The Pennsylvania State University

The Graduate School

Interdisciplinary Program in Immunology and Infectious Disease

MEMBERS OF THE *PLASMODIUM YOELII* CAF1/CCR4/NOT COMPLEX ARE CRUCIAL OR ESSENTIAL FOR PROPER GAMETOCYTE DEVELOPMENT

A Dissertation in

Immunology and Infectious Disease

by

Kevin J. Hart

Submitted in Partial Fulfillment

of the Requirements

For the Degree of

Doctor of Philosophy

December 2018

The dissertation of Kevin J. Hart was reviewed and approved* by the following:

Scott E. Lindner

Assistant Professor of Biochemistry

Dissertation Advisor

Chair of Committee

Paul Babitzke

Professor of Biochemistry

Liwang Cui

Professor of Entomology

Manuel Llinás

Professor of Biochemistry

Professor of Chemistry

Joseph C. Reese

Professor of Biochemistry

Margherita T. Cantorna

Distinguished Professor of Molecular Immunology

*Signatures are on file in the Graduate School.

ABSTRACT

The transmission of the malaria parasite between mosquitoes and mammals requires translational repression to ensure that only the proper proteins are expressed at the right time, while still allowing the parasite to generate the mRNAs it will need for the next developmental stage. With relatively few known sequence specific transcription factors that may regulate initiation of transcription, *Plasmodium* also regulates the stability and turnover of transcripts to provide more comprehensive gene regulation. In model eukaryotes, the CAF1/CCR4/NOT complex has been shown to be crucial for not only mRNA degradation, but also translational control through its deadenylases CAF1 and CCR4. These deadenylases are thought to act primarily through association with NOT1 of the CAF1/CCR4/NOT complex. I have identified and characterized CCR4-1 in *Plasmodium yoelii*, and found that it plays a role in the activation and development of male gametocytes, and regulates host-to-vector transmission. With the deletion of *ccr4-1*, there is a loss in the coordination of male gametocyte activation, and a reduction in the ability of the parasite to productively infect the mosquito that is independent of its effect upon male activation. While gene deletion of the major deadenylase CAF1 is lethal, expression of the N-terminal CAF1 domain maintains survival, but prevents proper complex assembly and phenocopies the *ccr4-1* deletion. Additionally, in stark contrast to virtually all other eukaryotes, I have bioinformatically identified that the *Aconoidasida* class of Apicomplexans encode two putative NOT1 proteins. Generally, NOT1 acts as the scaffold for the CAF1/CCR4/NOT complex. I found that neither NOT1 nor its paralog NOT1-G are essential for *Plasmodium yoelii* asexual blood stage development, but that NOT1-G is essential for complete gametocyte development. Moreover, *pynot1-g*$^{-}$ parasites only produce female/immature gametocytes that, by genetic cross experiments, are sterile and cannot be transmitted to mosquitoes. Comparative transcriptomics of parasites with a *pynot1-g* deletion show massive dysregulation of the female gametocyte transcriptome and implicate NOT1-G as an early regulator of RNA metabolism for gametocytogenesis. Thus, I demonstrated that CCR4-1, CAF1, and NOT1-G play intertwined roles in, and are essential for, proper gametocyte development.

Table of Contents

List of Figures

Chapter 1:

Chapter 3:

Chapter 4:

Chapter 5:

List of Tables

List of Abbreviations

4E-T	EIF4E transport
ACP	Acyl-carrier protein
ALS	Amyotrophic lateral sclerosis
ApiAP2	Apicomplexan Apetala 2 family
BIP	Binding immunoglobulin protein
BLASTp	Basic local alignment search tool program.
CAF	CCR4-associated factor
CCR4	Carbon catabolite repressor 4
CELF2	CUGBP Elav-like family member 2
CITH	Homolog of worm CAR-I and fly Trailer Hitch
CK1	Casein kinase 1
CRISPRi	Clustered regularly interspaced short palindromic repeat interference
cRT-PCR	Circular-reverse transcriptase polymerase chain reaction
cryo EM	Cryo-electron microscopy
CSP	Circumsporozoite protein
DAPI	4',6-diamidino-2-phenylindole
dCCR4-1	Catalytically dead carbon catabolite repressor 4
DCP2	mRNA Decapping 2
DEDD	Aspartic acid-glutamic acid- aspartic acid- aspartic acid
DEseq2	A differential expression software package
DOZI	Development of zygote inhibited
EBF1	Early B-cell factor 1
EEP	Exonuclease-endonuclease-phosphatase
EF-1a	Elongation Factor 1 alpha
eIF2a	Eukaryotic translation initiation factor 2a
EMBL	European molecular biology laboratory
EMSA	Electromobility shift assay
FDR	False discovery rate
FRAP	Fluorescence recovery after photobleaching
gDNA	Genomic DNA
GDV1	Gametocyte development gene 1
GEST	Gametocyte egress and sporozoite traversal
GFP	Green fluorescent protein
glmS	Glucosamine-6-phosphate riboswitch ribozyme
GO term	Gene ontology term
H3K4me3	Histone three lysine four tri-methylation
H3K9me3	Histone three lysine nine tri-methylation
HeLa	A adenocarcinoma line derived from the cervix of Henrietta Lacks
HMGB2	High mobility group box protein 2

HP1	Heterochromatin Protein 1
HsDHFR	*Homo sapiens* dihydrofolate reductase
IMC	Inner membrane complex
IP	Immunoprecipitation
LAP2	Lamin- and chromatin-binding nuclear protein
LRR	Leucine-Rich Repeat Region
MHC	Major histocompatibility complex
miRNA	MicroRNA
NOT	Negative on TATA-less
PABP	Poly(A)-binding protein
P-bodies	Processing body
PCR	Polymerase chain reaction
PlasmoDB	Plasmodium genome database
*Plasmo*GEM	Plasmodium genetic modification database
Poly (A)	Poly Adenosine
PTMs	Post-translational modifications
PUF	Pumilio and FBF
Py17XNL	*Plasmodium yoelii* strain 17x non-lethal
PyCAF1ΔC	*Plasmodium yoelii* CCR4 associated factor 1 c-terminal deletion
PyDD	*Plasmodium yoelii* dynein heavy chain delta
PyDDD	*Plasmodium yoelii* dynein heavy chain delta domain
RBP	RNA-binding protein
RCK/DHH1/DDX6	SF2 DEAD-box RNA helicase ortholog of DOZI
RNAPII	RNA Polymerase II
RNA-seq	RNA sequencing
RT-PCR	Reverse transcriptase polymerase chain reaction
SAINT	Significance analysis of interactome algorithm
SEM	Standard error of the mean
TDP-43	TAR DNA-binding protein of 43 kDa
TetR	Tet Repressor proteins
TRiC	TCP-1 Ring Complex
TTP	Tristetraprolin
TTP-BD	TTP-binding domain
XRN1	5’-3’ Exoribonuclease 1

Preface

While I am the first author of these works, others have contributed significantly to it, and I should mention them. I will list the individuals by their order of contribution to this work and what role they played. I performed nearly all of the experimental work except what is mentioned below. Scott Lindner provided the funding as well as performed the circular RT PCR experimentation for Chapter 3 and helped edit and submit the resulting pre-print and manuscript. Scott also assisted in the design of experimentation. Mike Walker did the processing of raw RNA-seq data for both Chapter 3 and Chapter 4; he also helped edit the manuscript resulting from the work described in Chapter 3. Allen Minns expressed protein to generate antibodies for the DDD-based flow cytometry sorting for both Chapter 3 and Chapter 4 and helped edit the manuscript resulting from the work described in Chapter 3. The lab of John Adams including Jenna Oberstaller, Ian Padykula, and John Adams contributed *P. falciparum* data towards the manuscript that resulted from Chapter 3. Much of that data is omitted here but can be found in the pre-print and manuscript.

Acknowledgments

In no way can I capture the love and support that I have gotten over this journey on this piece of paper. Here, I hope to convey special thanks to some of those who have helped me grow as a person and as a scientist.

First and foremost, to my wife Sara, and my son Carson. You have supported me from the beginning and let me take you away from Pittsburgh and bring you to State College. Here we had a fantastic son who has brought unrelenting joy to my life. Thank you both for supporting me throughout this adventure. To Carson, thank you for being the best little boy you could be and always being ready to make a bad day better with hugs.

To my parents and grandparents: Thank you for the support throughout my life allowing me to choose what I wanted to do, even if it meant that I moved across the state. You have taught me many life lessons that I continue to use and develop as I grow. I appreciate your support and willingness to support me as much as possible.

Scott, what can I say but that you are a fantastic mentor! I received overwhelming support from you from the beginning of this program and appreciate that you have let me explore some (crazy) ideas, experimentally. You have been available at every turn to answer my questions and help me grow not only as a scientist but as a person.

Allen! We have been together since the beginning of this adventure, and while so much has changed and evolved throughout this program, you have remained a constant. Always there to listen to my problems and provide support when I need it. Always there to help me think through a difficult concept or problem and always there to help with math. I cannot put into words how much you have helped both the lab and myself grow.

My fellow grad students Mike and Kelly, it has been quite a trip! Thanks for providing support and feedback throughout our time here together. From road trips to MPM to going out to lunch to procrastinate writing this thesis, we worked through this together! Thanks for your support!

Where would we be without our many technicians over the years! From Hannah to Brooke and from Mark to Dean to Tyler to Joann, you have helped me learn and grow whether I was learning from you (Mark) or teaching you, you have each helped me grow more as a scientist and a person.

My Undergrads, Laura Bowman, Steve Griffin, and Logan Finger. By allowing me to train you, you have helped me become a better mentor, a better scientist and a better person. Each of you has provided me with separate challenges and joys as you completed your undergraduate degrees. Each of you has a bright and promising future, and I cannot wait to hear where science takes you next!

To the many Llinas Lab members over the years. Thank you for helping to develop how I critically think about science. Your insight into various aspects about *Plasmodium* biology has helped me grow as a scientist.

I want to thank my committee for helping me grow. Each of you has provided a unique perspective on how I should learn and think and while tough at times; I am a better scientist because of it.

My lifelong friends Travis Strouse and Ryan Dietrich. You two have stuck with me for what feels like millennia! You have always supported me, and I appreciate all the work it takes keeping in contact with someone who lives so far away and who you never see.

Sarah Sumner and Dave Williamson. You two have been fantastic friends and fierce supporters of my work for as long as I can remember from being my peer mentor (Dave) and since we met such a short time ago (Sarah). You both have influenced me both scientifically and socially. You have supported my incessant need to play soccer on a weekly basis, and without your help and support, we may not have been able to grow this team into what it is now.

A special thanks to Doug Smith. While he did not directly impact this work, his training and belief in me allowed me to become a scientist. Without him, I likely would not have ended up at Clarion University where my scientific journey began, and without his support and education, I would not be where I am today. I hope one day to be loved by my students/employees even ½ as much as he was by everyone who worked under him.

Ben Allen has been a fantastic friend and supporter. He has helped me think critically about my science and supported me in all the times where I felt like it was not going anywhere. This friendship is one of my most treasured at Penn State.

Awtum, Jeong, Micha, and Beng, we have been through a lot! From meeting at our first Huck institutes retreat through recently, we have all been around to support and help each other grow. While we may leave at different times and go to different places, I will always remember you all as my original cohort.

Thanks to all those who have played soccer with me over time. We have developed something extraordinary that has blasted through the borders between scientific fields. Playing soccer has been a welcome distraction over the years that has helped us meet new people and expand our connections throughout the university.

I would like to thank the Huck Institutes generally. They gave me the opportunity to be a graduate student here at Penn State and funded a grant that allowed me to pursue some more risky ideas. I also appreciate all of the support that they gave to my peers especially those who were in environments that were not conducive to a good quality of life.

I would like to thank my collaborators John Adams, and Liwang Cui. You have provided me not only valuable insight on my primary project but also the opportunity to spread out and do more work that was more tangentially related.

A special thanks to all of the Animal Care workers but especially Eric, Bob, Dan, and Brad. By caring about our animals and knowing that what you do affects how the animals respond to my experiments, you have helped keep my work on track. I appreciate the hard work all of you do to keep the animals healthy and the facility in order.

Lastly, I would like to thank the granting institutions who have provided financial support to Scott so that I may continue in his lab. Especially the NIAID whom Scott received a K22 and R01 through and Penn State University for providing Start-Up funds to get the lab off the ground. “Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the NIAID.”

Chapter 1: INTRODUCTION

Malaria is Still a Major Global Health Problem

Malaria, a prehistoric disease, remains one of the significant global health problems today, with 216 million new infections and 445,000 deaths attributed to it annually (Figure 1.1[1]). The *Plasmodium* parasite causes malaria and is a eukaryotic Apicomplexan parasite in the class Aconoidasida. Five species of *Plasmodium* are known to infect humans, *P. falciparum, P. vivax, P. malariae*, *P. ovale* and recently a fifth species *P. knowlesi* [2]. *P. falciparum* is the most lethal *Plasmodium* parasite in humans and *P. falciparum* and *P. vivax* account for the majority of human infections. The majority of deaths occur in sub-Saharan Africa and are caused by *P. falciparum*. Symptoms of malaria typically include; fever, chills, and body aches. However, if the infection goes untreated, it can lead to severe anemia, cerebral malaria, kidney failure, and death. Understanding the development and transmission of the malaria parasite is vital to efforts to reduce or eliminate deaths due to this infection. Moreover, resistance to frontline anti-malarial drugs is becoming ever-more prevalent with the spread of resistance or decreased sensitivity to all known antimalarials. There is thus an urgent need to devise novel prevention and treatment strategies to combat malaria.

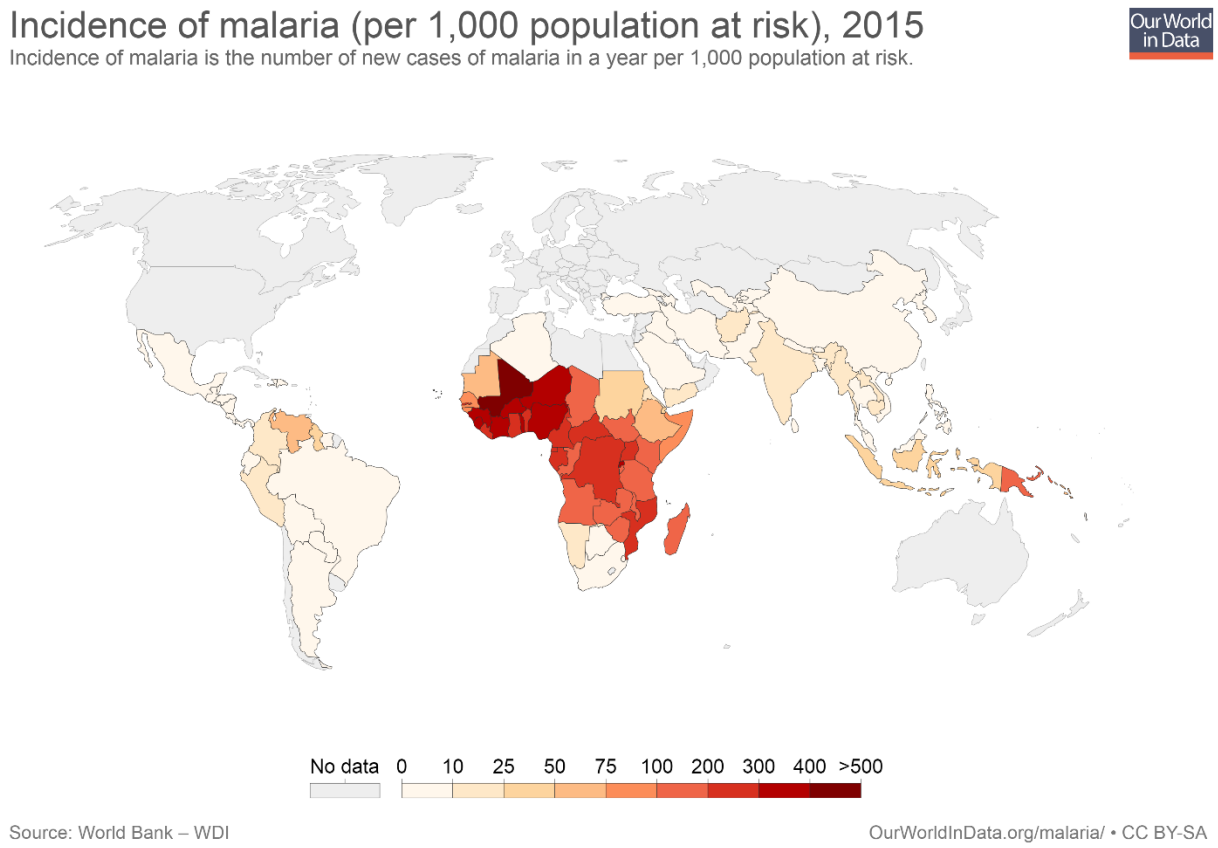


Figure 1.1: Malaria is Still a Major Global Health Problem: Malaria incidence in 2015 by country, per 1,000 population at risk. Dark red indicates higher incidence per country [3].

In the effort to develop vaccines and drugs, transmission events have been identified as prime targets because they are population bottlenecks in the parasite life cycle. Population bottlenecks allow for the targeting of ten to hundreds of parasites in either transmission stage instead of targeting millions of actively replicating parasites in the blood. Rodent model species of malaria, *P. berghei* and *P. yoelii* are often used to study transmission because completion of the *Plasmodium* lifecycle with the reinfection of a new host requires a compatible host liver. While there are humanized mice that can be used to study this in *P. falciparum,* the high cost of these mice places it out of reach for many labs [4]. *P. berghei* and *P. yoelii* are closely related to one another but more distantly related to human infectious species. Despite this, the parasite has the same basic lifecycle, with some changes in developmental timing. Additionally, 4518 genes out of 6257 genes are syntenic orthologues of *P. falciparum* 3D7 [5]. Many genes that

are different are *P. falciparum* erythrocyte membrane proteins (PfEMP's) which are antigenically variant genes in *P. falciparum*. *Plasmodium yoelii* has a distinct set of antigenically variant genes that are more closely related to *P. vivax*. Many proteins that have been studied in both human and rodent infectious *Plasmodium* parasites have similar functions, and while some proteins have different timing or may be absent, the general lifecycle processes are conserved. To appreciate why certain stages are targeted for drug for vaccine therapy, it is important first to consider how the parasite develops through both the mammalian host and mosquito vector.

The *Plasmodium* Life Cycle Requires Transmission between Host and Vector

A malarial infection of a mammal starts with a bite from an infected female *Anopheles* mosquito (Figure 1.2A). The sporozoites are deposited into the skin and will traverse the endothelium to invade a blood vessel and enter the bloodstream (Figure 1.2B). Once at the liver, they will traverse through multiple hepatocytes before invading a single hepatocyte, within which it develops (Figure 1.2C). The complete development of a single sporozoite in the hepatocyte results in tens of thousands of merozoites. Merosomes that are formed from the host cell membrane allow for the release of packaged merozoites into the bloodstream (Figure 1.2D). These packaged merozoites allow the parasite to evade the immune system, masked by the host cell membrane [6-8]. When the merosomes reach the lung micro-capillaries, they burst and release merozoites into the bloodstream. Each merozoite can then seek out and invade a single red blood cell, initiating the blood stage of infection (Figure 1.2E [9, 10]).

During the blood stage of infection, each parasite can become an asexual parasite or a sexual stage gametocyte. Asexual development will allow the parasite to progress through ring, trophozoite and finally schizont stages when daughter merozoites will be released to invade new red blood cells (Figure 1.2F). The asexual parasites cause the clinical symptoms of *Plasmodium* infection due to repeated cycles of bursting red blood cells and sequestration of parasites into tissues. Sequestration is a method for the parasite to prevent clearance from the bloodstream but causes severe damage in the afflicted tissues. A small proportion of these blood stage asexual parasites will make a transcriptional switch and develop into gametocytes (Figure 1.2G). The sexual stage

gametocytes are the only stages that can productively infect a mosquito. *Plasmodium* gametocytes mostly form from committed asexual blood stage parasites; however, gametocytes can be released directly from liver stage schizonts [11-13]. Tight transcriptional and translational controls are essential for proper gametocyte development and transmission [14].

Interestingly, all merozoites from a single committed schizont will become either male or female gametocytes [15-17]. These data indicate that sexual commitment, as well as commitment to a particular sex, occurs during development in the previous cycle. Sexual commitment is, at least in part, controlled by epigenetic regulation [11]. Additionally, stress-inducing conditions like anemia can induce increased sexual commitment in *Plasmodium* cultures [18]. Additionally, sexual commitment, as it's currently understood, is controlled through a variety of mechanisms that revolve around ApiAP2-G. ApiAp2-G is essential for commitment to sexual development in *Plasmodium* species [19]. The ApiAP2-G locus is controlled by histone deacetylase 2 (HDA2) which will remove histone acetylation and allow for histone methylation [20]. Then heterochromatin protein (HP1) can bind and induce silencing [20]. HP1 can then be evicted by gametocyte development gene 1 (GDV1) to allow for ApiAP2-G expression [21]. While commitment has been studied, sex determination is not well understood, but it is thought to be controlled by modulating gene expression.

When a mosquito takes a blood meal from an infected host, it can result in a small number of gametocytes being taken up into the mosquito's midgut (Figure 1.2G [22]). In the midgut, male gametocytes will develop and exflagellate. Once male gametes are able to egress out of the red blood cell individual gametes will seek out an activated female gamete (Figure 1.2H). Activation of gametocytes is initiated in the mosquito midgut because of the temperature shift and an increase in xanthurenic acid concentration. If activation of both male and female gametocytes is successful, the male and female gametes that are released will fuse and form a zygote in the mosquito midgut (Figure 1.2I). This zygote will develop into a motile ookinete that will burrow through the midgut epithelium and form cysts, which are attached beneath the midgut basal lamina (Figure 1.2J[23]). These cysts will grow in size and develop thousands of

sporozoites within them. The sporozoites contained in the cyst are not fully infectious, but after traveling through the hemocoel to reach the salivary gland, they become highly infectious. The sporozoites remain in the salivary glands until the mosquito takes a blood meal and transmits them to another vertebrate host in the anticoagulant saliva (Figure 1.2K).

At both host-to-vector and vector-to-host transmission points, very few parasites are transmitted; thus these bottlenecks are excellent windows of therapeutic intervention. Because of this, we and others have focused upon the transmitted gametocyte and sporozoite stages of *Plasmodium* parasites to identify and exploit any identifiable weaknesses. The identification of molecular processes that are important to the transmission of the parasite in one or both of these events are thus a top priority for the development of new therapeutics.

One such molecular process that is essential to both transmission points is the production, stabilization, and silencing of mRNAs that are needed to establish the infection post-transmission ([9, 14, 19, 24-36], Lindner and Swearingen *et al*. Manuscript in Preparation). The malaria parasite requires tight transcriptional and translational control to navigate these complex transmission events, yet few specific transcription factors are encoded in their genome [34, 37-40] (reviewed in [41]). Perhaps because of this, other post-transcriptional strategies to regulate gene expression have increased importance in malaria parasites. Among these are strategies that control translation, repression, stabilization, and degradation of mRNAs [24, 38]. While more intricate control is achieved through the regulation of translation, transcription factors play an essential role in *Plasmodium* transmission.

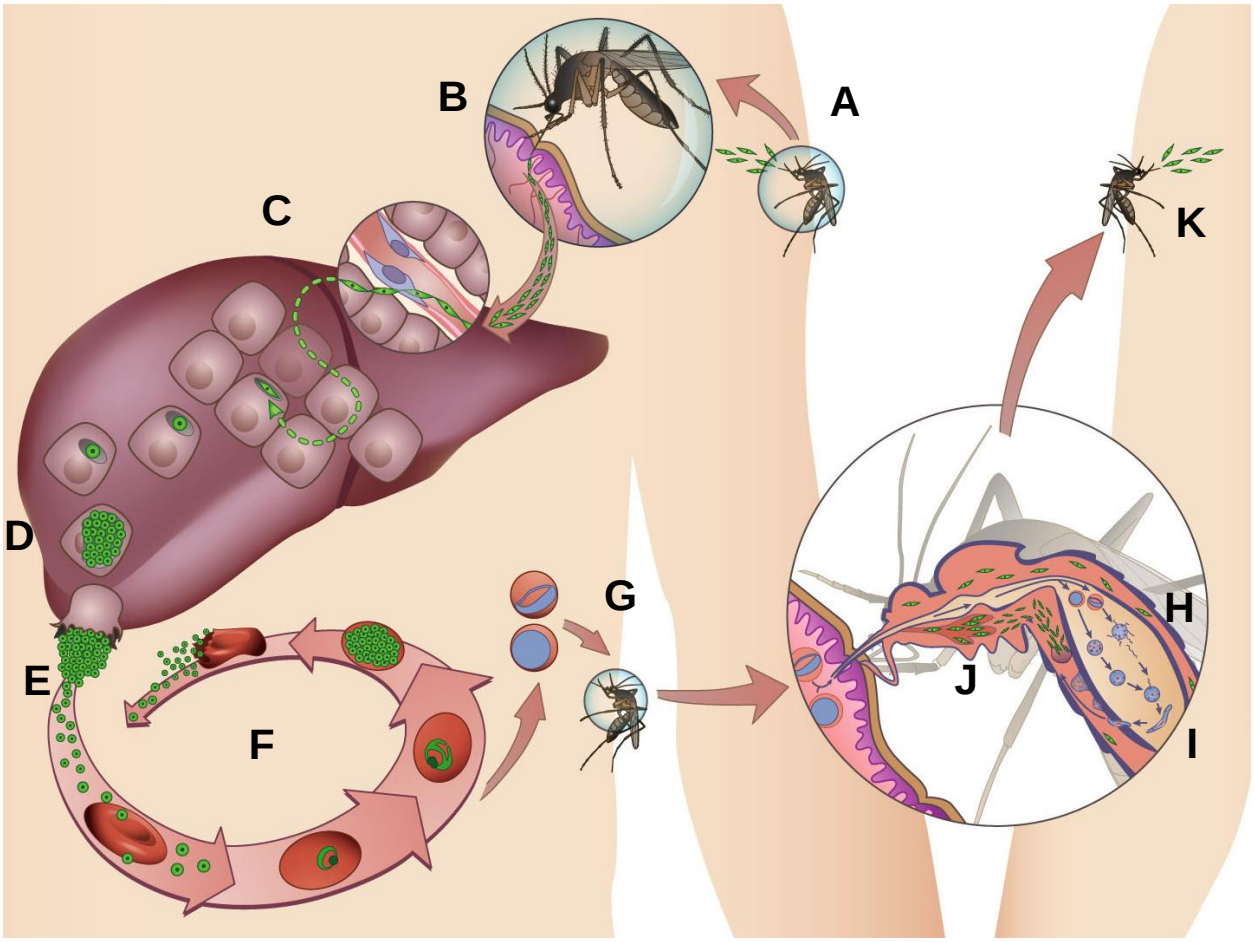


Figure 1.2: The *Plasmodium* Life Cycle is Complex and Requires a Host and a Vector. An illustrated representation of the *Plasmodium* life cycle is shown courtesy of Maria Mota. A) an infected mosquito takes a blood meal and B) injects sporozoites which are able to find the blood stream and migrate to the liver. C) in the liver the sporozoites invade a hepatocyte and develop into liver stage schizonts. D) Merozoites will bud off the infected hepatocyte as merosomes and enter the blood stream where they will enter the lung micro-capillaries and burst. E) Individual merozoites will invade red blood cells and start (F) asexual replication. G) A small proportion of these asexual parasites will make the switch to sexual development and can be taken up by a mosquito, and productively infect it. H) Upon entering the midgut, gametocytes will activate and release gametes which will I) fuse to form a zygote and then a motile ookinete. This ookinete will migrate through the mosquito midgut and form a cyst. Within this cyst sporozoites will develop and J) migrate to the salivary gland where they will await K) another blood meal and initiate another host infection.

Transcriptional Regulation of *Plasmodium* Transmission

Plasmodium has only one known class of specific transcription factors, the ApiAP2 (Apicomplexan Apetala 2) family [42]. In comparison, humans have over 30 defined

families of transcription factors and 1107 individual transcription factors with a known DNA-binding motif [43]. Many members of the ApiAP2 family are expressed in gametocytes, but only three of them have been implicated thus far in gametocyte development, ApiAP2-G (*P. falciparum, P. berghei*, *P. yoelii*), ApiAP2-G2 (*P. berghei, P. yoelii*) and ApiAP2-G3 (*P. yoelii*) [11, 19, 30, 44, 45]. ApiAP2-G is a transcription factor that has been shown to be essential for gametocytogenesis [30]. Gametocytes are not produced in *Apiap2-g* gene deletions, while parasite strains with higher levels of ApiAP2-G expression have a higher proportion of cells committing to gametocytogenesis [19, 30]. Because of this, the activity of ApiAP2-G must be tightly regulated.

There are multiple lines of evidence suggesting that ApiAP2-G is epigenetically silenced in the majority of cells. The *apiap2-g* gene is located on the nuclear periphery, and H3K9me3 marks the AP2-G locus in asexual *P. falciparum* parasites [46]. H3K9me3 (histone three lysine nine tri-methylation) histone modification and localization on the nuclear periphery are associated with silenced genes [47, 48] . H3K9me3 is a strongly repressive histone modification that can block transcription factors from binding to the DNA [46, 49, 50]. HP1 (Heterochromatin Protein 1), a transcriptional repressor that associates with 40% of the known gametocyte gene loci also associates with the *apiap2-g* locus [20, 51]. Once HP1 binds to the H3K9me3 it can cause oligomerization of nucleosomes into heterochromatin and result in gene silencing [52]. When the HP1 protein is depleted, gametocyte conversion increases, providing additional evidence that HP1 acts as a transcriptional repressor [51]. Gametocyte development gene 1 (GDV1) also controls ApiAP2-G expression [21, 53]. GDV1 associates with HP1 and is present on the *apiap2-g* genomic locus. GDV1 was recently shown to displace HP1 to allow for *ap2-g* expression [21]. In a parasite line overexpressing GDV1, gametocyte conversion rates increase to 60% compared to ~10% in cells expressing only endogenous levels of GDV1 [21]. It is not yet understood what leads to GDV1 expression, and further dissection of gametocyte commitment is necessary to determine what ultimately leads to the induction of gametocytogenesis.

ApiAP2-G2 is another transcription factor that when deleted has severe effects on gametocyte development; however, these effects are not as extreme as ApiAP2-G [30]. ApiAP2-G2 deficient *P. berghei* parasites still commit to gametocytogenesis but stop their development early before sex can be determined [44]. ApiAP2-G2 targets mostly asexual genes, indicating that ApiAP2-G2 is likely promoting sexual conversion by acting as a repressor and preventing the parasite from continuing along an asexual path of development [44]. Lastly, parasites with an *apiap2-g3* gene deletion have a reduced number of gametocytes [45]. This is likely because parasites with an *apiap2-g3* gene deletion have reduced expression of *apiap2-g* [45]. Additional work has begun to characterize epigenetic regulation in gametocytes. It was found through transcriptomics and proteomics that histone variants can fluctuate throughout asexual blood stage and gametocytogenesis [54]. Interestingly, certain variants peak in mature stage five gametocytes. There are also certain PTMs associated with early vs late gametocytes. This includes H3K36me2 which was previously shown to be a repressive histone mark in *P. falciparum* [55].

Less is known about specific transcriptional regulation in the context of transmission. A genetic screen in *P. berghei* showed that seven *apiap2* genes play varying but essential roles in mosquito stage development [56]. ApiAP2-O, ApiAP2-O2, ApiAP2-O3, and ApiAP2-O4 all play roles in oocyst development in *P. berghei*. These were confirmed in *P. yoelii*, and additionally ApiAP2-O5 was identified as a gene that could not be disrupted in *P. berghei*, put had no effect upon asexual development in *P. yoelii* and *apiap2-o5* deletion resulted in reduced motility of the oocyst [45]. In the context of vector-to-host transmission, ApiAP2-SP and ApiAP2-SP2 are both essential for sporozoite development, and gene deletions of *apiap2-sp3* produce non-motile sporozoites that do not leave the oocyst in *P. berghei*. ApiAP2-SP was found to bind a specific 8 nt sequence (TGCATGCA) by EMSA (Electromobility shift assay) and regulates sporozoite-specific genes [57-59]. Interestingly, the orthologue of AP2-SP in *P. falciparum* is AP2-EXP and was characterized to play an essential role in asexual development, unlike that of AP2-SP in *P. berghei [60]*. ApiAP2-O and ApiAP2-SP were shown to have functions in regulating transcription in asexual blood stage as well, but are not required for blood stage development [56]. An additional gene, ApiAP2-L plays

an important role in liver stage development, as many cells deficient in ApiAP2-L do not complete liver stage development and those that do develop have a much longer time until a blood stage infection is detectable post infection with sporozoites [56, 61]. While transcription factors can induce expression of a large number of genes, tighter temporal control is necessary for efficient development and transmission. This is demonstrated when translational repressors like *dozi, cith,* and *puf2* are absent [14, 35]. *Plasmodium* parasites can accomplish this through tight translational control.

Translational Regulation of *Plasmodium* Transmission

Post-transcriptional regulation can be used to modulate gene expression. For example, inhibition of PK4, a eukaryotic translation initiation factor 2 alpha (eIF2a) kinase, causes global translational repression and results in the death of parasites [62]. Beyond this specific example cells can increase the stability of transcripts by binding by certain RNA-binding proteins and by lengthening the poly (A) tail, these transcripts can also be stored in granules that do not contain RNA degradation enzymes. Conversely, other RNA-binding proteins can bind to transcripts and target them for decay, through either the usual decay pathway which involves deadenylation, decapping and 5'-to-3' exonuclease activity or through various other exonucleases and endonucleases. Because the timing of transmission is unpredictable, rapid modulation of gene expression is necessary. The proactive generation of transcripts allows for the parasite to sequester these transcripts, and then translate them upon transmission. The *P. falciparum* proteome contains anywhere from 3% (Bioinformatically) to 18% (mRNA-bound proteome) RNA-binding proteins (RBPs) and utilizes these RBPs to control protein expression [63, 64]. In comparison, *Saccharomyces cerevisiae* encodes 600 RBPs, out of the 6000 protein coding genes in the genome, accounting for about 10% of its proteome [65]. Humans encode approximately 30,000 proteins, and ~1000 of them are RBPs accounting for 3% of the human genome [66]. These data demonstrate that *Plasmodium* parasites contribute approximately the same percentage of their genome to encoding RBPs as other eukaryotes in stark contrast to the very few specific transcription factors identified in the parasite genome. We know that *Plasmodium* utilizes translational repression, a process to repress the translation of an mRNA, to

control translation in gametocytes [14, 28, 31-33, 36]. Translational repression in *Plasmodium* is most appreciated in female gametocytes where it is essential for host-to-vector transmission. Translational repression can be identified through highly abundant transcripts whose protein levels are low or absent; however, this may not apply to all of the transcripts identified in this manner.

The parasite has evolved to use RNA-binding proteins (e.g., DOZI, CITH) to impose translational repression and mRNA stabilization in female gametocytes as it prepares for the host-to-vector transmission event [14, 31, 36, 67]. DOZI (Development of Zygote Inhibited) is a DDX6 RNA helicase that is homologous with yeast DHH1 and human rck/p54, and CITH is a homolog of worm CAR-I and fly Trailer Hitch, a Lsm14 orthologue. Parasites lacking DOZI do not complete DNA replication that is important for the formation of the zygote. This causes the parasite to remain 2N and the zygote never fully develops. Whereas, parasites lacking CITH do replicate their DNA normally to become 4N, they cannot develop from the circular, immobile zygote to the sickle-shaped, motile ookinete [36]. Aside from DOZI and CITH, PUF (Pumilio and Fem3 RNA-Binding Factor) proteins have also been shown to play an important role in gametocyte development. Genetic disruption of *pfpuf2* results in the misregulation of gametocyte transcripts and causes a male-biased gametocyte ratio in both *P. falciparum* and *P. berghei* [33, 35, 39].

A current model invokes translational repression as a means for the parasite to always be ready to respond to external stimuli that indicate that transmission has occurred. Translational repression then enables the rapid translation of the preserved mRNAs and successful establishment of the new infection [68]. Similar regulatory events happen in the other transmitted stage (sporozoites) via the PUF2 protein, as gene deletion of *pypuf2* results in the gradual loss of infectivity and subsequent premature dedifferentiation into a liver stage-like form while in the salivary gland of the mosquito, even prior to transmission [34, 39, 40]. As in model eukaryotes, these regulatory functions of RNA metabolism are predicted to occur in cytosolic granules within the parasite as well [14, 68-70]. Defining the composition of different cytosolic granules

within *Plasmodium*, and comparing them to analogous granules in model eukaryotes, is important to better understand how the parasite regulates gene expression.

Cytosolic RNA Processing/Storage Granules

Cytosolic granules are sites of translational repression and RNA processing. These cytosolic granules exist as membrane-less, liquid-liquid phase separated entities that use the affinity of certain protein-protein and protein-RNA interactions to remain separated from the rest of the cytoplasm [71]. The nucleolus was the first discovered liquid-liquid phase separated entity in the cell. There are examples of cytosolic granules with dynamic outer shells and less dynamic core structures implying higher level interactions can occur within these phase separations, as well as examples that are single layered. Multiple smaller granules can move around on microtubules, come together, and merge to share components and to become larger [72-74]. While there is no evidence that these granules that are identified in *Plasmodium* are phase separations, similar granules have been defined as such in model organisms. These have been defined as liquid-liquid because they act as a liquid in the cell. Cytosolic granules appear to drip off of larger membrane bound organelles. Additionally, the exchange and mixing of materials within cytosolic granules is quick as shown by FRAP experiments [71]. Similar experimentation could be done to formally demonstrate phase separation in *Plasmodium* species. These act differently than protein complexes because of the multivalent interactions that occur that allow for the phase separation to occur and for the exclusion of certain larger proteins that cannot diffuse into them.

There are several subsets of cytosolic granules, including stress granules and processing bodies (Figure 1.3). Stress granules are an example of such cytoplasmic membrane-less organelles that form in response to cellular stress (Stress granules reviewed in [75]). Processing bodies (P-Bodies) are cytosolic granules that serve as hubs for the processing of RNA, including processes such as decapping and deadenylation that destabilize the RNA (reviewed in [76]). Depending on the organism and the type of cytosolic granule, different components can be present, which provide a diversity of functions to the granule. In addition to P-bodies and stress granules, there are germ granules, RNA transport granules, U bodies, signaling puncta, and likely more

undescribed cytosolic granules [71]. In yeast, P-bodies and the related stress granules overlap in many of their components and functions, and they progress between these two types along a continuum of different compositions [68, 77].

