PV) generated from parasite the host cell plasma membranes. The parasite is however able to export effector proteins across the PV membrane (PVM) to subvert the host cellular functions. Some of these effector proteins include rhoptry organelle proteins (ROPs) that can function as kinases to modulate host cell protein functions. Within the cytoplasm, a ROP5/ROP18/ROP17 complex has been shown to phosphorylate host immunity-related GTPases (IRGs) to prevent GTP hydrolysis of the PV[6,7]. Elevated ROP18 is thought to be responsible for the pathogenicity of Type I parasites due to its early expression immediately after host cell invasion, allowing the parasite to evade autonomous host cell immunity[8,9]. Once the parasite is established within the PV, ROP16 is transported into the nucleus and phosphorylates STAT3 and STAT6, two transcription factors that control the suppression of pro-inflammatory cytokine IL-12 and the Th1 cytokine response[10–12]. Th1 cells produce vast amounts of IFN-γ, which activates STAT1, inducing expression of several genes that aid in destruction of the parasite. STAT1-dependent transcription is inhibited by the *T. gondii* inhibitor of STAT1 transcriptional activity (TgIST) protein, which does so by recruiting the Mi-2 nucleosome remodeling and deacetylase complex (Mi-2/NuRD) to STAT1-dependent promoters[13,14]. *T. gondii* NCoR/SMRT modulator (TgNSM) serves a similar function by targeting a corepressor complex (NCoR/SMRT) and recruiting histone deacetylase enzymes such as HDAC3 to promoter regions resulting in transcriptional repression of PKR and MLKL, proteins necessary for initiation of necroptosis, a form of inflammatory programmed cell death[13]. *Toxoplasma* E2F4-associated EZH2-inducing gene regulator (TEEGR) also manipulates transcriptional regulation by binding E2F/DP transcription factors, inducing expression and recruitment of EZH2, a catalytic subunit of the Polycomb repressive complex 2 (PRC2) that inhibits expression of NF-κB regulated genes through H3K27me3 marks[15].

What makes TgIST, TgNSM, TEEGR, and other similar effector proteins unique is that they manipulate host cell signaling through epigenetic changes by altering expression of chromatin regulatory factors or by directly targeting host chromatin **(Fig. 5.1A)**[14]. Epigenetic reprogramming has been studied extensively during viral infection. Epstein-Barr virus, vaccinia virus, Human T-lymphotropic virus type I and SARS-CoV-2 are shown to modulate host gene expression through chromatin remodeling via triggering loop formation, disrupting higher order chromatin structure and widespread weakening of A compartments[16–19]. Along with hepatitis B virus and hepatitis C virus, EBV upregulates DNA methyltransferases resulting in CpG methylation silencing DNA promoter regions, thus viruses utilize a multi-faceted approach to controlling host transcription[20–22].

Direct evidence that demonstrate that *T. gondii* can directly alter 3D chromatin structure and epigenetic features to control host gene expression is still limited. In this study we aimed to expand our understanding of host-pathogen interactions during *T. gondii* infections by identifying global changes in the host chromatin architecture, histone modifications and DNA methylation to suppress the host immune response allowing parasite survival and proliferation. We compared uninfected human foreskin fibroblasts (HFFs) to those infected by *T. gondii* ME49 at 6- and 24-hours post-infection to identify the series of changes to the chromatin and epigenetic landscape that will ultimately lead to parasite survival.

RESULTS

T. gondii suppresses the host immune response and disrupts cell signaling

To identify change in gene expression during infection, we infected human foreskin fibroblasts (HFFs) with an RFP-expressing line of *T. gondii*. Infected and non-infected cells were then then carefully isolated using flow cytometry and fluorescence activated cell sorting (FACS) at 6 and 24 hours post inoculation (hpi) to mitigate cell lysis **(Fig. 5.1B)**. Immediately following collection, a homogeneous infected population was separated from cellular lysate and free parasites using FACS **(Fig. 5.1C-D)**. Because the parasite undergoes multiple rounds of endodyogeny within a 24-hour period, tighter gating had to be utilized at the 24 hpi time point to ensure that only infected HFFs were being collected.

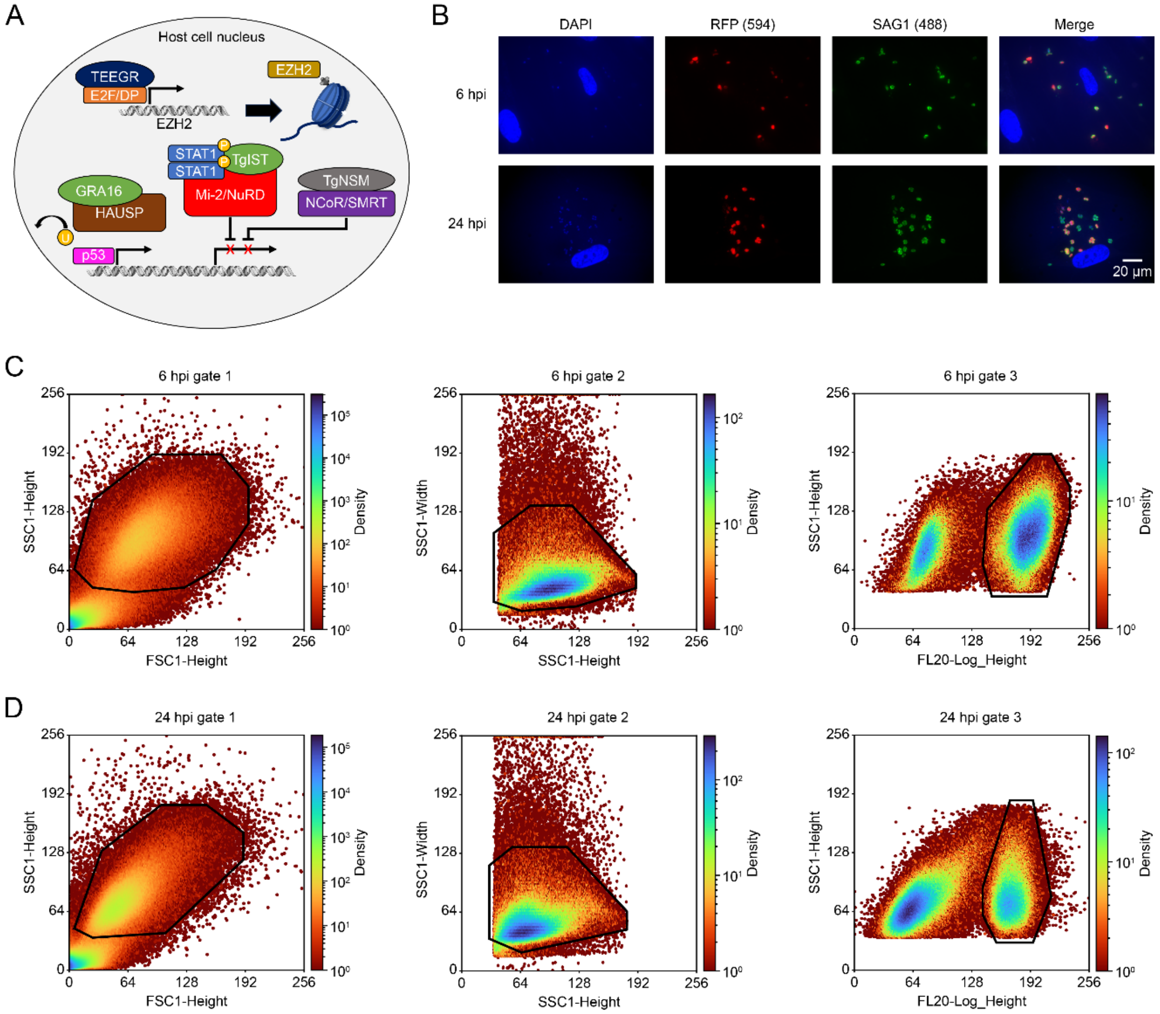

Figure 5.1. Collection of infected population of HFFs. (A) *T. gondii* effector proteins exported into the nucleus inhibit the innate immune response, promote cell cycle arrest and block programmed cell death. **(B)** Immunofluorescence assay of *T. gondii* ME49 displaying widespread RFP expression. HFF and *T. gondii* nuclei are stained with DAPI and the *T. gondii* plasma membrane is stained with SAG1, both of which colocalize with RFP inside *T. gondii* tachyzoites.**(C, D)** Collecting cells at 6 hpi **(C)** and 24 hpi **(D)** via FACS using a series of gates to filter debris from lysed cells and remove doublets and non-fluorescent cells ensures a highly homogeneous population and enhances biological interpretability of RNA-seq and epigenetic analyses. Data is binned down to a 256x256 grid with color indicating the $\log_{10}$ scaled point density. Most axes are scaled linearly, but the fluorescent intensity (FL20-Log_Height) is $\log_{10}$ scaled due to the wide range of fluorescent signals.

RNA-seq analysis of infected HFFs revealed strong spearman correlation between infected cells at both time points likely due to the large number of differentially expressed genes in response to parasite invasion **(Supp. Fig. 5.1A)**. Transcriptomic analysis identified 5,436 genes (FDR < 0.05, $\log_2$FC > 0.5) differentially expressed at 6 hpi (3,930 upregulated and 1,506 downregulated) and 11,673 genes at 24 hpi (8,118 upregulated and 3,555 downregulated) versus uninfected HFFs **(Supp Table 5.1-5.3)**. Gene set enrichment analysis of highly variable genes identified an over-representation of genes involved in the immune system, signal transduction and cellular response **(Fig. 5.2A, Supp. Fig. 5.2A)**. There appears to be a direct correlation between time point and expression of highly variable genes, which is expected considering the number of differentially expressed genes at 24 hpi. Further investigation of these three biological processes reveals that while the number of upregulated genes for each pathway increases by 50% from 6 hpi to 24 hpi, the number of downregulated genes increases more than 300% **(Fig. 5.2B)**.

Pathway analysis of differentially expressed genes at 6 hpi indicated a robust early immune response through upregulation of pro-inflammatory genes involved in interleukin signaling, including IL-6 and IL-11 (4.5-4.7 $\log_2$FC) **(Supp. Table 5.2, 5.4, Supp. Fig. 5.2C)**[23–25]. Several pro-inflammatory chemokines were also significantly upregulated (7.0-9.7 $\log_2$FC) at 6 hpi, such as CXCL1, CXCL2 and CXCL8, the latter of which is a potent chemoattractant for neutrophils and important for rapid immune responses[26–29]. Although the pro-inflammatory response was strong at 6 hpi, there was elevated expression of anti-inflammatory genes as well. Five of the six core members of the IL-10 family of cytokines (IL-10, IL-19, IL-20, IL-24 and IL-26) were highly upregulated (4.4-6.8 $\log_2$FC), working to suppress production of pro-inflammatory cytokines[30,31]. Similar expression profiles of pro-inflammatory cytokines/chemokines is seen in fibroblasts affected by immune-mediated inflammatory diseases such as IBD, rheumatoid arthritis and Crohn's disease[32,33]. Due to their annotations and functional overlap, most of the upregulated pathways at 6 hpi can be broadly categorized as cell signaling or immune system **(Supp. Table 5.4)**. Strangely, although there are over 1,500 genes that are downregulated at 6 hpi, pathway analysis did not reveal any significantly enriched biological processes.

The immune response was not as pronounced at 24 hpi as IL-6 and IL-11 are not differentially expressed compared to the uninfected HFFs and are thus downregulated from 6 hpi indicating a return to the homeostatic environment more akin to uninfected cells **(Supp. Table 5.3)**. This could be the result of overexpression of IL-10 cytokines counteracting production of IL-6 and decreased STAT1 and STAT2 production (-1.1 to -2.3 loc_2FC) as these crucial transcription factors are required for stimulating transcription within the nucleus via interferon gamma signaling[34,35]. Reduced STAT1 signaling also coincides with the transcriptional repression of several guanylate-binding proteins (GBPs), inhibiting cell-autonomous immunity[36–38], and MHC class I human leukocyte antigens (HLAs), preventing recognition of infected fibroblasts by T lymphocytes[39]. Pathway analysis of differentially expressed genes at 24 hpi deviated from 6 hpi as many cell signaling pathways were no longer upregulated and there is activation of DNA organization and repair pathways **(Supp. Table 5.5, Supp. Fig. 5.2D)**. Furthermore, downregulated genes at 24 hpi were associated with metabolic pathways, likely to conserve energy, as well as genes associated with extracellular matrix organization and maintenance **(Supp. Fig. 5.2E)**. Some evidence suggests that the ECM improves immune cell motility and some ECM-derived proteins act as neutrophil chemoattractants, thus a loss in ECM gene expression could be another mechanism by which the parasite evades the innate immune response[40–42]. Time course plots demonstrate three possible transcriptional fates for these immune genes during the course of the infection: (1) rapid transcriptional upregulation following pathogen invasion followed by sustained or moderately reduced expression over time, (2) initial upregulation followed by a return to baseline expression similar to levels detected in uninfected cells, or (3) a small initial increase or slight decrease in expression that is sustained over time resulting in less expression at 24 hpi than uninfected cells **(Fig. 5.2C)**.