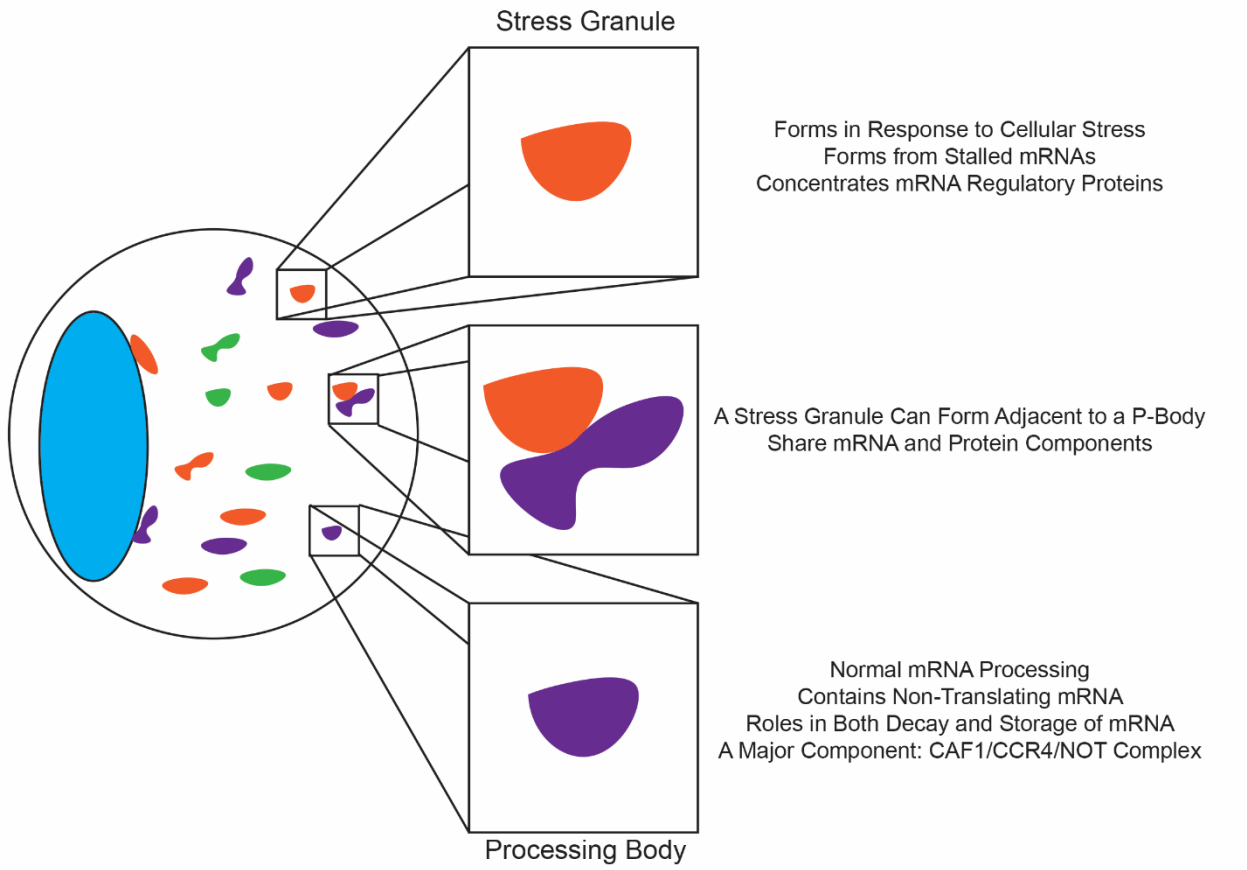


Figure 1.3: Stress Granules and Processing Bodies Play Major Roles in RNA Regulation in the Cell. Stress granules (Orange) and P-bodies (Purple) are capable of docking and sharing components. While stress granules typically form due to cellular stress, P-bodies are a part of normal mRNA processing and are a home for non-translation mRNA.

During the induction of stress, resulting in the phosphorylation of eIF2a, stress granules form from transcripts stalled in translation [77, 78]. Multiple stresses have been shown to cause stress granule formation including heat shock, oxidative, and drug-associated stresses [77]. Heat shock granules were initially found in tomato plants and then later in mammalian cells [79, 80]. In mammalian cells, these granules dissipated slowly upon the removal of the heat stress [80]. Stress granules have mRNAs as well as RNA-binding proteins that mediate their formation and dissolution, and are dynamic, with

proteins and RNA moving in and out of the outer shell quickly [77]. The dynamic nature of stress granules has been demonstrated through FRAP (Fluorescence recovery after photobleaching) experiments. Interestingly, only 10% of total RNA accumulates in stress granules, and only 185 genes in mammalian stress granules account for 50% of their total transcript abundance [81]. Many more transcripts have been identified as translationally repressed in *Plasmodium* transmission stages; however, whether or not this repression occurs in granules for many transcripts and what percentage of these transcripts exist in granules has not been characterized.

Stress granules can function in multiple ways. One way is to bring transcripts and their translation factors in close proximity and in high density to encourage the formation of translation initiation complexes [82]. Another is to sequester mRNAs and proteins away from the others in the cytosol [83-85]. Mutations, post-translational modifications (PTMs) in the components of stress granules, and disruption in stress granule component recycling can result in a multitude of neurological and muscular diseases [86-90]. For example, TDP-43 is a protein involved in normal RNA splicing in the nucleus, yet can have amino acid changes that cause it to be transported to the cytosol instead. Here TDP-43 interfaces with stress granules that are thought to add PTMs that cause its accumulation in the cytoplasm. This accumulation in the cytoplasm results in plaques that are not able to be cleared and causes Amyotrophic lateral sclerosis (ALS) in human patients [90]. A current model for these diseases suggests that stress granule proteins/RNAs build up and result in the death of the cell [78]. While under certain conditions they can cause disease, stress granules play a vital role in regulating transcript abundance in stressed cells. Stress granules do not function alone and can dock with P-bodies to share components [91].

The entire composition of P-bodies and their transient factors have not yet been defined; however, some core components are known, including many RNA-binding proteins (Reviewed in [68, 76]). Formation of P-bodies requires the presence of mRNA and these granules are sites of mRNA storage [71, 76]. mRNA decay in the cell can begin with the removal of the poly(A) tail, known as deadenylation [92]. This process typically occurs through the action of the carbon catabolite repressor 4 associated factor

1/carbon catabolite repressor 4/ negative on TATA less (CAF1/CCR4/NOT) complex, which is known to be important for P-body formation and function in mammalian cells [93].The 5’ 7-methylguanosine cap, along with the cap-binding protein complex, make a transcript resistant to 5’-3’ exonucleases [94]. DCP2 (mRNA DeCaPping 2) is the major decapping enzyme in the canonical P-body complex that functions to remove the cap, and thus decrease the stability of the RNA [76]. Once deadenylation and decapping are complete, XRN1 (exoribonuclease 1), a 5’ to 3’ exonuclease, will degrade the transcript [76]. P-bodies can function in translational repression and require repressed transcripts for their formation [76]. Whether a transcript is degraded or repressed depends on context. If repressors bind the transcript (ex. 4E-T), it may be resistant to decapping and degradation [68]. Transcripts are not only required for P-body formation, but also for their stability as treatment with cycloheximide, which traps transcripts in polysomes, causes a loss of P-bodies [95]. RCK/DHH1/DDX6, the orthologue or DOZI in *Plasmodium*, is a DEAD box helicase that is present in P-bodies, and is required for translational repression through its interaction with the decapping and deadenylase complexes [96, 97].

Composition and Structure of the CAF1/CCR4/NOT Complex

The CAF1/CCR4/NOT complex has been detected as a component of P-bodies, and itself is megadalton-sized complex [68]. The canonical CAF1/CCR4/NOT complex is made up primarily of nine conserved proteins: NOT1, NOT2, NOT3, NOT4, NOT5, CCR4, CAF1, CAF40, and CAF130 [98]. This complex is localized in both the nucleus and the cytoplasm and many of the components are able to play roles in both compartments. The spatial organization of the CAF1/CCR4/NOT complex has been well described. Recently, a cryo EM structure of the *S. pombe* CAF1/CCR4/NOT complex was resolved using immunoprecipitated material [99]. This work confirmed previous studies that used recombinant proteins and binding assays to show that the CAF1/CCR4/NOT complex is L-shaped and that NOT1 (Negative on TATA-less) acts as the scaffold [100]. CCR4 binds to the complex through bridging interactions with CAF1, and the NOT proteins interact with NOT1 on the opposite end from CAF1 and CCR4 [99-103].

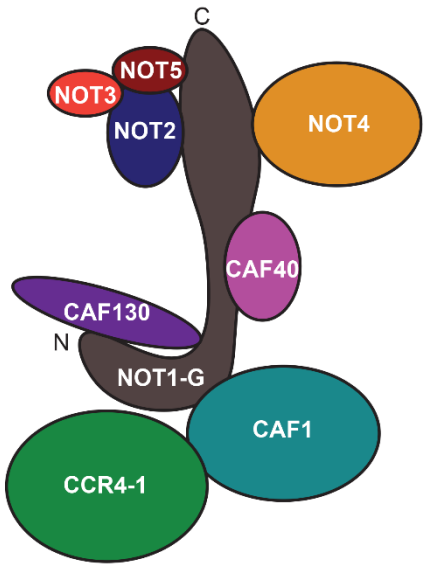


Figure 1.4: The CAF1/CCR4/NOT complex of S. cerevisiae. A model of the CAF1/CCR4/NOT complex, based on structural reconstructions is shown (28271483).

The Nuclear CAF1/CCR4/NOT Complex in Model Eukaryotes

The CAF1/CCR4/NOT complex has been shown to play both direct and indirect roles in the regulation of transcription. The CAF1/CCR4/NOT complex can regulate histone modifications, is directly associated with transcription initiation, and can rescue backtracked RNAPII and can enhance the activity of TFIIS [104, 105].

Histone modifications play an essential role in gene regulation, allowing or prohibiting transcription factors to access and bind to DNA, thereby activating or repressing expression of the gene. Mutations in CAF1/CCR4/NOT complex components have been shown to have significant effects upon histone H3 and H4 modifications [104]. In yeast cells lacking NOT4 or NOT5, histones H3 and H4 are globally hypoacetylated. These data indicate that NOT4 and NOT5 are likely regulating levels of histone acetylases or histone deacetylases [104, 106]. NOT4 and NOT5 also play an integral role in H3K4 tri-methylation (H3K4me3). NOT4 was shown to be able to polyubiquitinate and control the level of JHD2, a demethylase that is responsible for the demethylation of histone H3 [107-109]. Once a histone has been modified, and a transcription factor has access to the DNA it can bind to and recruit RNA polymerases to initiate transcription.

Transcription initiation requires promoter sequences that allow specific proteins to bind and activate transcription. Other proteins can also bind to these promoters and prevent (repress) transcription. Certain elements of the CAF1/CCR4/NOT complex can be crosslinked to promoter sequences, and when tethering individual components to these sequences, they are capable of repressing transcription [110-114]. NOT1 and CAF40 can also associate directly with specific transcription factors, and have been shown to prevent these transcription factors from activating transcription [115, 116]. While acting

as a general regulator, this complex can also play specific roles in regulating gene expression. The CAF1/CCR4/NOT complex contributes to repression of MHC Class II, which binds to antigens and presents them to activate immune cells [117]. NOT3 is also able to interact with EBF1 (Early B-cell Factor 1) and regulate B-cell development [118]. Additionally, NOT4 has been shown to ubiquitinate a specific transcription factor and a subunit of the polymerase associated factor complex leading to their decay [119, 120]. Thus, proteins that belong to the core CAF1/CCR4/NOT complex play important roles in transcription initiation.

Transcription elongation is the next step to gene expression, and the CAF1/CCR4/NOT complex also plays a role in its regulation as well. The CAF1/CCR4/NOT complex has been shown to stimulate elongation of arrested RNAPII *in vitro,* but does not affect the rate of transcription of RNAPII that is not halted [121]. NOT3 and NOT5 can associate directly with the RBP4 subunit of RNA polymerase II (RNAPII), and NOT5 is required for the presence of RBP4 on polysomes [122, 123]. It has been described that upon transcription, certain transcripts may be bound by NOT1 along with other members of the CAF1/CCR4/NOT complex. The CAF1/CCR4/NOT complex can then control the fate of these transcripts [124].

The Cytosolic CAF1/CCR4/NOT Complex in Model Eukaryotes

The CAF1/CCR4/NOT complex plays significant roles in RNA metabolism and in model eukaryotes is the central complex responsible for deadenylation [125]. This complex typically contains two deadenylases, with CAF1 (CCR4-Associated Factor 1) serving as the major deadenylase and CCR4 (Carbon Catabolite Repressor 4) playing additional or specialized roles, except for in yeast where the functions are reversed [125]. Recently it was demonstrated in yeast and humans that CCR4 could displace Poly(A) binding protein (PABP) from transcripts but CAF1 cannot [126, 127]. Though the two deadenylation enzymes are functionally similar, CCR4 is a part of the Exonuclease-Endonuclease-Phosphatase (EEP) family of enzymes while CAF1 is a part of the DEDD nuclease family of proteins. These two enzyme classes contain different residues in their active sites where EEP includes Asp and His, and DEDD has Asp and Glu. The

CAF1/CCR4/NOT complex has been implicated in translational repression and commonly uses accessory proteins to help bind to specific sites on RNA [128-130].

In other species including *Drosophila, Caenorhabditis elegans,* and *Homo sapiens*, CAF1/CCR4/NOT is known to interact with Pumilio and Nanos proteins, which recruit them to a specific transcript and lead to repression or degradation of specific mRNAs [130-132]. Additionally, CAF40 has been shown to interact with *Drosophila* Bag-of-Marbles. Bag-of-Marbles can bind to and use CAF40 and the rest of the complex to regulate expression of its target mRNAs through translational repression [133].

CCR4 and CAF1 are responsible for the majority of deadenylation in model eukaryotes. The removal of the poly(A) tail will allow for decapping factors to remove the cap from the RNA and promote for 5' to 3' degradation. While CAF1 and CCR4 play roles in general RNA decay, their functions in targeted mRNA decay are more appreciated. The complex is recruited to transcripts by RNA-binding proteins including tristetraprolin, BTG/TOB, NANOS, and PUF proteins. These allow for the complex to degrade transcripts that they are recruited to specifically. It has also been demonstrated that CAF1 and CCR4 can repress the translation of target transcripts by tethering the protein to the RNA [125]. Poly (A)-binding protein (PABP) occupancy on a transcript affects whether it is degraded by CCR4 or by CAF1 [126, 127]. In yeast and humans, both CCR4 and CAF1 can control translation when a poly (A) tail is present; however, only CAF1 was able to when the poly (A) tail was absent [126, 127] . Repression by PUF proteins in humans appears to require CAF1/CCR4 and a poly (A) tail on the mRNA for maximal effect; however, it is possible for PUF proteins to repress translation with CAF1/CCR4 but without a poly (A) tail [131].

NOT1 is essential for RNA metabolism, but miRNA has been used to deplete *not1* in human cells, leading to a reduction in the deadenylase activity of CCR4 and CAF1 [134]. This depletion also results in a decrease in the number of P-bodies and results in apoptotic cell death [134]. NOT1 acts as the scaffold to recruit transcripts for deadenylation; it brings together RNA-binding proteins, like the zinc finger RNA-binding protein tristetraprolin (TTP), with the CAF1 and CCR4 deadenylases [135]. The NOT module of the complex includes NOT 2, 3, 4, and 5 which bind to NOT1 along the long

portion of its L-shaped structure [99]. NOT2 variants in yeast have growth defects and a similar loss of polysomes to those of NOT4 and NOT5 mutants. However, it is unclear if the decrease in polysomes is a direct consequence of NOT2/4/5, or if it is due to the slow growth that is a result of their gene deletion. Additionally, NOT4 is required for proper assembly of the proteasome lid [136]. Gene duplication of *not5* has resulted in the gene *not3,* which is absent in many organisms. While NOT 2, 3, 5 associate on one end of NOT1, CAF1 and CCR4 associate on the other. NOT4 can associate in the middle of NOT1 [99]. While all of the above complex members have been shown to play important roles, CAF130 is not well characterized other than its defined association with the complex. While these associations and activities have been well described in model eukaryotes, little is known about the CAF1/CCR4/NOT complex's form, function, and importance in malaria parasites.

The CAF1/CCR4/NOT Complex in *Plasmodium*

The CAF1/CCR4/NOT complex has not been well characterized in *Plasmodium*. Only one manuscript has focused on any members of the putative CAF1/CCR4/NOT complex in *Plasmodium* [137]. In this work, the Adams lab demonstrated that CAF1 is crucial for asexual blood-stage growth [137, 138]. Interestingly, insertion of a *piggyBac* transposon into the coding sequence revealed that CAF1 is responsible for regulation of invasion and egress-related transcripts in asexual blood-stage parasites [137]. Multiple independent attempts to delete the *caf1* gene in the rodent-infectious species *P. berghei* failed, suggesting that it is essential for parasite development [137].

Bioinformatically, seven members of the CAF1/CCR4/NOT complex (NOT1, NOT2, NOT4, NOT5, CCR4, CAF1, and CAF40), have been identified [5, 64]. Interestingly, an additional protein termed NOT-family protein was found that has a significant homology of its identifiable domains to canonical NOT1 domains [64]. While no specific work has addressed the function of any of these members except CAF1, information can be gleaned from two forward genetic screens, *Plasmo*GEM in *P. berghei,* and *piggyBac* transposon mutagenesis in *P. falciparum* [139-141]. These screens are immensely useful in determining what proteins may be essential; however, targeted gene deletion is necessary to confirm that these genes are truly essential. A summary of the results of

these two screens on members of the CAF1/CCR4/NOT complex can be found in Table 1.1.

	P. yoelii Gene ID	*P. berghei* Gene ID	*P. falciparum* Gene ID	PlasmoGEM Data (Pb)	Transposon Mutagenesis (Pf)
NOT1 (NOT1G)	PY17X_0945600	PBANKA_094310	PF3D7_1103800	Not Phenotyped	Dispensable
NOT Family Protein (NOT1)	PY17X_1027900	PBANKA_102550	PF3D7_1417200	Essential	Dispensable
CCR4-1	PY17X_1237700	PBANKA_123430	PF3D7_0519500	Dispensable	Dispensable
CAF1	PY17X_1428300	PBANKA_142620	PF3D7_0811300	Essential	Dispensable
NOT5	PY17X_1207500	PBANKA_120430	PF3D7_1006100	Not Phenotyped	Slow Growth
NOT2	PY17X_0921700	PBANKA_091960	PF3D7_1128600	Not Phenotyped	Dispensable
NOT4	PY17X_1452400	PBANKA_144990	PF3D7_1235300	Not Phenotyped	Slow Growth
CAF40	PY17X_1108300	PBANKA_110720	PF3D7_0507600	Essential	Essential

Table 1.1: A Comparison of Forward Genetic Screen Results on Members of the CAF1/CCR4/NOT Complex [139-141]. “Essential” indicates that the gene could not be deleted or disrupted. “Dispensable” indicates that the gene was able to be deleted or disrupted and “Slow Growth” indicates that the parasite containing the deletion/disruption had slower growth when compared to a pool of other mutants.

*Plasmo*GEM utilized a library of over 2,000 plasmids each designed to delete a distinct gene from the genome. Comparison of the growth of the resulting mutants could then be performed. A few of the NOT complex proteins were not phenotypically characterized in the *Plasmo*GEM screen (NOT1, NOT2, NOT4, and NOT5). The remaining four (CCR4-1, CAF1, CAF40, and NOT1-G) were targeted. Unsurprisingly, CAF1 was shown to be essential, and CCR4 was shown to be dispensable. However, CAF40, which is ordinarily dispensable in model eukaryotes, and NOT-family protein, a protein with good domain homology to NOT1, were shown to be essential in this screen. With CAF40’s function in model eukaryotes to recruit the CAF1/CCR4/NOT complex to mRNAs for degradation, further characterization of this protein should be done in the future [133].

In the human malaria parasite *P. falciparum,* the *piggyBac* transposon mutagenesis screen was able to achieve saturation of insertion sites within the genome, including all of the genes of the CAF1/CCR4/NOT complex [139]. Growth competition assays were used to determine whether any of the mutants displayed an asexual growth defect. This study showed that both NOT1 proteins, CCR4, CAF1, and NOT2 were dispensable for asexual blood stage development. NOT4 and NOT5 mutants had slow competitive growth compared to a pool of other mutants, and CAF40 was found to be essential. Differences in deletion/mutation technique can likely explain the differences in results

between these two screens, but differences between the human and rodent infectious species cannot be ruled out.

The deletion of *pypuf2* in *Plasmodium* sporozoites led to significant changes in the transcript abundance of several members of the CAF1/CCR4/NOT complex by total comparative RNA-sequencing [34, 142]. One possible explanation is that deadenylases, like those in the CAF1/CCR4/NOT complex, were upregulated in an attempt to degrade transcripts that were no longer being repressed. This effort to prevent mis-timed protein expression was not able to compensate for the loss of PUF2, and *puf2*$^{-}$ sporozoites progressively become less infectious with time [34, 39, 40]. Among the affected transcripts, two mRNAs encoding CCR4 domain-containing proteins were increased in abundance. As the deadenylase proteins of the CAF1/CCR4/NOT complex have been shown to be specialized regulators of transcript abundance in other species, I have investigated the possibility that CCR4 domain-containing proteins may be acting in this capacity in *Plasmodium* as well [129, 143].

While the CAF1/CCR4/NOT complex has been well studied in model eukaryotes, there are still new functions and targets of this complex being discovered. Only one protein in this complex, CAF1, has been characterized in *Plasmodium*, and just in a single stage of parasite development, asexual parasites. I show here that multiple components of this complex are important or essential for gametocyte development and host-to-vector transmission.

The work that I have completed expands our knowledge of the CAF1/CCR4/NOT complex in *P. yoelii*. In Chapter 3, I describe how the CCR4 and CAF1 putative deadenylases affect *P. yoelii* development and transmission. CCR4-1 activity is shown to play a role in male gametocyte development and activation. CAF1 plays a role in asexual blood stage development as previously described. However, I found that CAF1 is also required for CCR4-1's function, and loss of CAF1's interaction with CCR4-1 reproduces the deletion of *ccr4-1* phenotypes. In Chapter 4, I describe how *P. yoelii* contains two *not1* genes and uses one paralog, which I have termed NOT1-G, as an essential component of gametocyte maturation and transmission. In Chapter 5, I integrate these findings into what is currently appreciated in the malaria research field,

and propose where future studies could be directed. Future work based on what I have completed will continue to address important questions about the roles of this complex in *Plasmodium*.

Chapter 2: MATERIALS AND METHODS

Ethics Statement

All animal care strictly followed the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) guidelines and was approved by the Pennsylvania State University Institutional Animal Care and Use Committee (IACUC# 42678-01). All procedures involving vertebrate animals were conducted in strict accordance with the recommendations in the Guide for Care and Use of Laboratory Animals of the National Institutes of Health with approved Office for Laboratory Animal Welfare (OLAW) assurance.

Experimental Animals

Six-to-eight week old female Swiss Webster mice from Harlan (recently acquired by Envigo) were used for all of the experiments in this work. *Anopheles stephensi* mosquitoes (obtained from the Center for Infectious Disease Research; Seattle, WA) were reared at 24 °C and 70% humidity and were used to cycle *P. yoelii* (17XNL strain) parasites.

Production of Transgenic Parasite Lines

Transgenic *P. yoelii* (17XNL strain) parasites were created using targeting sequences to incorporate sequence into the target gene using double homologous recombination as previously described [34]. Targeting sequences corresponding to ~750 bp regions immediately 5' and 3' of the open reading frame of the gene-of-interest for the gene deletions and the DDDprom::GFP, or 5' and 3' of the stop codon for the C-terminal tag and dCCR4-1, were generated by PCR using purified *P. yoelii* 17XNL genomic DNA using Phusion polymerase (NEB) and specific primers (Appendix A). These PCR products were combined by Sequence Overlap Extension (SOE) PCR to create a final targeting sequence [144]. This targeting sequence was gel extracted (QIAquick Gel Extraction Kit, Qiagen, Cat# 28706), precipitated with ethanol, and ligated into an intermediate vector (pCR-Blunt, Life Technologies) for sequence verification. The targeting sequence was enzymatically digested out of the intermediate vector and inserted into a pDEF final vector (pSL0444). This vector contains a Green Fluorescent

Protein mutant 2 (GFPmut2) expression cassette for detection and an HsDHFR expression cassette for drug selection for the gene deletions. The final vector (pSL0442) contains a GFPmut2 tag for detection and a HsDHFR expression cassette for drug selection for the C-terminal tag and dCCR4-1 mutant. The DDDprom::GFP was similar to pSL0444 except that the promoter for GFP was replaced with a ~1500bp sequence upstream of DDD. Plasmid DNA was linearized using a unique restriction site between the 5' and 3' targeting sequences and was transfected as previously published [145]. Briefly, the *P. yoelii* 17XNL strain was grown in mice to 1% parasitemia and cultured *in vitro* to the schizont stage. Parasites were purified by an Accudenz discontinuous gradient and were electroporated in Lonza T-cell solution or cytomix in the presence of a total of 5-10 ug of linearized plasmid DNA. Transfected parasites were injected intravenously into mice, and drug cycling was initiated one-day post-transfection by supplying pyrimethamine (0.007% w/v, Fisher Scientific, Cat# ICN19418025) in acidified drinking water (pH~4.0) for three days. Following this, mice were placed on untreated drinking water and parasites were allowed to reach 1% parasitemia. One hundred microliters of infected blood was transferred to a new mouse, and drug cycling was repeated as above. Upon reaching 1% parasitemia, the mouse was euthanized, and the parasites were isolated. Parasite genomic DNA was purified (QIAamp DNA Blood Kit, Qiagen, Cat# 51106) and genotyping PCR was performed to assess the ratio of WT to transgenic parasites present. Clonal parasite populations were produced using limiting dilution cloning by injecting one parasite in 100 uL RPMI media into groups of mice. Mice where <37% of the group became blood stage patent were considered to be infected with a single parasite (as per the Poisson distribution), and genotyping PCR was used to confirm that they were transgenic and clonal. The PfCAF1ΔC line was previously generated as described [137]. Validation of CAF1 transcript expression was performed via RT-PCR on 100ng DNase-treated RNA from PfCAF1ΔC and NF54-control parasites using primer sets supplied in S7 Table.

Production and Accudenz Purification of *P. yoelii* Schizonts and Gametocytes

To produce schizonts in culture, infected Swiss Webster mice were exsanguinated by cardiac puncture, and the blood was placed in 5 mL complete RPMI (cRPMI) (20% v/v

fetal bovine serum in RPMI 1640 with 14.2 ug/mL final concentration of gentamycin) as previously described [37]. The blood was spun at 200 *xg* for 8 min to remove the serum. Blood was then suspended in 30 mL cRPMI and placed in a plugged T75 flask, gassed with a 5% CO_2, 10% Oxygen, 85% Nitrogen gas mixture and cultured for 12 hr at 37 °C on a slight incline while on an orbital shaker at 50-60rpm. Parasites were checked by Giemsa-stained thin blood smears to ensure that parasites had developed fully into mature schizonts/gametocytes. Once confirmed, aliquots of 30 mL of culture were underlayed with 10 mL of 17% w/v Accudenz in 5 mM Tris-HCl (pH 7.5@RT), 3 mM KCl, 0.3 mM disodium EDTA, 0.4x PBS (without calcium and magnesium) and spun at 200 *xg* for 20 min with no brake [146]. Upon completion of the spin, parasites were collected from the interface between the Accudenz and cRPMI layers, transferred to a fresh conical tube, additional cRPMI was added to double the volume of the collected parasites, and then the solution was spun again for 10 min at 200 *xg*. The supernatant was then removed and the parasite pellet processed for downstream applications.

Gametocytes were produced by treatment of the mice at 1% parasitemia with 10 mg/L sulfadiazine (VWR, Cat# AAA12370-30) in their drinking water for two days before exsanguination. The blood was placed in 30 mL of pre-warmed cRPMI to prevent activation of gametocytes and was purified as described above for schizonts.

P. falciparum Gametocyte Production (Performed by John Adams Lab, USF)

Gametocyte-producing cultures were established as described previously [147] with some modification. Briefly, starter cultures of wild-type *P. falciparum* NF54 and PfCAF1ΔC were grown to ~5% parasitemia in 25 cm^3 flasks via standard culture conditions in complete medium [RPMI 1640 medium (Sigma-Aldrich) supplemented with 25 mM HEPES (Sigma-Aldrich), 0.2 % D-glucose (Sigma-Aldrich), 200 µM hypoxanthine (Sigma-Aldrich), 0.2 % sodium bicarbonate, and 10% heat-inactivated human serum] at 6% hematocrit in a tri-gas incubator (5%CO_2, 5%O_2) at 37 ℃. On Day 0, starter cultures were then used to inoculate 75 cm^2 flasks (15 ml culture volume at 6% hematocrit) in technical duplicate for each line at 0.5% parasitemia. Parasites were cultured for 17 days with daily media changes and no fresh addition of blood. Samples were taken to monitor parasite development starting at D3 post-infection and then every

48 hr until D13 post-infection, determined through Giemsa-stained smears. At seven days post-infection, technical replicate flasks were combined into one flask, which was maintained for the duration of the experiment. Cultures were grown in biological triplicate.

DDD Recombinant Protein Expression and Purification

A codon-optimized gene block for expression in *E. coli* containing the PyDDD (PY17X_0418900, AA1845-2334) coding sequence was generated by IDT and inserted into an expression vector with GST and HIS tags (pSL0220) [148]. To help recombinant protein expression, codon optimization was used to reassign codons more commonly used in *E. coli* utilizing the IDT codon optimization tool. These plasmids were transformed into BL21 (DE3) pLysS CodonPlus cells and expressed in a 20L fermenter at 23 °C with 0.5 mM IPTG for 20 hr. The cells were then pelleted and resuspended in 500 mL of low-imidazole buffer (25 mM Tris-Cl [pH 7.5] at room temperature [RT], 500 mM NaCl, 10 mM imidazole, 1 mM dithiothreitol [DTT], 1 mM benzamidine, and 10% v/v glycerol). The solution was then sonicated (4 times, 60% amplitude, 30-second exposure, 50% duty cycle) and then spun at 10,000 rpm for 30 min at 4 °C to pellet insoluble material. The lysate was then incubated with 5 mL of Ni-NTA agarose resin for 1hr at 4C then washed with 50 mL low-imidazole buffer, 20 mL of mid-imidazole buffer (25 mM Tris-Cl [pH 7.5] at RT, 500 mM NaCl, 50 mM imidazole, 1 mM DTT, 1 mM benzamidine, and 10% v/v glycerol) and then eluted with 40 mL of high-imidazole buffer (25 mM Tris-Cl [pH 7.5] at room temperature [RT], 500 mM NaCl, 300 mM imidazole, 1 mM dithiothreitol [DTT], 1 mM benzamidine, and 10% v/v glycerol). The pool was then dialyzed into GST lysis buffer and then incubated with 2 mL of glutathione agarose for 1 hr at 4 °C. The resin was washed with 20 mL GST lysis buffer (50 mM Tris-Cl [pH 8.0] at RT, 150 mM NaCl, 1 mM DTT, 1 mM benzamidine, and 10% glycerol) and eluted with 30 mL GST elution buffer (50 mM Tris-Cl [pH 8.0] at RT, 150 mM NaCl, 1 mM DTT, 1 mM benzamidine, 30 mM reduced glutathione, and 10% v/v glycerol). Purity was assessed by SDS-PAGE stained with Coomassie and confirmed to be >90%. Pocono Rabbit Farm and Laboratory, Inc performed antibody generation as a 84 day fusion protein protocol as approved by The Pennsylvania State University IACUC.

NOT1-G TTP-Binding Domain Recombinant Protein Expression and Purification

A codon-optimized gene block containing the TTP-binding domain from NOT1-G (PY17X_0945600, AA1-199) was generated by IDT and inserted into an expression vector with GST and HIS tags (pSL1075). These plasmids were transformed into pLysS CodonPlus and then expressed in an 8 L culture at 18 °C with 0.5 mM IPTG overnight. The cells were pelleted then suspended in 100 mL of GST lysis buffer (50 mM Tris-Cl [pH 8.0] at RT, 150 mM NaCl, 1 mM DTT, 1 mM benzamidine, and 10% v/v glycerol). The lysate was then sonicated (4 times, 60% amplitude, 30-second exposure, 50% duty cycle) and then spun at 10,000 rpm for 30 min at 4 °C to pellet insoluble material. The supernatant was incubated with 5 mL of glutathione agarose for 1 hr at 4 °C, and then the resin was washed with 50 mL of GST lysis buffer. Elution off of the glutathione-agarose was with 25 mL of GST elution buffer (50 mM Tris-Cl [pH 8.0] at RT, 150 mM NaCl, 1 mM DTT, 1 mM benzamidine, 30mM reduced glutathione, and 10% v/v glycerol). To remove the GST tag, thrombin was added at 10 units/mL for 16 hours at 4 °C. Cation exchange was then performed. The sample was bound to a 20 mL S-Sepharose column and washed with 30 mL of Buffer A (H75 pH 6.76 (20mM HEPES pH 6.76, 75 mM NaCl, 1 mM DTT, and 10% v/v glycerol)) and then eluted with a linear gradient to 100% Buffer B (H1000 pH 6.76 (20 mM HEPES pH 6.76, 1000 mM NaCl, 1 mM DTT, and 10% v/v glycerol)). Purity was assessed by SDS PAGE stained with Coomassie stain and was confirmed to be >90%. Pocono Rabbit Farm and Laboratory, Inc performed antibody generation as a 84 day fusion protein protocol as approved by The Pennsylvania State University IACUC.

Live Fluorescence and IFA Microscopy

CCR4-1, CCR4-2, CAF1, and NOT1-G expression in blood stages, oocyst sporozoites, salivary gland sporozoites and liver stages was observed by an indirect immunofluorescence assay (IFA), and expression in day seven oocysts was observed by live fluorescence of the GFP conjugated to each of these proteins. All samples for IFA were prepared as previously described [149]. Parasites were stained with the following primary antibodies: rabbit anti-GFP (1:1000, Invitrogen, Cat# A11122; 1:1000,

Pocono Rabbit Farm & Laboratory, Custom PAb), rabbit anti-PyACP (1:1000, Pocono Rabbit Farm & Laboratory, Custom PAb), mouse anti-GFP (1:1000, DSHB, Clone 4C9), rabbit anti-HsDDX6 that cross-reacts with DOZI (1:1000, gift from Joe Reese, Custom PAb), mouse anti-alpha-tubulin (Clone B-5-1-2) (1:1000, Sigma, Cat# T5168), and mouse anti-PyCSP (1:1000, Clone 2F6 [34]). Secondary antibodies used for all stages were Alexa Fluor-conjugated (AF488, AF594) and specific to rabbit or mouse (1:1000, Invitrogen, Cat# A11001, A11005, A11008, A11012). 4′,6-diamidino-2-phenylindole (DAPI) was used to stain nucleic acids and samples were covered with VectaShield anti-fade reagent (Vector Laboratories, VWR, Cat# 101098-048) and a coverslip, then sealed with nail polish before visualization. Fluorescence and DIC images were taken using a Zeiss fluorescence/phase contrast microscope (Zeiss Axioscope A1 with 8-bit AxioCam ICc1 camera) using a 40X or 100X oil objective and processed by Zen imaging software.

Measurement of Blood Stage Growth Kinetics

Cryopreserved blood infected with either wild-type (Py17XNL), *ccr4-1*$^{-}$*,* dCCR4-1, CAF1ΔC, or *not1-g*$^{-}$ parasites were injected intraperitoneally into Swiss Webster starter mice, and parasitemia was allowed to increase to 1%. This blood was extracted via cardiac puncture and diluted in RPMI to 10,000 parasites per 100 uL (*ccr4-1*$^{-}$, CAF1ΔC) or 1,000 parasite per 100 ul microliter (dCCR4-1, *not1-g*$^{-}$). One hundred microliters was injected intravenously (IV) into three mice per replicate for each parasite line. Three biological replicates were conducted, each with three technical replicates. Parasitemia was measured daily by Giemsa-stained thin blood smears. Centers-of-movement/exflagellation centers were also measured daily via wet mount of the blood incubated at room temperature for 10 min by counting the number of exflagellating male gametocytes in a confluent monolayer per 400x field (40x objective x 10x eyepiece).

Flow Cytometry Gametocyte Counts

Cryopreserved blood infected with either control PY17XNL with *ccr4-1*, DDDprom::GFP, dCCR4-1, CAF1ΔC or *not1-g*$^{-}$ parasites was injected intraperitoneally into starter mice and transferred as above (10,000 parasites/100ul). On D5 gametocytes were selected for by treatment of the mice with 10 mg/L sulfadiazine (VWR, Cat# AAA12370-30) in

their drinking water for two days. The blood was then placed in 10 mL of pre-warmed cRPMI to prevent activation of gametocytes and spun at 37°C. Blood was then fixed, passed through a cellulose column and stained as if for IFA as described above. Parasites were stained with the following primary antibodies: mouse anti-PvBIP Clone 7C6B4 (1:1000; [150]) and rabbit anti-Dynein Heavy Chain (Amino Acids: 1845 to 2335)(PY17X_0418900) (1:1000, Pocono Rabbit Farm & Laboratory, Custom PAb), Secondary antibodies used for all stages were Alexa Fluor-conjugated secondary antibodies Goat anti-mouse conjugated to AF594 (Fisher Scientific,A11012) Goat anti-rabbit conjugated to AF647(Fisher Scientific, PIA32733). The stained parasites were then run on an LSRFortessa (BD) in tube mode and collected samples were analyzed in FlowJo.

Mosquito Transmission Studies

Cryopreserved blood infected with either wild-type (Py17XNL), *ccr4-1*$^{-}$, dCCR4-1, CAF1ΔC or *not1-g*$^{-}$ parasites was injected intraperitoneally into starter mice and transferred as above [37]. Centers-of-movement were checked daily as above, and mice were fed to mosquitoes on the peak day of exflagellation (day 5). Mice were anesthetized by IP injection of a ketamine/xylazine cocktail (100 mg/kg ketamine, 10 mg/kg xylazine in sterile 1xPBS without Calcium and Magnesium) and 200 mosquitoes were allowed to feed on two mice per cage for 15 min. The positions of the mice were switched every 5 min to allow for more even feeding. Mosquito midguts were dissected at D7 post feed and analyzed for the prevalence of infection and oocyst numbers by microscopy. Mosquito midguts were dissected on day 10, and their midguts ground and oocyst sporozoite numbers counted on a Hausser Bright-Line Phase hemocytometer (Fisher Scientific, Cat# 02-671-6). Similarly, day 14 mosquito salivary glands were dissected, and ground, and salivary gland sporozoite numbers were similarly counted.

Genetic Cross

Cryopreserved blood infected with either wild-type (Py17XNL), or *not1-g*$^{-}$ (Expressing GFP off of a constitutive promoter) parasites were injected intraperitoneally into starter mice, and the parasites were allowed to grow until parasitemia reached ~1%. Mice were then exsanguinated by cardiac puncture, and the infected blood was diluted to 5,000

parasites per 50 uL. Blood was then combined either wild-type with wild-type, *not1-g*$^{-}$ with *not1-g*$^{-}$ or wild-type with *not1-g*$^{-}$ for 100 uL total. These parasites were each then injected intravenously into separate mice, and centers-of-movement were measured daily until the peak of exflagellation centers was reached. On the peak day, mice were fed to mosquitoes and dissected on D7. Dissected midguts were analyzed on a fluorescence microscope as above, and numbers of oocysts and their color were recorded (Green/Transgenic or colorless/wild-type).

Immunoprecipitations, Western Blotting, and Mass Spectrometric Proteomics

Parasite pellets (schizonts) were crosslinked in 1% v/v formaldehyde and lysed using a combination of RIPA lysis buffer (50 mM Tris-HCl (pH 8.0@RT), 0.1% w/v SDS, 1 mM EDTA, 150 mM NaCl, 1% v/v NP40, 0.5% w/v sodium deoxycholate) with a 1x protease inhibitor cocktail (Roche, VWR, Cat# PI88266) and 0.5% v/v SUPERase In (Life Technologies, Cat# AM2694)) for one hr, dounce homogenization (tight) for 30 s, and sonication using a single pulse 0.5 s at 10% amplitude with a Branson Model 102C Sonicator fitted with the microtip. The parasite lysate was then precleared using streptavidin-coated dynabeads (Life Technologies, Cat# 65601) for one hr at 4 °C with rotation. Flow through material that did not bind non-specifically to the beads was immunoprecipitated using a biotin-conjugated anti-GFP antibody (Abcam, Cat# ab6658) attached to the streptavidin-coated dynabeads for three hr at 4 °C with rotation. The beads were washed with modified RIPA wash buffer (50 mM Tris-HCl (pH 8.0 at RT), 1 mM EDTA, 150 mM NaCl, 1% v/v NP40) once, and then transferred to a new tube. The beads were washed 3 more times with modified RIPA wash buffer and then eluted using elution buffer (50 mM Tris-HCl (pH 6.8 at RT), 5% w/v SDS, 5% v/v glycerol, 0.16% w/v bromophenol blue, 200 mM NaCl , 5% v/v β-mercaptoethanol added just before use) at 45 °C overnight in a heat block.