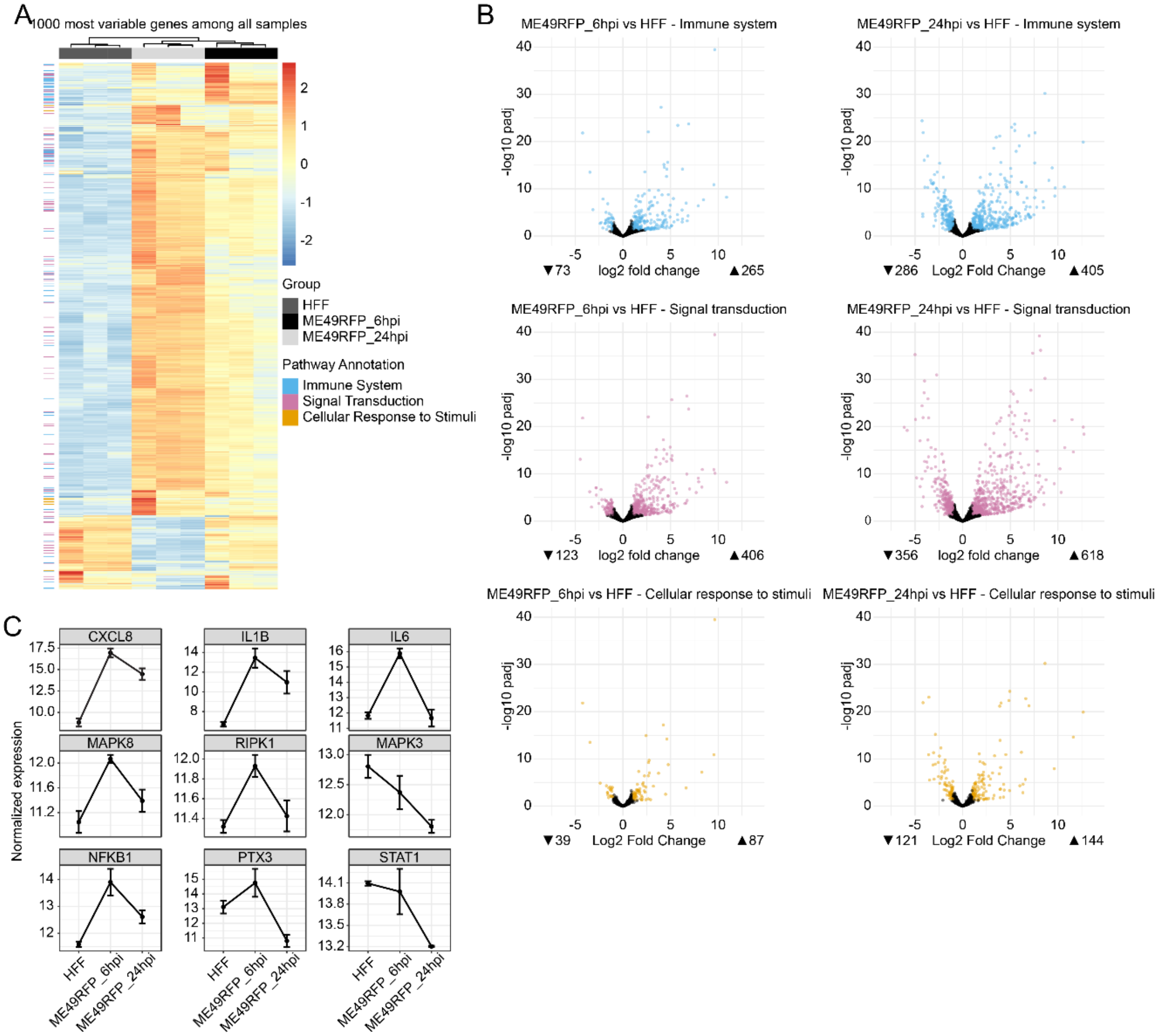

Figure 5.2. *T. gondii* alters the host stress response and disrupts cell signaling. (A) Gene set enrichment analysis of rlog transformed expression values indicates high variability among key pathways necessary for the functional immune response, including genes directly involved in the immune system. Of the 1000 genes with the greatest transcriptional variability, a small subset of 25 immune system and signal transduction genes show suppressed expression following 24 hours of infection. **(B)** Volcano plots emphasize the disruption in tightly regulated pathways as many immune regulators and signaling genes are both up and downregulated over the course of the infection. **(C)** Time course expression plots show varying patterns of transcription in important cell signaling genes and transcriptional activators.

T. gondii alters host chromatin compartmentalization

To better understand the molecular mechanisms controlling changes in gene expression during infection, we investigated global change in chromatin organization using Hi-C, a high-throughput genomic and epigenomic technique to capture chromatin conformation[43]. We collected ~500-630,000 infected HFFs per biological replicate using FACS and also collected ~890-910,000 uninfected cells per biological replicate via an automatic cell sorter. Because high resolution data is required to inspect small scale changes (<25 kb) within the chromatin, such as intradomain modifications, and due to the high Pearson correlation ($\rho > 0.8$) between biological replicates, we elected to merge read counts from replicates (~50-210 million reads) into a single representative sample containing ~351-462 million reads per condition (uninfected HFF, 6 hpi, and 24 hpi).

Following parasite invasion, we were able to detect a significant global decrease of long-range interactions and loss of chromatin compartmentalization as the defined borders of chromatin compartments are disrupted. Distance decay analysis was performed on ICED normalized matrices binned at 25 kb resolution and used to generate a log-scale contact probability curve of the observed number of interactions between loci less than 100 Mb apart **(Fig. 5.3A)**[44]. Contact frequency of chromatin interactions at 6 hpi does not significantly deviate from uninfected HFFs, unlike the 24 hpi cells which show a considerable decrease in long-range interactions. This is also reflected by the ratio of short versus long-range interactions per chromosome, indicated by the increasing number of short-range contacts over time **(Fig. 5.3B)**. Data was then converted to observed/expected (O/E) matrices before calculating the Pearson correlation between genomic loci. A PCA analysis of the correlation values was performed, and the first eigenvector polarity was corrected using GC-content. At 6 hpi there is an overall shift toward B compartmentalization as A compartments have weakened, B compartments have strengthened, and there are some A to B compartment transition **(Fig. 5.3C)**[45]. Normalized intrachromosomal interaction matrices binned at 100 kb were plotted to identify broad differences between condition and to map the interaction pattern within these differential compartments **(Fig. 5.3D)**. The largest changes in compartmentalization at each time point are found near interacting loci ~2-30 Mb apart (**Fig. 5.3D**, black box).

Pearson correlation matrices were generated to emphasize the shift in interactions identified by compartment analysis **(Fig. 5.3E)**. At 6 hpi, A compartments are notably weaker while B compartments strengthen as indicated by the E1 principal component values (PC1), which is likely a stress response to the parasite invasion leading to changes in transcriptional regulation. While the strength of A compartments at 24 hpi returns to levels closer to that of uninfected cells, B compartment strength does not significantly decrease from the changes that occurred prior to 6 hpi. A gradual decrease in Pearson correlation within established compartments over time (red arrow) indicates higher order chromatin restructuring and loss of mid- to long-range interactions **(Fig. 5.3E)**.