To determine if the immunoprecipitation was successful, one-quarter of the eluted material was subjected to western blotting. Samples were electrophoresed through a 4% stacking and a 10% resolving sodium dodecyl sulfate polyacrylamide gel layers at 200 V for 40 minutes. The proteins were then transferred to a PVDF membrane overnight at 40 V at 4 °C. Membranes were blocked for two hr with 5% w/v dried milk

suspended in 1x PBS and 0.1% v/v Tween-20, and then were probed with a rabbit anti-GFP primary antibody (1:1000, Invitrogen, Cat# A11122), washed with the milk blocking solution, and then were probed with an anti-rabbit secondary antibody (1:1000, Invitrogen, Cat# A16104) conjugated to horseradish peroxidase. SuperSignal West Pico chemiluminescent substrate (VWR, Cat# PI34080) was used for detection by exposure to x-ray film.

For mass spectrometric identification of proteins bound to the protein-of-interest, the remaining three-quarters of the proteins eluted from the beads were electrophoresed one-third of the way through a commercially prepared 4-20% gradient polyacrylamide gel (ThermoScientific, Cat# PI25204). Proteins were stained with Imperial Stain (Fisher Scientific, Cat# PI-24615) for one hr and then repeatedly destained in dH2O. Each lane was cut into four equal slices and was diced into ~ 1 mm^3 pieces. The slices were destained further in 300 uL 50 mM ammonium bicarbonate (AmBic) and 50% v/v Acetonitrile (ACN) solution by washing them three times for 10 min each round. The gel slices were then dehydrated in 300 uL 100% ACN, and the supernatant was aspirated and residual supernatant was evaporated in a speed vac for 10 min (no heat, no pulsed vacuum). Disulfide bonds were reduced by addition 300 uL DTT to 10 mM (final concentration) and incubation at 37 °C in a thermomixer at 800 rpm for 30 min. The cysteine residues were then alkylated by addition of 300 uL 50 mM iodoacetamide and incubation at 37 °C in a thermomixer at 800 rpm for 30 min. The gel slices were then washed with 300 uL 50 mM AmBic, 50% v/v ACN twice for 10 min each round, and then dehydrated again in 300 uL 100% ACN and dried in a speed vac as above. Digestion with trypsin was performed overnight by addition 50-75 uL 6 ng/uL trypsin gold (Thermo Cat# 90055) in 50 mM AmBic at 37 °C in a thermomixer at 700 rpm. The supernatant, containing tryptic peptides, was transferred to a fresh microfuge tube. Peptides remaining in the gel slices were extracted at 37 °C in a thermomixer at 800 rpm for 30 min first with 100 uL 1% v/v formic acid and 2% v/v ACN in nuclease-free water, secondly with 0.5% v/v formic acid and 60% v/v ACN for 30 min, and finally with 100% ACN. Supernatants containing peptides from all extractions were combined with the supernatant from the tryptic digestion in a single microfuge tube for final drying in the

Speed-vac. Dried peptides were then submitted to the Harvard Proteomics Core or the Indiana University Proteomics Core for nano-LC/MS/MS analysis.

For each sample, peptides were resuspended in 0.1% v/v formic acid, and a portion was loaded onto an Acclaim PepMap100 trapping column (column (100 μm × 2 cm, C18, 5 μm, 100 Å, Thermo) at a flow rate of 20 uL/min using 4% v/v aqueous acetonitrile (ACN), 0.1% v/v formic acid (FA) in dH_2O as a mobile phase. The peptides were separated on an Acclaim PepMap RSLC column (75 um × 15 cm, C18, 2 um, 100 Å, Thermo) with a 90 min 4% - 60% v/v linear gradient of acetonitrile in dH2O containing 0.1% v/v formic acid. The gradient was delivered to the column by a Dionex Ultimate 3000 nano-LC system (Thermo) at 300 nL/minute. An LTQ Orbitrap Velos mass spectrometer (Thermo) was set up for a '2nd Order Double Play' type of experiment with the following parameters: full positive-ion 1000-ms FT-MS scan at R 60,000 over 350 – 1700 m/z followed by ten ion-trap MS2 scans on most intense precursors with collision-induced dissociation (CID) activation. Precursor ion signal threshold was set at 5000, isolation width 2 m/z, normalized collision energy 35.0 V, activation Q 0.250, activation time 10.0 ms. The precursors were selected using an FT master scan preview mode with charge states less than +2 rejected and monoisotopic precursor selection enabled. Dynamic exclusion repeat duration was 25 s, exclusion duration 13 s, exclusion list size was 200, and the exclusion mass width was +/- 10 ppm relative to the excluded m/z. Polysiloxane signal (m/z 445.1200) was used as a lock mass.

The data was processed using the Trans-Proteomic Pipeline (TPP) [151] as described previously with few modifications [34]. Raw data were converted to .mzml format using msconvert [152] and searched using both X!Tandem [153] and Comet [154]. Spectra were searched against reference sequences downloaded in February 2016 from *P. yoelii* 17X (PlasmoDB, v27), mouse (UniProt), and common contaminants (Common repository of adventitious protein sequences, [155] and randomized decoys generated through TPP. iX!Tandem and Comet searches were combined in iProphet [156], and Peptide Prophet determined protein identifications. Only proteins with a highly stringent false positive error rate of less than 1% were reported. To combine replicate proteomics datasets, SAINT version 2.5.0 was used [157]. The algorithm utilizes the total spectral

abundance for each protein and the protein length to normalize and determine the probability of interaction with the bait (CCR4-1::GFP, CAF1ΔC::GFP). The algorithm was run with the following settings: lowMode=1, minFold=0, normalize=1. Only proteins with SAINT scores below 0.1 (most stringent) or 0.35 (stringent) were considered significant hits and included in the analyses, as used previously [22, 34, 37].

Total Proteomics by Mass Spectrometry

Mixed blood-stage parasites from one mouse infected with *P. yoelii* wild-type (Py17XNL strain) (~4% parasitemia) or *ccr4-1*$^{-}$ (~6% parasitemia) transgenic parasites were purified by cellulose column purification (to remove WBC's) and saponin lysis followed by a 4000 *xg* spin for 10 min. The pellets were washed with 1x PBS and spun again same as before. Parasite pellets were placed in 2x sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer. Samples were then sonicated with a Branson Model 102C Sonicator fitted with the microtip (10% amplitude; 50% duty cycle; 10 bursts) three times. Samples were allowed to cool and were spun to remove bubbles between each round. Beta-mercaptoethanol was added to a final concentration of 5% to reduce disulfide bonds. Samples were then heated to 70 °C for 5 min and loaded on a commercial 4-20% gradient polyacrylamide gel (Thermo, Cat# PI25204). Samples were electrophoresed until the dye front reached near the bottom of the gel, and individual sample lanes were separated and stained with Imperial Stain (Thermo Scientific, Cat#: 24615) for 1.5 hr and then destained for 1.5 hr in dH2O. Each lane was cut into 24 separate slices, and each slice was submitted for independent mass spectrometric analysis at the Penn State Mass Spectrometry Core using the same run parameters as described above for proteomic analyses conducted at the Harvard and Indiana University Proteomic Core.

Total and Comparative RNA-seq

Gametocyte samples were collected as described above. Following purification by an Accudenz gradient, mouse RBCs were subjected to lysis with 0.1% w/v saponin in 1x PBS, then washed with 1x PBS. Released parasites were then lysed immediately. RNA from the samples was isolated by the QIAgen RNeasy Kit (QIAgen, Cat No. 74104) using the manufacturer's protocol with the additional on-column DNaseI digestion. RNA

yields were quantified spectrophotometrically (NanoDrop 2000c, Thermo Scientific), and RNA samples were submitted to the Penn State Genomics Core Facility. The quality of all samples was confirmed by measuring RNA Integrity Number (RIN) using the Agilent Bioanalyzer. A barcoded library was made from each sample by using the Illumina TruSeq Stranded mRNA Library Prep Kit (Illumina, Cat# RS-122-2101) according to the manufacturer's protocol. Quantatative polymerase chain reaction (qPCR) was performed to determine the concentration of each library, and an equimolar pool was made of all libraries. The library pool was sequenced on an Illumina HiSeq 2500 in Rapid Run mode according to the manufacturer's protocol. For gametocyte samples, 100 nt single read sequencing was performed. The resulting data was mapped to the *P. yoelii* 17XNL strain reference genome (plasmodb.org, v32 using Tophat2 in a local Galaxy instance (version .9). Gene and transcript expression profiles for both WT-GFP and *ccr4-1*$^{-}$ assemblies were generated using htseq-count (Galaxy version 0.6.1galaxy3) [158] using the union mode for read overlaps. Count files were merged and compared using DESeq2 (Galaxy version 2.11.39 [159]). Six biological replicates were used for the WT transcriptomic profile, while four replicates were used in for the *ccr4-1*$^{-}$ profiles. These data were analyzed by a mean fit type with outlier replacement turned on to normalize the variance between the count files. The P-adjusted value was used for all analysis.

Circular Reverse Transcription PCR (cRT-PCR) (Work performed by Scott Lindner)

RNA was isolated from purified *P. yoelii* wild-type or *pyccr4-1*$^{-}$ gametocytes by TRIzol/chloroform extraction and extensive DNaseI digestion. The 7-methylguanosine cap was removed from 10 ug of total RNA using 2.5 U Cap-Clip Acid Pyrophosphatase (CellScript, Madison, WI, #C-CC15011H) in 1x Cap-Clip Buffer supplemented with 10U Murine RNase Inhibitor (NEB, #M0314) at 37 °C for 1 hr. The decapping reaction was stopped by TRIzol/chloroform extraction, and RNA was precipitated with an equal volume of isopropanol and 10 ug glycogen, washed twice in 80% v/v ethanol, vacuum aspirated and dried at 42 °C for 5 min. RNA was quality controlled by NanoDrop to ensure no residual chaotropic salts (A230) or phenol (A270) remained. Decapped RNA heated to 65 °C for 5 min, and was circularized using 10 U T4 RNA Ligase (NEB,

#M0204) in 100 ul 1x T4 DNA Ligase Buffer supplemented with 10% w/v PEG8000 and 10U Murine RNase Inhibitor at 16 °C for 24 hr. The circularization reaction was stopped by TRIzol/chloroform extraction, and RNA was precipitated by isopropanol/ethanol and quality controlled as above.

Reverse transcription of specific mRNAs was performed using SuperScript IV (Invitrogen, #18090010) and a gene-specific primer (Appendix A). Fifty nanograms of circularized RNA (cRNA) was denatured at 65 °C for 5 min in 14.25 ul Solution 1 (9.5 ul DEPC-treated double distilled water (ddH2O), 2 ul 1 uM reverse primer, 1 ul 10 mM dNTP mix, 1ul cRNA) with chilling on ice for 1 min. First-strand synthesis was performed by addition of 5.75 ul Solution 2 (4 ul 5x SuperScript IV Buffer, 1 ul 100mM DTT, 0.5 ul 20 U/ul Murine RNase Inhibitor, 0.25 ul SuperScript IV) and incubation at 50 °C for 10 min. Reverse transcription reactions were stopped by heating to 80 °C for 10 min.

Specific PCR amplification of *gapdh* and *p28* sequences was conducted using Phusion polymerase (NEB) and gene-specific primers (Appendix A), with products assessed by a 2% w/v agarose gel with 100 bp molecular weight ladder (NEB).

Statistical Analyses

Statistical differences between *P. yoelii* wild-type and transgenic parasites were assessed via a two-tailed t-test on GraphPad Prism. Statistical differences between *P. falciparum* wild-type and PfCAF1ΔC parasites were assessed via a paired Wilcoxon test using R v. 3.3.1 [160] with $p < 0.05$ indicating significance.

Data Availability Statement

Proteomics Data

Proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD007042 [161].

Dataset Identifiers: SEL31: Wild-type parasites expressing GFPmut2 from a disrupted *p230p* locus (WT-GFP) replicate one, control for SEL32; SEL32: *ccr4-1*$^{-}$ experimental replicate one; SEL33: WT-GFP replicate two, control for SEL35; SEL34: WT-GFP replicate three, control for SEL36; SEL35 *ccr4-1*$^{-}$ experimental replicate two; SEL36:

ccr4-1^{-} experimental replicate three; SEL80: WT-GFP replicate one, control for SEL82; SEL82: *CAF1ΔC* Experimental replicate one; SEL83: WT-GFP replicate two, control for SEL85; SEL84: WT-GFP replicate three, control for SEL86; SEL85: *CAF1ΔC* Experimental replicate two; SEL86: *CAF1ΔC* Experimental replicate three.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD007042[161].

Dataset Identifiers: SEL1: PY17XNL total proteomics, replicate one; SEL2: *ccr4-1*$^{-}$ total proteomics, replicate one.

RNA Sequencing

The DESeq2 output can be found in the GEO depository (Accession number GSE101484).

RAW and processed transcriptomic files have been deposited at GEO (Accession # GSE101484)

Chapter 3: PYCCR4-1 PLAYS AN IMPORTANT ROLE IN MODULATING TRANSCRIPT ABUNDANCES FOR PROPER GAMETOCYTE DEVELOPMENT AND TRANSMISSION

Introduction:

The CAF1/CCR4/NOT complex is a multi-protein complex that plays roles from modulating transcription to the decay of transcripts. The complex consists of a group of NOT proteins and a group of deadenylases, CCR4 and CAF1 [104, 125]. While members and functions of this complex have been defined in model eukaryotes, only one member of this complex, CAF1, has been characterized in *Plasmodium* [104, 125, 137]. In *Plasmodium*, CAF1 has previously been characterized as being likely essential in *P. berghei* because the gene was unable to be deleted from the genome. Additionally, CAF1 was able to be disrupted by a transposon in its coding sequence in *P. falciparum*. This disruption resulted in dysregulation of genes involved in merozoite invasion [137]. In this chapter, I have characterized both the CCR4 and CAF1 deadenylases and their roles throughout the *P. yoelii* life cycle. I found that PyCCR4-1 plays an important role in modulating transcript abundances for proper gametocyte development and transmission. This effect is dependent on PyCCR4-1's catalytic residues, as when the sequence that encodes these is mutated it phenocopies a *pyccr4-1* gene deletion. I have also demonstrated that CAF1 is likely essential in *P. yoelii* and showed that the interaction of PyCCR4-1 with PyCAF1, as well as the rest of the complex, is required for its ability to promote proper gametocytogenesis and host-to-vector transmission. Finally, I show that PyCCR4-1 binds directly to some of the affected transcripts but not all and does not affect poly (A) tail length of one unaffected and one affected transcript. These data indicate that PyCCR4-1 could be playing a role other than deadenylation, such as assisting in translational repression, in the regulation of transcripts or that the effects of *ccr4-1* deletion are mostly indirect.

PyCCR4-1 and PyCCR4-2 Localize to Discrete Cytosolic Granules

In yeast, *ccr4* and *caf1* deletion phenotypes are similar to those when their interaction with the complex is disrupted [162]. This data indicates that CCR4 and CAF1 require the association with the complex for their major functions. The CAF1/CCR4/NOT complex is found in nuclear and cytosolic granular structures in both yeast and human cells [104]. To determine the localization of PyCCR4-1 and PyCCR4-2, I C-terminally tagged these proteins with GFP and confirmed these integrations with genotyping PCR

(gPCR) (Figure 3.1. Figure 3.2). Using immunofluorescence and live fluorescence assays with transgenic parasites, PyCCR4-1::GFP was shown to localize to cytoplasmic puncta in asexual blood stage parasites, as well as male and female gametocytes (Figure 3.2). Moreover, this expression pattern extends to oocysts, oocyst sporozoites, and salivary gland sporozoites, where PyCCR4-1 was seen both in cytosolic puncta and located diffusely throughout the parasite (Figure 3.2). However, PyCCR4-1 was not detected above background in liver stage parasites. In agreement with this expression pattern of CCR4-1, transgenic parasites expressing CCR4-2::GFP also localize to cytosolic puncta in the same stages as CCR4-1::GFP (Figure 3.3). Thus, the near constitutive expression and localization of PyCCR4-1 and PyCCR4-2 in cytoplasmic foci in *Plasmodium* resembles that of its orthologues in model eukaryotes.

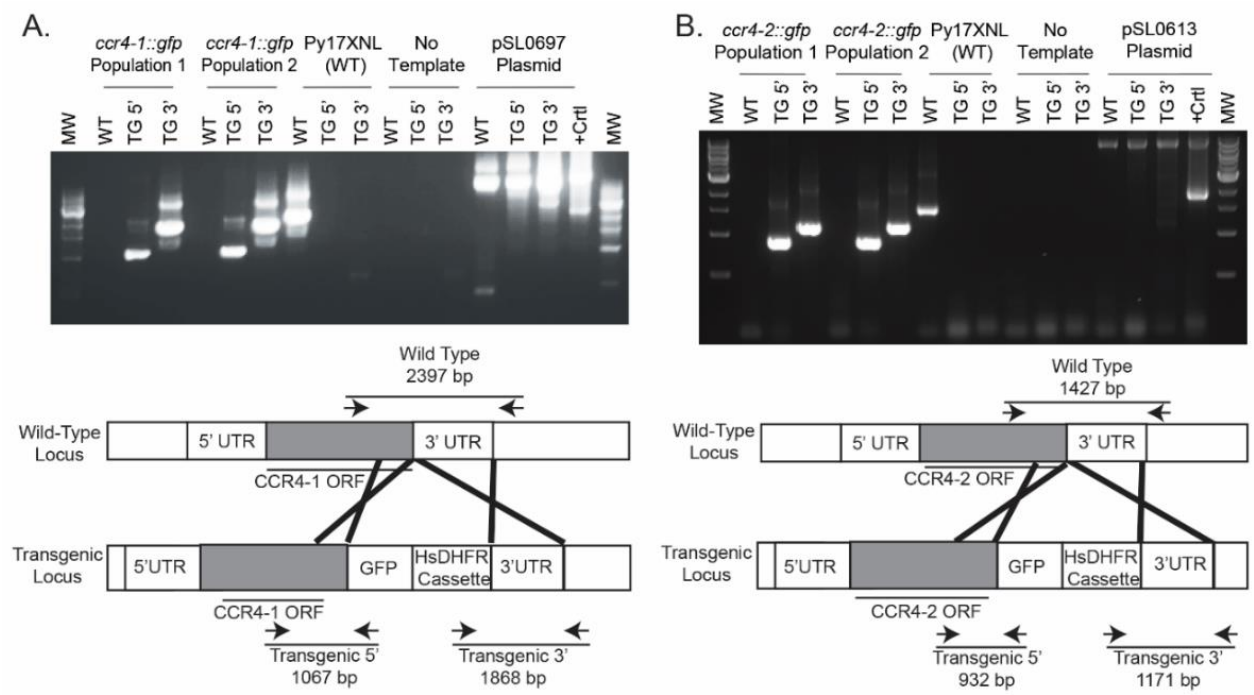
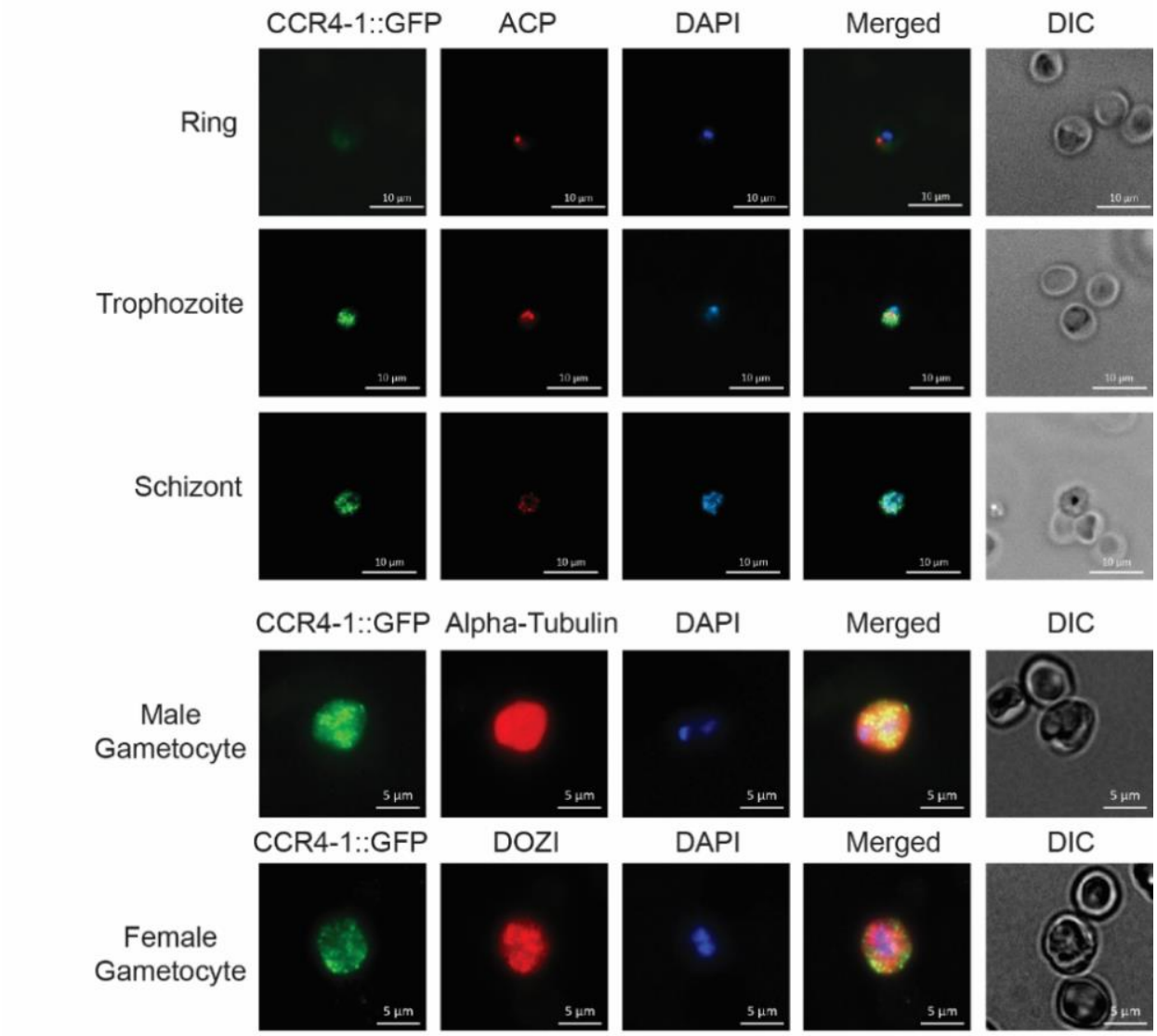


Figure 3.1: Genotyping PCR of (A) *pyccr4-1::gfp* and (B) *pyccr4-2::gfp.* Insertion of a C-terminal GFP tag was created using double homologous recombination at the 3' end of the respective coding sequence. Genotyping was performed by PCR on parasites using the primers indicated (listed in Appendix A and illustrated on the genome schematic). Independent clones were analyzed with a Py17XNL wild-type control, a no template control, and a plasmid positive control in parallel. pSL0697 plasmid positive control size is 2108bp. pSL0613 plasmid positive control size is 1922bp. Ladder is from NEB (N3232L). WT=Wild-type TG=Transgenic


CCR4-1::GFP
ACP
DAPI
Merged
DIC
Ring
Trophozoite
Schizont
CCR4-1::GFP
Alpha-Tubulin
DAPI
Merged
DIC
Male Gametocyte
CCR4-1::GFP
DOZI
DAPI
Merged
DIC
Female Gametocyte

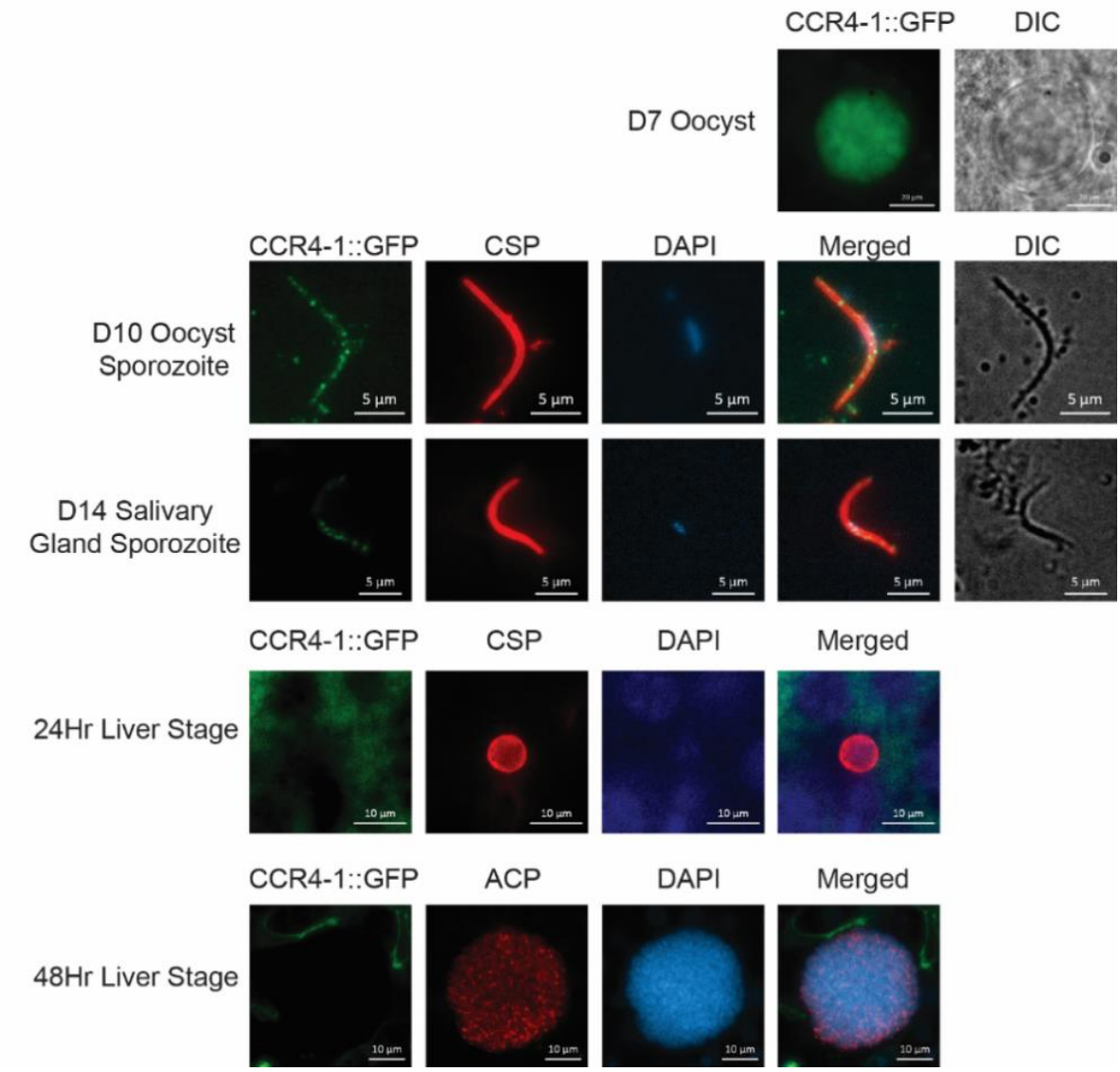
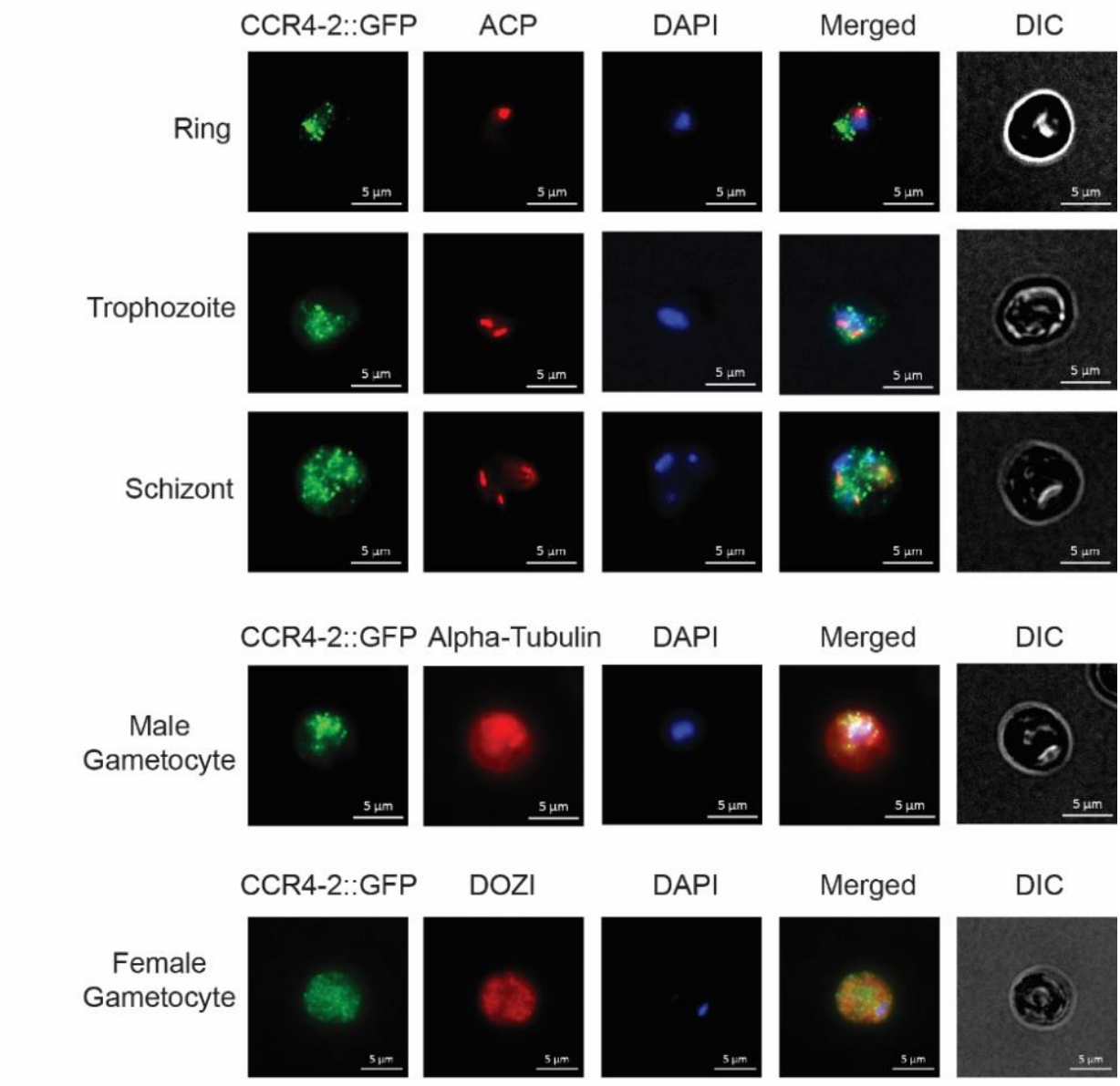


Figure 3.2: PyCCR4-1::GFP is Expressed in Asexual, Sexual, and Mosquito Stage Parasites but is Not Detectable in Liver Stage Parasites. Representative images of asexual blood stage parasites, sexual blood stage parasites, oocyst sporozoites, salivary gland sporozoites and 24-hour and 48-hour liver stage parasites treated with a nucleic acid stain (DAPI) and antibodies to GFP (to detect PyCCR4-1::GFP) or to stage-specific cellular markers (e.g., CSP, alpha-tubulin, DOZI ACP) are shown. Live fluorescence was used to image oocysts. Scale bars are either 20 microns (oocysts), 5 microns (sexual blood stage, sporozoites), or 10 microns (asexual blood stage and liver stage parasites).


CCR4-2::GFP
ACP
DAPI
Merged
DIC
Ring
Trophozoite
Schizont
CCR4-2::GFP
Alpha-Tubulin
DAPI
Merged
DIC
Male Gametocyte
CCR4-2::GFP
DOZI
DAPI
Merged
DIC
Female Gametocyte
5 µm

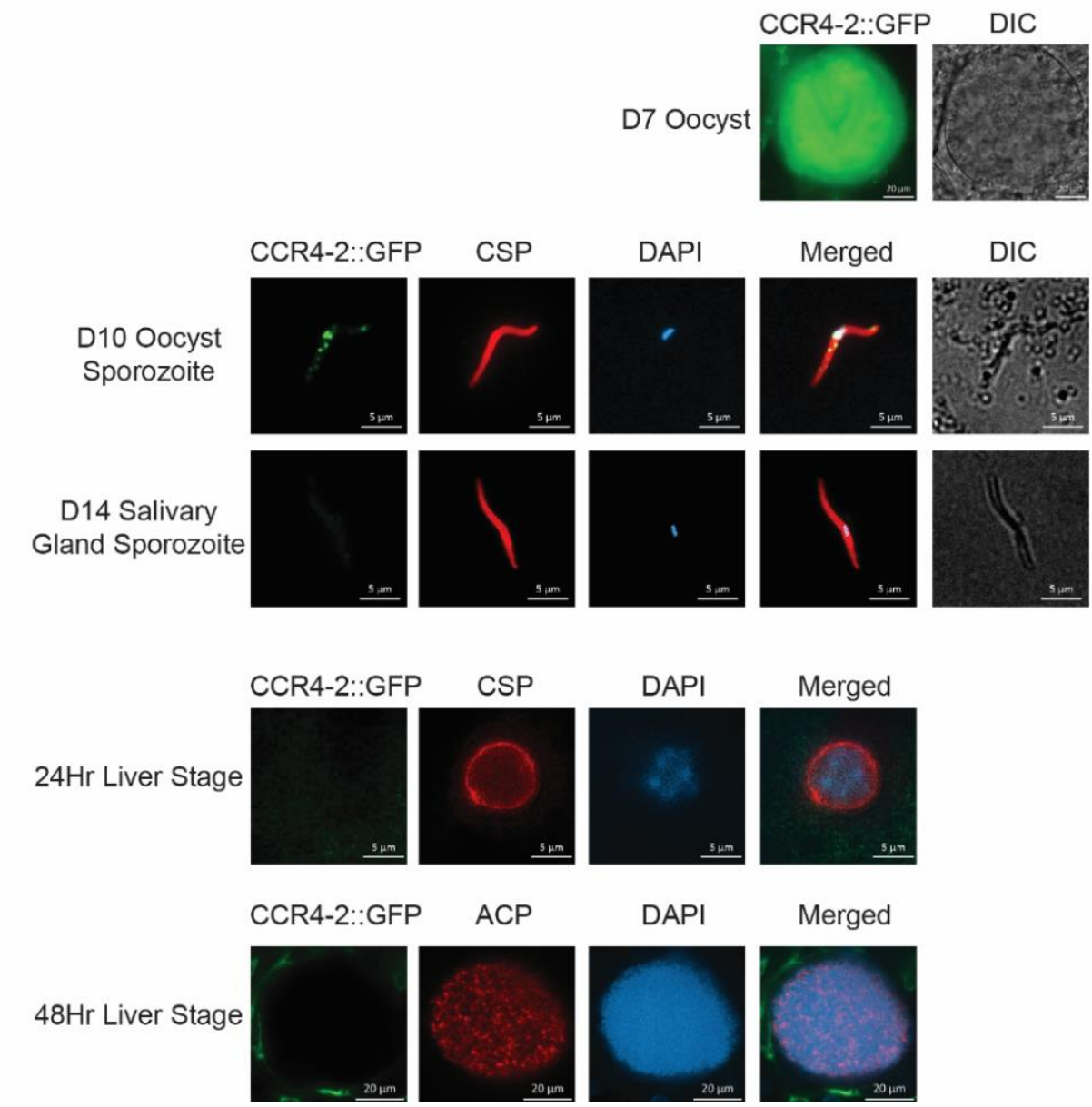


Figure 3.3: PyCCR4-2::GFP is Expressed in Asexual, Sexual, And Mosquito Stage Parasites but is Not Detectable in Liver Stage Parasites. Representative images of asexual blood stage parasites, sexual blood stage parasites, oocyst sporozoites, salivary gland sporozoites and 24-hour and 48-hour liver stage parasites treated with a nucleic acid stain (DAPI) and antibodies to GFP (to detect PyCCR4-2::GFP) or to stage-specific cellular markers (e.g., CSP, ACP, alpha-tubulin, DOZI) are shown. Live fluorescence was used to image oocysts. Scale bars are either 20 microns (oocysts, 48hr liver stage), 5 microns (asexual and sexual blood stage, sporozoites, and 24hr liver stage).

PyCCR4-1 Associates with a Canonical CAF1/CCR4/NOT Complex

Bioinformatically using BlastP searches, I identified the genes encoding all members of a canonical *S. cerevisiae* CAF1/CCR4/NOT complex in *P. yoelii* (Py)*, P. berghei* (Pb), *P. falciparum* (Pf), *and P. vivax* (Pv), except for *not3* and *caf130* (7/9) [163]. The absence of these two particular genes is not surprising, as they are also absent in humans and *Drosophila* [104]. I also identified three other CCR4 domain-containing proteins (PyCCR4-2, PyCCR4-3, and PyCCR4-4) (Figure 3.4A). While the EEP domain of these proteins appears to be well conserved, there is no conservation outside of these domains between the three CCR4-domain containing proteins. The catalytic residues of this domain are all conserved across paralogs and across *Plasmodium*. This pattern of conservation within domains and not outside of the domains is consistent not only for the CCR4-domain containing proteins but also for the rest of the identified core CAF1/CCR4/NOT complex. This includes comparisons with humans where the amino acids are conserved within the domain, especially within the active site, but not beyond that. Total proteomics of mixed blood stage samples of *P. yoelii* wild-type and *pyccr4-1* gene deletion parasite lines indicated that many of these bioinformatically defined members (NOT1, NOT4, NOT5, NOT1-G, CAF40) are expressed at levels that permit their detection, whereas PyCCR4-1 and PyCAF1 were not sufficiently abundant to be detected in either dataset using highly stringent thresholds (See materials and methods for proteomics repository).

To experimentally determine the composition of this complex in *P. yoelii*, the C-terminal GFP tag on PyCCR4-1::GFP was utilized to immunoprecipitate the CAF1/CCR4/NOT complex from synchronized schizonts when PyCCR4-1 is most abundant (by live fluorescence) and is localized to cytoplasmic granules. I did not enrich for the cytosolic fraction for this experiment, thus we would expect contributions from any CCR4-1 that is present in the nucleus. Crosslinking with formaldehyde was used to preserve these interactions. As seen in model organisms, PyCCR4-1 associates, directly or indirectly, with most members of the canonical CAF1/CCR4/NOT complex in *P. yoelii* (Table 3.1). Specifically, I found that PyCCR4-1 associates with CAF1, NOT1, CAF40, NOT2 and a NOT family protein above our most stringent SAINT (Significance Analysis of

INTeractome) threshold (0.1), and with NOT5 using a less stringent threshold (0.1 to 0.35). A small number of peptide spectral matches for NOT4 were also observed but were not sufficiently enriched to be included confidently. This low abundance of NOT4 is consistent with its known transient association with the CAF1/CCR4/NOT complex in other eukaryotes [164]. I also found that PyCCR4-1 interacts with proteins involved in the nuclear pore complex and RNA export (e.g., karyopherin-beta 3, exportin-1, UAP56), proteins involved in translation initiation (e.g., eIF2A, EF-1a, EIF3D, PABP), and translational repression (e.g., CELF2/Bruno, DOZI, CITH, PABP) (Table 3.1). All of these interactions are consistent with known CAF1/CCR4/NOT interactions in model eukaryotes such as yeast, human cells, and/or *Drosophila* [104, 165]. Recently, a proteome of stress granule components in *S. cerevisiae* defined several cytosolic granule regulators, and several of the orthologues of these regulators are also found associated with PyCCR4-1 [166]. Specifically, I identified that multiple CCT proteins of the TRiC complex (e.g. CCT4, CCT5, CCT8) and HSP40-A associate with the CAF1/CCR4/NOT complex (SAINT score < 0.1), and this list expands to include the remainder of the TRiC core complex (e.g., TCP1, CCT2, CCT3, CCT6, CCT7) and casein kinase 1 when considering SAINT scores between 0.1 to 0.35. Also, recent work has implicated karyopherins/nuclear-import receptors in the regulation of proteins found within liquid-liquid phase separations/cytosolic granules. Nuclear-import receptors can potentially help mediate the dissolution of these cytosolic granules [167]. Here, I identified that karyopherin beta three associates with the *P. yoelii* CAF1/CCR4/NOT complex, and perhaps indicates that similar regulatory processes are at work. These data indicate that the composition of the CAF1/CCR4/NOT complex, including the presence of cytosolic granule regulators, are likely conserved from model eukaryotes to *Plasmodium*.