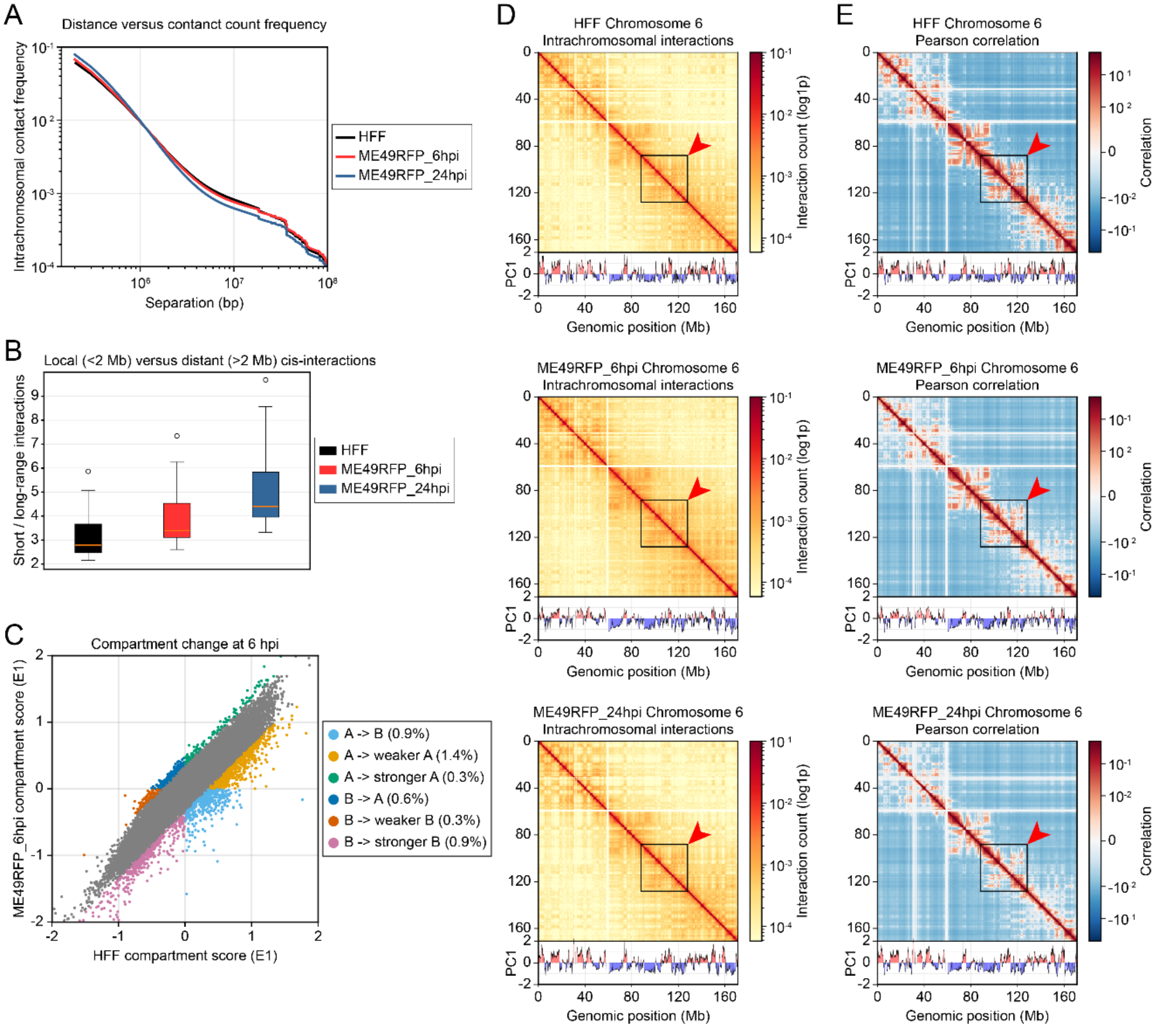

Figure 5.3. Infected cells exhibit decreased large-scale compartmentalization. (A) Intrachromosomal interactions are plotted against genomic distance and infected HFFs at 24hpi exhibit decreased mid and long-range interactions (>1 Mb) and slightly increased short-range interactions (<1 Mb). **(B)** At 24 hpi, the ratio of short versus long-range interactions is elevated. **(C)** Comparison of E1 score for genome-wide 50kb bins in uninfected HFF and 6hpi showing transition to more prominent B compartmentalization. **(D)** Intrachromosomal interactions are not distinctly different among the three conditions; however, E1 values of the compartment analysis (PC1) indicate an A to B compartment transition at 6 hpi. By 24 hpi the A compartments have recovered but B compartments remain stable. **(E)** Pearson correlation analysis of chromosome 6 shows the loss of mid to long-range interactions at 24 hpi as highlighted by the region within the box and red arrow.

Local changes associated with loss of intra-TAD interactions

In addition to the loss of long-range interactions within the chromatin of infected cells, there is an increase in short-range interactions (<1 Mb) between regions that exhibit TAD-like structures (black arrow) with sharp defined borders implying an increase in inter-domain contacts **(Fig. 5.4A)**. These same regions also show decreased mid-range interactions (dark blue arrow), with some demonstrating an inversion of identified compartments at 24 hpi. Some inter-domain alterations also show an asymmetric stripe pattern (light blue arrow) within a narrow region signifying an attempt to bring cis-regulatory enhancer elements to the promoter regions of target genes to initiate or upregulate transcription[46,47].

Enhancer-promoter interactions between distal regulatory elements are sometimes favored over proximal interactions, possibly to prevent the expression of nearby genes as these can remain sequestered within tightly regulated domains[47]. While there is tighter regulation of genes within TADs via enhancer-promoter interactions, evidence suggests that perturbations of TAD boundaries result in enhancer mediated transcriptional upregulation at a much higher rate[48]. Consequently, the expanded inter-TAD enhancer-promoter interactions likely leads to the rapid immune response within infected cells. However, one would expect that the changes in chromatin structure that lead to immune upregulation would occur prior to 6 hours after an invasive pathogen has entered the cell, whereas in our infected HFFs it is clear that local chromatin restructuring results in a more disorganized chromatin architecture as many TAD boundaries are either weakened or fully eliminated **(Fig. 5.4B-C)**. The TAD boundary downstream of NLRP1 (chr17:5.50Mb-5.58Mb)—a well characterized inflammasome protein and pyroptosis mediator that is constitutively expressed in the intestinal epithelium[49, 50]—is weakened at 6 hpi and merges with a neighboring TAD **(Fig. 5.4B, middle)**. The disrupted TAD boundary is accompanied by a slight downregulation of NLRP1 at 6 hpi ($Log_2FC = -1.02$, **Supp. Table 5.2**). This boundary is then reinstated at 24 hpi, however it has shifted toward the NLRP1 locus **(Fig. 5.4B, bottom)**, correlating with further downregulating NLRP1 expression ($Log_2FC = -2.37$, **Supp. Table 5.3**).

Not all differential expression correlates with a loss or shift in major TAD boundaries; some genes are upregulated due to weakening TAD boundaries and loss of sub-TAD boundaries, resulting in inter-TAD regulatory interactions and up- or downregulation of transcription. A 2 Mb region on chromosome 10 **(Fig. 5.3C)** shows weakening of existing major TAD boundaries, while some weaker sub-TAD boundaries are entirely lost. However, to fully elucidate the regulatory effects of shifting TAD and sub-TAD boundaries, higher resolution data via deeper sequencing will be required. Nevertheless, insulation scores measured at genome-wide TAD boundaries show that the mean distance between TAD boundaries increases by 4-5% and median boundary strength decreases by 8% and 14% at 6 hpi and 24 hpi, respectively. Together these results imply diminished chromatin organization due to shifting and weakened TAD boundaries are likely one determinant for the transcriptional dysregulation exhibited by the infected fibroblasts.