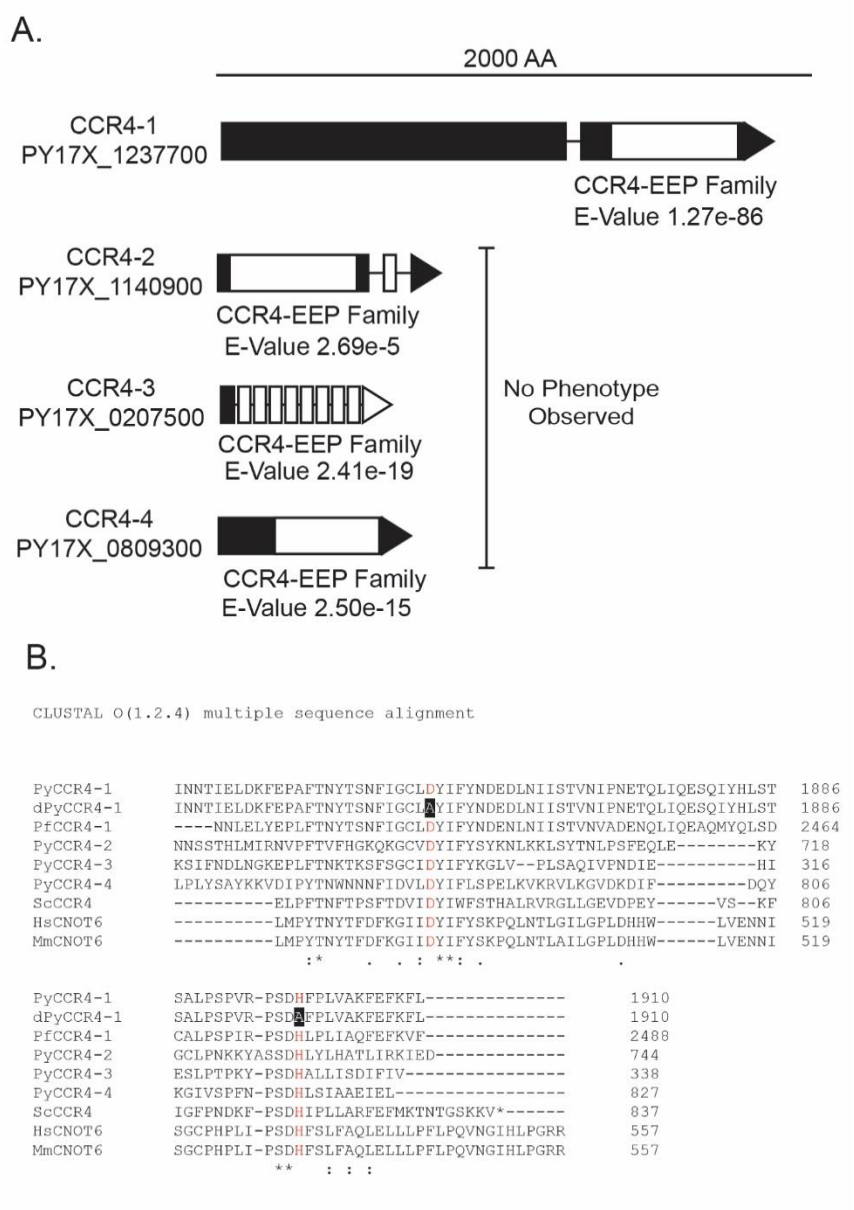


Figure 3.4: *Plasmodium* Species have Four Bioinformatically Predictable CCR4 Domain-Containing Proteins. (A) The four proteins with identified exonuclease-endonuclease-phosphatase domains (EEP) are shown to scale with their domain architecture, introns (gaps) and exons (rectangles). Also shown are E-values for their EEP domain based upon their alignment with CCR4 (PLN03144) via the Conserved Protein Domain Database. (B) The four bioinformatically predictable CCR4 domain-contains proteins from *P. yoelii*, the catalytically dead PyCCR4-1 (dPyCCR4-1), CCR4-1 from *P. falciparum*, and examples from *S. cerevisiae*, human, and mouse were aligned using EMBL Clustal Omega. Shown is the region around the catalytic residues of CCR4. Highlighted in red font are the two catalytic residues, and the two residues that were changed to create the catalytically dead variant are highlighted in black with white font.

Gene ID	Name	PyCCR4-1::GFP			PyCAF1ΔC::GFP		
		Spectra	Control Spectra	SAINT Score	Spectra	Control Spectra	SAINT Score
CAF1/CCR4/NOT Complex							
PY17X_1237700	CCR4-1	127|112|22	N/A	N/A	0|0|0	0|0|0	N/A
PY17X_1027900	NOT family protein putative	97|132|19	0|0|0	0	1|3|0	0|0|0	0.4873
PY17X_0945600	NOT1	50|46|7	0|0|0	0	1|2|3	0|0|0	0.2637
PY17X_1428300	CAF1	47|58|11	0|0|0	0	26|19|17	N/A	N/A
PY17X_1108300	CAF40	17|20|7	0|0|0	0.0004	0|2|1	0|0|0	0.3506
PY17X_0921700	NOT2	8|12|3	0|0|0	0.002	0|0|0	0|0|0	N/A
PY17X_1207500	NOT5	7|2|0	0|0|0	0.2196	0|0|0	0|0|0	N/A
PY17X_1452400	NOT4	0|1|0	0|0|0	0.8736	0|0|0	0|0|0	N/A
Translational Repressors							
PY17X_1220900	DOZI	12|13|4	0|0|0	0.0011	3|9|4	1|2|7	0.4504
PY17X_1035100	CELF2	17|21|5	1|0|0	0.0037	1|12|9	0|0|6	0.2238
Stress Granule							
PY17X_0311400	CCT8	10|3|3	0|0|0	0.0131	3|19|15	1|0|8	0.2019
PY17X_0613400	HSP40 subfamily A	15|15|5	0|3|0	0.0368	2|17|12	0|1|6	0.2069
PY17X_1135600	CCT4	12|5|1	1|0|0	0.0735	0|7|6	0|0|2	0.2134
PY17X_0913800	TSN	7|10|1	1|0|0	0.0774	3|21|12	0|2|6	0.2097
PY17X_1221400	CCT5	4|2|1	0|0|0	0.0847	0|8|4	1|1|4	0.3536
PY17X_0822200	HSP70-2	17|8|5	3|0|0	0.0882	8|26|35	6|21|24	0.5228
Nuclear Transport							
PY17X_0307400	UAP56	9|4|6	0|0|0	0.0071	12|17|22	10|14|13	0.5413
PY17X_0403700	exportin-1	8|12|2	0|1|0	0.033	8|20|10	2|1|8	0.168
PY17X_1242000	karyopherin beta-3	39|64|16	6|6|0	0.0525	45|86|89	20|37|59	0.476
PY17X_1416100	exportin-1	3|4|2	0|0|0	0.0694	2|9|8	1|0|5	0.256
Translation							
PY17X_1209300	EIF3D	9|9|3	0|0|0	0.0028	2|18|15	2|4|8	0.2705
PY17X_0926900	EF-1 subunit alpha	4|6|3	0|0|0	0.0205	1|3|4	0|0|1	0.128
PY17X_0504300	eIF2A	15|33|2	0|0|0	0.026	0|6|4	0|0|0	0.3075
PY17X_0309700	peptide chain release factor subunit 1	3|2|2	0|0|0	0.057	2|3|4	0|1|0	0.0499
PY17X_1034300	eIF2 gamma	3|3|1	0|0|0	0.0612	1|12|8	0|2|6	0.279
PY17X_1369900	60S ribosomal protein L17	6|6|2	1|0|0	0.0653	1|9|4	2|2|6	0.4393
Other							
PY17X_0933200	HSP101	15|13|10	0|0|0	0.0002	14|25|22	17|6|21	0.8288
PY17X_0104900	putative anonymous antigen-1*	17|19|3	0|0|0	0.0044	1|15|10	0|0|4	0.1823
PY17X_0814600	Ran-binding protein	8|7|4	0|0|0	0.0058	7|13|14	4|8|8	0.4982
PY17X_1402200	cytoadherence linked asexual protein	12|8|3	1|0|0	0.0155	18|38|35	10|17|15	0.3512
PY17X_0109000	OAT	5|4|3	0|0|0	0.0131	8|12|7	2|1|8	0.3583
PY17X_1139200	glideosome-associated connector*	34|30|5	0|4|0	0.0181	30|70|55	19|15|33	0.3619
PY17X_1016400	coatamer protein beta subunit	5|9|2	0|0|0	0.0226	3|10|11	0|4|5	0.2887
PY17X_1322200	6-phosphogluconate dehydrogenase decarboxylating	6|8|1	0|0|0	0.0294	2|7|7	0|1|2	0.1208
PY17X_1313800	M17 leucyl aminopeptidase putative	7|2|2	0|0|0	0.0405	4|8|8	5|2|9	0.5377
PY17X_0833500	RhopH2	18|12|13	1|3|0	0.0441	25|58|53	16|22|31	0.4018
PY17X_0418800	RhopH3	20|9|6	3|0|0	0.0479	12|46|34	10|14|18	0.3455
PY17X_0708300	SEC23	2|6|2	0|0|0	0.0811	13|11|12	5|4|12	0.5213
PY17X_0525700	tryptophan/threonine-rich antigen	2|2|2	0|0|0	0.0915	0|6|4	9|0|4	0.534
PY17X_1411400	meiosis-specific nuclear structural protein 1	2|2|3	0|0|0	0.0949	0|0|0	0|0|0	N/A
PY17X_0621600	putative hydrolase/p36 like protein*	5|1|1	0|0|0	0.0988	3|14|11	1|2|4	0.1966

Table 3.1: PyCCR4-1::GFP Associates with the CAF1/CCR4/NOT Complex while PyCAF1ΔC::GFP is Unable to Associate Strongly with the Complex. Identified proteins are categorized based upon known or putative functions. Those that are currently unannotated are marked with an asterisk and described with the name of the closest protein identified by BLASTp alignment and/or a Phyre2 search [168, 169]. Within each category, proteins are listed from lowest to highest SAINT score based on the PyCCR4-1::GFP data. The most stringent SAINT Score is unshaded, SAINT scores that fall between 0.1 and 0.35 are shaded in light gray and above 0.35 are shaded in dark gray. The strength of interactions in the other categories differs profoundly between the two immunoprecipitations. Average P is the average P value of all three biological replicates, and SAINT Score is SAINT's representation of FDR. A total number of experimental and control spectra (peptide spectral matches) for each replicate is also shown for comparison.

PyCCR4-1 is Important for the Development and Transmission of Male Gametocytes

As deadenylases are also known to act as translational regulators in specific and temporal manners, I investigated the role of all members of the CCR4 domain-containing protein family throughout the *P. yoelii* life cycle [104]. Utilizing BLASTp alignments, I identified four high-confidence CCR4 domain-containing proteins in *Plasmodium* spp. that all have homology to deadenylases in other model eukaryotes (e.g., yeast, human, mouse) (Figure 3.4B). The domain architecture of CCR4-like proteins involves a Leucine-Rich Repeat Region (LRR) and an Endonuclease/Exonuclease/Phosphatase (EEP) domain. The LRR mediates the interaction of CCR4 with CAF1 and the rest of the NOT complex, while the EEP domain contains active site residues required for deadenylation activity [162]. Of these, I found that the EEP domain of PyCCR4-1 aligns most closely with the consensus CCR4 domain-containing proteins from model eukaryotes. However, beyond the CCR4-EEP domain, there is no significant homology between other regions from PyCCR4-1, 2, 3, and 4 to each other, or to homologs from model species (Figure 3.4B) [170]. While a canonical LRR was not bioinformatically detectable in any of the CCR4 domain-containing proteins, crosslinking immunoprecipitation of PyCCR4-1::GFP demonstrated that it retains the ability to associate with its complex (Table 3.1).

As our recent RNA-seq data from *P. yoelii* shows that all four genes are expressed in asexual blood stages and gametocytes [171], I sought to determine if any of the CCR4 domain-containing proteins played an important, stage-specific role in the parasite life cycle. To this end, I replaced coding sequences of all four CCR4-domain containing proteins with a GFP-expression cassette and a human dihydrofolate reductase (HsDHFR)-expression cassette via double homologous recombination in the *P. yoelii* 17XNL strain (Figure 3.5). These lines were cloned via limiting dilution before characterization, and their transgenic genotypes were confirmed using PCR across both homology regions. These clonal parasites revealed that deletion of any one of these genes individually was not lethal in asexual blood stages.

As CCR4 proteins are known to play specialized roles in eukaryotes, transgenic parasites lacking one of the four genes (*pyccr4-1, pyccr4-2, pyccr4-3,* and *pyccr4-4*) were also compared to wild-type parasites in all stages of the parasite's life cycle. Phenotypic observations were performed by measuring parasite numbers in mice, the prevalence of mosquito infection, and developmental timing/completion throughout the *Plasmodium* life cycle (Figure 3.6ABCD, Table 3.2). Deletion of *pyccr4-2, pyccr4-3,* or *pyccr4-4* resulted in transgenic parasites that behaved as wild-type in all life cycle stages (Table 3.2).

However, while deletion of *pyccr4-1* did not affect asexual blood stage growth (Figure 3.6A), it led to significant defects in male gametocyte maturation and host-to-vector transmission (Figure 3.6BCDE). First, to assess gametocytogenesis and the number of mature male gametocytes present, I developed an antibody-based flow cytometry assay based in part upon the effective reporter system (820cl1m1cl1) commonly used in *P. berghei* [36]. Allen Minns and I generated antibodies against a recombinant domain variant ("PyDDD" = AA1845-2334) of dynein heavy chain delta (PyDD, PY17X_0418900), together with anti-PvBiP antibodies to counterstain RBCs containing a parasite. I then confirmed by flow cytometry and Giemsa staining that PyDD is a marker for mature male gametocytes in *P. yoelii*, as it is in *P. berghei.* The antibody was further validated using a transgenic parasite line with a PyDDprom::GFPmut2 cassette integrated into the p230p safe harbor locus, where the population positive for both anti-

PyDD and anti-GFP signals were approximately 80% overlapped (Figure 3.6F). This assay allows for the measurement of male gametocytes (PyDD+) as one group and then pools immature gametocytes and female gametocytes(PyDD-) together. Antibody staining against BiP allows for the seperation of infected and uninfected red blood cells and for gametocytemia counts. This flow cytometric assay is an improvement over the fluorescent protein system of *P. berghei* because this approach allows these measurements to be done without the need to conduct reverse genetics in a base fluorescent reporter line and frees GFP and RFP for other purposes.

Using this flow cytometric method, I found that transgenic *pyccr4-1* deletion parasites produce fewer mature male gametocytes compared to wild-type parasites (Figure 3.6BC). Additionally, there is an increase in the number of immature/female gametocytes that do not express PyDD, which in combination with a decrease in male gametocytes, results in an altered male/female gametocyte ratio. Secondly, in contrast to wild-type parasites that have a semi-synchronous wave of gametocyte development, *pyccr4-1* gene deletion parasites lose this coordination and instead develop fewer male gametocytes that can form gametes, and do so in an asynchronous manner (Figure 3.6D).

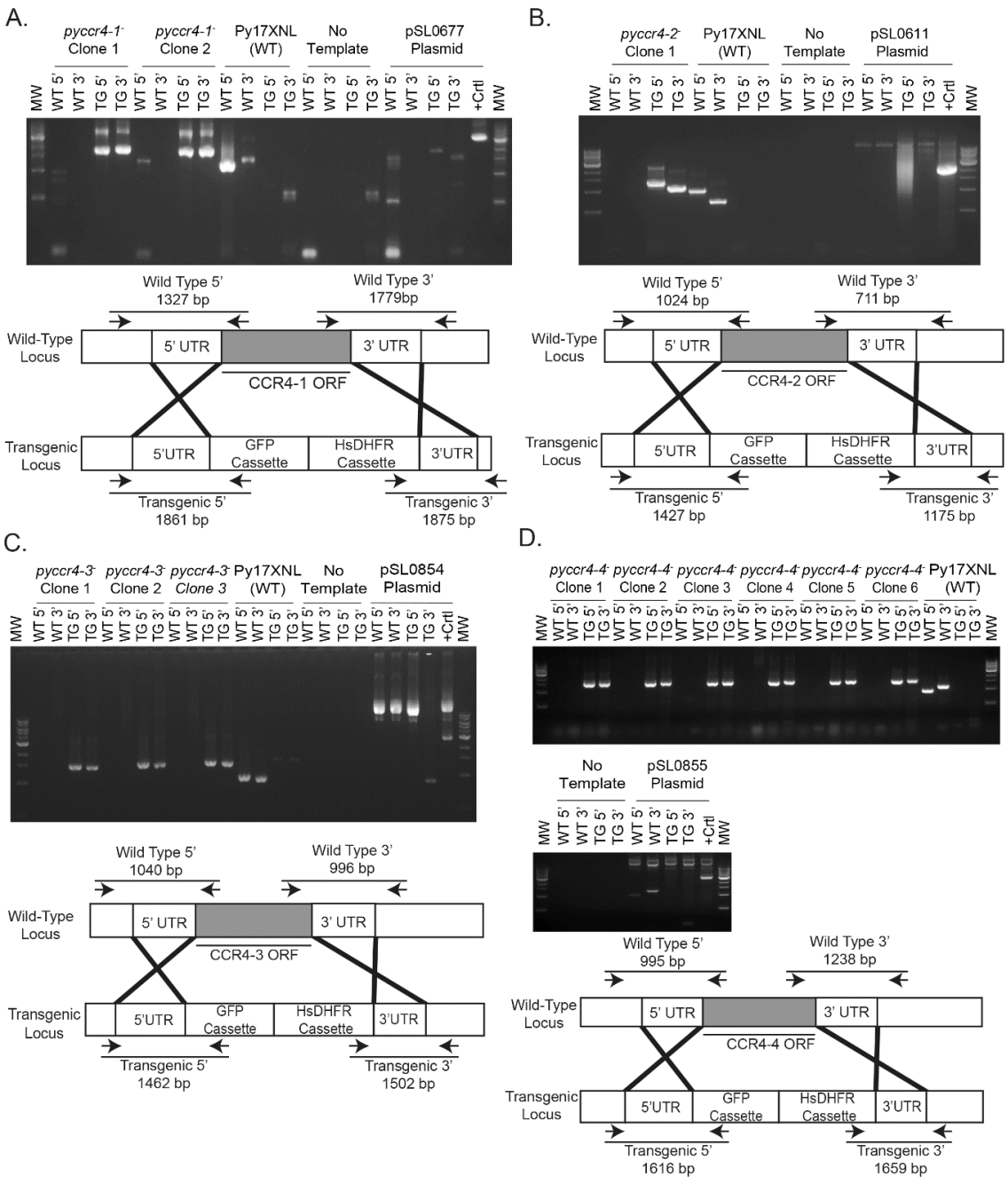
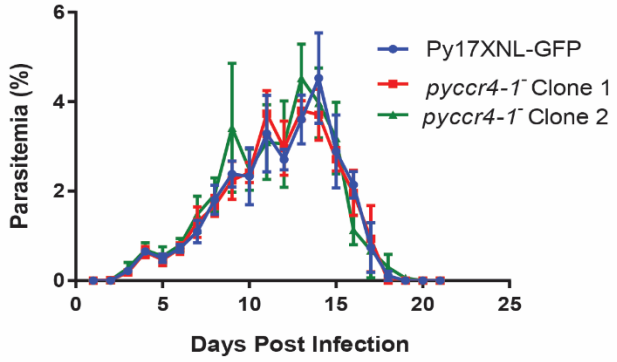
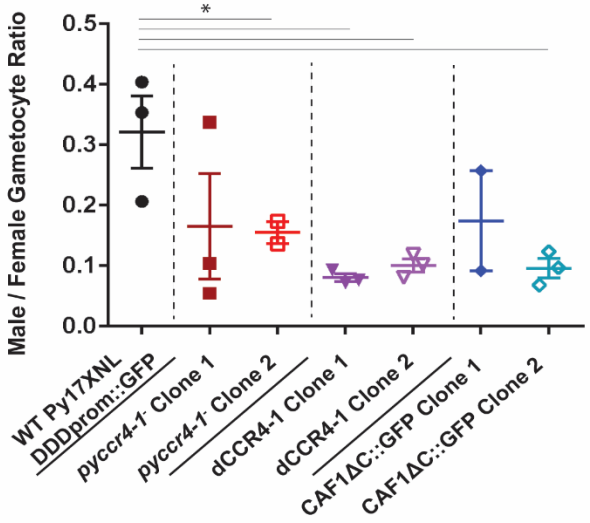


Figure 3.5: Genotyping PCR of (A) *pyccr4-1*⁻ , (B) *pyccr4-2*⁻, (C) *pyccr4-3*⁻, and (D) *pyccr4-4*⁻ Transgenic Parasites. Successful gene deletions were created using double homologous recombination of the targeting sequence consisting of ~750 bp on either side of the ORF. Genotyping was performed by PCR on parasites cloned by limiting dilution using the primers indicated (listed in Appendix A). Independent clones were compared to a Py17XNL wild-type control, a no template control, and a plasmid positive control in parallel. pSL0677 plasmid positive control size is 3270 bp. pSL0611 plasmid positive control size is 2421 bp. pSL0854 plasmid positive control size is 2694 bp. pSL0855 plasmid positive control size is 2845 bp. Molecular weight ladder is from NEB (N3232L).

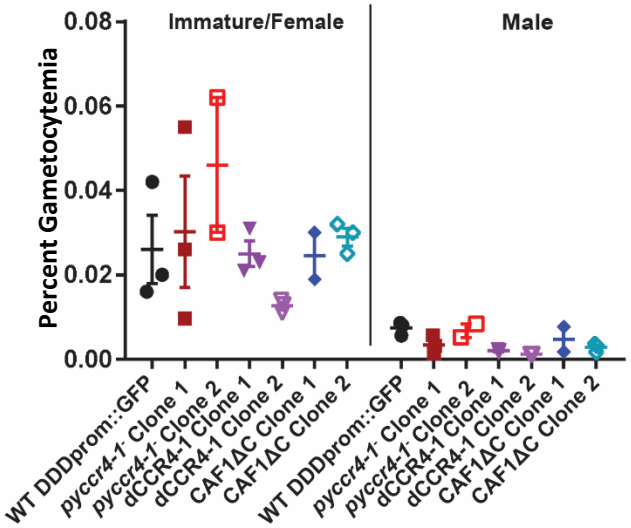
A.


Parasitemia (%)
Days Post Infection
Py17XNL-GFP
pyccr4-1- Clone 1
pyccr4-1- Clone 2

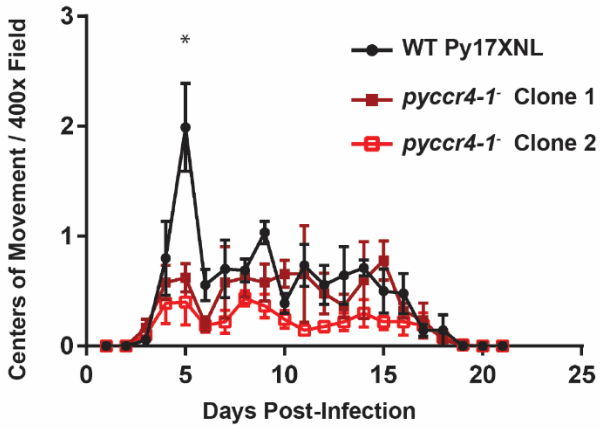
B.


Male / Female Gametocyte Ratio
*
WT Py17XNL DDDprom::GFP
pyccr4-1- Clone 1
pyccr4-1- Clone 2
dCCR4-1 Clone 1
dCCR4-1 Clone 2
CAF1ΔC::GFP Clone 1
CAF1ΔC::GFP Clone 2

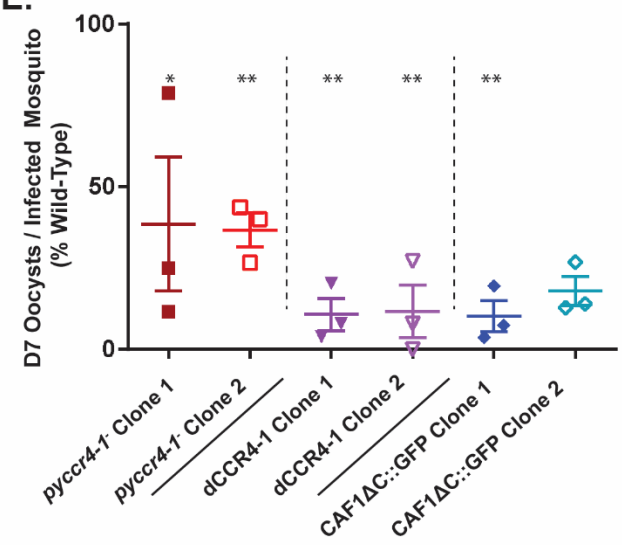
C.


Percent Gametocytemia
Immature/Female
Male
WT DDDprom::GFP
pyccr4-1- Clone 1
pyccr4-1- Clone 2
dCCR4-1 Clone 1
dCCR4-1 Clone 2
CAF1ΔC Clone 1
CAF1ΔC Clone 2

D.


Centers of Movement / 400x Field
Days Post-Infection
WT Py17XNL
pyccr4-1- Clone 1
pyccr4-1- Clone 2
*

E.


D7 Oocysts / Infected Mosquito (% Wild-Type)
pyccr4-1- Clone 1
pyccr4-1- Clone 2
dCCR4-1 Clone 1
dCCR4-1 Clone 2
CAF1ΔC::GFP Clone 1
CAF1ΔC::GFP Clone 2

F.

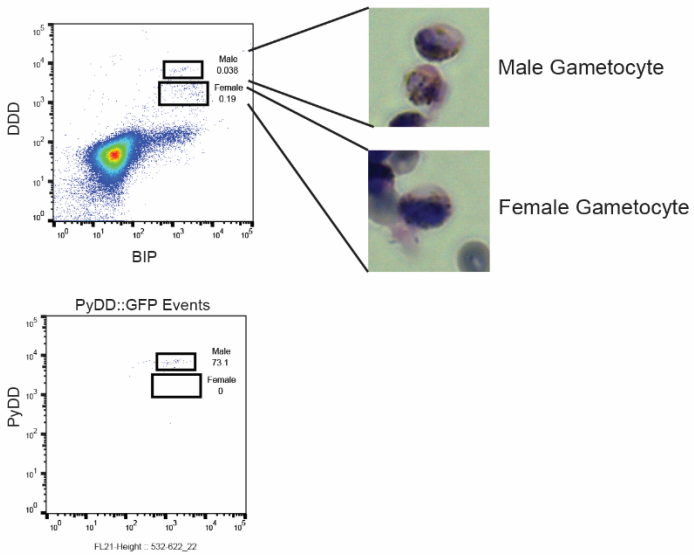

DDD
BIP
Male
0.038
Female
0.19
Male Gametocyte
Female Gametocyte
PyDD::GFP Events
PyDD
Male
73.1
Female
0

Figure 3.6: **Phenotyping of *ccr4-1*$^{-}$ Transgenic Parasites** A) Asexual blood stage growth was monitored for two *pyccr4-1*$^{-}$ transgenic clonal lines compared to a WT-GFP control line over the entire course of an infection. No significant difference in growth kinetics was observed. B) The number of mature male gametocytes was determined by flow cytometry of sulfadiazine-treated and DD/BIP-stained *P. yoelii* parasites. C) Raw gametocytemia values are plotted. Error bars are standard error of the mean. D) The number of centers-of-movement/exflagellation centers were quantified daily by microscopy on a 400x field upon injection with 10,000 blood stage parasites. Plotted are three biological replicates with three technical replicates each. Error bars represent the standard error of the mean. E) The number of oocysts per infected mosquito on day seven post-infectious blood meal are plotted. Data represent at least 20 dissected mosquitoes per biological replicate conducted in triplicate. Error bars represent the standard error of the mean. F) Gametocyte counts were performed using flow cytometry. Asexual stage parasites were killed with two days of sulfadiazine treatment and WBC's were removed using a cellulose column. PyDD high and BIP + cells were scored as mature male gametocytes, and PyDD mid, and BIP+ cells were scored as immature gametocytes and female gametocytes. No red blood cells were excluded in this analysis, and thus permitted measurement of gametocytemia. A PyDD promoter driving GFP was used to establish gating of mature male gametocytes. PyDD+ cells were FACS selected and observed to be male gametocytes by Giemsa staining to determine their identity and could undergo gametogenesis (exflagellation assay). * Significant below P=0.05, ** significant below P=0.01

	Centers of Movement			D7 Oocyst Prevelance		
ccr4-1	Wild-Type	*ccr4-1⁻ Clone 1*	*ccr4-1⁻ Clone 2*	Wild-Type	*ccr4-1⁻ Clone 1*	*ccr4-1⁻ Clone 2*
Average	1.22	0.42	0.35	0.96	0.78	0.69
SEM	0.11	0.11	0.18	0.00	0.05	0.09
ccr4-2	Wild-Type	*ccr4-2⁻ Clone 1*	*ccr4-2⁻ Clone 2*	Wild-Type	*ccr4-2⁻ Clone 1*	*ccr4-2⁻ Clone 2*
Average	1.45	1.25	1.37	0.92	0.89	0.96
SEM	0.27	0.13	0.22	0.03	0.04	0.00
ccr4-3	Wild-Type	*ccr4-3⁻ Clone 1*	*ccr4-3⁻ Clone 2*	Wild-Type	*ccr4-3⁻ Clone 1*	*ccr4-3⁻ Clone 2*
Average	1.15	1.60	1.32	0.93	0.95	0.85
SEM	0.20	0.41	0.18	0.03	0.03	0.02
ccr4-4	Wild-Type	*ccr4-4⁻ Clone 1*	*ccr4-4⁻ Clone 2*	Wild-Type	*ccr4-4⁻ Clone 1*	*ccr4-4⁻ Clone 2*
Average	1.53	1.17	1.57	0.97	0.92	0.93
SEM	0.20	0.27	0.32	0.02	0.04	0.03

	D7 Oocysts / Infected Mosquito			D10 Oocyst Sporozoites / Infected Mosquito		
ccr4-1	Wild-Type	*ccr4-1⁻ Clone 1*	*ccr4-1⁻ Clone 2*	Wild-Type	*ccr4-1⁻ Clone 1*	*ccr4-1⁻ Clone 2*
Average	84.99	32.43	31.14	62083.33	19055.33	23786.00
SEM	0.76	13.92	3.34	15081.76	1409.86	7863.93
ccr4-2	Wild-Type	*ccr4-2⁻ Clone 1*	*ccr4-2⁻ Clone 2*	Wild-Type	*ccr4-2⁻ Clone 1*	*ccr4-2⁻ Clone 2*
Average	91.07	80.60	97.22	56009.09	56941.43	51805.56
SEM	7.09	4.44	9.19	3078.89	9334.98	7644.15
ccr4-3	Wild-Type	*ccr4-3⁻ Clone 1*	*ccr4-3⁻ Clone 2*	Wild-Type	*ccr4-3⁻ Clone 1*	*ccr4-3⁻ Clone 2*
Average	57.86	75.51	42.98	31208.83	43528.28	35620.91
SEM	11.51	12.96	13.16	2570.33	11487.64	16010.21
ccr4-4	Wild-Type	*ccr4-4⁻ Clone 1*	*ccr4-4⁻ Clone 2*	Wild-Type	*ccr4-4⁻ Clone 1*	*ccr4-4⁻ Clone 2*
Average	140.27	115.66	81.79	47713.00	55307.50	43344.83
SEM	15.17	16.89	19.02	3792.14	9345.86	11252.48

	D14 Salivary Gland Sporozoites / Infected Mosquito			% of Mice Blood Stage Patent Within 3 Days		
ccr4-1	Wild-Type	*ccr4-1⁻ Clone 1*	*ccr4-1⁻ Clone 2*	Wild-Type	*ccr4-1⁻ Clone 1*	*ccr4-1⁻ Clone 2*
Average	32409.72	11845.69	6911.44	100	100	100
SEM	9961.09	1696.13	2067.79			
ccr4-2	Wild-Type	*ccr4-2⁻ Clone 1*	*ccr4-2⁻ Clone 2*	Wild-Type	*ccr4-2⁻ Clone 1*	*ccr4-2⁻ Clone 2*
Average	15672.73	15506.64	16259.06	100	100	100
SEM	1807.31	1266.41	4776.30			
ccr4-3	Wild-Type	*ccr4-3⁻ Clone 1*	*ccr4-3⁻ Clone 2*	Wild-Type	*ccr4-3⁻ Clone 1*	*ccr4-3⁻ Clone 2*
Average	13561.86	17847.98	13124.03	100	100	100
SEM	278.81	2617.91	947.14			
ccr4-4	Wild-Type	*ccr4-4⁻ Clone 1*	*ccr4-4⁻ Clone 2*	Wild-Type	*ccr4-4⁻ Clone 1*	*ccr4-4⁻ Clone 2*
Average	27094.20	21505.45	13018.97	100	100	100
SEM	7258.13	3620.55	1761.26			

Table 3.2 The Transmission of parasites with deletions of ccr4 Domain-Containing genes.

Measurements of transmission-related phenomena for each *pyccr4-x⁻* clone is shown as averages for each replicate. Centers-of-movement/exflagellation centers are shown as an average of the number of exflagellating male gametocytess per field in ten 400x fields. Prevalence of mosquito infections, oocyst counts, and sporozoite counts are all averages from at least 20 mosquitoes per replicate conducted in biological triplicate.

The Putative Catalytic Residues of PyCCR4-1 are Required for its Roles in Gametocytogenesis and Transmission

CCR4 proteins have well defined, conserved catalytic residues in model eukaryotes that are also conserved in *Plasmodium* species (Figure 3.4B) [172]. I showed that CCR4-1 plays a role in male gametocyte activation and host-to-vector transmission. To determine if the putative active site residues contribute to PyCCR4-1's functions in male gametocytes, I created transgenic parasites with alanine substituted for two of the putative catalytic residues (D1852A, H1898A) of PyCCR4-1 (dCCR4-1). Like *pyccr4-1* gene deletion parasites, dCCR4-1 transgenic parasites also produce fewer mature male gametocytes, and also lacked a synchronous wave of male activation (Figure 3.6BC, 3.7B).

Both male and female gametocytes are required for productive host-to-vector transmission. Because some male gametocytes retained the ability to mature and exflagellate in both the *pyccr4-1* gene deletion and dCCR4-1 lines, I assessed whether they were transmissible to mosquitoes. In both transgenic lines, I observed a corresponding decrease of similar scale in the number of day seven oocysts compared to wild-type parasites when transmitted to *Anopheles stephensi* on the peak day of male gametocyte activation into gametes (in this case day five post-infection) (Figure 3.6E). Moreover, although there is no statistical difference in the number of male gametocytes that can activate between wild-type and *pyccr4-1* gene deletion parasites after the peak day (Figure 3.6D, days six and beyond), a significant decrease ($p<0.05$) in the number of oocysts in the mosquito was still observed when parasites were transmitted two days post-peak. This data suggests that a defect beyond male activation, potentially in female gametocyte development, is also present in *ccr4-1* deletion parasites that prevent efficient host-to-vector transmission. These data indicate that the catalytic residues of PyCCR4-1 are required for normal male gametocyte development and host-to-vector transmission. Phenotypes are stronger in this mutant and that could be because other CCR4-domain containing proteins can bind to and make up for some of the defect in knock-out parasites, but in this case the dCCR4-1 could prevent the others from interacting by binding to CAF1 instead.

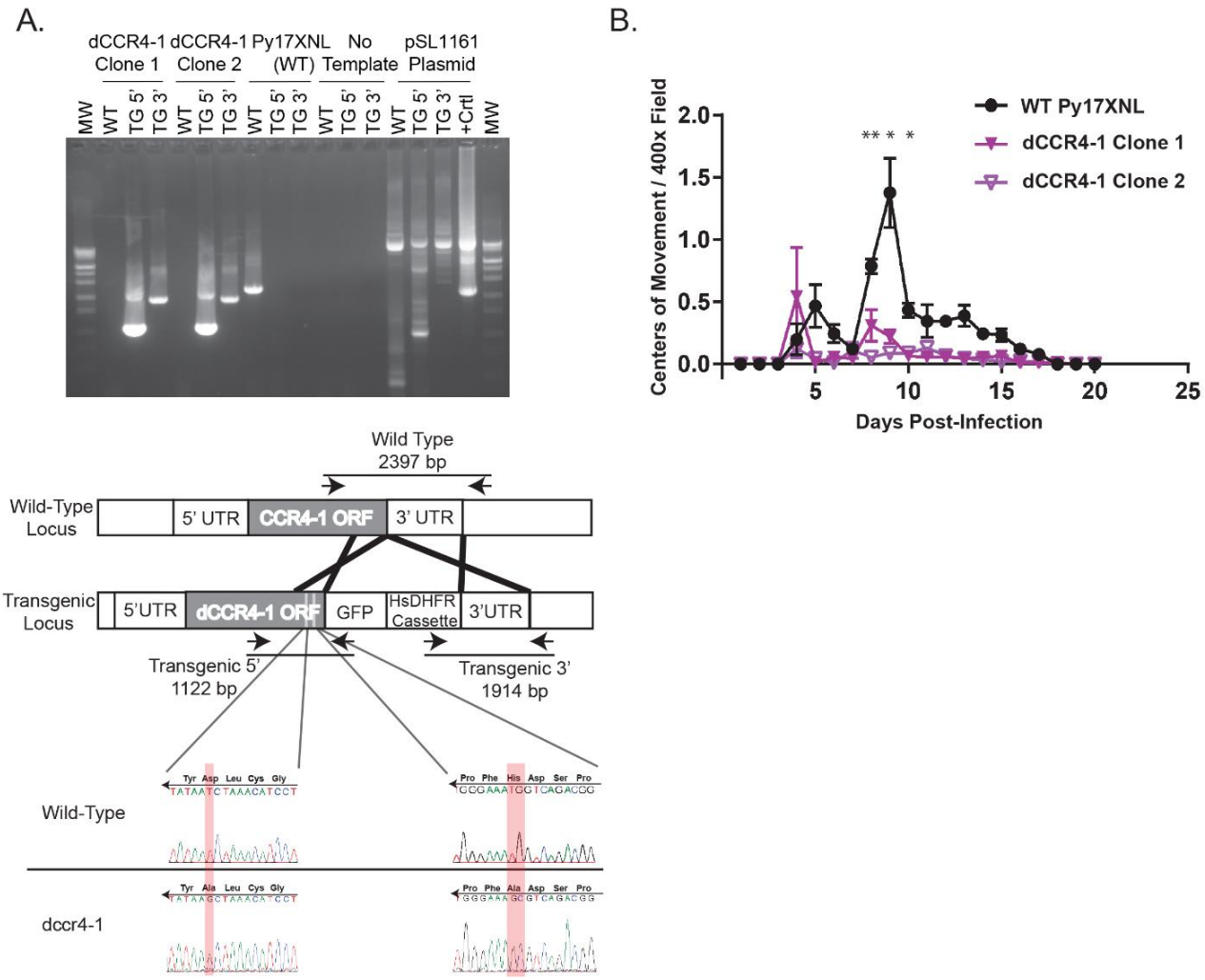


Figure 3.7: dCCR4-1 Transgenic Parasites Phenocopy *pyccr4-1* Gene Deletion Parasites. A) The creation of dCCR4-1 transgenic parasites is confirmed by genotyping PCR. A successful replacement of the PyCCR4-1 catalytic residues were created using double homologous recombination to insert a C-terminal GFP tag and stop codon following the PyCCR4-1 coding sequence but before the stop codon. Genotyping was performed by PCR on parasites cloned by limiting dilution cloning using the primers indicated (listed in S7 Table). Independent clones were analyzed with a Py17XNL wild-type control, a no template control, and a plasmid positive control in parallel. Sequencing results are shown demonstrating the appropriate base change has occurred. These base changes mutates the putative active residues to alanine residues. B) The number of centers-of-movement/exflagellation centers were quantified daily by microscopy on a 400x field upon injection with 1,000 blood stage parasites. Fewer WT and dCCR4-1 parasites were injected to better preserve animal health over the course of the experiment in accordance with our IACUC protocol

Truncation of PyCAF1 Prevents Full Assembly of the CAF1/CCR4/NOT Complex and Phenocopies the Deletion of *pyccr4-1*

In model eukaryotes, CCR4 is able to associate with the CAF1/CCR4/NOT complex through CAF1. I sought to determine if the typical association of PyCCR4 with the rest of its complex was required for its role in gametocyte development and host-to-vector transmission. As CCR4 domain-containing proteins associate with the NOT1 scaffold of the CAF1/CCR4/NOT complex indirectly by binding CAF1 in model eukaryotes, gene deletion of *pycaf1* would theoretically dissociate CCR4-1 from its complex [162]. However, complete deletions of the *caf1* gene have been unsuccessful in both a conventional targeted attempt and in the *Plasmo*GEM broad-scale genetic screen in *P. berghei*, indicating that it is likely essential [137, 140]. I have also attempted to delete the *P. yoelii caf1* coding sequence entirely and similarly was unable to delete this gene (Figure 3.9A). Instead, as the insertion of the *piggyBac* transposon into the *P. falciparum caf1* gene occurred in the coding sequence downstream of the CAF1 domain, I hypothesized that this portion of PfCAF1 may still be expressed and may be necessary for proper asexual blood stage development [137]. In support of this hypothesis, work performed by John Adams lab demonstrated that the CAF1 domain was still transcribed up to the insertion site, but not after in the *P. falciparum* CAF1 disruptant line (PfCAF1ΔC) (Figure 3.8) [173].

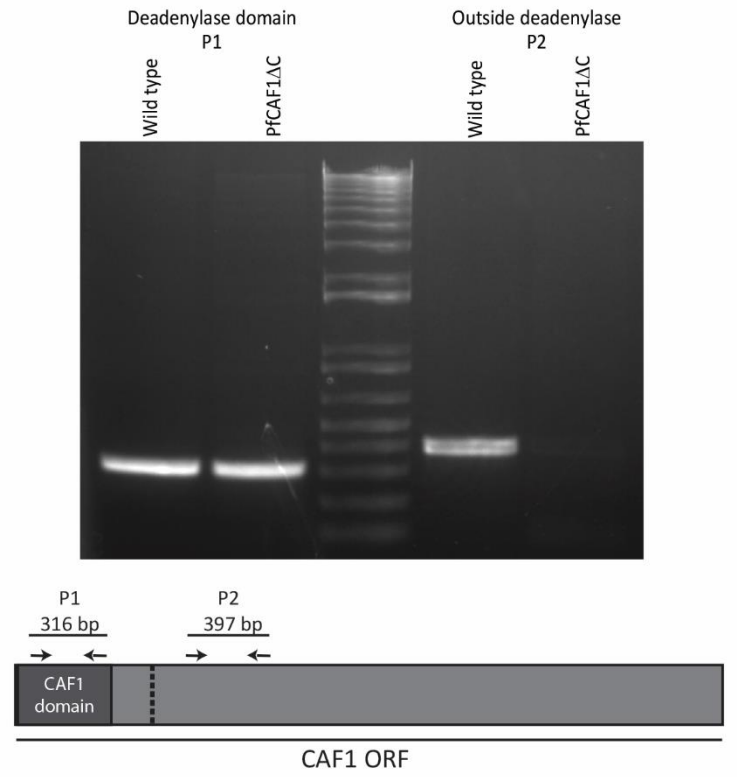


Figure 3.8: A *P. falciparum* line carrying a *piggyBac* transposon inserted after the CAF1 deadenylase domain makes a truncated transcript. Schematic of RT-PCR primers aligned to the CAF1 ORF. The site of the *piggyBac* disruption is indicated by dotted line. RT-PCR results indicate wild-type NF54 and PfCAF1D parasites both make transcript from the deadenylase domain. PfCAF1D parasites do not make a full-length transcript. Truncation made and experiment performed by the lab of John Adams at USF.

Based upon these expression data, I created a *P. yoelii* transgenic line that mimics this transposon insertion by inserting a C-terminal GFP tag and stop codon in the *P. yoelii caf1* gene in a comparable location following the CAF1 domain, thus creating a PyCAF1ΔC (AA 1-335) variant (Figure 3.9B) [137]. I found that expression of the PyCAF1ΔC::GFP variant resulted in viable parasites, but importantly, that these parasites exhibit a similar growth attenuation as was observed for the *P. falciparum* PfCAF1ΔC line (Figure 3.9C) [137]. To further assess the impact of the PyCAF1ΔC variant upon parasite growth and transmission, I observed comparable, but more pronounced, effects upon the activation of male exflagellation and parasite transmission as was seen with *pyccr4-1* deletion parasites (10-fold decrease in male activation on peak day and >4-fold reduction in transmission to mosquitoes, respectively) (Figure 3.6BCE, 3.9D). These exacerbated phenotypes may be caused by the combined effects of a reduction in total parasite numbers due to the deletion of portions of PyCAF1 and a PyCCR4-1-dependent defect in male gametocyte development. Additionally, CAF1 may play its own, uncharacterized roles in gametocyte development.

To determine if the PfCAF1ΔC variant in human-infectious *P. falciparum* similarly impairs gametocytogenesis as was seen in rodent-infectious *P. yoelii*, the PfCAF1ΔC piggyBac-insertion parasite line was also assessed for effects upon parasitemia, gametocytogenesis, as well as male gametocyte activation by John Adams' lab. They demonstrated that the PfCAF1ΔC line exhibited significant decreases in gametocyte conversion, total gametocytemia, and exflagellation on the peak day [173]. These data support the observed *P. yoelii* phenotype and indicate that this conserved complex is important for sexual development across *Plasmodium* species.

As I hypothesized that expression of only the CAF1 domain would prevent proper assembly of the CAF1/CCR4/NOT complex, I used IP-MS to determine if PyCAF1ΔC::GFP can still associate normally with the CAF1/CCR4/NOT complex. I found that when *pycaf1* is disrupted, the resulting PyCAF1ΔC protein does not associate with most of the components of the CAF1/CCR4/NOT complex, including PyCCR4-1 (Table 3.1). Utilizing the same stringent SAINT score (<0.1), I found that only two proteins (PyCAF1ΔC itself, and subunit one of peptide chain release factor) were detected. By expanding the SAINT threshold to 0.35, only 57 (41%) of PyCCR4-1::GFP's 139 protein interactions were detected. Importantly, I did not detect any association of PyCAF1ΔC::GFP with PyCCR4-1 (no peptide spectral matches), and observed only a greatly reduced association with NOT1. These data suggest that PyCAF1ΔC only weakly associates with other members of the CAF1/CCR4/NOT complex and it is no longer able to interact with PyCCR4-1. Full-length PyCAF1::GFP and PyCAF1ΔC::GFP were found localized in cytosolic puncta, further indicating that PyCAF1ΔC::GFP still partially associates with the complex (Figure 3.10AB).

These changes in complex assembly caused by expression of the truncated PyCAF1ΔC variant, which phenocopies the deletion of *pyccr4-1*, indicate several potential mechanisms for its function. First, only the CAF1 domain of PfCAF1 and PyCAF1 is essential for parasite viability. Second, while portions of PfCAF1 and PyCAF1 C-terminal to the CAF1 domain are dispensable, they are critical for CAF1 association with its complex, and essential for recruiting CCR4-1. Additionally, these data indirectly suggest that PyCCR4-1 requires its association with PyCAF1 and the NOT complex to

promote coordinated gametocytogenesis and efficient transmission to the mosquito. Finally, the deadenylase function of CAF1 itself may play an important role in gametocyte development, similarly to CCR4-1, and this disruption limits its activity.

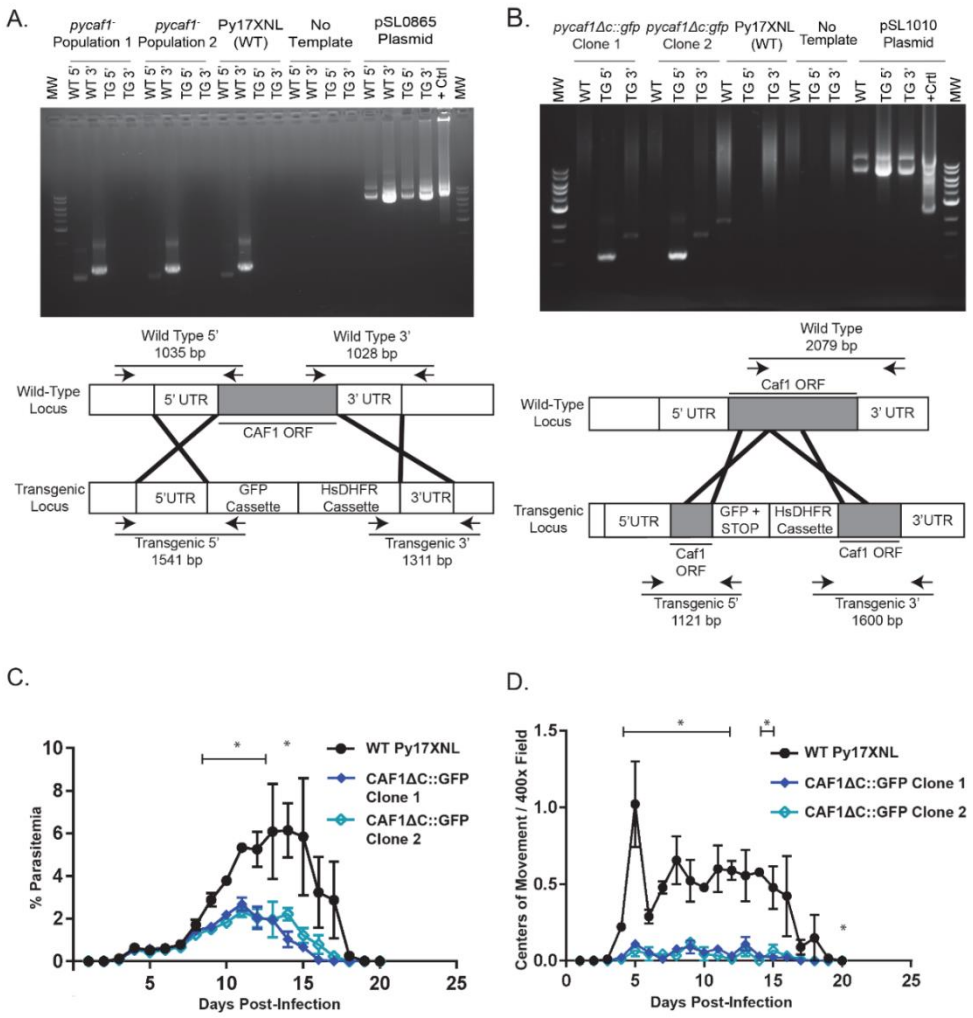
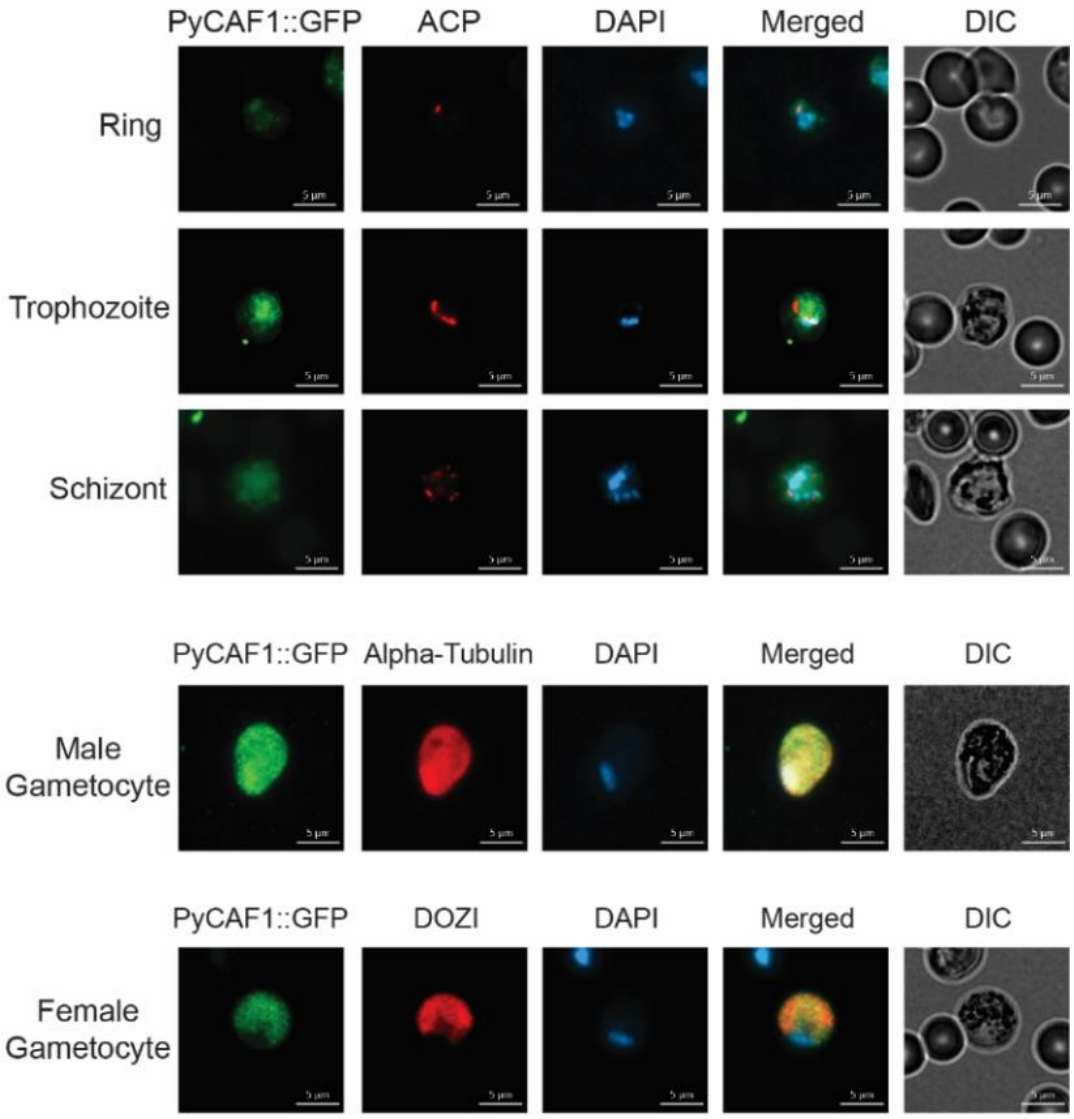


Figure 3.9: Phenotyping of PyCAF1ΔC::GFP Transgenic Parasites (A) Genotyping PCR of *pycaf1*- transgenic parasites. An unsuccessful attempt at gene deletion by double homologous recombination using targeting sequences consisting of ~750bp on either side of the ORF is depicted. Genotyping was performed by PCR on parasites subjected to two rounds of drug selection using the primers indicated (listed in Appendix A). These parasites were compared to a Py17XNL wild-type control, a no template control, and a plasmid positive control in parallel. pSL0865 plasmid positive control size is 2588 bp. (B) Genotyping PCR of *pycaf1* disruptant transgenic parasites is shown. Successful disruption of *pycaf1* was created using double homologous recombination to insert a C-terminal GFP tag and stop codon following the CAF1 domain. Genotyping was performed by PCR on parasites cloned by limiting dilution cloning using the primers indicated (listed in Appendix A). Independent clones were analyzed with a Py17XNL wild-type control, a no template control, and a plasmid positive control in parallel. pSL1010 plasmid positive control size is 2186 bp. Molecular weight ladder is from NEB (N3232L). C) Parasitemia was measured microscopically by giemsa-stained thin blood smears. Plotted are three biological replicates with three technical replicates each. Error bars represent the standard error of the mean. D) The number of centers-of-movement/exflagellation centers were quantified daily by microscopy on a 400x field upon injection with 10,000 blood stage parasites.

A.


PyCAF1::GFP
ACP
DAPI
Merged
DIC
Ring
Trophozoite
Schizont
PyCAF1::GFP
Alpha-Tubulin
DAPI
Merged
DIC
Male Gametocyte
PyCAF1::GFP
DOZI
DAPI
Merged
DIC
Female Gametocyte

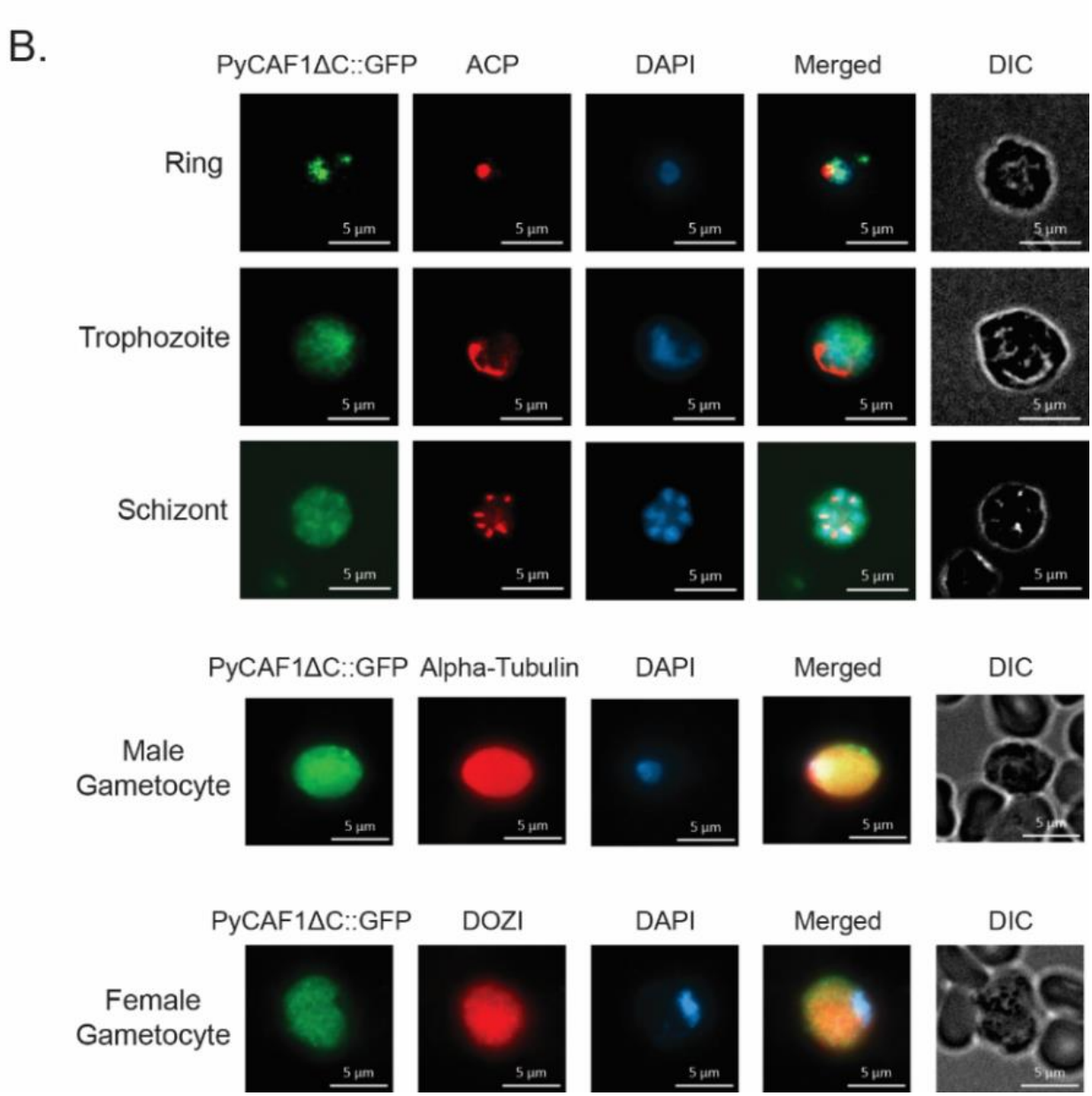


Figure 3.10: PyCAF1::GFP and PyCAF1ΔC::GFP are Expressed in Cytosolic Granules. (A) PyCAF1::GFP is diffusely expressed with some areas of higher intensity in asexual blood stages and cytosolic granules in gametocytes. Representative images of asexual blood stages and sexual stages treated with DAPI and antibodies to GFP (to detect PyCAF1::GFP) or to stage-specific cellular markers (ACP, alpha-tubulin, or human DDX6 that cross-reacts with DOZI) are shown. Scale bars are 5 microns. (B) PyCAF1ΔC::GFP is diffusely expressed in asexual blood stages and diffusely expressed with some areas of higher intensity in gametocytes. Representative images of asexual blood stages and sexual stages treated with DAPI and antibodies to GFP (to detect PyCAF1::GFP) or to stage-specific cellular markers (ACP, alpha-tubulin, or human DDX6 that cross-reacts with DOZI) are shown. Scale bars are 5 microns.

PyCCR4-1 Affects Important Gametocyte and Mosquito Stage Transcripts

Because PyCCR4-1 is a putative deadenylase that may act to degrade specific transcripts, I hypothesized that these phenotypes in male and possibly female gametocytes might be attributed to PyCCR4-1 affecting the abundance of specific transcripts important to gametocytogenesis, gamete activation, and/or parasite transmission to mosquitoes. To determine the role of PyCCR4-1 in the regulation of transcripts in gametocytes, total comparative RNA sequencing (RNA-seq) was performed. Gametocytes from a wild-type line expressing GFP from the p230p dispensable locus (WT-GFP) and the *pyccr4-1* gene deletion transgenic line were selected using sulfadiazine treatment, purified on an Accudenz gradient, and their RNA extracted for RNA-seq which was then sequenced on an Illumina HiSeq. Mike Walker assessed the differential abundance of transcripts via DEseq2, and I utilized the p-adjusted value for all analyses (Figure 3.11A, complete datasets are publically available, see Chapter 2 for data availability statement and accession numbers) [174].

Globally only 175 significantly affected transcripts (P-adjusted < 0.05, > 2 fold change) transcripts changed in abundance out of 5704 total transcripts detected, indicating that CCR4-1 may be acting specifically or that gene buffering is occurring to correct for the loss of CCR4-1. This gene buffering may allow for other transcriptional or RNA metabolic functions to partially make up for the loss of CCR4-1 function. Nearly all (172 of 175) of the significantly affected transcripts between WT-GFP and *pyccr4-1* gene deletion parasites decreased in abundance in the *pyccr4-1* gene deletion parasites, while only three transcripts increased in overall abundance. The transcripts that increased encode two antigenically variant proteins and a CoQ10 homolog. Many of these decreases in transcript abundance are for mRNAs that encode male-enriched proteins [175], and thus these changes can likely be attributed to the production of fewer mature male gametocytes in the *pyccr4-1*$^{-}$ transgenic line. However, the effect upon other transcripts indicates that a female gametocyte defect may be present as well (Figure 3.11B). Most notably, transcripts that encode proteins involved in gamete function (e.g., GEST) and early mosquito stage development (e.g., p28, CITH, AP2-O, HMGB2, LAP2) decreased in abundance significantly in the absence of PyCCR4-1 [36,

56, 175-179]. Moreover, many of these transcripts are known to be translationally repressed in *P. falciparum* female gametocytes [28]. Interestingly, an ApiAP2 protein (PY17X_1417400) that may be important for gene expression during gametocytogenesis decreased in abundance 10-fold in *pyccr4-1* gene deletion gametocytes [180]. Disruption of this ApiAP2 gene in *P. falciparum* by *piggyBac* transposon insertion showed dysregulation of transcript abundances and may indicate that it plays a role in gametocyte development [180]. Effects on this ApiAP2 may explain some of the decreased abundance of transcripts in the *ccr4-1* deletion. Other transcripts of interest that decreased in abundance are those that encode for multiple uncharacterized RNA-binding proteins (PY17X_1203900, PY17X_1457300, and PY17X_0923600), a second ApiAP2 protein (PY17X_1317000) and BDP2 (PY17X_1431000), and an uncharacterized putative transcriptional activator. Together, I conclude that these differences in transcript abundance cause a reduction in the number of mature male gametocytes, and also indicate that PyCCR4-1 may directly or indirectly act to preserve specific transcripts important for the gametocyte and early mosquito stage parasite.

Interestingly, transcripts that decrease in abundance between *pyccr4-1* gene deletion gametocytes and gametocytes with a catalytically dead CCR4-1 protein have different changes in transcript abundances (Figure 3.11C). There are only two transcripts that decrease in abundance in both parasite lines, both uncharacterized. One is a HORMA domain protein (PY17X_1409700) which may bind to chromatin and another is a serine/threonine protein kinase (PY17X_0905200). This may explain why the dCCR4-1 phenotypes are stronger than that of the gene deletion. If the other CCR4-domain containing proteins (CCR4-2, CCR4-3, or CCR4-4) are able to bind to CAF1 and modulate transcript abundances when CCR4-1 is absent then the dCCR4-1 protein would likely prevent this association. If a different CCR4-domain containing protein is allowed to bind and has a different program for initiating RNA decay or different specificity factors than this may cause these differences. However, this does not explain how the *ccr4-1* gene deletion and catalytic mutant result in similar phenotypes.

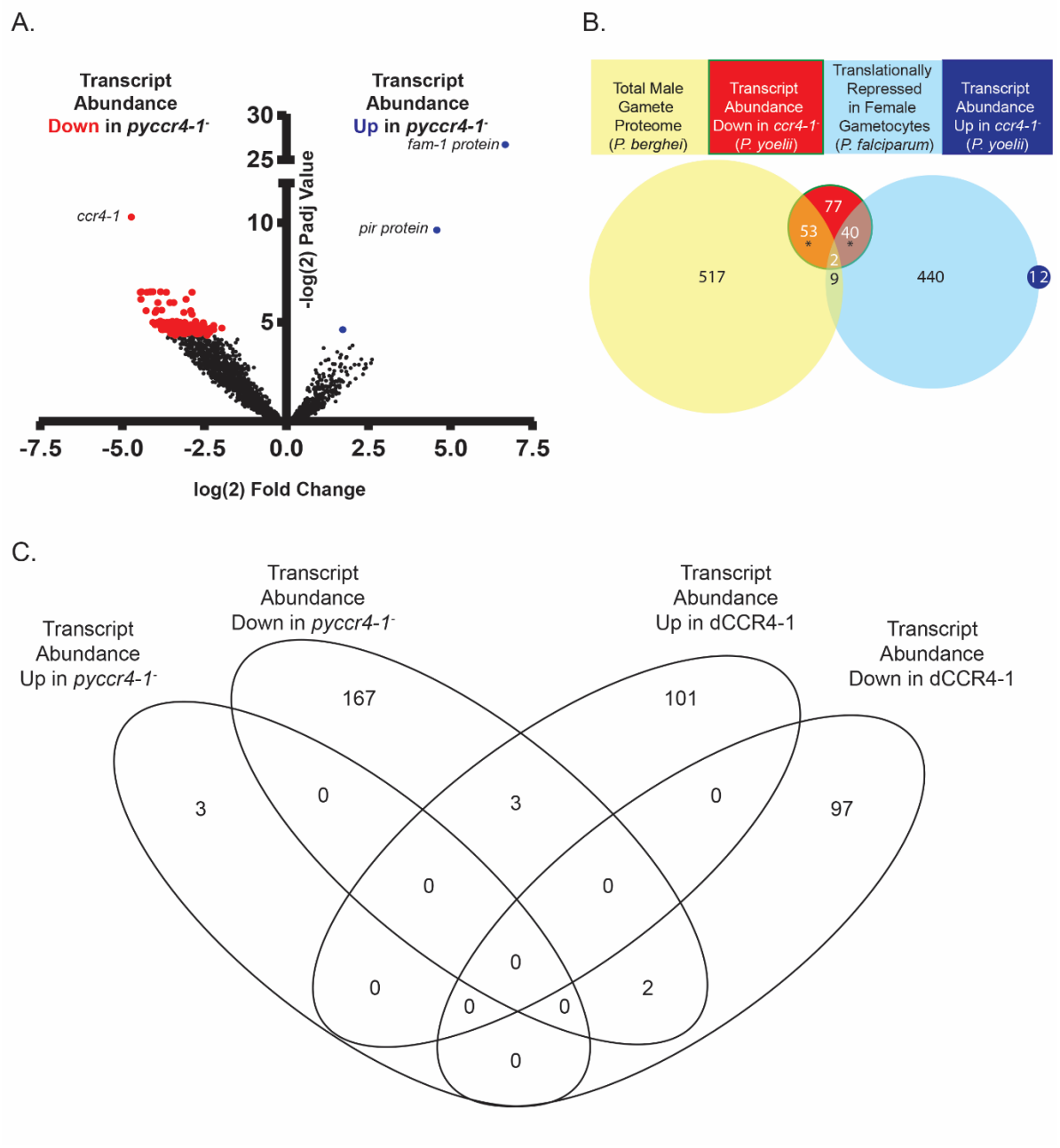


Figure 3.11: PyCCR4-1 Modulates Transcripts with Sex/Transmission-Related Functions. A) A volcano plot showing changes in transcript abundance up (blue), down (red) or unchanged (black). While few transcripts go up in abundance (from antigenically variant genes), nearly all affected transcripts decrease in abundance, thus indicating that PyCCR4-1 may play a role in preserving these mRNAs. A pseudo-count was added for *ccr4-1* transcript levels to visualize it on the plot because it was absent from the gene deletion transcriptomics as expected. B) In *pyccr4-1* gene deletion gametocytes, twenty-four percent of differentially abundant transcripts are translationally repressed in female gametocytes, and another one-third of the transcripts are enriched in the male gamete proteome [28, 175]. * = p<0.01 by Fisher test C. A comparison of *pyccr4-1* gene deletion gametocytes and catalytically dead PyCCR4-1

gametocytes (dCCR4) demonstrates that the transcripts that change in abundance are not shared between these two lines.

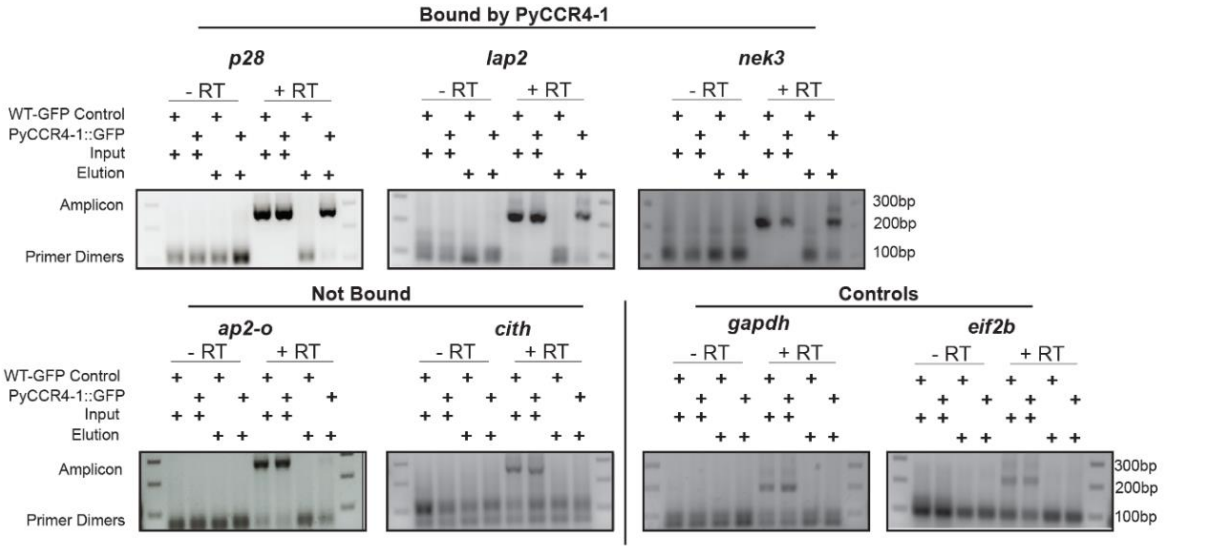
The CAF1/CCR4/NOT Complex Specifically Binds Transcripts that are Dysregulated in *pyccr4-1* Gene Deletion Parasites.

As transcript abundances could be affected directly by the binding of PyCCR4-1 and its complex, or indirectly through compensatory mechanisms such as gene buffering when the *pyccr4-1* gene is deleted [181], I determined if these dysregulated mRNAs were bound by the CAF1/CCR4/NOT complex using non-crosslinking immunoprecipitation. To this end, unfused GFP (expressed in WT-GFP parasites) and PyCCR4-1::GFP were immunoprecipitated from purified, transgenic gametocytes, and the association of co-precipitated transcripts was detected by RT-PCR. I found that PyCCR4-1::GFP interacted specifically with several selected transcripts that substantially decrease in abundance in *pyccr4-1* gene deletion parasites, including *p28* (PY17X_0515900), *lap2* (PY17X_1304300), and *nek3* (PY17X_0603200) (Figure 3.12A, top row). These transcripts are notable, as they are all important/essential for gametocytogenesis or transmission, and include transcripts known to be important to male (*nek3*) and/or female (*p28*, *lap2*) gametocytes [28, 182-185]. However, not all dysregulated transcripts (*cith* and *ap2-o*), nor unaffected housekeeping transcripts (*gapdh*, *eif2b*), were found specifically associated with PyCCR4-1 (Figure 3.12A, bottom row), suggesting that these gross changes in transcript abundance are likely the result of a combination of both direct and indirect effects.

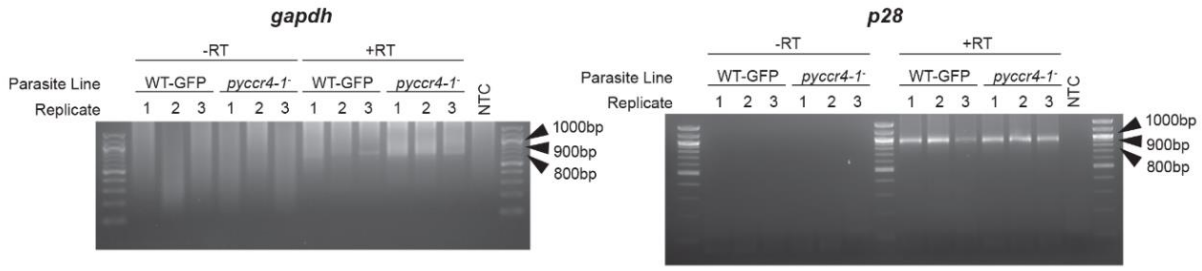
The direct effects that CCR4 can have on transcript abundance in model eukaryotes have resulted from deadenylation of a target transcript, or through translational repression by binding/tethering to its complex [129]. To investigate if the poly(A) tail length was impacted by *ccr4-1* gene deletion, Scott Lindner used circular RT-PCR (cRT-PCR) against both a control transcript (gapdh, not affected by *pyccr4-1* gene deletion, does not interact with PyCCR4-1) and an affected/bound transcript (p28). cRT-PCR allows one to sequence the 5'UTR, 3' UTR and poly (A) tail of a transcript by circularizing it and amplifying across the UTRs. This can be measured by running cRT-PCR products on a gel or by sequencing them. cRT-PCR demonstrated that, upon

pyccr4-1 deletion, there were no gross effects on UTR/poly(A) tail lengths of these transcripts using primers that anneal near the start and stop codons (Figure 3.12BC, oligonucleotides provided in Appendix A). I performed sequencing of cloned PCR products from both wild-type and *pyccr4-1* gene deletion samples that revealed the consistent composition of the 5' and 3'UTRs, as well as the presence of a poly(A) tail of the *p28* transcript (Figure 3.12D). However, none of these sequencing runs could resolve the precise length of the poly (A) tail because they could not read through the length of repeating adenine residues to the other UTR, but did permit design of primers near to the poly(A) tail to assess the distribution of poly(A) tail lengths in the population. Using these primers, we did not observe any differences in poly(A) length between the wild-type and *pyccr4-1*$^{-}$ populations (Figure 3.12C). These data indicate that the direct effect of PyCCR4-1 on these transcripts in gametocytes does not impact the poly (A) tail/UTR length of these control and bound/dysregulated transcripts suggesting that the complex may be acting indirectly to preserve these transcripts.

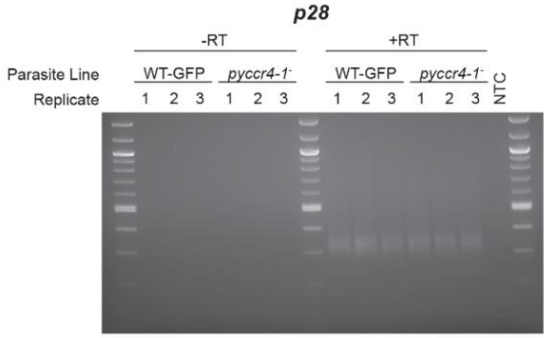
A.



B.



C.



D.

>*P. yoelii* p28 transcript, cRT-PCR product

GGTACAGGTAGTGGTACTGGAACACCAGCAAATAGTAGTATTATGAACGGAATGTCAATATTCAGCATTA
TTGCATTACTTGTTATTTATATAGCAGTAATGTAAtatatccaattgttatcgcatattgtaggaatatc
tatatagagagagacaaaaaaaaattaaaacagatttgacatttaaataattgatatttgtcgtggctaa
gtttacaagaatgaaataactatttttttttttttttgtctatttaaatatgttaaaaatatgtaaattt
ttttttccgtttaaagggaaaaggaatacataaaaatggcattaatatgatatacaatgaggagaacaag
agggaaataataaaataaaataaaaatatacatttaattgtatatatttttattatgatattatttttc
tcttccaacttgtatttcatttgttgttttcaatatttttttttgttttgttttttttttgttttgtttt
tttcattattcctttaagattaatgacttaaatttaactatattgaaaattaaaaaaaaaaaagtataaa
gtataaatagatttatAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA----------
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAttttgatttcactttatacatttttttaa
atttaacattatttatattctcataatttccgtaaaaaaaacaaaacaatttcataaaattatactcata
acaacagttatttaacaaaaatattttatactaaattttcacgaaaATGAATTTTAAATACAGTTTTAT
TTTTTTATTTTTTATCCAACTTGCGATAAGATATAATAATGCAAAAGTCACTGTAGACACG

Figure 3.12: The CAF1/CCR4/NOT complex associates with some dysregulated transcripts but doesn't affect UTR/poly(A) tail length of two select transcripts. A) Immunoprecipitation of PyCCR4-1::GFP allowed detection of the association of three selected transcripts that are affected by *pyccr4-1*$^{-}$ (top), whereas other affected transcripts do not associate with PyCCR4-1::GFP (bottom, left). Two control transcripts that do not change upon deletion of *pyccr4-1* also do not interact with PyCCR4-1::GFP (bottom right). Shown are input and elution samples of a constitutive GFP-expressing clone and a PyCCR4-1::GFP clone. Amplicons and primer dimer bands are indicated. B and C) Circularized RT-PCR (cRT-PCR) of *p28* was conducted in both wild-type and *pyccr4-1*$^{-}$ parasites to detect effects by PyCCR4-1 upon UTR and poly(A) tail length in purified gametocytes. Primers were designed within the coding sequences to define the 5' UTR, 3' UTR and poly(A) tail (B) or near the poly(A) tail to determine the length of the tail itself (C) as defined by sequencing of cRT-PCR products from panel B, which allow observations of UTR and poly(A) tail lengths respectively. Three biological replicates and a No Template Control (NTC) are shown with NEB 100 bp molecular weight ladder in parallel. D) Sanger sequencing of cRT-PCR products from the circularized p28 transcript using primers that anneal within the coding sequence (upper case) permitted identification of the 5' (red font, lower case) and 3' UTRs (blue font, lower case), as well as the poly(A) tail (black font, lowercase, underlined). Sequencing could not extend robustly through the poly(A) tail to provide an exact length from either forward or reverse sequencing primers (denoted by dashes). Sequences of the circularized gene product are provided from a cloned PCR product that is representative of UTRs from both wild-type and *pyccr4-1*$^{-}$ samples.