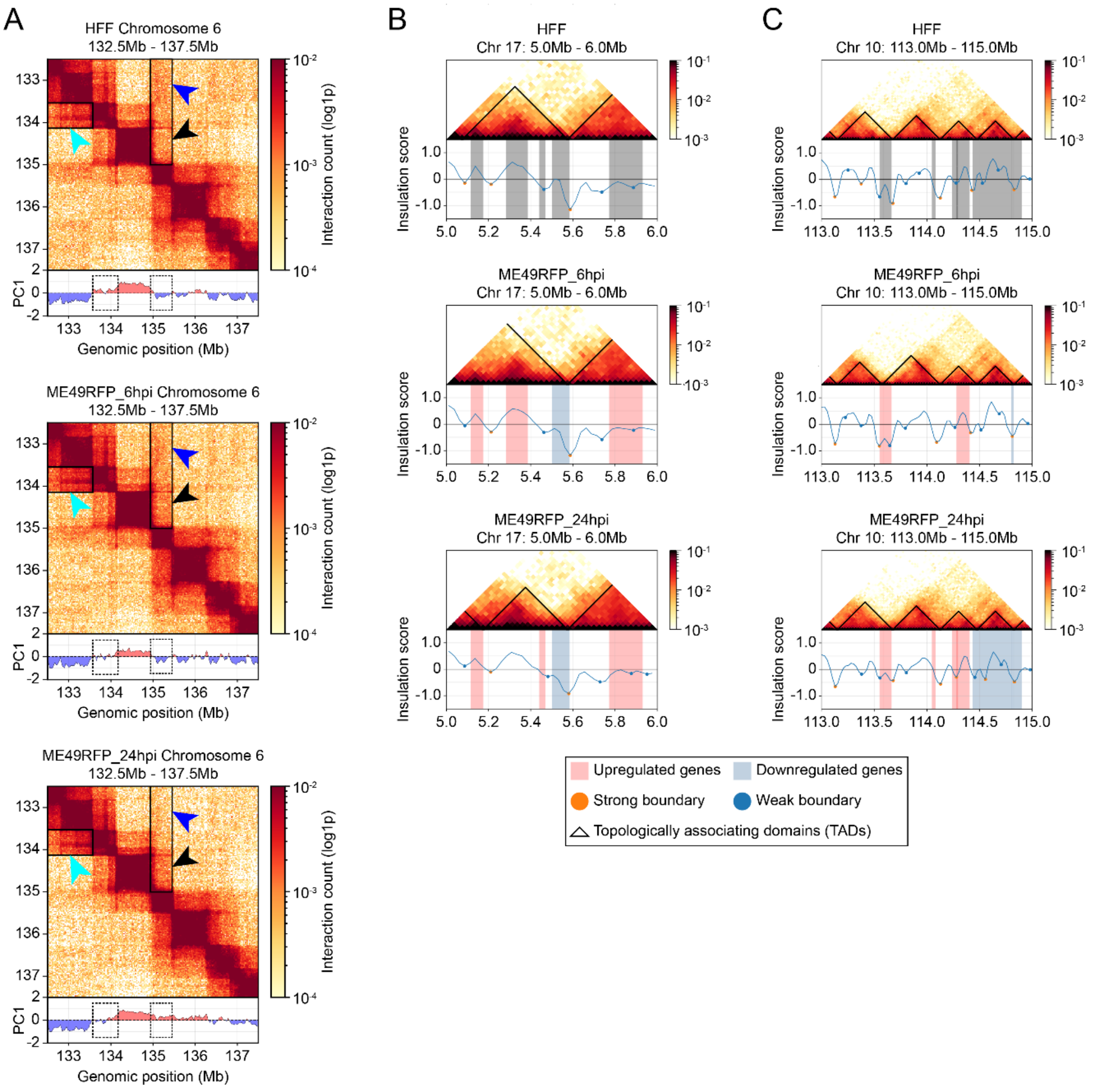

Figure 5.4. Weak TAD boundaries results in elevated inter-TAD interactions. (A) 2 Mb window of chromosome 6 shows compartment switching and enhancer-promoter interactions. The decrease of mid-range interactions (blue arrow) and increase of short-range interactions (black arrow) for loci within the window near 135 Mb. This change is emphasized by the compartment inversion (B to A) seen in the compartment tracks below each heatmap. The other highlighted region (blue arrow) focuses on the narrow strip of increased interactions that indicate enhancer-promoter contact, thus modifying local chromatin interactions resulting in compartmental changes. **(B)** 1 Mb window (5 Mb-6 Mb) of chromosome 17 provides a more detailed analysis of intra and inter-TAD contacts. There is a merger and subsequent reformation of two TADs to the left of NLRP1 (highlighted in blue) at 6 hpi. Domain boundary insulation trends toward 0 indicating a weakening of TAD boundaries and increased likelihood of inter-TAD interactions. **(C)** 2 Mb window (113 Mb-115 Mb) on chromosome 10 displays a similar loss of defined TAD boundaries as the number and strength of the boundaries decreases at each time point.

DISCUSSION

Toxoplasma gondii must successfully invade a host cell and establish an environment conducive to proliferation, while simultaneously suppressing the host immune response and programmed cell death pathways to ensure survival. Secreted effector proteins exported by the parasite into the host cell nucleus provide the means for survival by either up- or downregulating key genes of immune and cell signaling pathways. However, direct evidence of epigenetic manipulation of host cell gene expression is limited. Furthermore, the evidence we have is narrow in scope and focuses on established effector proteins and known targets. Our goal was to identify global molecular mechanisms regulating changes in gene expression. We therefore investigated global epigenetic changes in infected versus non-infected host cells across time provide evidence worthy of investigation in further studies. This was achieved using a mutli-omics approach to analyze chromatin structure and epigenetic modifications over the course of an infection.

Our RNA-seq results revealed a robust innate immune response through the upregulation of several pro-inflammatory cytokines and chemokines at 6 hours post-infection. By 24 hours however, we saw a decrease in expression of many of these pro-inflammatory genes and sustained upregulation of anti-inflammatory cytokines. Over the course of the infection there is a significant amount of DNA damage and dysregulation of chromatin structure because expression of DNA maintenance genes is noticeably upregulated at 24 hpi.

We investigated changes in the global chromatin structure using Hi-C to ascertain whether large-scale chromatin modifications are responsible for the changes in immune gene transcription. The dysregulation of DNA repair mechanisms is likely an outcome of rapid large-scale changes to the chromatin architecture as chromatin structure is typically static and maintained in tightly regulated compartments and domains. Disruption of TAD boundaries and chromatin rearrangements affected fibroblast cellular homeostasis prompting the recruitment of DNA repair mechanisms by 24 hpi to aid survival. It is likely that these large-scale changes are a combination of direct and indirect effects of parasite manipulating host cell chromatin. Additional analyses of the epigenome of infected cells using various molecular biology techniques have already been performed, however the data has not been fully investigated and therefore have not been included. Further investigation using *Toxoplasma* mutant strains will be required to fully explore the breadth of epigenetic changes that occur during *T. gondii* infection and provide a more comprehensive picture of the manipulations that ensure parasite survival.