Discussion:

Specific post-transcriptional control of gene expression has been shown to be essential for *Plasmodium* transmission. Here I demonstrate that PyCCR4-1 of the CAF1/CCR4/NOT complex plays a role in the expression/degradation of transcripts that are important for *Plasmodium* transmission. I also demonstrate that when CCR4-1 is absent or inactive, many male gametocytes do not develop properly and parasites do not transmit to the mosquito vector efficiently. CAF1/CCR4/NOT also binds directly to some of the transcripts that are reduced in abundance when the *ccr4-1* gene is deleted, but we do not see any evidence of a decrease in poly (A) tail length, indicating that this effect may be poly (A) independent and independent of the CCR4-1 deadenylase activity. Previous work has demonstrated that deletion of *pbdoz*i (A dhh1 orthologue) had similar effects on transcript abundances as a deletion of PyCCR4-1 [14]. In yeast, CCR4 and DHH1 play roles as elongation factors and CAF1/CCR4/NOT was shown to bind directly to RNAPII [121, 186]. However, there is evidence that PyCCR4-1 is not acting this way in *Plasmodium*, as I did not detect this interaction with RNA polymerase

within either of our association significance thresholds by crosslinking immunoprecipitation. Instead in *Plasmodium,* DOZI can assist in preserving translationally repressed transcripts and allow them to be stabilized in gametocytes before host-to-vector transmission [14]. PyCCR4-1 may be acting similarly, as recent experiments in *Xenopus* and *Drosophila* have shown that 4E-T interacts with CAF1/CCR4/NOT and DDX6/DHH1 by creating a chain of proteins linking the 5' and 3' ends of a transcript, and can mediate deadenylation-independent translational repression by the complex [187, 188]. This is an interesting hypothesis of how CCR4-1 may play a role in the preservation of transcripts as I see in gametocytes. Currently, a 4E-T cannot be bioinformatically identified in *Plasmodium;* however, further work may uncover how CCR4-1 can promote the preservation of transcripts.

These analyses demonstrate that CCR4-1 plays a specific role in the development of male gametocytes, as well as the regulation of transcript abundance in gametocytes. Deletion of *pyccr4-1*, or expression of a variant with alanine substitutions of its putative catalytic residues, results in fewer mature male gametocytes and the loss of a coordinated wave of gametocyte development. These effects ultimately manifest as an overall reduction in transmission from host-to-vector. Moreover, the general CAF1 deadenylase also participates in these functions in both *P. yoelii* and *P. falciparum*, minimally because CCR4 typically associates with the complex through CAF1. Together, I find that both a specialized and a general deadenylase have been adapted as specific regulators of *Plasmodium* host-to-vector transmission through their concerted effects upon specific, critical gametocyte transcripts.

Chapter 4: NOT1-G IS A NOVEL REGULATOR OF GAMETOCYTE DEVELOPMENT IN *PLASMODIUM YOELII*

Introduction

The CAF1/CCR4/NOT complex typically consists of nine conserved members. I have shown previously that the complex assembles six of the seven bioinformatically predictable members in *P. yoelii* (Table 3.1, Chapter 3, [173]). While the function of NOT1 is well conserved, it has not yet been defined in *Plasmodium*. NOT1 has been described in other eukaryotes to act as a scaffold, onto which the rest of the CAF1/CCR4/NOT complex members and other proteins including RNA-binding proteins assemble [104]. Once assembled, the proteins can act on their mRNA targets. The NOT1 protein has been shown to be an essential component of P-bodies, which are major sites of mRNA processing. When NOT1 is depleted, these cytosolic granules dissolve [134]. NOT1 also allows for the CAF1 and CCR4 components to assemble with RNA-binding proteins to function as deadenylases and/or translational repressors [133, 189]. Therefore, NOT1 is a protein in model eukaryotes that plays an essential role in RNA metabolism and may be playing similar roles in *Plasmodium* species.

In most eukaryotes there is only a single *not1* gene; however, I identified two *not1* genes in *Plasmodium* through BlastP searches against *S. cerevisiae*, *Drosophila*, and *Homo sapien* CAF1/CCR4/NOT complex core components. RNA metabolism is utilized extensively to control proper development of *Plasmodium* transmission stages and the stages following transmission and thus may be why *Plasmodium* has duplicated this gene. This duplication is only seen in the Aconoidasida class, which include *Plasmodium* species, and has not been seen any other sequenced eukaryotes. These proteins, which I have named PyNOT1 and PyNOT1-G, have only been studied previously through bioinformatic analyses and in forward genetic screens [64, 139, 141]. These studies are immensely valuable; however, care needs to be taken in utilizing the data, and independent verification is always recommended. Because the forward genetic screens were done with pooled parasites, mutations that result in a substantial growth defect may be missed. Additionally modifications like transposon insertions that do not entirely prevent expression may be missed. The *plasmo*GEM study *in P. berghei* targeted PbNOT1, but not PbNOT1-G, and PbNOT1 was determined to be essential [140, 141]. The *piggyBac* transposon mutagenesis of *P. falciparum* resulted in insertions

into both of these genes, and indicates that neither was essential. However, PfNOT1 had a growth defect in competition assays [139]. This disparity could be explained by differences in human and rodent infectious species, that these genes may be required *in vivo* or *in vitro* but not in both, or by the method of disruption.

I found that neither NOT1 nor NOT1-G alone are essential for asexual parasite development in *P. yoelii;* however, parasites deficient in either of these genes exhibited severe defects at different stages during the life cycle. I focus here on PyNOT1-G as PyNOT1 was challenging to delete genetically and grew slowly. Asexual parasites with a *pynot1-g* gene deletion have a mild growth defect, indicating it plays some role in asexual blood stage development. I found that the primary defects in parasites with a *pynot1-g* gene deletion are in proper male and female gametocyte development. Few mature male gametocytes were produced, and all of the produced female or immature gametocytes were unable to be productively transmitted. Parasites with a *pynot1-g* gene deletion, do not transmit to mosquitoes even when crossed with wild-type PY17XNL parasites. In summary NOT1-G is essential for proper gametocyte maturation and may play roles in gametocyte commitment.

not1 is Duplicated in all Parasites in the Aconoidasida Class

In most eukaryotes, only one copy of a *not1* gene is encoded in the genome; however, *Plasmodium* and closely related parasites are different. Through bioinformatics analysis, I identified that all sequenced species in the Aconoidasida class, but no other sequenced species, encode two NOT1 proteins using BLASTp searches. While expansions in members of the CAF1/CCR4/NOT complex are not uncommon in eukaryotes (e.g., *ccr4*, *caf1*) an expansion in *not1* has not been described previously and may allude to parasite-specific functions related to its biology. NOT1 proteins typically contain a tristetraprolin-binding domain (TTP binding domain), a CAF1-binding domain, a domain of unknown function (DUF3819), and a NOT1 Domain (Figure 4.1). The composition of domains is conserved between PyNOT1-G (PY17X_0945600) and a typical model eukaryote NOT1 (*S. cerevisiae* shown in Figure 4.1). However, PyNOT1 (PY17X_1027900) is missing a bioinformatically predictable tristetraprolin-binding domain in all sequenced Aconoidasida species (Figure 4.1). The domain compositions

of these two genes are conserved across the Aconoidasida. I have previously shown that the CAF1-binding domain is likely functional on both proteins through detection of these proteins interacting with PyCCR4-1 (Table 3.1, Chapter 3, [173]) because CCR4 must associate with the complex through CAF1 but the other domains have not yet been characterized.

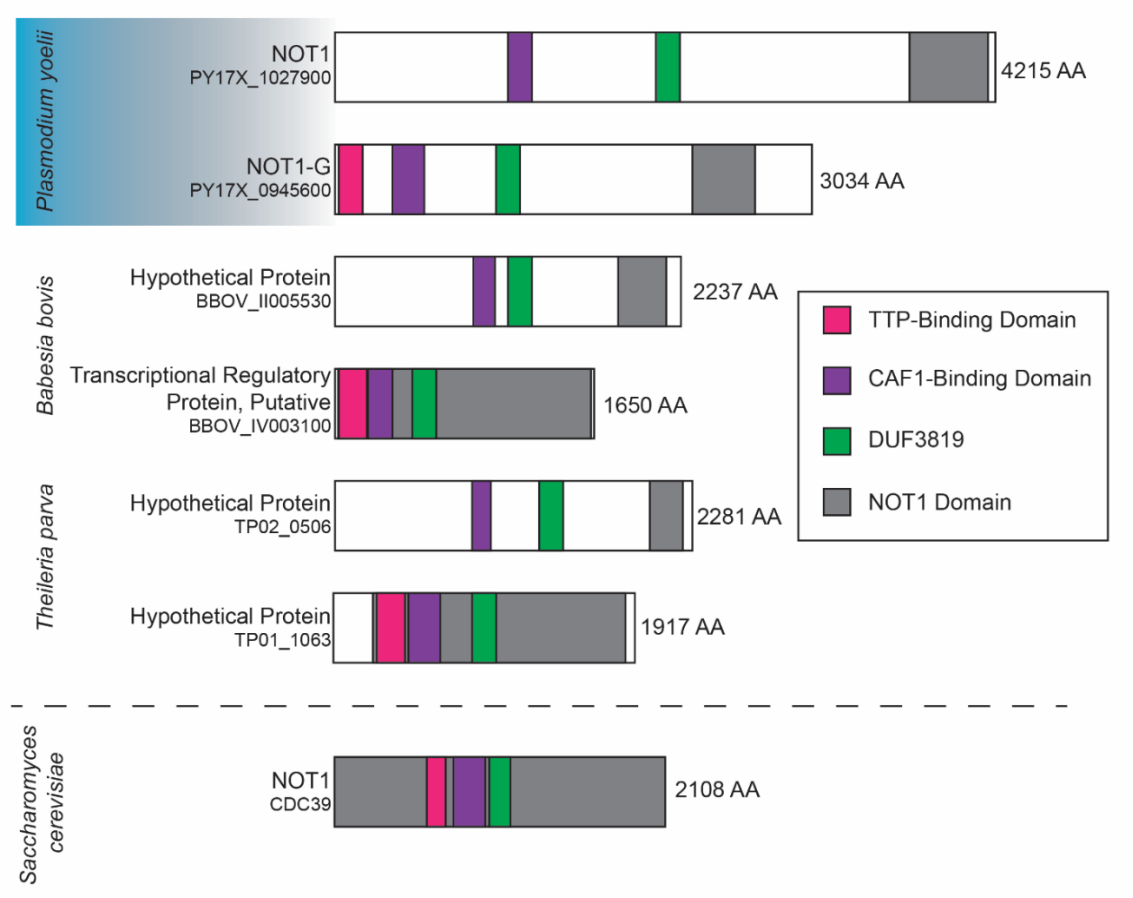


Figure 4.1: NOT1 is Duplicated in the Aconoidasida Class of Apicomplexans. Shown here are examples of NOT1 proteins from the Aconoidasida class of Apicomplexans with a representative model eukaryote NOT1 protein from *S. cerevisiae*. Protein and domain sizes are drawn to scale. Domains identified through BlastP amino acid alignments [163]. Sequences used in these searches were procured through PlasmoDB and the Saccharomyces genome database [190, 191].

PyNOT1-G localizes to discrete cytosolic foci

NOT1 is known to interact with a variety of components in model eukaryotes and is critical for assembly of the CAF1/CCR4/NOT complex [162]. It is also known to be present in cytosolic granules in these species [192]. To determine if this was the case in *P. yoelii* I fused a C-terminal GFP tag to the *not1-g* gene (Figure 4.2). Utilizing a GFP-tagged PyNOT1-G protein, I demonstrated that PyNOT1-G::GFP localizes to cytosolic puncta in all erythrocytic stages of *Plasmodium* development (Figure 4.3). This localization is consistent with studies from other eukaryotes that demonstrate the NOT1 associates in cytosolic granules, and may allude to model eukaryote-like functions in *Plasmodium* (Chapter 3, [134, 173]).

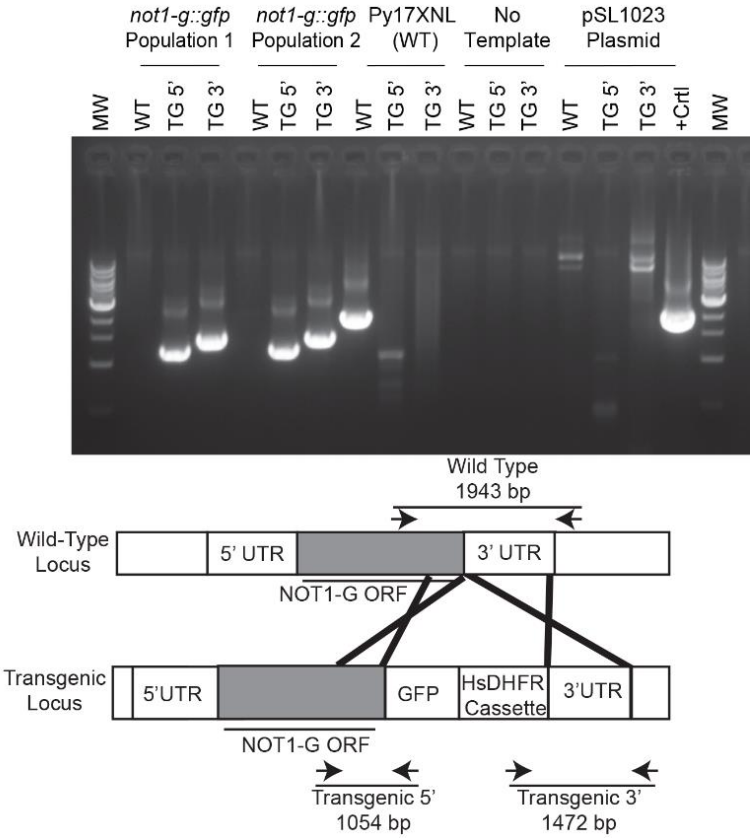
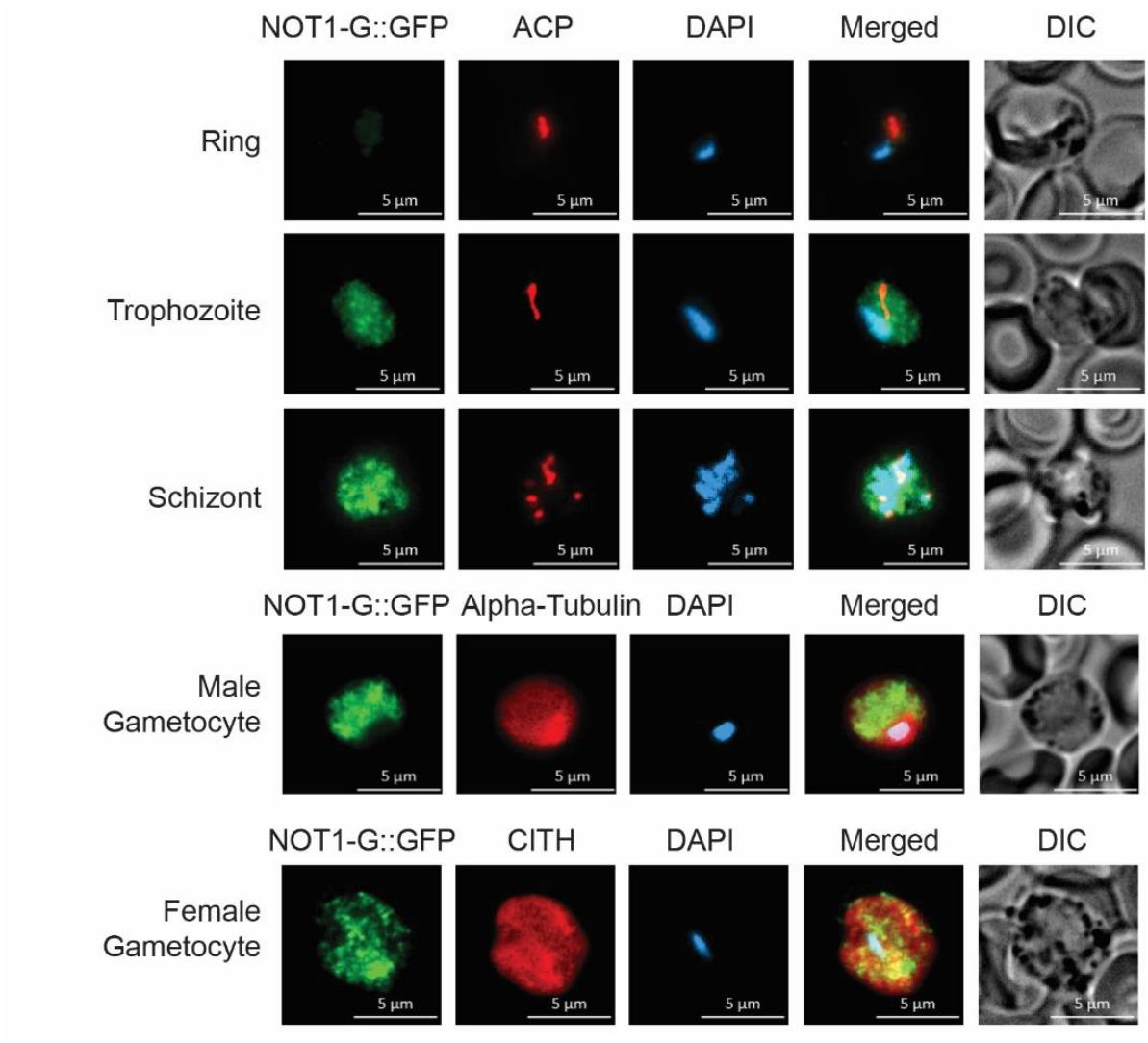


Figure 4.2: Genotyping PCR of *pynot1-g::gfp* Insertion of a C-terminal GFP tag was created using double homologous recombination at the 3' end of the *pynot1-g* coding sequence. Genotyping was performed by PCR on parasites using the primers indicated (listed in Appendix A). Independent clones were analyzed with a Py17XNL wild-type control, a no template control, and a plasmid positive control in parallel. pSL1023 plasmid positive control size is 2019 bp. Molecular weight ladder is from NEB (N3232L).


NOT1-G::GFP
ACP
DAPI
Merged
DIC
Ring
Trophozoite
Schizont
5 µm
NOT1-G::GFP
Alpha-Tubulin
DAPI
Merged
DIC
Male Gametocyte
NOT1-G::GFP
CITH
DAPI
Merged
DIC
Female Gametocyte

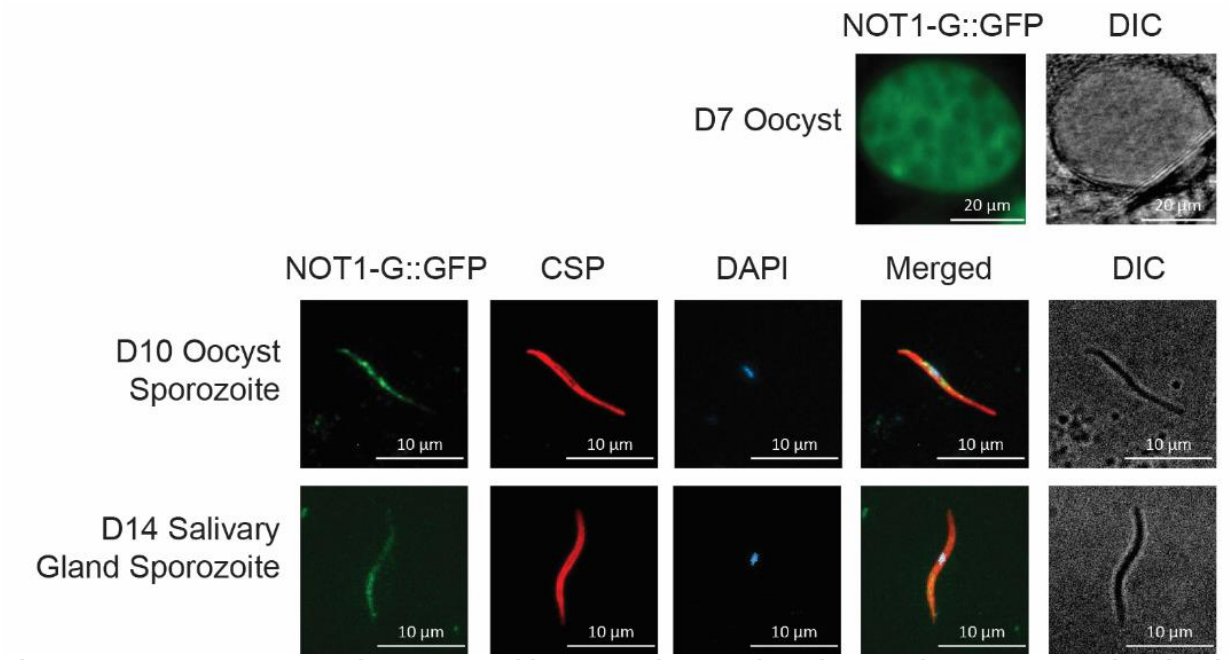


Figure 4.3: PyNOT1-G::GFP is Expressed in Asexual, Sexual, and Mosquito Stage Parasites but is not Detectable in Liver Stage Parasites. Representative images of asexual blood stage parasites, sexual blood stage parasites, oocyst sporozoites, and salivary gland sporozoites treated with DAPI and antibodies to GFP (to detect PyNOT1-G::GFP) or to stage-specific cellular markers (CSP, ACP) are shown. Live fluorescence was used to image oocysts. Scale bars are either 20 microns (oocysts), 5 microns (asexual and sexual stage parasites), or 10 microns (sporozoites).

not1-g^{-} Parasites have a Moderate Asexual Blood Stage Growth Defect

NOT1 is known to be essential for development in model organisms [104]. It can be depleted from HeLa cells, but it prevents the formation of P-bodies, and results in the death of the cells by apoptosis [134]. Because *Plasmodia* have two NOT1 proteins, I sought to determine if either *Plasmodium* NOT1 protein is essential in *P. yoelii*. While the PyNOT1-G gene was easily deleted, initial attempts to disrupt PyNOT1 were unsuccessful. On the sixth attempt, a gene deletion was achieved however the parasites took longer than normal to come up after transfection than typical (Figure 4.4AB). These data indicate that either the PyNOT1 gene is important for asexual blood stage development of the parasite.

I chose to focus my initial studies on PyNOT1-G because bioinformatically it is most similar to the canonical NOT1, and genetic manipulation of this gene was more efficient

than PyNOT1. The resulting *pynot1-g* gene deletion asexual parasites grew well during the asexual blood stage, thus facilitating experimentation. To determine if PyNOT1-G plays a vital role in blood stage development, I compared the growth kinetics of wild-type parasites and *pynot1-g*$^{-}$ parasites in mice. First, I infected mice to establish an infection, and then diluted parasites from this mouse to infect another with 1,000 mixed-blood stage parasites, then parasitemia was measured daily by microscopy. Measurements continued until the parasitemia exceeded 10% when the mouse was sacrificed (per the IACUC protocol), or the mouse had successfully cleared the infection. Parasites with a *pynot1-g* gene deletion were not able to achieve the same peak of parasitemiaas wildtype PY17XNL parasites (Figure 4.5A). These data indicate that while PyNOT1-G is not essential, it may either play a role during blood stage development, or in gametocyte commitment leading to this phenotype. These data also demonstrate that PyNOT1 cannot completely compensate for the lack of PyNOT1-G in asexual blood stage development. Together, these data show that NOT1 and NOT1-G likely play separate important roles in asexual blood stage development.

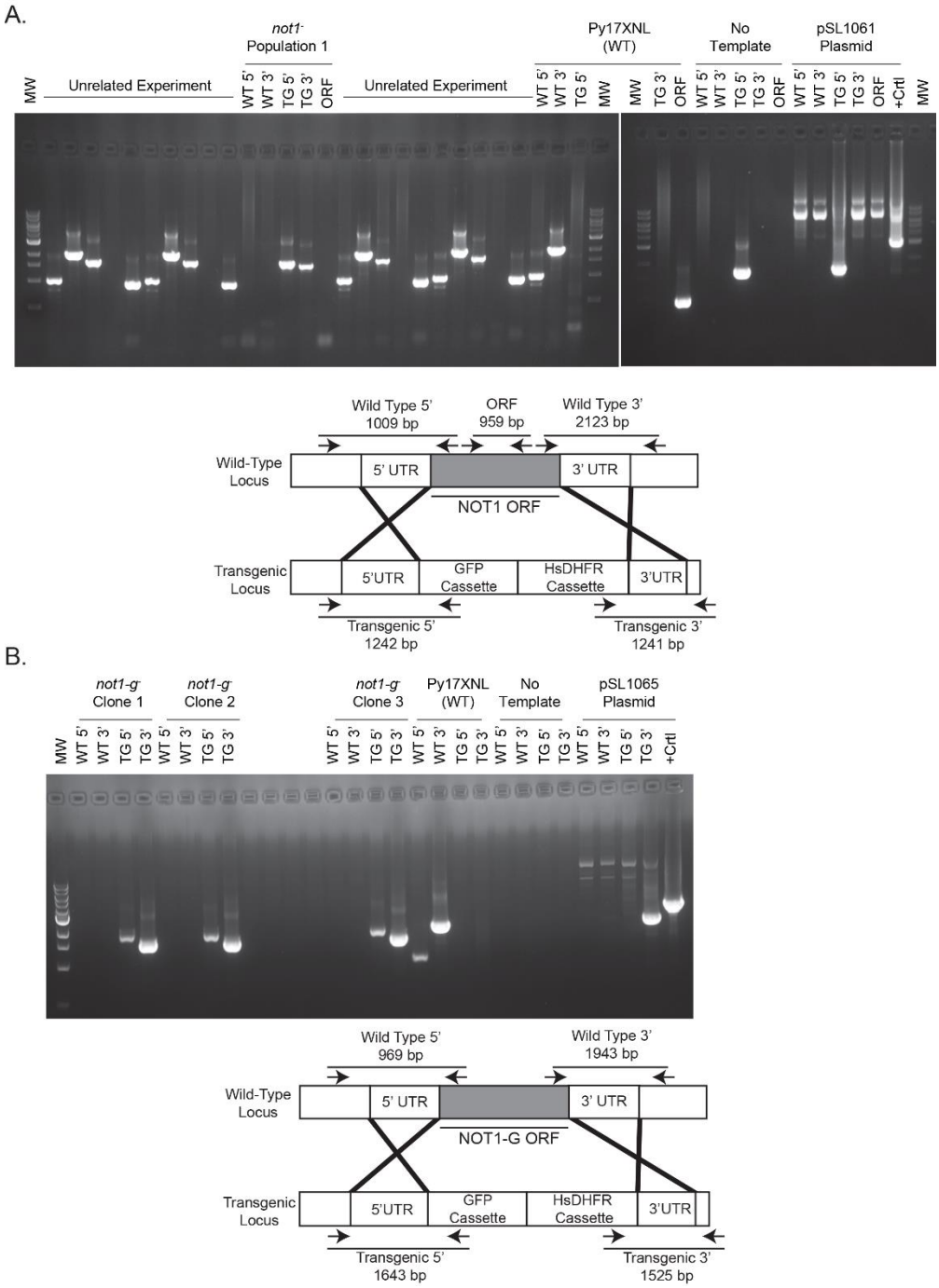


Figure 4.4: Genotyping PCR of (A) *pynot1-*, and (B) *pynot1-g-* Transgenic Parasites. Successful gene deletions were created using double homologous recombination of the targeting sequence consisting of ~750bp on either side of the ORF. Genotyping was performed by PCR on parasites cloned by limiting dilution using the primers indicated (listed in Appendix A). Populations or Independent clones were compared to a Py17XNL wild-type control, a no template control, and a plasmid positive control in parallel. pSL1061 plasmid positive control size is 2230 bp. pSL1065 plasmid positive control size is 2845 bp. Molecular weight ladder is from NEB (N3232L).

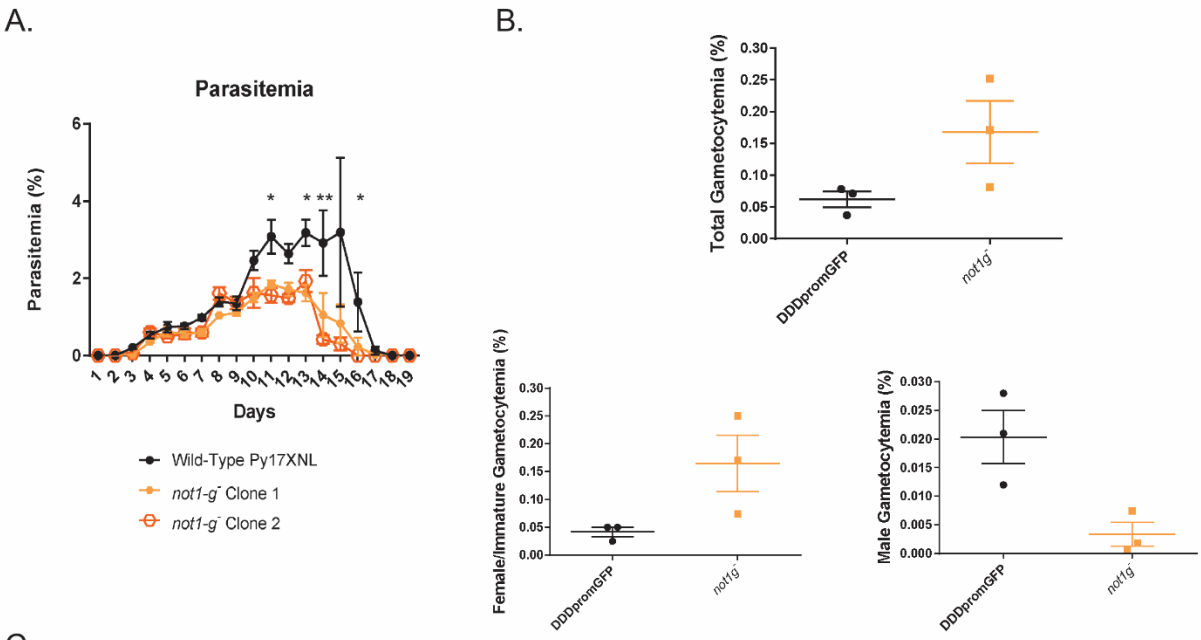


Figure 4.5: **Phenotyping of *not1-g*⁻ Transgenic Parasites**. (A) Asexual blood stage growth was monitored for two *pynot1-g*⁻ transgenic clonal lines compared to a WT-GFP control line over the entire course of an infection. Parasites with a *pynot1-g* gene deletion did not reach the same maximum parasitemia as wild-type. B) The number of mature male and female/immature gametocytes was determined by flow cytometry of sulfadiazine-treated and DDD/BIP-stained *P. yoelii* parasites. C) The number of centers-of-movement/exflagellation centers were quantified daily by microscopy on a 400x field upon injection with 1,000 blood stage parasites. Plotted are three biological replicates with three technical replicates each. Error bars represent the standard error of the mean.

NOT1-G is Essential for Proper Gametocyte Development

The presence of an additional NOT1 protein is unable to compensate for the loss of *pynot1-g*. PyNOT1 and PyNOT1-G appear to play distinct, non-redundant roles. Therefore, I hypothesized that one of the two NOT1 genes is likely to play an essential role during parasite development. Because of the conservation of interaction domains and essential functions, I hypothesize that there are two distinct NOT complexes in *Plasmodium*. The parasite requires sophisticated control of gene expression to navigate

the life cycle, thus two distinct NOT complexes may allow for more intricate control of gene expression.

To determine what effect PyNOT1-G may have on gametocyte development, I utilized a flow cytometry assay that measures total gametocytes and mature male gametocytes. Sulfadiazine treatment was used to select for sexual stage parasites from mice before analysis. To this end, I found a severe reduction in the number of mature male gametocytes that express PyDD (dynein heavy chain delta) (Figure 4.5B), which explains the lack of observed male activation (Figure 4.5C). Interestingly, I also observed an increase in the number of total gametocytes. Because of this, I hypothesize that PyNOT1-G may dampen gametocyte commitment and in the absence of PyNOT1-G, more parasites commit to sexual development. However, given the data that has been generated thus far, it is unclear what role of this complex may underlie this function. The NOT1-G complex could affect transcription of, or regulate transcript abundance of, commitment genes.

A small number of PyDD-expressing male gametocytes were produced. To determine if these male gametocytes could activate into gametes, male activation was quantified during the growth curve by counting exflagellation centers daily until the clearance of parasitemia (Figure 4.5C). While I observed male activation in wild-type PY17XNL parasites, parasites with a *pynot1-g* gene deletion were unable to produce any exflagellation centers. These parasites were also unable to be transmitted to mosquitoes (Figure 4.6). These data demonstrate that parasites with a *pynot1-g* gene deletion cannot transmit from the rodent host to the anopheline mosquito vector.

To determine if PyNOT1-G also plays a role in the development of female gametocytes, I performed a genetic cross of parasites with a *pynot1-g* gene deletion with wild-type PY17XNL parasites. In this cross, four possible outcomes can result: 1) Wild-type male and female (colorless) gametocytes fuse to form colorless oocysts, 2) wild-type male (colorless) and *pynot1-g*$^{-}$ female (GFP+) gametocytes fuse to form GFP+ oocysts, 3) *pynot1-g*$^{-}$ male (GFP+) and wild-type female (colorless) gametocytes fuse to form GFP+ oocysts, and 4) *pynot1-g*$^{-}$ male and *pynot1-g*$^{-}$ female (GFP+) gametocytes fuse to form GFP+ oocysts.

As no male gametes were observed in *pynot1-g*$^{-}$ parasites, outcomes 3 and 4 should not be possible. Thus, the presence of colorless (GFP-) oocysts represents the outcome of a WTxWT cross, while the presence of GFP+ oocysts represents the outcome of WT male x *pynot1-g*$^{-}$ female gametocytes.

When parasitemia reached ~1% mice were infected with either 10,000 wild-type, 10,000 *not1-g*$^{-}$, or a mixture of 5,000 wild-type and 5,000 *not1-g*$^{-}$ parasites. Oocyst numbers were then counted on day seven post-infectious feed under a fluorescence microscope to determine if any parasites WT x WT, WT x *not1-g*$^{-}$, or *not1-g*$^{-}$ x *not1-g*$^{-}$were able to infect the mosquito. As expected, mosquitoes that fed on mice infected with only parasites with a *pynot1-g* gene deletion did not have any infection, and those fed only wild-type parasites produced a robust infection in the mosquito. However, mosquitoes that fed on mice with the mixed population of *not1-g*$^{-}$ and wild-type parasites were able to be infected, but the only oocysts that resulted from the infection were wild-type (colorless (GFP-)) (Figure 4.6). The infection that resulted from the mixed parasite feed also produced only 50% of the oocysts when compared to the wild-type only infection. These data indicate that PyNOT1-G plays a critical role in both male and female gametocyte development and is essential for the production of transmissible male and female gametocytes. It is unclear what role NOT1 plays in regulating gametocyte development. The NOT1-G complex could affect transcription of, or regulate transcript abundance of, transcripts important for gametocyte development.

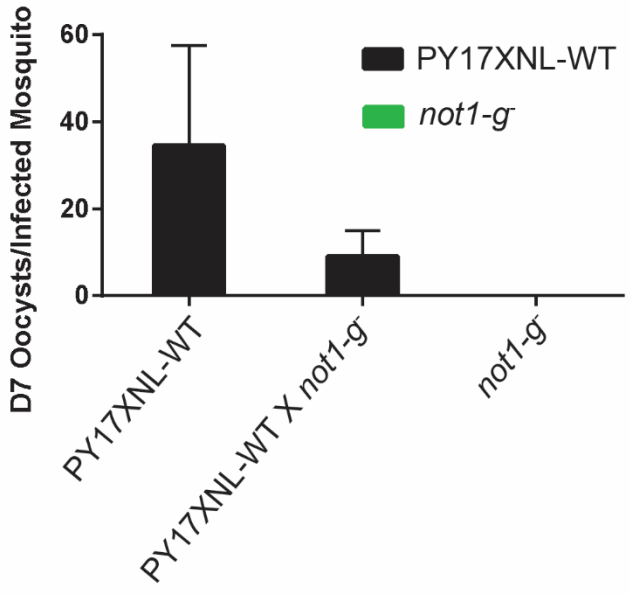


Figure 4.6: Both Male and Female *not1-g*$^{-}$ Gametocytes are Infertile. Colorless wild-type parasites were mixed equally with GFP-expressing *pynot1-g* gene deletion parasites in mice and fed to mosquitoes on the peak day of wild-type exflagellation. Oocyst counts, D7 after mosquito feed, are shown here as the average of three replicates. Error bars represent Standard error of the mean (SEM).

While male and female gametocytes are infertile, a large number of gametocytes are still produced (Figure 4.5B). I hypothesize that NOT1-G may increase commitment to gametocyte development. NOT1-G could play this role through modulating gene expression by affecting the transcription or RNA abundance of specific factors involved in this process. This phenotype would lead to an increase in total gametocytemia even if they do not develop past an immature stage.

pynot1-g Deletion results in Genome-wide Changes in mRNA abundance in gametocytes

The CAF1/CCR4/NOT complex in known to play essential roles in RNA metabolism in eukaryotes. Data shown here in *P. yoelii* demonstrate that while there are two NOT1 proteins, each PyNOT1 protein plays non-redundant roles. I hypothesize that PyNOT1 and PyNOT1-G each have specific effects on transcripts. The slow growth and difficult transfection of pynot1 deletion parasites indicates that PyNOT1 is playing an important role in asexual blood stage development. While there are other explanations for this phenotype, only characterization of this parasite line will fully explain this defect. By determining which transcripts are affected by PyNOT1-G, it may be possible to assign functions attributed to this protein. To determine what effects deletion of *pynot1-g* may have on gametocytes, I performed total comparative RNA-seq on wild-type and *pynot1-g* gene deletion gametocytes as performed in chapter 3**.** To determine significance, we utilized a P-adjusted value of 0.05. My results demonstrate that a large number of transcripts are dysregulated in the *pynot1-g* gene deletion gametocytes (Figure 4.7, Table 4.1). Because fewer male gametocytes form in parasites with a *not1-g* gene deletion, it is not straightforward to assess the dysregulation of male-enriched transcripts. This is because these effects could be attributed to the presence of a lower number of male gametocytes (Figure 4.5B). However, this data may provide and explanation for the observation that female gametocytes are infertile. While roughly the same number of transcript abundances increased (613) and decreased (804) out of the 5720 detected, the scale of the fold change between the two is substantially different (Figure 4.7A).