MATERIALS AND METHODS

Parasite culturing

Human foreskin fibroblast (HFF) were maintained in Dulbecco's modified Eagle medium supplemented with 10% heat-inactivated fetal bovine serum (Corning, 35-016-CV), a combination of 100 µg/ml streptomycin and 100 units/ml penicillin (GeneClone Penicillin-Streptomycin Solution 100X, Cat no. 25-512), and 10 µg/ml gentamycin (Quality Biological, 120-098-661). Parasite strain ME49-RFP[51] was cultured in DMEM supplemented with 5% heat-inactivated fetal bovine serum (Corning, 35-016-CV), 100 µg/ml streptomycin and 100 units/ml penicillin, and 10 µg/ml gentamicin at 37°C in a humidified incubator with 21% O2 and 5% CO2 as described elsewhere[52].

Cell sorting infected HFFs

At 6 and 24 hours post inoculation, cells were lifted from the culture flask using pre-warmed trypsin-EDTA then washed and resuspended in ice-cold FACS buffer (50 mg EDTA and 4 g BSA (bovine serum albumin) dissolved in 1 L 1x PBS). Infected cells were FACS sorted using a MoFlo Astrios EQ Cell Sorter with the following gating strategy: gate 1 forward scatter (FSC) and side scatter (SSC) to exclude cellular debris, gate 2 FSC-H (height) and FSC-A (area) to remove doublets, and gate 3 filters by the presence of RFP.

Counting uninfected HFFs

Uninfected HFFs were lifted from the culture flask using pre-warmed trypsin-EDTA then washed and resuspended in 1X PBS. The number of cells was determined using a Bio-Rad TC10™ automated cell counter.

Immunofluorescence assay

Parasites grown in HFF monolayer were fixed in 4% paraformaldehyde for 15 min at room temperature followed by three washes in 1X PBS. Fixed cells were permeabilized with 100% acetone for 10 min at room temperature and washed three times with 1X PBS then blocked with blocking buffer (5% Donkey serum) for 30 min, then incubated with primary antibodies in blocking buffer at 4°C overnight. Cells were washed with 1X PBS and incubated with DAPI (Invitrogen, 1:1000) and Alexa-fluor secondary antibodies coupled with desired fluorophore for 1 hr at room temperature. Cells were washed and mounted with Prolong Gold antifade reagent with DAPI before analysis. Dilutions for primary antibodies: *T. gondii* mice α-SAG1 (Invitrogen, MA1-83499) 1:1000, rabbit α-RFP (Rockland immunochemicals, 600-401-379) 1:1000.

RNA-seq library preparation

Total RNA was extracted from sorted cells by incubating briefly at room temperature in 1 ml TRIzol LS Reagent (Invitrogen, 10296010) then adding 0.2 ml chloroform. RNA was precipitated at 4°C for 10 min in 0.5 ml isopropanol followed by 1 ml 75% ethanol wash. Bulk RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit (NEB, E7760) and PCR enriched with NEBNext Multiplex Oligos (NEB, E6440). Libraries were then quantified by Agilent 4200 TapeStation System and sequenced using the NovaSeq X Plus (Illumina) with 1.5B flow cell for 150 bp paired-end reads. In total, three replicates for the uninfected HFF control, 6 hpi ME49RFP infected line, and 24 hpi ME49RFP infected line were collected.

RNA-seq data processing and differential expression analysis

Raw reads from the RNA-seq libraries were analyzed with FastQC (v0.12.1) to assess sequence quality and content for determining overall library quality and trimming length. Adapters and low-quality bases (PHRED < 30) were trimmed and reads paired using Cutadapt (v4.6)[53] before aligning to both the *H. sapien* (GRCh38.p14) genome with HISAT2 (v2.2.1)[54] . Samtools (v1.19.2) was used to filter low-quality, unpaired, and unmapped reads[55]. Finally, reads that were high quality, properly paired, and aligned were mapped to all protein coding genes in the target genomes using HTseq (v.2.0.5)[56]. Differentially expressed genes (padj < 0.05) within the infected HFFs were identified between the uninfected control and 6 hpi or 24 hpi time points using DESeq2 (v.1.42.0)[57].

Hi-C library preparation

The Arima-HiC+ Kit (Arima Genomics, A510008) was used to collect proximity-ligated genome-wide chromatin interactions as described in the user guide for low cell input. DNA was then sonicated with the Covaris S220 Ultrasonicator (Duty factor: 10%, PIP: 140 W, cycles per burst: 200, Temperature: 4°C, Time: 55 sec) and libraries prepared using the Accel-NGS 2S Plus DNA Library Kit (Swift Biosciences, 21024) while following the Arima Genomics user guide for library preparation of bead-bound DNA. DNA libraries were amplified using KAPA HiFi HotStart Ready Mix (2X) (KAPA, KK2620) and NEBNext Multiplex Oligos (NEB, E6440). DNA concentrations were measured using an Agilent 2100 Bioanalyzer and libraries sequenced using the NovaSeq X Plus with 10B flow cell for 150 bp paired-end reads

Hi-C data processing

Paired-end HiC libraries were processed using the HiC-Pro suite while aligning to the *H. sapien* (GRCh38.p14) genome with mapping quality cutoff set to 30 and removing multimapped and duplicated read pairs using default parameters set by HiC-Pro[44]. Interaction matrices were generated, and replicates merged to improve the resolution of downstream visualizations. HiC-Pro matrices were converted to cool format using hicConvertFormat from the HiCExplorer package and sex chromosomes removed[58]. The Cooltools API was used to perform data normalization using on cis contacts as well as the downstream statistical and feature analysis where specified[59]. Chromatin compartmentalization was investigated by computing compartment eigenvectors (cis_eigs) ignoring the first two diagonals and using GC density to orient the eigenvectors to reflect expected compartments. Intrachromosomal/Pearson correlation heatmaps were generated with max values set to the 95th percentile to improve visualization of more sparsely interacting regions (Fig. 5.3). Expected contacts per genomic distance were calculated (expected_cis) by narrowing the intervals of interest to cis-contacts within chromosome arms due to trans-centromeric contacts having less interactions per genomic distance. Data smoothing was applied prior to plotting aggregated expected contacts over logarithmic distance (Fig. 5.4A). Short versus long-range interactions per chromosome were calculated by dividing the total number of interactions between bins less than 2 Mb apart by those greater than 2 Mb (Fig. 5.4B). TADs were called by hicFindTADs with FDR correction ($q < 0.05$) (Fig. 5.4C-D)[58]. To quantitatively measure the strength of TAD boundaries, insulation scores were calculated (cooltools.insulation) within a three-bin sliding window, differentiating weak from strong boundaries between TADs. Chromatin compartments were differentiated between infected and uninfected samples by plotting the z-score of the E1 values for bins found in both datasets (Fig. 5.4E).

ACKNOWLEDGEMENTS

We thank Michael Grigg for the RFP-tagged *T. gondii* ME49 parasites[46]. We also want to thank Mary Hamer at the UCR School of Medicine flow cytometry core facility for providing guidance and assistance with FACS experiments.

FUNDING

This work was funded by the National Institutes of Health and the National Institutes of Allergy and Infectious Diseases (NIH/NIAID R01 AI158417 to E.W. and NIH/NIAID R21 AI142506 to K.L.R.).

AUTHOR CONTRIBUTIONS

Todd Lenz: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – Original draft, Visualization, Project administration. **Sandeep Srivastava**: Methodology, Investigation, Resources, Writing – Original draft. **Edward Vizcarra**: Methodology, Investigation, Resources. **Emma Wilson**: Supervision, Funding acquisition. **Karine G Le Roch**: Conceptualization, Validation, Writing – Original draft, Supervision, Project administration, Funding acquisition

COMPETING INTERESTS

The authors declare no competing interests.

CODE AVAILABILITY

Packages or tools used for bioinformatic analysis and plotting of our data are provided in the Materials and Methods. All custom pipelines and scripts are available within repositories at https://github.com/tlenz88.

REFERENCES

1. Sibley, L. D. & Boothroyd, J. C. Virulent strains of Toxoplasma gondii comprise a single clonal lineage. *Nature* **359**, 82-85 (1992).

2. Howe, D. K. & Sibley, L. D. Toxoplasma gondii Comprises Three Clonal Lineages: Correlation of Parasite Genotype with Human Disease. *J. Infect. Dis.* **172**, 1561-1566 (1995).

3. Minot, S. et al. Admixture and recombination among Toxoplasma gondii lineages explain global genome diversity. *Proc. Natl Acad. Sci. USA* **109**, 13458-13463 (2012).

4. Su, C. et al. Globally diverse Toxoplasma gondii isolates comprise six major clades originating from a small number of distinct ancestral lineages. *Proc. Natl Acad. Sci. USA* **109**, 5844-5849 (2012).

5. Galal, L., Hamidović, A., Dardé, M. L. & Mercier, M. Diversity of Toxoplasma gondii strains at the global level and its determinants. *Food Waterborne Parasitol.* **15**, e00052 (2019).

6. Fentress, S. J. et al. Phosphorylation of Immunity-Related GTPases by a Toxoplasma gondii-Secreted Kinase Promotes Macrophage Survival and Virulence. *Cell Host Microbe* **8**, 484-495 (2010).

7. Etheridge, R. D. et al. The Toxoplasma Pseudokinase ROP5 Forms Complexes with ROP18 and ROP17 Kinases that Synergize to Control Acute Virulence in Mice. *Cell Host Microbe* **15**, 537-550 (2014).

8. Yamamoto, M. et al. ATF6β is a host cellular target of the Toxoplasma gondii virulence factor ROP18. *J. Exp. Med.* **208**, 1533-1546 (2011).

9. Torres-Morales, E. et al. Th1 and Th2 immune response to P30 and ROP18 peptides in human toxoplasmosis. *Med. Microbiol. Immunol.* **203**, 315-322 (2014).

10. Ong, Y.-C., Reese, M. L. & Boothroyd, J. C. Toxoplasma Rhoptry Protein 16 (ROP16) Subverts Host Function by Direct Tyrosine Phosphorylation of STAT6. *J. Biol. Chem.* **285**, 28731-28740 (2010).

11. Butcher, B. A. et al. Toxoplasma gondii Rhoptry Kinase ROP16 Activates STAT3 and STAT6 Resulting in Cytokine Inhibition and Arginase-1-Dependent Growth Control. *PLoS Pathog.* **7**, e1002236 (2011).

12. Zhu, W., Li, J., Pappoe, F., Shen, J. & Yu, L. Strategies Developed by Toxoplasma gondii to Survive in the Host. *Front. Microbiol.* **10**, 899 (2019).

13. Rosenberg, A. & Sibley, L. D. Toxoplasma gondii secreted effectors co-opt host repressor complexes to inhibit necroptosis. *Cell Host Microbe* **29**, 1186-1198 (2021).

14. Huang, Z. et al. The intrinsically disordered protein TgIST from Toxoplasma gondii inhibits STAT1 signaling by blocking cofactor recruitment. *Nat. Commun.* **13**, 4047 (2022).

15. Braun, L. et al. The Toxoplasma effector TEEGR promotes parasite persistence by modulating NF-κB signalling via EZH2. *Nat. Microbiol.* **4**, 1208-1220 (2019).

16. Murray-Nerger, L. A. et al. The DNA loop release factor WAPL suppresses Epstein-Barr virus latent membrane protein expression to maintain the highly restricted latency I program. Preprint at bioRxiv https://doi.org/10.1101/2024.05.09.593401 (2024).

17. Venu, V. et al. Multi-omics analysis reveals the dynamic interplay between Vero host chromatin structure and function during vaccinia virus infection. *Commun. Biol.* **7**, 721 (2024).

18. Satou, Y. et al. The retrovirus HTLV-1 inserts an ectopic CTCF-binding site into the human genome. *Proc. Natl Acad. Sci. USA* **113**, 3054-3059 (2016).

19. Wang, R. et al. SARS-CoV-2 restructures host chromatin architecture. *Nat. Microbiol.* **8**, 679-694 (2023).

20. Zheng, D.-L. et al. Epigenetic modification induced by hepatitis B virus X protein via interaction with de novo DNA methyltransferase DNMT3A. *J. Hepatol.* **50**, 377-387 (2009).

21. Arora, P., Kim, E.-O., Jung, J. K. & Jang, K. L. Hepatitis C virus core protein downregulates E-cadherin expression via activation of DNA methyltransferase 1 and 3b. *Cancer Lett.* **261**, 244-252 (2008).

22. Luo, X. et al. DNMT1 mediates metabolic reprogramming induced by Epstein-Barr virus latent membrane protein 1 and reversed by grifolin in nasopharyngeal carcinoma. *Cell Death Dis.* **9**, 619 (2018).

23. Heinrich, P. C. et al. Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem. J.* **374**, 1-20 (2003).

24. Hirano, T. IL-6 in inflammation, autoimmunity and cancer. *Int. Immunol.* **33**, 127-148 (2021).

25. Lian, B. S. X. et al. Regulation of Il6 expression by single CpG methylation in downstream of Il6 transcription initiation site. *iScience* **25**, 104118 (2022).

26. Zhou, C., Gao, Y., Ding, P., Wu, T. & Ji, G. The role of CXCL family members in different diseases. *Cell Death Discov.* **9**, 212 (2023).

27. Burke, S. J. et al. NF-κB and STAT1 control CXCL1 and CXCL2 gene transcription. *Am. J. Physiol. Endocrinol. Metab.* **306**, E131-E149 (2014).

28. Cambier, S., Gouwy, M. & Proost, P. The chemokines CXCL8 and CXCL12: molecular and functional properties, role in disease and efforts towards pharmacological intervention. *Cell. Mol. Immunol.* **20**, 217-251 (2023).