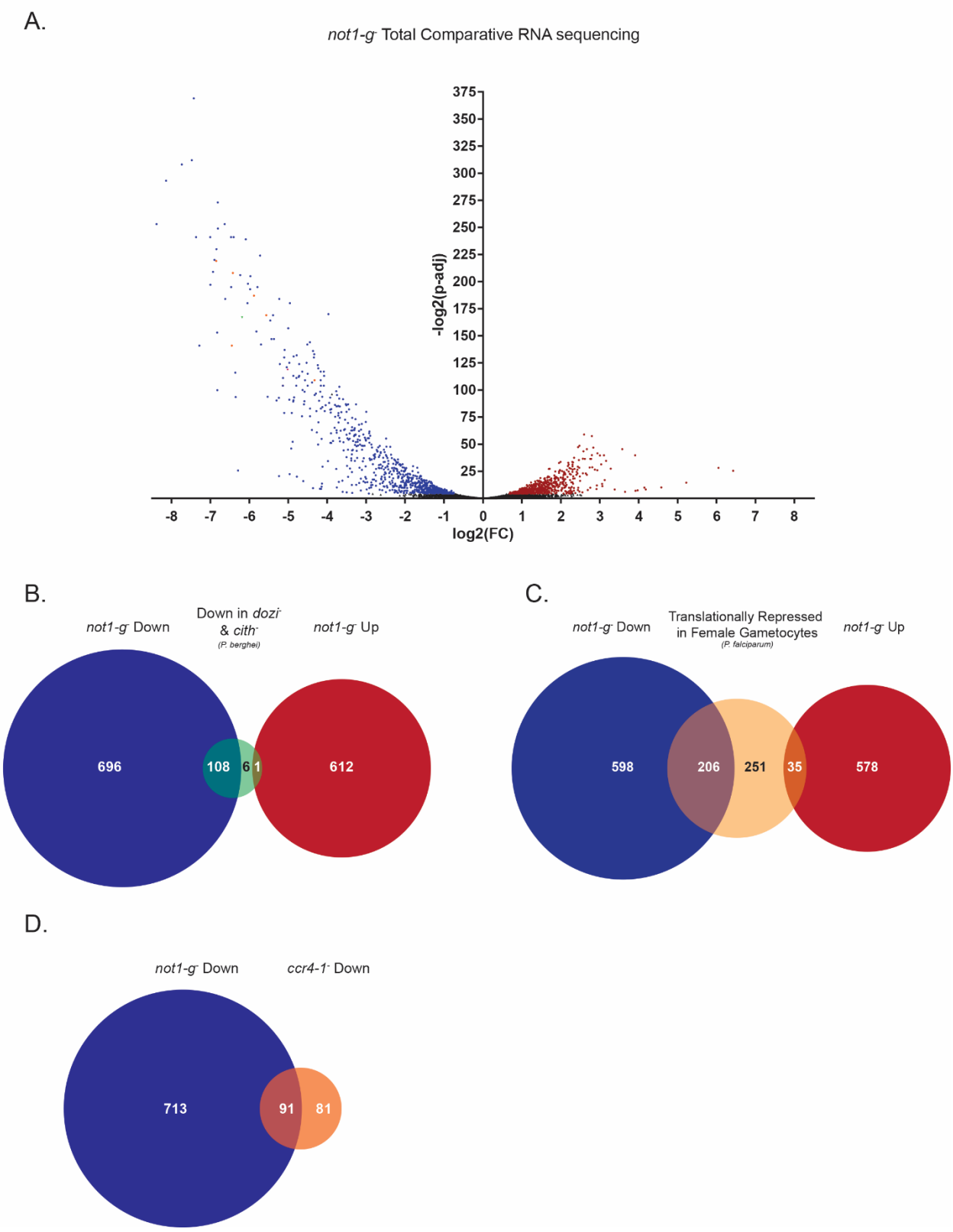

A.
not1-g- Total Comparative RNA sequencing
-log2(p-adj)
log2(FC)
B.
not1-g- Down
Down in dozi- & cith-
(P. berghei)
not1-g- Up
696
108
6
1
612
C.
not1-g- Down
Translationally Repressed in Female Gametocytes
(P. falciparum)
not1-g- Up
598
206
251
35
578
D.
not1-g- Down
ccr4-1- Down
713
91
81

Figure 4.7: Translationally Repressed Transcripts and those Modulated by DOZI and CITH are affected by Gene Deletion of *pynot1-g*. A) A volcano plot showing changes in transcript abundance up (red), down (blue) or unchanged (black). Additionally, labeled are IMC complex transcripts in orange, P28 in green and p25 in pink. A large number of transcripts increase or decrease in abundance, indicating that PyNOT1-G plays an essential role in regulating them in gametocytes. B) In gametocytes with a *pynot1-g* deletion, the vast majority of transcripts that are affected by *pbdozi gene deletion* and *cith* gene deletion decrease in abundance. C) Almost half of identified translationally repressed transcripts in female gametocytes decrease in abundance with *pynot1-g* gene deletion indicating that PyNOT1-G plays an essential role in female gametocytes. D) More than half of the transcripts that decrease in abundance in *ccr4-1*$^{-}$ gametocytes also decrease in abundance in *not1-g*$^{-}$ gametocytes.

Top Transcripts Decreasing in Abundance with *not1-g*$^{-}$					
Gene ID	Wild-Type-GFP Avg	*not1-g*$^{-}$ Avg	Fold Change	P-adj Value	Gene Name
PY17X_0945600	40.66666667	0	0	1.88627E-08	CCR4-NOT transcription complex subunit 1, putative (not1-g)
PY17X_0807200	9	0	0	0.002210063	tRNA Histidine
PY17X_1009100	5.666666667	0	0	0.015966827	tRNA Serine
PY17X_0929600	2753.333333	6.666666667	0.002421308	8.48506E-77	conserved Plasmodium protein, unknown function
PY17X_1463900	3870.333333	12.33333333	0.003186633	7.70914E-89	conserved Plasmodium protein, unknown function
PY17X_0909300	3762.666667	16	0.004252303	2.4116E-93	conserved Plasmodium protein, unknown function
PY17X_1323300	7240.333333	35.66666667	0.004926108	1.19596E-94	LCCL domain-containing protein
PY17X_1136400	8510.333333	42.66666667	0.005013513	6.372E-112	zinc finger protein, putative
PY17X_1432500	759.6666667	4	0.005265467	4.97823E-43	conserved Plasmodium protein, unknown function
PY17X_1321000	2900.666667	15.33333333	0.005286141	3.652E-73	pyridine nucleotide transhydrogenase, putative
PY17X_1451800	383.3333333	2.666666667	0.006956522	8.30365E-31	upregulated in late gametocytes ULG8, putative
PY17X_1140000	1633	11.66666667	0.007144315	1.32879E-63	conserved Plasmodium protein, unknown function
PY17X_0705100	2045.666667	14.66666667	0.007169627	4.0333E-60	conserved Plasmodium protein, unknown function
PY17X_1411400	688.3333333	5	0.007263923	8.0057E-47	meiosis-specific nuclear structural protein 1, putative
PY17X_1360600	3989	29.66666667	0.007437119	9.43548E-67	inner membrane complex protein 1k, putative
PY17X_1359200	7706.333333	59	0.00765604	3.43982E-73	conserved Plasmodium protein, unknown function
PY17X_1133800	2756.666667	21.33333333	0.007738815	8.14957E-76	conserved Plasmodium protein, unknown function
PY17X_1339500	2982.333333	23.33333333	0.007823852	7.34145E-70	conserved Plasmodium protein, unknown function
PY17X_0814000	6132	48.66666667	0.007936508	4.35342E-67	subpellicular microtubule protein 1, putative
Top Transcripts Increasing in Abundance with *not1-g*$^{-}$					
Gene ID	Wild-Type-GFP Avg	*not1-g*$^{-}$ Avg	Fold Change	P-adj Value	Gene Name
PY17X_0622300	6.666666667	90.66666667	13.6	1.08884E-12	PelOta protein homologue, putative
PY17X_0111900	9	78.33333333	8.703703704	5.58601E-09	conserved Plasmodium protein, unknown function
PY17X_1308600	2.666666667	23	8.625	0.000248491	conserved Plasmodium protein, unknown function
PY17X_1411700	12.66666667	100.3333333	7.921052632	3.75787E-11	AP-1 complex subunit mu-1, putative
PY17X_0215500	3.333333333	26	7.8	9.24482E-05	conserved Plasmodium protein, unknown function
PY17X_1220500	22.33333333	161.3333333	7.223880597	1.0537E-12	UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminephosphotransferase, putative
PY17X_0513400	15.66666667	113	7.212765957	1.18406E-12	RING zinc finger protein, putative
PY17X_1445800	242.3333333	1677	6.920220083	9.44328E-13	merozoite surface protein 9, putative
PY17X_0307000	9.333333333	63	6.75	9.51901E-09	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, putative
PY17X_0111800	42.33333333	279	6.590551181	3.04399E-13	lsm12, putative
PY17X_1003600	11.33333333	73.33333333	6.470588235	1.47382E-08	6-cysteine protein, putative
PY17X_1147000	1.666666667	10.66666667	6.4	0.03440506	U2 small nuclear ribonucleoprotein A', putative
PY17X_0623200	1.666666667	10.33333333	6.2	0.0202613	tRNAHis guanylyltransferase, putative
PY17X_0208100	55.33333333	334	6.036144578	5.17803E-18	conserved Plasmodium protein, unknown function
PY17X_1428100	45.33333333	272.3333333	6.007352941	3.43871E-10	histone-arginine methyltransferase CARM1, putative
PY17X_0611500	5.666666667	33	5.823529412	0.001078724	conserved Plasmodium protein, unknown function
PY17X_1206400	58	331.6666667	5.718390805	7.71013E-15	methionine--tRNA ligase, putative
PY17X_1450700	207.3333333	1145	5.522508039	1.05368E-11	asparagine/aspartate rich protein, putative
PY17X_1210000	3	16.33333333	5.444444444	0.012169022	zinc finger (CCCH type) protein, putative
PY17X_1015000	527.3333333	2866.666667	5.436156764	1.05368E-11	large subunit GTPase 1, putative

Table 4.1: The Top Transcripts that Increased or Decreased in Abundance with the Deletion of *pynot1-g*. Transcripts that represent antigenically variant proteins are removed. The full dataset will be available publicly upon publication.

ApiAP2 Transcription Factors					
Gene ID	Wild-Type-GFP Avg	*not1-g*⁻ Avg	Fold Change	P-adj Value	Gene Name
PY17X_1317000	407.67	12.33	0.03	0.00000	AP2 domain transcription factor, putative
PY17X_0907300	3341.33	224.33	0.07	0.00000	AP2 domain transcription factor AP2-O, putative
PY17X_1323500	31.67	4.67	0.15	0.00034	AP2 domain transcription factor, putative
PY17X_0941600	521.00	199.67	0.38	0.00006	AP2 domain transcription factor, putative
PY17X_1440000	125.67	51.33	0.41	0.00222	AP2 domain transcription factor AP2-G, putative
PY17X_0215800	153.00	359.33	2.35	0.00017	AP2 domain transcription factor AP2-L, putative
PY17X_1003200	21.67	84.00	3.88	0.00000	AP2 domain transcription factor AP2-SP2, putative
PY17X_1017000	80.33	414.33	5.16	0.00000	AP2 domain transcription factor AP2-O3, putative
Known Function in Males					
Gene ID	Wild-Type-GFP Avg	*not1-g*⁻ Avg	Fold Change	P-adj Value	Gene Name
PY17X_0935700	107.67	20.67	0.19	0.00000	mitogen-activated protein kinase 2, putative
PY17X_1434500	504.33	213.00	0.42	0.00102	male development gene 1, putative
PY17X_0617900	226.67	126.00	0.56	0.03788	calcium-dependent protein kinase 4, putative
Associated or thought to be associated with Flagellum					
Gene ID	Wild-Type-GFP Avg	*not1-g*⁻ Avg	Fold Change	P-adj Value	Gene Name
PY17X_0418900	134.00	35.67	0.27	0.00006	dynein heavy chain, putative
PY17X_0505400	104.33	30.67	0.29	0.00008	flagellar outer arm dynein-associated protein, putative
PY17X_0505500	2409.00	54.67	0.02	0.00000	conserved Plasmodium protein, unknown function
PY17X_0508400	92.33	29.67	0.32	0.00145	dynein heavy chain, putative
PY17X_0512900	134.00	39.00	0.29	0.00039	centrin-3, putative
PY17X_0524100	713.33	286.67	0.40	0.00103	alpha tubulin 2
PY17X_0826200	129.33	42.00	0.32	0.00501	dynein intermediate chain, putative
PY17X_0927400	148.33	22.00	0.15	0.00000	dynein heavy chain, putative
PY17X_0944800	17.67	3.00	0.17	0.00888	radial spoke head protein, putative
PY17X_1126700	155.33	14.33	0.09	0.00000	conserved Plasmodium protein, unknown function
PY17X_1127300	18.33	3.67	0.20	0.01804	conserved Plasmodium protein, unknown function
PY17X_1333900	96.33	26.67	0.28	0.00005	dynein beta chain, putative
PY17X_1433800	90.00	15.67	0.17	0.00000	conserved Plasmodium protein, unknown function
Identified as Translationally Repressed in Gametocytes					
Gene ID	Wild-Type-GFP Avg	*not1-g*⁻ Avg	Fold Change	P-adj Value	Gene Name
PY17X_1451800	383.33	2.67	0.01	8.3036E-31	upregulated in late gametocytes ULG8, putative
PY17X_0515900	16633.67	233.33	0.01	6.4592E-51	28 kDa ookinete surface protein
PY17X_1237000	1150.67	32.33	0.03	8.8651E-29	secreted ookinete protein, putative
PY17X_1423700	353.00	23.67	0.07	7.9107E-22	secreted ookinete protein, putative
PY17X_1130500	1110.00	95.33	0.09	2.1138E-18	secreted ookinete protein, putative
PY17X_1411000	984.33	112.00	0.11	7.5049E-17	RNA-binding protein, putative
PY17X_1020600	141.67	24.00	0.17	1.2671E-08	DEAD/DEAH box helicase, putative

Table 4.2: A Selected List of Transcripts Affected by the Deletion of *pynot1-g*. Transcripts that encode ApiAP2 transcription factors were affected by *pynot1-g* gene deletion and some increased or decreased in abundance. Effects on ApiAP2 transcription factors can lead to indirect effects of the *pynot1-g* gene deletion. Shown are transcripts that have effects on both male and female gametocytes. Those with a known function in male gametocytes and that are important for flagellar development are likely decreased because of the lack of mature male gametocytes. A selection of transcripts identified as translationally repressed are shown and represent severe effects on the female gametocyte [28, 175].

The transcripts that changed the most are those that encode inner membrane complex (IMC) proteins (*inner membrane complex protein 1k, d, h, e, j, l, c, b, f, g* in order of greatest decrease to least, orange color), and highly translationally repressed transcripts such as *p28* (green)*, p25* (pink) *and ap2-o* (Figure 4.7A, Table 4.1). It is known that blocking the assembly of the IMC prevents *P. falciparum* gametocyte development [193]. While IMC proteins are important for the shape of *P. falciparum* gametocytes, which are elongated rather than round like *P. yoelii* gametocytes, this data could indicate that the IMC complex may be playing similar roles in gametocyte formation *P. yoelii*.

Some translationally repressed transcripts have been shown to be essential for host-to-vector transmission. Transcripts that decreased in abundance in the gametocytes with a deletion of *pynot1-g*, (*p28, p25,* and *ap2-o*) are known to be important or essential for early mosquito stage development, and are translationally repressed in DOZI/CITH containing cytosolic granules (Table 4.1) [31, 194]. Because *pbdozi*$^{-}$ and *pbcith*$^{-}$ parasites fail to produce oocysts in the mosquito, I compared transcripts that decreased in abundance in *pynot1-g*$^{-}$ gametocytes with those that decreased in abundance in both *pbdozi*$^{-}$ and *pbcith*$^{-}$ gametocytes. I found that 108 of the 115 *P. yoelii* orthologues of the *P. berghei* data also decrease in abundance with the deletion of *pynot1-g* (Figure 4.7B). Interestingly, DHH1 and DDX6 (DOZI orthologues in model eukaryotes) are known to associate with NOT1 to impose translational repression, and I previously identified that DOZI associates with the CAF1/CCR4/NOT complex (Table 3.1, [173]). Translationally repressed transcripts in *P. falciparum* gametocytes are also affected by the *not1-g* gene deletion. 206 of the 492 *P. yoelii* orthologues that were identified as being translationally repressed in *P. falciparum* were affected by the *not1-g* gene deletion [28]. Interestingly, neither *pydozi* nor *pycith* were significantly affected by the *pynot1-g* deletion in gametocytes. In addition to the effects on the above translationally repressed transcripts, multiple ApiAP2 transcripts decreased in abundance in *pynot1-g* gene deletion parasites. These transcripts that decrease include three that are known to be translationally repressed through association with DOZI/CITH (PY17X_1317000), or were identified as translationally repressed in *P. falciparum* (PY17X_0907300, PY17X_0941600) (Table 4.2). Additionally, two others including ApiAP2-G

(PY17X_1440000, PY17X_1323500) decreased in abundance. The effects on ApiAP2-G are particularly important given the phenotype of *not1-g*$^{-}$ parasites. ApiAP2-G is essential for gametocyte commitment and thus a decrease in ApiAP2-G should result in less gametocytes; however, I observed an increase in the total number of gametocytes. This could indicate that there is a difference in transcript dynamics in these parasite lines that affects expression of ApiAP2-G. These experiments need to be performed in asexual blood stage parasites to confirm *apiap2-g* transcript abundances are similarly affected. Three transcripts that encode ApiAP2 proteins (*apiap2-l, apiap2-sp2,* and *apiap2-o3)* increase in abundance in the *pynot1-g* gene deletion. The effects of *pynot1-g* gene deletion on specific transcription factors in gametocytes may induce a change in transcription that causes additional, indirect, and widespread changes in transcript abundance. These data demonstrate that gametocytes with a *not1-g* deletion have dysregulated transcriptomes affecting transcripts that are important for gametocyte and early mosquito stage development. It also demonstrates that the CAF1/CCR4/NOT complex may act with DOZI and CITH to mediate translational repression of particular transcripts. Additionally, this data overlaps well with transcript abundance changes observed in the *ccr4-1* deletion gametocytes (Figure 4.7D). This suggests that some of the effects of CCR4-1 on transcript abundances may be specific to this scaffold. NOT1 and NOT1-G were detected in association with CCR4-1 and that the functions of NOT1 and NOT1-G appear so different. Thus it is not surprising that CCR4-1 may affect different transcripts when in association with a different scaffold and presumably different targets but is an interesting finding that requires further experimentation to understand.

Bioinformatically the only identifiable difference between PyNOT1 and PyNOT1-G is that PyNOT1-G has a bioinformatically predictable TTP (tristetraprolin)-binding domain (Figure 4.1A). However, in *Plasmodium* there is not a bioinformatically identifiable protein that aligns with model eukaryote TTP. Thus, to determine if a TTP could be found interacting with NOT1-G, I overexpressed the TTP-binding domain of NOT1-G with a C-terminal GFP tag (AA1-199) (TTP-BD) to perform immunoprecipitation and mass spectrometry (IP-MS). Overexpression of the TTP-BD has no effect on blood stage parasite development (Figure 4.8AB). I performed a crosslinking

immunoprecipitation (IP), and a western blot demonstrated that TTP-BD could be immunoprecipitated (Figure 4.8C). I then utilized mass spectrometry to analyze samples, and identified a conserved *Plasmodium* protein (PY17X_0708000) that has no identifiable domains. An additional protein “zinc finger protein, putative” (PY17X_0417500) was also identified from the PyCCR4-1::GFP IP-MS data. This protein has a C3H1 zinc finger but lacks the tandem zinc finger domains of typically tristetraprolin proteins (Chapter 3, [173]). Further characterization of a tristetraprolin in *Plasmodium* may be key to differentiating the roles of NOT1 and NOT1-G.

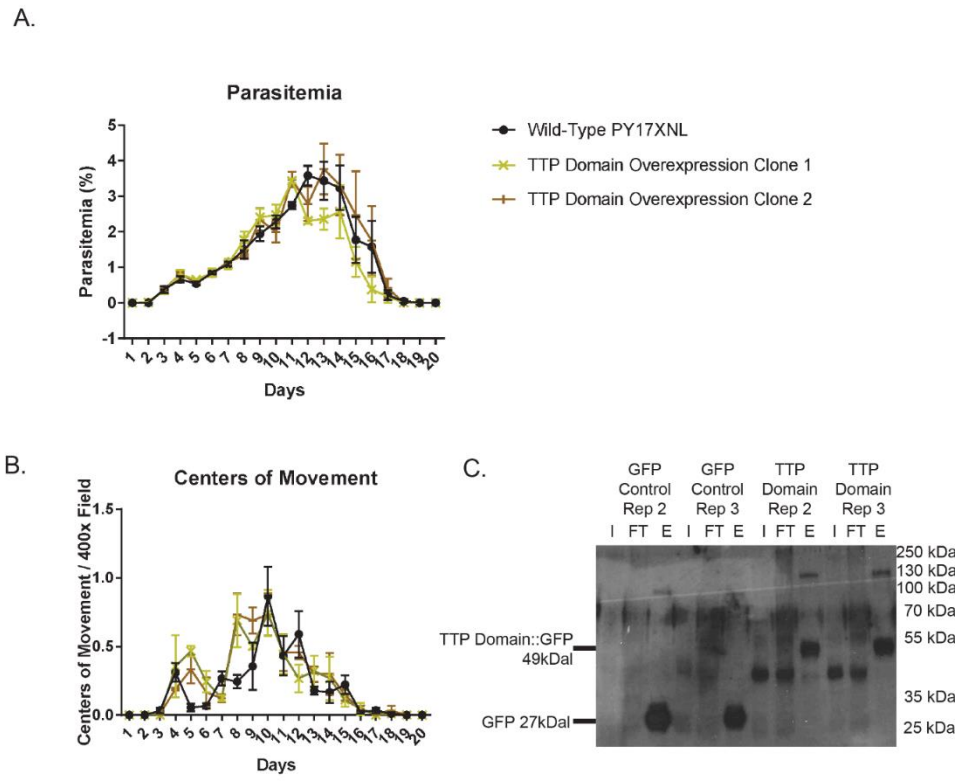


Figure 4.8: Phenotyping of the TTP-Binding Domain Overexpression Parasites (A) Asexual blood stage growth was monitored for two PyNOT1-G TTP-Binding Domain overexpression transgenic clonal lines compared to a WT-GFP control line over the entire course of an infection. No significant difference in growth kinetics was observed. B) The number of centers-of-movement/exflagellation centers were quantified daily by microscopy on a 400x field upon injection with 1,000 blood stage parasites. Plotted are three biological replicates with three technical replicates each. Error bars represent the standard error of the mean. C) A western blot demonstrating the successful Immunoprecipitation of the PyNOT1-G TTP-binding domain (49 kDa), and the GFP control (27kDa) for two replicates. I=Input, FT= Flow-through, and E=Elution.

Discussion

NOT1 of the CAF1/CCR4/NOT complex was shown to be essential for growth in model eukaryotes. However, in the Aconoidasida class of parasites (*Plasmodium*, *Theileria* and *Babesia*), there is a duplication of the NOT1 gene. The data presented here demonstrate that PyNOT1 and PyNOT1-G play roles in asexual blood stage development, and that PyNOT1-G plays an essential role in gametocyte development. Blood stage asexual parasites with a *pynot1-g* genetic deletion are not able to grow as efficiently in the blood. This blood stage growth phenotype could potentially be linked to an increased number of total gametocytes in these parasites. Sexual gametocytes originate from the differentiation of asexual blood stage parasites to a sexual path. Because sexual gametocytes cannot differentiate back into asexual blood stage parasites, the total number of asexual parasites in the blood decreases upon differentiation. If commitment is affected in parasites with a *pynot1-g* gene deletion this may explain the asexual development phenotype. It has not yet been determined if the increased number of gametocytes is due to gametocyte immaturity, or if female gametocyte conversion is higher. These remaining questions can be explored by enhancing the flow assay with the inclusion of a female specific marker, and additionally by using a *pfnot1-g* genetic deletion *in vitro* where parasites can be synchronized and more easily staged.

While neither NOT1 protein in *P. yoelii* is essential for normal asexual blood stage growth, PyNOT1-G is essential for proper gametocyte development and transmission. I predict that this essentiality is due to a conserved domain difference between the two NOT1 genes. PyNOT1-G contains a bioinformatically predictable tristetraprolin (TTP)-binding domain whereas PyNOT1 does not (Figure 4.1A). In model eukaryotes, the TTP-binding domain interacts with TTP, an RNA-binding protein that helps to determine what transcripts are recruited to the complex. While I cannot bioinformatically predict any proteins that align with high similarity to TTP in any *Plasmodium*, TTP could have evolved away from the canonical protein in model eukaryotes to conform its amino acids to the AT bias in all *Plasmodium* or to evolve to parasite specific biology. Since this protein is not bioinformatically identifiable but its interaction domain is conserved, it may

be possible to experimentally determine if this protein exists in any *Plasmodium*. I have identified two proteins, one that associates with high confidence to an overexpressed TTP-binding domain in *P. yoelii* and the other, a zinc finger protein that associates with PyCCR4-1. These two candidates are candidates for being TTP proteins in *Plasmodium* and require further characterization.

These analyses demonstrate that PyNOT1-G has several important and essential roles in *P. yoelii* parasite development. While the role of PyNOT1-G in asexual blood stage development has not been thoroughly examined here, deletion of *pynot1-g* does not permit parasites to reach the same peak parasitemia as seen with wild-type parasites. This phenotype could result from many different defects; a reduction in the number of merozoites produced, an increase in time to ring-schizont development, a change in immunogenicity of the parasites such that the immune system of the mouse begins clearing them sooner, or an increase in gametocyte conversion. Additionally, PyNOT1-G was found to be essential for male gametocyte development, and few PyDD expressing male gametocytes are produced in lines with a *pynot1-g* deletion (Figure 4.5BC). These data may reflect that either male gametocytes are not being produced, there is a defect in male gametocyte conversion, or that male gametocytes that are produced remain immature and do not express PyDD. Because of the increased number of gametocytes seen in *pynot1-g* genetic deletion clones, I hypothesize that male gametocytes are likely produced that do not develop to the point of PyDD expression. Additionally, I hypothesize that the increase in gametocytes is because of an increase in gametocyte commitment. In parasites with a *not1-g* deletion there is no transmission to the mosquito, and when crossed with wild-type parasites, no parasites with a *pynot1-g* genetic deletion could be transmitted. These data indicate that PyNOT1-G is also essential for female gametocyte fertility. Additional effects on the female gametocyte can be measured using female-enriched or specific transcripts from total comparative RNA-seq. Female parasites with a *pynot1-g* gene deletion do not have the same abundance of certain transcripts that are essential for transmission, similar to but more extreme than what was defined in the *P. berghei dozi* and *cith* gene deletion parasites (Figure 4.7ABC, Table 4.1). Additionally, multiple ApiAP2 transcripts were dysregulated. Because of this, I hypothesize that female gametocytes are also

produced but remain immature, and are unable to infect the mosquito productively. Taken together, these data demonstrate the immense importance of the NOT1 proteins found within the CAF1/CCR4/NOT complex for *P. yoelii* development and transmission.

Chapter 5: DISCUSSION AND FUTURE DIRECTIONS

My Contributions to the Field: CCR4/CAF1

The CAF1/CCR4/NOT complex has been shown to be an essential component of RNA metabolism in the cell. Perhaps because *Plasmodium* species have only one known class of specific transcription factors (ApiAP2), the parasite utilizes RNA metabolic control more extensively to regulate gene expression. Bioinformatically in *Plasmodium*, seven canonical members of the CAF1/CCR4/NOT complex (NOT1, NOT2, NOT4, NOT5, CCR4, CAF1 and CAF40) could be identified, but whether they interact had not been tested (Table 1.1) [64]. The conservation of members of the CAF1/CCR4/NOT complex indicates that this complex will likely also have similar functions to the CAF1/CCR4/NOT complex of model eukaryotes. In most eukaryotes CAF1 is the most important deadenylase, and previous data demonstrated the essentiality of CAF1 in *P. falciparum*. Previously, Scott Lindner and others had demonstrated that, when the gene of an RNA-binding protein *pypuf2* was genetically deleted in sporozoites, the parasites became progressively less infectious in the salivary gland [34, 39, 40]. One of the four identified CCR4-domain containing proteins, CCR4-2, was one of the top transcripts that increased in abundance in salivary gland sporozoites with a deletion in *pypuf2* [34]. I hypothesized that PyCCR4-2 might be upregulated as a response to degrade transcripts that PyPUF2 would generally preserve, so that liver stage proteins were not expressed too early. Upon investigation and comparison of *pyccr4-2*$^{-}$ parasites with a wild-type reference, no differences in development throughout the life cycle were found. However, as *Plasmodium* has expanded this gene family to encode three other bioinformatically identifiable CCR4-domain containing proteins, it is possible that another of these proteins may play a role in the development or transmission of the parasite. The characterization of all of these CCR4-domain containing proteins, along with the characterization of CAF1, are presented in chapter three. CCR4-like proteins are known to have some specialized functions, and thus I thought that the four identifiable CCR4-domain-containing proteins might also have specialized functions [195]. While, all four of these CCR4 proteins are dispensable for blood stage development, and none are essential for progression through the *P. yoelii* life cycle, CCR4-1 plays important roles in sexual blood stage (Table 3.2, Chapter 3, [173]). I demonstrated that the bioinformatically predictable members of the CAf1/CCR4/NOT

complex can associate and form a core complex reminiscent to that of model eukaryotes (Table 3.1) [64]. The CAF1/CCR4/NOT complex is known to be present in the nucleus, and in P-bodies in cytosolic granules [104]. Utilizing the GFP tags on PyCCR4-1, and PyCCR4-2, I was able to show that both localize in the cytoplasm, and are found mostly in granules. These data demonstrate that PyCCR4-1 and PyCCR4-2 have similar localizations to that of CCR4 proteins in model eukaryotes.

Plasmodium host-to-vector transmission also requires intricate control of transcript abundance, including their translational repression and storage [14, 28, 31-33, 36]. Thus, the CAF1/CCR4/NOT complex may play an important role in gametocytes because, in model eukaryotes, it is intricately involved in RNA metabolism from the birth (Transcription) to death (Degradation) of a transcript [104]. I found that PyCCR4-1 is important for the development of the male gametocyte, as well as for the efficient transmission of gametocytes to the mosquito vector (Figure 3.6). These data suggest that PyCCR4-1 may be playing an important role in regulating gene expression in gametocytes. To address if PyCCR4-1 is playing this role, I performed total comparative RNA sequencing of gametocytes with a deletion of *pyccr4-1* and compared them to wild-type gametocytes. These data revealed that of the transcripts that changed, a majority decreased in abundance when *pyccr4-1* was absent. These data were counter to the expectation that PyCCR4-1 was acting primarily as a deadenylase to hasten their decay (Figure 3.11).

CCR4 proteins have an EEP (endonuclease, exonuclease, phosphatase) domain that has two defined catalytic residues that perform the nuclease activity [195]. The same experiments as above were carried out to determine if PyCCR4-1's putative catalytic residues had any effect on gametocyte development, with parasites that have both catalytic residues of PyCCR4-1 changed to alanine. These mutations should prevent the EEP domain from deadenylating any transcripts. I observed similar phenotypes in these catalytic dead variants (Figure 3.7). These data allude to another function of CCR4's in model eukaryotes, to repress translation, or that these effects may be from indirect effects of the *pyccr4-1* genetic deletion such a gene buffering [196].

CCR4 associates with the CAF1/CCR4/NOT complex through CAF1, and thus I wanted to understand if the association of PyCCR4-1 with PyCAF1, and by proxy, the remainder of the complex was important for its function. Work by others (*P. berghei*), and I (*P. yoelii*) showed that CAF1 was not able to be deleted from the genome [137, 173]. However, in the *P. falciparum PiggyBac* transposon forward genetic screen, a transposon inserted just after the CAF1 domain produced viable parasites. I appended a GFP tag onto the C-terminus of PyCAF1 where the *P. falciparum* transposon insertion had occurred. These data demonstrated that only the CAF1 domain is essential for asexual blood stage development. However, as defined in *P. falciparum* by the Adams lab and by my work in *P. yoelii*, parasites with truncated CAF1 proteins are still viable but have a slow asexual blood stage growth phenotype (Chapter 3, [137, 173]). Our collaborators demonstrated that only the CAF1 domain was expressed. Immunoprecipitation of this truncated PyCAF1 variant showed that it no longer associates with PyCCR4-1, and only weakly associates with PyNOT1. Interestingly, all that is necessary for CAF1 function in model eukaryotes appears to be this domain, yet in *P. falciparum and P. yoelii,* a phenotype is still present in these parasites. These data indicated that *Plasmodium* species might require this sequence C-terminal to the CAF1 domain to allow for CAF1 association with the remainder of the complex, or these sequences may be regulatory regions that affect its expression. If CCR4 can no longer associate with the complex, it may no longer be able to target the appropriate transcripts because many regulators of CCR4 activity do this through association with NOT1 [104]. Therefore, if CCR4 requires association with the complex to properly function, the CAF1 truncation will have a *pyccr4-1* deletion-like phenotype in addition to the already defined CAF1 disruption phenotypes (Chapter 3, [137, 173]). Interestingly, this parasite line also phenocopies parasites with a *pyccr4-1* deletion. These data indicate that PyCCR4-1 likely requires its association with CAF1 (and the rest of the complex) to retain some of its functions in gametocytes. This data adds to what was known already about the role of CAF1 in asexual blood stage development in *P. falciparum and P. yoelii*.

The influence that PyCCR4-1 has on transcript levels in gametocytes can be the result of direct effects, indirect effects, or both. It is known that CCR4 can translationally

repress transcripts directly and specifically with the aid of other components of the CAF1/CCR4/NOT complex [128, 129]. To determine if PyCCR4-1 associated directly with transcripts affected by its deletion, I performed immunoprecipitation followed by RT-PCR. I was able to demonstrate that PyCCR4-1 associates with some, but not all, of the transcripts that decrease in abundance in gametocytes with a *pyccr4-1* deletion**.** These data indicate that the effect PyCCR4-1 has on transcript abundance levels is likely a mixture of direct effects and indirect effects through its putative deadenylation and putative repressive functions. In model eukaryotes, the most well-known function of CCR4 is its ability to initiate the degradation of transcripts through the removal of the poly (A) tail. Tethering of CCR4 to a transcript can also cause translational repression. Indirect effects of CCR4 could include any of PyCCR4-1, or the complexes, direct effects on transcription factors/RNA metabolic factors that cause a domino effect upon downstream transcripts. Additionally, the parasite could modulate transcription/decay to reduce the effects of *pyccr4-1* deletion on parasite growth. Likely what was measured in my total comparative transcriptomic dataset is a combination of all of the above. To determine if any of the effects on transcript abundance could be direct, one can measure poly (A) tail length in the presence or absence of PyCCR4-1, and determine if the loss of PyCCR4-1 has any effect. To determine if PyCCR4-1 has an effect upon poly(A) tail length in these transcripts, Scott Lindner and I selected one control (GAPDH) that was not affected by the deletion of pyccr4-1 and does not associate with the complex by IP-RT PCR, and one that was affected and does associate with PyCCR4-1::GFP (P28). Utilizing circular RT PCR (cRT-PCR), Scott and I found that both of these transcripts retained their wild-type poly (A) tail lengths in gametocytes with a deletion of *pyccr4-1*. These data indicate that PyCCR4-1 is not acting on the poly (A) tail of these transcripts in gametocytes length to regulate its length.

In summary, in the absence of PyCCR4-1, fewer male gametocytes were produced, and host-to-vector transmission of parasites was decreased. I also observed that transcripts were not being correctly produced or stored in female gametocytes, and I observed no change in poly (A) length. Multiple different hypotheses could explain the observed reduction in transcript abundance with the deletion of *pyccr4-1*. First, that PyCCR4-1 is responsible for regulating a transcriptional activator or repressor, whose role is to

regulate expression of these transcripts. PyCCR4-1 could also be directly responsible for stabilizing some of these transcripts, and in its absence, they are degraded. Alternatively, gene buffering could be occurring to reduce the effect of the deletion of *pyccr4-1*. Together, I found that both a specialized and general member of the CAF1/CCR4/NOT complex have been adapted as specific regulators of *Plasmodium* host-to-vector transmission through their concerted effects upon specific, critical gametocyte transcripts. Given these data, I hypothesize that PyCCR4-1 is likely playing direct and indirect roles on transcripts, and could be altering both their stability and transcription to modulate gene expression. Currently, it is unclear if these phenotypes are due to direct involvement of PyCCR4-1 in gametocyte development, or just an effect of *pyccr4-1* deletion from the genome. However, data from my studies of PyNOT1-G (Chapter Four) suggest that the complex plays a direct role in gametocyte development.

My Contributions to the Field: NOT1-G

NOT1 is expressed from a single gene in model eukaryotes. However, I discovered that *Plasmodium* species have two genes that encode for two NOT1 paralogs (PY17X_0945600 (NOT1-G), PY17X_1027900 (NOT1)). Upon further investigation, this potential gene duplication event occurred within the common ancestor of the Aconoidasida class of apicomplexan parasites. These two *not1* genes are present in *Plasmodium*, *Theileria,* and *Babesia,* but absent from closely related species like *Toxoplasma*. All three of these species have similar host-to-vector transmission schemes, with the formation of transmissible stages in the red blood cell followed by ingestion by a vector (mosquito or tick). Interestingly, other closely related species like *Toxoplasma*, *Neospora* and *Cryptosporidium* do not have this insect vector transmission. This gene duplication may have evolved as a consequence of evolution to a differential transmission scheme. Inside the vector, fertilization and then development occurs. While host-to-vector transmission is very similar, all three of these species will invade different cell types when transmitted from vector-to-host: *Plasmodium* will invade hepatocytes, *Theileria* will invade lymphocytes, and *Babesia* will invade red blood cells. Specific control of RNA metabolism has been shown to be essential for *Plasmodium* host-to-vector transmission [14, 28, 31-33, 36]. Because NOT1 typically functions as a

scaffold to bring together core components with accessory factors to promote different RNA fates, this gene duplication may have occurred to add another layer of specificity to RNA metabolic control in the *Aconoidasida.* Because of the difficulty in deleting the coding sequence for PY17X_1027900 in asexual blood stage parasites, and the phenotype seen in gametocytes lacking the coding sequence for PY17X_0945600, I have named these two proteins PyNOT1 and PyNOT1-G, respectively (Figure 4.1).

PyNOT1 does not have a bioinformatically predictable TTP-binding domain, however PyNOT1-G does. Otherwise, they contain the same set of domains as canonical NOT1 proteins. While NOT1 proteins are typically essential, *pynot1-g* was deleted readily in *P. yoelii*. This data indicated that PyNOT1-G does not play an essential role in normal asexual blood stage development. When comparing parasites with a deletion of PyNOT1-G to wild-type parasites, I identified a mild asexual blood stage growth phenotype, where the parasites are unable to reach the same maximum parasitemia as wild-type parasites during a growth curve. This data indicates that PyNOT1-G likely plays some role in asexual blood stage growth; however, this has not yet been investigated further. Utilizing flow cytometry to count gametocytes, I was able to demonstrate that parasite clones with a deletion of *pynot1-g* produce virtually no mature male gametocytes (Figure 4.4). Moreover, these data were corroborated by a daily sampling of infected blood, where no mature male gametocytes that could activate to form gametes (as measured by exflagellation assays) were observed. I also attempted to transmit parasites with a deletion of *pynot1-g* and found that they cannot form oocysts in the mosquito. Thus, PyNOT1-G plays an essential role in the proper development of male gametocytes and is essential for host-to-vector transmission.