29. Metzemaekers, M. et al. Truncation of CXCL8 to CXCL8(9-77) enhances actin polymerization and in vivo migration of neutrophils. *J. Leukoc. Biol.* **107**, 1167-1173 (2020).

30. Iyer, S. S. & Cheng, G. Role of Interleukin 10 Transcriptional Regulation in Inflammation and Autoimmune Disease. *Crit. Rev. Immunol.* **32**, 23-63 (2012).

31. Butcher, B. A. et al. Cutting Edge: IL-10-Independent STAT3 Activation by Toxoplasma gondii Mediates Suppression of IL-12 and TNF-α in Host Macrophages. *J. Immunol.* **174**, 3148-3152 (2005).

32. Martin, J. C. et al. Single-Cell Analysis of Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell* **178**, 1493-1508 (2019).

33. Zhou, Y. et al. Fibroblasts in immune-mediated inflammatory diseases: The soil of inflammation. *Clin. Immunol.* **258**, 109849 (2024).

34. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14**, 36-49 (2014).

35. Cheon, H. & Stark, G. R. Unphosphorylated STAT1 prolongs the expression of interferon-induced immune regulatory genes. *Proc. Natl Acad. Sci. USA* **106**, 9373-9378 (2009).

36. Kravets, E. et al. Guanylate binding proteins directly attack Toxoplasma gondii via supramolecular complexes. *eLife* **5**, e11479 (2016).

37. Fisch, D. et al. Human GBP1 is a microbe-specific gatekeeper of macrophage apoptosis and pyroptosis. *EMBO J.* **38**, e100926 (2019).

38. Yamamoto, M. et al. A Cluster of Interferon-γ-Inducible p65 GTPases Plays a Critical Role in Host Defense against Toxoplasma gondii. *Immunity* **37**, 302-313 (2012).

39. Lüder, C. G. K. & Seeber, F. Toxoplasma gondii and MHC-restricted antigen presentation: on degradation, transport and modulation. *Int. J. Parasitol.* **31**, 1355-1369 (2001).

40. Gaggar, A. et al. A Novel Proteolytic Cascade Generates an Extracellular Matrix-Derived Chemoattractant in Chronic Neutrophilic Inflammation. *J. Immunol.* **180**, 5662-5669 (2008).

41. Sorokin, L. The impact of the extracellular matrix on inflammation. *Nat. Rev. Immunol.* **10**, 712-723 (2010).

42. Sutherland, T. E., Dyer, D. P. & Allen, J. E. The extracellular matrix and the immune system: A mutually dependent relationship. *Science* **379**, eabp8964 (2023).

43. Lieberman-Aiden, E. et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**, 289-293 (2009).

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

45. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl Acad. Sci. USA* **115**, E6697-E6706 (2018).

46. Kraft, K. et al. Serial genomic inversions induce tissue-specific architectural stripes, gene misexpression and congenital malformations. *Nat. Cell Biol.* **21**, 305-310 (2019).

47. Balasubramanian, D. et al. Enhancer-promoter interactions can form independently of genomic distance and be functional across TAD boundaries. *Nucleic Acids Res.* **52**, 1702-1719 (2024).

48. Cavalheiro, G. R., Pollex, T. & Furlong, E. E. To loop or not to loop: what is the role of TADs in enhancer function and gene regulation? *Curr. Opin. Genet. Dev.* **67**, 119-129 (2021).

49. Martinon, F., Burns, K. & Tschopp, J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Mol. Cell* **10**, 417-426 (2002).

50. Martinon, F. & Tschopp, J. NLRs join TLRs as innate sensors of pathogens. *Trends Immunol.* **26**, 447-454 (2005).

51. Sangaré, L. O. et al. In Vivo CRISPR Screen Identifies TgWIP as a Toxoplasma Modulator of Dendritic Cell Migration. *Cell Host Microbe* **26**, 478-492 (2019).

52. Vizcarra, E. A. et al. An ex vivo model of Toxoplasma recrudescence reveals developmental plasticity of the bradyzoite stage. *mBio* **14**, e01836-23 (2023).

53. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10-12 (2011).

54. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357-360 (2015).

55. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

56. Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E. & Zanini, F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* **38**, 2943-2945 (2022).

57. Love, M., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

58. Ramírez, F. et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).

59. Fudenberg, G. Cooltools: enabling high-resolution Hi-C analysis in Python. (2019); https://github.com/open2c/cooltools

CONCLUDING STATEMENTS

The global impact of malaria remains one of humanity's greatest public health challenges, affecting millions of people annually and causing significant morbidity and mortality, particularly in developing regions. My research focused on understanding the fundamental molecular mechanisms that enable *Plasmodium* and *Toxoplasma* parasites to successfully invade and proliferate within host cells while evading immune responses. Specifically, I investigated the complex interplay between nuclear organization, chromatin structure, and gene regulation that allows these parasites to coordinate their developmental programs and maintain virulence.

Through the application of genomic and molecular approaches, I made several key discoveries that advance our understanding of parasite biology. I revealed how the spatial organization of chromosomes and epigenetic modifications work together to enable the placental binding phenotype in *P. falciparum*. I demonstrated that *var2csa* plays a crucial role in coordinated *var* gene switching, as its deletion leads to increased heterochromatin compaction and disrupts the parasite's ability to undergo antigenic variation. I identified novel protein factors like *Pf*MORC that maintain chromatin structure and transcriptional control. I also aided in investigating ROP55, a major *T. gondii* virulence factor that prevents host cell death by suppressing pyroptosis and necroptosis pathways, thereby enabling parasite survival. I uncovered previously unknown mechanisms by which the parasite manipulates host cell chromatin architecture to facilitate its survival and proliferation. These findings provide valuable insights into host-pathogen interactions and identify potential therapeutic targets.

My work establishes experimental frameworks and analytical approaches that will enable future studies investigating chromatin biology in these parasitic organisms. As we continue to face evolving challenges in controlling malaria and toxoplasmosis globally, understanding how these parasites regulate virulence gene transcription and interact with host cells will be essential for developing more effective interventions.

This research highlights the broader importance of studying basic biological processes in parasitic organisms, as it reveals both conserved and unique aspects of chromatin regulation that have emerged through host-pathogen co-evolution. The molecular mechanisms we uncovered may have parallels in other pathogens, potentially informing therapeutic approaches beyond apicomplexan diseases. This work provides a foundation for further investigation into the complex relationship between nuclear organization, gene regulation, and parasite virulence—knowledge that will be crucial for combating these persistent global health threats.

Throughout my journey, I have developed extensive expertise in genomics, bioinformatics, and molecular biology while gaining a deep appreciation for the complexity of host-pathogen interactions. Working on these challenging projects has strengthened my analytical and problem-solving abilities, particularly in processing and interpreting large-scale genomic datasets. The collaborative nature of this work has taught me the value of scientific teamwork and effective communication across disciplines. Most importantly, this work has reinforced my commitment to conducting rigorous, impactful research. The skills and knowledge I have gained throughout these last six years have prepared me well for future scientific endeavors.