While male gametocytes were found to be either immature (PyDD-negative) or not produced, female gametocytes were not able to be physically characterized. Thus, I crossed colorless (no fluorescent proteins expressed) wild-type PY17XNL parasites with parasites with a deletion of *pynot1-g* that constitutively express GFP and do not develop mature male gametocytes. In this experiment, only colorless wild-type parasites productively infected the mosquito in both the wild-type x wild-type cross (control) and wild-type x *pynot1-g* $^{-}$ cross (experimental), indicating that both male and female

gametocytes with a *pynot1-g* gene deletion are sterile (Figure 4.5). These data demonstrate that PyNOT1-G plays an essential role in both male and female gametocyte maturation and/or fertility.

To determine how PyNOT1-G affects RNA metabolism in gametocytes, I performed total comparative RNA-seq on gametocytes with a deletion of PyNOT1-G and compared them to wild-type PY17XNL gametocytes. In the absence of PyNOT1-G, there was a massive dysregulation of transcript abundances, notably, those transcripts also affected by *dozi*$^{-}$ and *cith*$^{-}$ in *P. berghei*. In model eukaryotes, DHH1 (DDX6, DOZI) works with the CAF1/CCR4/NOT complex to impose translational repression. I propose that some of the effects on transcript abundance that I saw in gametocytes with a *pynot1-g* deletion is because DOZI cannot interact with PyNOT1-G to impose translational repression on this subset of transcripts [197].

The major difference between PyNOT1 and PyNOT1-G is the presence of a bioinformatically identifiable TTP-binding domain in PyNOT1-G. This domain could be interacting with RNA-binding proteins, as it does in model eukaryotes, and directing PyNOT1-G to specific transcripts. Interestingly, PyNOT1 lacks an identifiable TTP-binding domain and no other domains are identifiable in this region of the protein. If PyNOT1 and PyNOT1-G form distinct complexes, the repression/degradation of specific transcripts directed by TTP could be lost in the PyNOT1 complex. More likely an unknown protein-protein interaction domain is present in NOT1 that mediates these interactions. This other RNA-binding protein would then play the role of TTP.

In the context of model eukaryotes, my data suggests that there could be two NOT complexes playing distinct functions. It is difficult to determine what individual proteins are doing in the cell and in *Plasmodium* there are two NOT1 proteins that allow us to see distinct functions of each one. I hypothesize that this gene duplication occurred and allowed the NOT1-G complex to specialize its functions, in this case around host-to-vector transmission.

A potential model that integrates all of my findings with published work suggests that the CAF1/CCR4/NOT complex plays a wide array of roles in transcription and translation. Here I propose a model that suggests that PyNOT1-G works with DOZI (DHH1, DDX6)

to repress the translation of a large number of transcripts in female gametocytes (Figure 5.1). My data demonstrates an interaction of the complex with DOZI and an affect upon transcript abundances, but does not link this directly to the CAF1/CCR4/NOT complex-DOZI interaction. An orthologue of DOZI is known to interact with the CAF1/CCR4/NOT complex in other eukaryotes to regulate transcripts [198]. To make this claim direct DOZI to CAF1/CCR4/NOT complex interactions would need to be disrupted. Additionally, the complex could be purified and tested for translational repression/stabilization activity. In *Plasmodium*, this repression could occur in storage granules that lack decapping machinery and 5'-3' exonucleases like in *S. cerevisiae* [199]. While the absence of proteomics data does not mean that these proteins are absent, I identified no decapping machinery or 5'-3' exonucleases in my PyCCR4-1::GFP immunoprecipitation, which suggests that granules containing PyCCR4-1 may lack these components. I propose that when the PyCCR4-1 gene is deleted or the protein is made catalytically inactive, that RNA decay machinery is recruited to the *Plasmodium* storage granule. However, there is no evidence to suggest this other than the decay (or reduction of transcription) of transcripts in *pyccr4-1*$^{-}$ and dCCR4-1 parasites. Because NOT1 is known to be essential for granule formation, I propose that with the deletion of *pynot1-g,* transcripts destined for this granule are no longer stored and decay is uninhibited.

In conclusion, PyNOT1-G is essential for host-to-vector transmission because PyNOT1-G deficient parasites do not produce mature male or fertile female gametocytes. Many of the effects on total transcript abundances that I see resulting from *pyccr4-1* deletion are also seen in these analyses on the *pynot1-g* deletion. These data lead me to conclude that PyCCR4-1 is likely working with PyNOT1-G in gametocytes to perform at least a subset of its functions. Together these works characterize the scaffold and putative deadenylase subunits of the CAF1/CCR4/NOT complex in both asexual and gametocyte stage *Plasmodium* parasites. In a broader sense, these data provide more evidence of specific functions of the CAF1/CCR4/NOT complex. I believe that more work on members of this complex and in *P. falciparum*, where work on gametocytes is

performed more easily, are warranted to better define the functions of this complex in *Plasmodium* species.

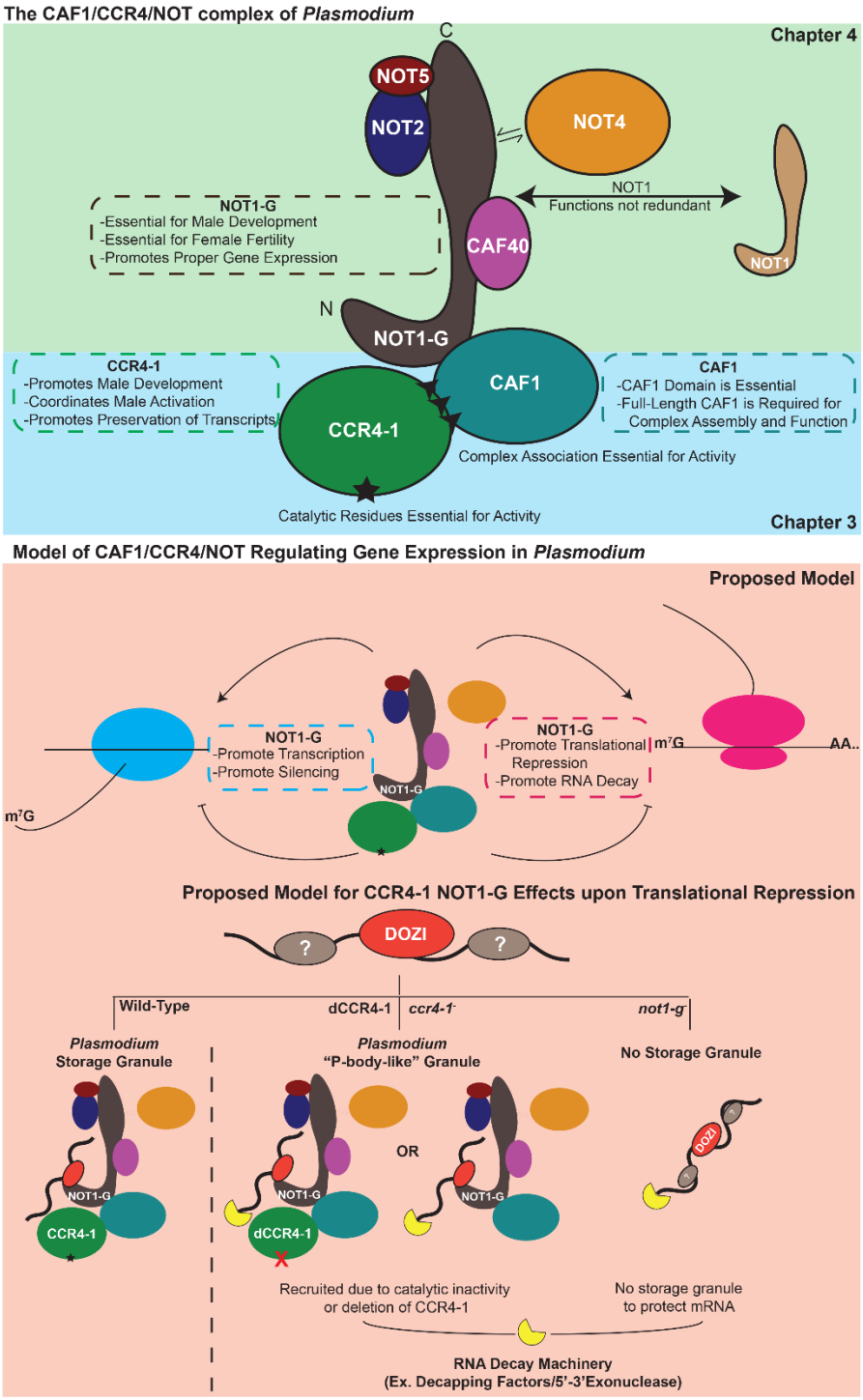


Figure 5.1: A Summation of Data from Chapters 3 and 4 with a Proposed Model Integrating My Findings to Data from Model Eukaryotes.

Future Directions of this work

Based on this work, several important lines of investigation should be pursued. First TTP, an RNA-binding protein that is responsible for directing the CAF1/CCR4/NOT complex to certain transcripts for translational repressor or decay, has a conserved binding site on NOT1-G. However, this protein is not bioinformatically identifiable in *Plasmodium*. Using IP/MS I have identified two candidate tristetraprolin proteins. Reverse genetics of these candidates will elucidate their functions as well as interaction partners to determine if they truly associate with this complex. If one of these candidates acts similarly to TTP from model eukaryotes I hypothesize it may act to regulate transcript abundances in gametocytes where NOT1-G is essential. Additionally, removal of the TTP-binding domain from NOT1-G will be useful to determine what affect this domain has on NOT1-G function and whether this interaction domain is essential for gametocyte development.

Second, work that is performed on the CAF1/CCR4/NOT complex in rodent infectious species of *Plasmodium* should be translated to human infectious strains. Additionally, many tools that cannot be used *in vivo* have been adapted for use in *in vitro P. falciparum* cultures. While *P. yoelii* is a great model for determining the effects of gene deletions on the entire life cycle, translation to *P. falciparum*, especially to characterize protein roles in the blood stages of development is critical to continued development of these projects. Some members of this complex have been identified to be essential (CAF40/CAF1), and the use of conditional gene expression systems that have been implemented in *P. falciparum* would be helpful to study these proteins. Additionally, *P. falciparum* takes much longer for gametocytes to develop. This feature of the parasite will allow for more precise dissection of the effects of CAF1/CCR4/NOT proteins that affect gametocyte commitment, development, and complete maturation. Study of NOT1-G in *P. falciparum* may elucidate its function in gametocyte maturation. If NOT1-G plays a similar role in *P. falciparum,* I would expect that gametocytes would not develop to maturity. I predict that male gametocytes with a deletion of *pfnot1-g* may not develop past a very early stage because I have no evidence of DDD expression in *P. yoelii* parasites with a *pynot1-g* deletion. DDD, according to data in *P. falciparum*, is

transcribed during the development of stage I gametocytes and is detected as a protein in gametocytes that are not staged [24, 28]. Female gametocytes may be able to develop longer but, because I do not measure female gametocytes directly, I cannot directly measure this. Instead, I have total transcriptomic data that demonstrate that either transcription (early in gametocyte development) or storage of these transcripts may be affected (anytime in gametocyte development). Utilizing methods to track transcription, which have been developed for use in *P. falciparum,* may further elucidate the function of the CAF1/CCR4/NOT complex with NOT1-G as the scaffold [24].

CAF40, which is usually dispensable, was shown to be essential in the *Plasmo*GEM screen and the *P. falciparum* transposon mutagenesis screen [104, 139-141]. CAF40's function in model eukaryotes is to recruit the CAF1/CCR4/NOT complex to mRNAs for degradation, and thus could be a central player in the function of this complex [104]. Because of the predicted essentiality of CAF40 during asexual development, the use of conditional knockdown systems in *P. falciparum* may help to determine what roles PfCAF40 has in development [200]. Additionally, methods to determine what RNA-binding proteins CAF40 may be interacting with will potentially further our knowledge of the essential proteins for asexual blood stage growth. I hypothesize that PfCAF40 is playing an essential role in controlling RNA metabolism during the normal asexual stage development and without it, proteins are expressed at the wrong time, resulting in the death of the parasites.

Using the advantages of working in the rodent model *P. yoelii,* one could characterize other components of the CAF1/CCR4/NOT complex that are not essential for blood stage development. While they were not characterized in the *P. berghei Plasmo*GEM dataset, *not2, not4*, and *not5* were targeted and determined to be non-essential in the *P. falciparum piggyBac* transposon mutagenesis forward genetic screen. NOT2, NOT4, and NOT5 are all known to play roles in transcriptional regulation. One of the significant advantages of working in *P. yoelii* is the ability to complete the entire parasite life cycle *in vivo* and cost-effectively. Characterization of genetic deletions of *pynot2*, *pynot4*, and *pynot5* can then be performed and the proteins function can be characterized across the entire life cycle, and it can be determined if these play an essential role at any point

in parasite development. Additional characterization of their effects on transcription could then be carried out in the identified life stage.

Additionally, using CRISPR/CAS9 technology we can now perform marker-less genetic deletions. Using CRISPR will allow us to perform multiple deletions in the same parasite line, for example attempting to delete both *not1* genes from *Plasmodium* species. We could also eliminate multiple CCR4-domain containing proteins at a time to determine if these proteins are playing redundant roles. Additionally, we could use this to add functionalities to proteins. The addition of the TTP-binding domain from NOT1-G to NOT1 may allow NOT1+TTP to play similar roles as NOT1-G in gametocytogenesis. These techniques can and should be used to further characterize this complex in the future.

The CAF1/CCR4/NOT complex is an essential complex that helps to regulate transcript abundances through transcription, translational repression, and decay. Presented here are just a few of the roles this complex plays in *Plasmodium* development. In chapter three, PyCCR4-1 was shown to be important for proper gametocyte development and host-to-vector transmission, while CAF1 is likely essential for asexual development and serves as a bridge to connect PyCCR4-1 to the CAF1/CCR4/NOT complex. In chapter four, PyNOT1-G was shown to be essential for host-to-vector transmission, and its deletion affects the abundance of hundreds of transcripts in gametocytes. The work that I have proposed in chapter five will deepen our understanding of critical members of this complex and will serve as a launch point for additional studies unraveling the complex web of functions that underlie its essentiality.

Bibliography

1. *World Malaria Report*. 2017, World Health Organization. p. 196.
2. Antinori, S., et al., *Biology of human malaria plasmodia including Plasmodium knowlesi.* Mediterr J Hematol Infect Dis, 2012. **4**(1): p. e2012013.
3. Ritchie, M.R.a.H. *Malaria*. 2018; Available from: https://ourworldindata.org/malaria.
4. Vaughan, A.M., et al., *Development of humanized mouse models to study human malaria parasite infection.* Future Microbiol, 2012. **7**(5): p. 657-65.
5. Aurrecoechea, C., et al., *PlasmoDB: a functional genomic database for malaria parasites.* Nucleic Acids Res, 2009. **37**(Database issue): p. D539-43.
6. Burda, P.C., R. Caldelari, and V.T. Heussler, *Manipulation of the Host Cell Membrane during Plasmodium Liver Stage Egress.* MBio, 2017. **8**(2).
7. Graewe, S., et al., *Hostile takeover by Plasmodium: reorganization of parasite and host cell membranes during liver stage egress.* PLoS Pathog, 2011. **7**(9): p. e1002224.
8. Sturm, A., et al., *Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids.* Science, 2006. **313**(5791): p. 1287-90.
9. Lindner, S.E., J.L. Miller, and S.H. Kappe, *Malaria parasite pre-erythrocytic infection: preparation meets opportunity.* Cell Microbiol, 2012. **14**(3): p. 316-24.
10. Baer, K., et al., *Release of hepatic Plasmodium yoelii merozoites into the pulmonary microvasculature.* PLoS Pathog, 2007. **3**(11): p. e171.
11. Liu, Z., J. Miao, and L. Cui, *Gametocytogenesis in malaria parasite: commitment, development and regulation.* Future Microbiol, 2011. **6**(11): p. 1351-69.
12. Killick-Kendrick, R. and M. Warren, *Primary exoerythrocytic schizonts of a mammalian Plasmodium as a source of gametocytes.* Nature, 1968. **220**(5163): p. 191-2.
13. James, S.P., W.D. Nicol, and P.G. Shute, *Clinical and Parasitological Observations on Induced Malaria: (Section of Tropical Diseases and Parasitology).* Proc R Soc Med, 1936. **29**(8): p. 879-94.
14. Mair, G.R., et al., *Regulation of sexual development of Plasmodium by translational repression.* Science, 2006. **313**(5787): p. 667-9.
15. Silvestrini, F., P. Alano, and J.L. Williams, *Commitment to the production of male and female gametocytes in the human malaria parasite Plasmodium falciparum.* Parasitology, 2000. **121 Pt 5**: p. 465-71.
16. Smith, T.G., et al., *Commitment to sexual differentiation in the human malaria parasite, Plasmodium falciparum.* Parasitology, 2000. **121 (Pt 2)**: p. 127-33.
17. Bruce, M.C., et al., *Commitment of the malaria parasite Plasmodium falciparum to sexual and asexual development.* Parasitology, 1990. **100 Pt 2**: p. 191-200.
18. Paul, R.E., et al., *Sex determination in malaria parasites.* Science, 2000. **287**(5450): p. 128-31.
19. Kafsack, B.F., et al., *A transcriptional switch underlies commitment to sexual development in malaria parasites.* Nature, 2014. **507**(7491): p. 248-52.
20. Coleman, B.I., et al., *A Plasmodium falciparum histone deacetylase regulates antigenic variation and gametocyte conversion.* Cell Host Microbe, 2014. **16**(2): p. 177-186.
21. Filarsky, M., et al., *GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing.* Science, 2018. **359**(6381): p. 1259-1263.
22. Smith, R.C., J. Vega-Rodriguez, and M. Jacobs-Lorena, *The Plasmodium bottleneck: malaria parasite losses in the mosquito vector.* Mem Inst Oswaldo Cruz, 2014. **109**(5): p. 644-61.
23. Arrighi, R.B. and H. Hurd, *The role of Plasmodium berghei ookinete proteins in binding to basal lamina components and transformation into oocysts.* Int J Parasitol, 2002. **32**(1): p. 91-8.

24. Painter, H.J., M. Carrasquilla, and M. Llinas, *Capturing in vivo RNA transcriptional dynamics from the malaria parasite Plasmodium falciparum.* Genome Res, 2017. **27**(6): p. 1074-1086.
25. Turque, O., et al., *Translational Repression in Malaria Sporozoites.* Microb Cell, 2016. **3**(5): p. 227-229.
26. Silva, P.A., et al., *Translational Control of UIS4 Protein of the Host-Parasite Interface Is Mediated by the RNA Binding Protein Puf2 in Plasmodium berghei Sporozoites.* PLoS One, 2016. **11**(1): p. e0147940.
27. Shrestha, S., et al., *The RNA-binding protein Puf1 functions in the maintenance of gametocytes in Plasmodium falciparum.* J Cell Sci, 2016. **129**(16): p. 3144-52.
28. Lasonder, E., et al., *Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression.* Nucleic Acids Res, 2016. **44**(13): p. 6087-101.
29. Josling, G.A. and M. Llinas, *Sexual development in Plasmodium parasites: knowing when it's time to commit.* Nat Rev Microbiol, 2015. **13**(9): p. 573-87.
30. Sinha, A., et al., *A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium.* Nature, 2014. **507**(7491): p. 253-257.
31. Guerreiro, A., et al., *Genome-wide RIP-Chip analysis of translational repressor-bound mRNAs in the Plasmodium gametocyte.* Genome Biol, 2014. **15**(11): p. 493.
32. Saeed, S., et al., *Translational repression controls temporal expression of the Plasmodium berghei LCCL protein complex.* Mol Biochem Parasitol, 2013. **189**(1-2): p. 38-42.
33. Miao, J., et al., *Puf mediates translation repression of transmission-blocking vaccine candidates in malaria parasites.* PLoS Pathog, 2013. **9**(4): p. e1003268.
34. Lindner, S.E., et al., *Perturbations of Plasmodium Puf2 expression and RNA-seq of Puf2-deficient sporozoites reveal a critical role in maintaining RNA homeostasis and parasite transmissibility.* Cell Microbiol, 2013. **15**(7): p. 1266-83.
35. Miao, J., et al., *The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite Plasmodium falciparum.* J Cell Sci, 2010. **123**(Pt 7): p. 1039-49.
36. Mair, G.R., et al., *Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development.* PLoS Pathog, 2010. **6**(2): p. e1000767.
37. Munoz, E.E., et al., *ALBA4 modulates its stage-specific interactions and specific mRNA fates during Plasmodium yoelii growth and transmission.* Mol Microbiol, 2017.
38. Cui, L., S. Lindner, and J. Miao, *Translational regulation during stage transitions in malaria parasites.* Ann N Y Acad Sci, 2015. **1342**: p. 1-9.
39. Muller, K., K. Matuschewski, and O. Silvie, *The Puf-family RNA-binding protein Puf2 controls sporozoite conversion to liver stages in the malaria parasite.* PLoS One, 2011. **6**(5): p. e19860.
40. Gomes-Santos, C.S., et al., *Transition of Plasmodium sporozoites into liver stage-like forms is regulated by the RNA binding protein Pumilio.* PLoS Pathog, 2011. **7**(5): p. e1002046.
41. Painter, H.J., T.L. Campbell, and M. Llinas, *The Apicomplexan AP2 family: integral factors regulating Plasmodium development.* Mol Biochem Parasitol, 2011. **176**(1): p. 1-7.
42. Campbell, T.L., et al., *Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite.* PLoS Pathog, 2010. **6**(10): p. e1001165.
43. Lambert, S.A., et al., *The Human Transcription Factors.* Cell, 2018. **172**(4): p. 650-665.
44. Yuda, M., et al., *Global transcriptional repression: An initial and essential step for Plasmodium sexual development.* Proc Natl Acad Sci U S A, 2015. **112**(41): p. 12824-9.
45. Zhang, C., et al., *Systematic CRISPR-Cas9-Mediated Modifications of Plasmodium yoelii ApiAP2 Genes Reveal Functional Insights into Parasite Development.* MBio, 2017. **8**(6).

46. Lopez-Rubio, J.J., L. Mancio-Silva, and A. Scherf, *Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites.* Cell Host Microbe, 2009. **5**(2): p. 179-90.
47. Becker, J.S., D. Nicetto, and K.S. Zaret, *H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes.* Trends Genet, 2016. **32**(1): p. 29-41.
48. Schneider, R. and R. Grosschedl, *Dynamics and interplay of nuclear architecture, genome organization, and gene expression.* Genes Dev, 2007. **21**(23): p. 3027-43.
49. Soufi, A., G. Donahue, and K.S. Zaret, *Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome.* Cell, 2012. **151**(5): p. 994-1004.
50. 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, 2009. **106**(24): p. 9655-60.
51. Flueck, C., et al., *Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors.* PLoS Pathog, 2009. **5**(9): p. e1000569.
52. Voss, T.S., Z. Bozdech, and R. Bartfai, *Epigenetic memory takes center stage in the survival strategy of malaria parasites.* Curr Opin Microbiol, 2014. **20**: p. 88-95.
53. Eksi, S., et al., *Plasmodium falciparum gametocyte development 1 (Pfgdv1) and gametocytogenesis early gene identification and commitment to sexual development.* PLoS Pathog, 2012. **8**(10): p. e1002964.
54. 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, 2017. **7**(1): p. 607.
55. Karmodiya, K., et al., *A comprehensive epigenome map of Plasmodium falciparum reveals unique mechanisms of transcriptional regulation and identifies H3K36me2 as a global mark of gene suppression.* Epigenetics Chromatin, 2015. **8**: p. 32.
56. Modrzynska, K., et al., *A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle.* Cell Host Microbe, 2017. **21**(1): p. 11-22.
57. Yuda, M., et al., *Transcription factor AP2-Sp and its target genes in malarial sporozoites.* Mol Microbiol, 2010. **75**(4): p. 854-63.
58. Yuda, M., et al., *Identification of a transcription factor in the mosquito-invasive stage of malaria parasites.* Mol Microbiol, 2009. **71**(6): p. 1402-14.
59. De Silva, E.K., et al., *Specific DNA-binding by apicomplexan AP2 transcription factors.* Proc Natl Acad Sci U S A, 2008. **105**(24): p. 8393-8.
60. Martins, R.M., et al., *An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite Plasmodium falciparum.* Sci Rep, 2017. **7**(1): p. 14042.
61. Iwanaga, S., et al., *Identification of an AP2-family protein that is critical for malaria liver stage development.* PLoS One, 2012. **7**(11): p. e47557.
62. Zhang, M., et al., *PK4, a eukaryotic initiation factor 2alpha(eIF2alpha) kinase, is essential for the development of the erythrocytic cycle of Plasmodium.* Proc Natl Acad Sci U S A, 2012. **109**(10): p. 3956-61.
63. Bunnik, E.M., et al., *The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum.* Genome Biol, 2016. **17**(1): p. 147.
64. Reddy, B.P., et al., *A bioinformatic survey of RNA-binding proteins in Plasmodium.* BMC Genomics, 2015. **16**: p. 890.
65. Tsvetanova, N.G., et al., *Proteome-wide search reveals unexpected RNA-binding proteins in Saccharomyces cerevisiae.* PLoS One, 2010. **5**(9).

66. Malhotra, S. and R. Sowdhamini, *Sequence search and analysis of gene products containing RNA recognition motifs in the human genome.* BMC Genomics, 2014. **15**: p. 1159.
67. Paton, M.G., et al., *Structure and expression of a post-transcriptionally regulated malaria gene encoding a surface protein from the sexual stages of Plasmodium berghei.* Mol Biochem Parasitol, 1993. **59**(2): p. 263-75.
68. Parker, R. and U. Sheth, *P bodies and the control of mRNA translation and degradation.* Mol Cell, 2007. **25**(5): p. 635-46.
69. Painter, H.J., et al., *Genome-wide real-time in vivo transcriptional dynamics during Plasmodium falciparum blood-stage development.* Nat Commun, 2018. **9**(1): p. 2656.
70. Lindner, S.E., et al., *Total and putative surface proteomics of malaria parasite salivary gland sporozoites.* Mol Cell Proteomics, 2013. **12**(5): p. 1127-43.
71. Banani, S.F., et al., *Biomolecular condensates: organizers of cellular biochemistry.* Nat Rev Mol Cell Biol, 2017. **18**(5): p. 285-298.
72. Nadezhdina, E.S., et al., *Microtubules govern stress granule mobility and dynamics.* Biochim Biophys Acta, 2010. **1803**(3): p. 361-71.
73. Chernov, K.G., et al., *Role of microtubules in stress granule assembly: microtubule dynamical instability favors the formation of micrometric stress granules in cells.* J Biol Chem, 2009. **284**(52): p. 36569-80.
74. Loschi, M., et al., *Dynein and kinesin regulate stress-granule and P-body dynamics.* J Cell Sci, 2009. **122**(Pt 21): p. 3973-82.
75. Sheinberger, J. and Y. Shav-Tal, *mRNPs meet stress granules.* FEBS Lett, 2017. **591**(17): p. 2534-2542.
76. Luo, Y., Z. Na, and S.A. Slavoff, *P-Bodies: Composition, Properties, and Functions.* Biochemistry, 2018. **57**(17): p. 2424-2431.
77. Buchan, J.R. and R. Parker, *Eukaryotic stress granules: the ins and outs of translation.* Mol Cell, 2009. **36**(6): p. 932-41.
78. Protter, D.S. and R. Parker, *Principles and Properties of Stress Granules.* Trends Cell Biol, 2016. **26**(9): p. 668-79.
79. Nover, L., K.D. Scharf, and D. Neumann, *Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves.* Mol Cell Biol, 1983. **3**(9): p. 1648-55.
80. Arrigo, A.P., J.P. Suhan, and W.J. Welch, *Dynamic changes in the structure and intracellular locale of the mammalian low-molecular-weight heat shock protein.* Mol Cell Biol, 1988. **8**(12): p. 5059-71.
81. Khong, A., et al., *The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules.* Mol Cell, 2017. **68**(4): p. 808-820 e5.
82. Buchan, J.R., D. Muhlrad, and R. Parker, *P bodies promote stress granule assembly in Saccharomyces cerevisiae.* J Cell Biol, 2008. **183**(3): p. 441-55.
83. Thedieck, K., et al., *Inhibition of mTORC1 by astrin and stress granules prevents apoptosis in cancer cells.* Cell, 2013. **154**(4): p. 859-74.
84. Takahara, T. and T. Maeda, *Transient sequestration of TORC1 into stress granules during heat stress.* Mol Cell, 2012. **47**(2): p. 242-52.
85. Arimoto, K., et al., *Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways.* Nat Cell Biol, 2008. **10**(11): p. 1324-32.
86. Cipolat Mis, M.S., et al., *Autophagy in motor neuron disease: Key pathogenetic mechanisms and therapeutic targets.* Mol Cell Neurosci, 2016. **72**: p. 84-90.
87. Li, S., et al., *Genetic interaction of hnRNPA2B1 and DNAJB6 in a Drosophila model of multisystem proteinopathy.* Hum Mol Genet, 2016. **25**(5): p. 936-50.

88. Walters, R.W., et al., *Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in Saccharomyces cerevisiae.* RNA, 2015. **21**(9): p. 1660-71.
89. Ramaswami, M., J.P. Taylor, and R. Parker, *Altered ribostasis: RNA-protein granules in degenerative disorders.* Cell, 2013. **154**(4): p. 727-36.
90. Li, Y.R., et al., *Stress granules as crucibles of ALS pathogenesis.* J Cell Biol, 2013. **201**(3): p. 361-72.
91. Kedersha, N., et al., *Stress granules and processing bodies are dynamically linked sites of mRNP remodeling.* J Cell Biol, 2005. **169**(6): p. 871-84.
92. Chen, C.Y. and A.B. Shyu, *Mechanisms of deadenylation-dependent decay.* Wiley Interdiscip Rev RNA, 2011. **2**(2): p. 167-83.
93. Chen, C.Y. and A.B. Shyu, *Deadenylation and P-bodies.* Adv Exp Med Biol, 2013. **768**: p. 183-95.
94. Muhlrad, D., C.J. Decker, and R. Parker, *Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript.* Genes Dev, 1994. **8**(7): p. 855-66.
95. Teixeira, D., et al., *Processing bodies require RNA for assembly and contain nontranslating mRNAs.* RNA, 2005. **11**(4): p. 371-82.
96. Sheth, U. and R. Parker, *Decapping and decay of messenger RNA occur in cytoplasmic processing bodies.* Science, 2003. **300**(5620): p. 805-8.
97. Coller, J.M., et al., *The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes.* RNA, 2001. **7**(12): p. 1717-27.
98. Collart, M.A. and O.O. Panasenko, *The Ccr4-Not Complex: Architecture and Structural Insights.* Subcell Biochem, 2017. **83**: p. 349-379.
99. Ukleja, M., et al., *The architecture of the Schizosaccharomyces pombe CCR4-NOT complex.* Nat Commun, 2016. **7**: p. 10433.
100. Xu, K., et al., *Insights into the structure and architecture of the CCR4-NOT complex.* Front Genet, 2014. **5**: p. 137.
101. Petit, A.P., et al., *The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex.* Nucleic Acids Res, 2012. **40**(21): p. 11058-72.
102. Bhaskar, V., et al., *Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex.* Nat Struct Mol Biol, 2013. **20**(11): p. 1281-8.
103. Collart, M.A. and O.O. Panasenko, *The Ccr4--not complex.* Gene, 2012. **492**(1): p. 42-53.
104. Collart, M.A., *The Ccr4-Not complex is a key regulator of eukaryotic gene expression.* Wiley Interdiscip Rev RNA, 2016. **7**(4): p. 438-54.
105. Dutta, A., et al., *Ccr4-Not and TFIIS Function Cooperatively To Rescue Arrested RNA Polymerase II.* Mol Cell Biol, 2015. **35**(11): p. 1915-25.
106. Peng, W., et al., *Regulators of cellular levels of histone acetylation in Saccharomyces cerevisiae.* Genetics, 2008. **179**(1): p. 277-89.
107. Mersman, D.P., et al., *Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression.* Genes Dev, 2009. **23**(8): p. 951-62.
108. Mulder, K.W., et al., *Regulation of histone H3K4 tri-methylation and PAF complex recruitment by the Ccr4-Not complex.* Nucleic Acids Res, 2007. **35**(7): p. 2428-39.
109. Laribee, R.N., et al., *CCR4/NOT complex associates with the proteasome and regulates histone methylation.* Proc Natl Acad Sci U S A, 2007. **104**(14): p. 5836-41.
110. Venters, B.J., et al., *A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces.* Mol Cell, 2011. **41**(4): p. 480-92.
111. Hu, G., et al., *A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal.* Genes Dev, 2009. **23**(7): p. 837-48.

112. Zwartjes, C.G., et al., *Repression of promoter activity by CNOT2, a subunit of the transcription regulatory Ccr4-not complex.* J Biol Chem, 2004. **279**(12): p. 10848-54.
113. Swanson, M.J., et al., *A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo.* Mol Cell Biol, 2003. **23**(8): p. 2800-20.
114. Deluen, C., et al., *The Ccr4-not complex and yTAF1 (yTaf(II)130p/yTaf(II)145p) show physical and functional interactions.* Mol Cell Biol, 2002. **22**(19): p. 6735-49.
115. Winkler, G.S., et al., *Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription.* EMBO J, 2006. **25**(13): p. 3089-99.
116. Haas, M., et al., *c-Myb protein interacts with Rcd-1, a component of the CCR4 transcription mediator complex.* Biochemistry, 2004. **43**(25): p. 8152-9.
117. Rodriguez-Gil, A., et al., *The CCR4-NOT complex contributes to repression of Major Histocompatibility Complex class II transcription.* Sci Rep, 2017. **7**(1): p. 3547.
118. Yang, C.Y., et al., *Interaction of CCR4-NOT with EBF1 regulates gene-specific transcription and mRNA stability in B lymphopoiesis.* Genes Dev, 2016. **30**(20): p. 2310-2324.
119. Sun, H.Y., et al., *Protein Degradation of RNA Polymerase II-Association Factor 1(PAF1) Is Controlled by CNOT4 and 26S Proteasome.* PLoS One, 2015. **10**(5): p. e0125599.
120. Gulshan, K., B. Thommandru, and W.S. Moye-Rowley, *Proteolytic degradation of the Yap1 transcription factor is regulated by subcellular localization and the E3 ubiquitin ligase Not4.* J Biol Chem, 2012. **287**(32): p. 26796-805.
121. Kruk, J.A., et al., *The multifunctional Ccr4-Not complex directly promotes transcription elongation.* Genes Dev, 2011. **25**(6): p. 581-93.
122. Villanyi, Z., et al., *The Not5 subunit of the ccr4-not complex connects transcription and translation.* PLoS Genet, 2014. **10**(10): p. e1004569.
123. Babbarwal, V., J. Fu, and J.C. Reese, *The Rpb4/7 module of RNA polymerase II is required for carbon catabolite repressor protein 4-negative on TATA (Ccr4-not) complex to promote elongation.* J Biol Chem, 2014. **289**(48): p. 33125-30.
124. Gupta, I., et al., *Translational Capacity of a Cell Is Determined during Transcription Elongation via the Ccr4-Not Complex.* Cell Rep, 2016. **15**(8): p. 1782-94.
125. Miller, J.E. and J.C. Reese, *Ccr4-Not complex: the control freak of eukaryotic cells.* Crit Rev Biochem Mol Biol, 2012. **47**(4): p. 315-33.
126. Yi, H., et al., *PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay.* Mol Cell, 2018. **70**(6): p. 1081-1088 e5.
127. Webster, M.W., et al., *mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases.* Mol Cell, 2018. **70**(6): p. 1089-1100 e8.
128. Duy, D.L., Y. Suda, and K. Irie, *Cytoplasmic deadenylase Ccr4 is required for translational repression of LRG1 mRNA in the stationary phase.* PLoS One, 2017. **12**(2): p. e0172476.
129. Cooke, A., A. Prigge, and M. Wickens, *Translational repression by deadenylases.* J Biol Chem, 2010. **285**(37): p. 28506-13.
130. Bhandari, D., et al., *Structural basis for the Nanos-mediated recruitment of the CCR4-NOT complex and translational repression.* Genes Dev, 2014. **28**(8): p. 888-901.
131. Van Etten, J., et al., *Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs.* J Biol Chem, 2012. **287**(43): p. 36370-83.
132. Joly, W., et al., *The CCR4 deadenylase acts with Nanos and Pumilio in the fine-tuning of Mei-P26 expression to promote germline stem cell self-renewal.* Stem Cell Reports, 2013. **1**(5): p. 411-24.
133. Sgromo, A., et al., *Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets.* RNA, 2018. **24**(3): p. 381-395.
134. Ito, K., et al., *The role of the CNOT1 subunit of the CCR4-NOT complex in mRNA deadenylation and cell viability.* Protein Cell, 2011. **2**(9): p. 755-63.

135. Sandler, H., et al., *Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin.* Nucleic Acids Res, 2011. **39**(10): p. 4373-86.
136. Panasenko, O.O. and M.A. Collart, *Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29.* Mol Cell Biol, 2011. **31**(8): p. 1610-23.
137. Balu, B., et al., *CCR4-associated factor 1 coordinates the expression of Plasmodium falciparum egress and invasion proteins.* Eukaryot Cell, 2011. **10**(9): p. 1257-63.
138. Balu, B., et al., *A genetic screen for attenuated growth identifies genes crucial for intraerythrocytic development of Plasmodium falciparum.* PLoS One, 2010. **5**(10): p. e13282.
139. Zhang, M., et al., *Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis.* Science, 2018. **360**(6388).
140. Schwach, F., et al., *PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites.* Nucleic Acids Res, 2015. **43**(Database issue): p. D1176-82.
141. Bushell, E., et al., *Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes.* Cell, 2017. **170**(2): p. 260-272 e8.
142. Reddy, B.P., et al., *A bioinformatic survey of RNA-binding proteins in Plasmodium.* BMC Genomics, 2015. **16**(1): p. 890.
143. Guntur, A.R., et al., *An essential role for the circadian-regulated gene nocturnin in osteogenesis: the importance of local timekeeping in skeletal homeostasis.* Ann N Y Acad Sci, 2011. **1237**: p. 58-63.
144. Mikolajczak, S.A., et al., *An efficient strategy for gene targeting and phenotypic assessment in the Plasmodium yoelii rodent malaria model.* Mol Biochem Parasitol, 2008. **158**(2): p. 213-6.
145. Jongco, A.M., et al., *Improved transfection and new selectable markers for the rodent malaria parasite Plasmodium yoelii.* Mol Biochem Parasitol, 2006. **146**(2): p. 242-50.
146. *Methods in Malaria Research*, A.K. Kristen Moll, Artur Scherf, Mats Wahlgren, Editor. 2013. p. 474.
147. Carter, R., *The Culture and Preparation of Gametocytes of Plasmodium falciparum for Immunochemical, Molecular, and Mosquito Infectivity Studies*, in *Protocols in Molecular Parasitology*, J.E. Hyde, Editor. 1993. p. 67-88.
148. Minns, A.M., et al., *Nuclear, Cytosolic, and Surface-Localized Poly(A)-Binding Proteins of Plasmodium yoelii.* mSphere, 2018. **3**(1).
149. Miller, J.L., et al., *Plasmodium yoelii macrophage migration inhibitory factor is necessary for efficient liver-stage development.* Infect Immun, 2012. **80**(4): p. 1399-407.
150. Mikolajczak, S.A., et al., *Plasmodium vivax liver stage development and hypnozoite persistence in human liver-chimeric mice.* Cell Host Microbe, 2015. **17**(4): p. 526-35.
151. Deutsch, E.W., et al., *Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics.* Proteomics Clin Appl, 2015. **9**(7-8): p. 745-54.
152. Kessner, D., et al., *ProteoWizard: open source software for rapid proteomics tools development.* Bioinformatics, 2008. **24**(21): p. 2534-6.
153. Craig, R. and R.C. Beavis, *TANDEM: matching proteins with tandem mass spectra.* Bioinformatics, 2004. **20**(9): p. 1466-7.
154. Eng, J.K., T.A. Jahan, and M.R. Hoopmann, *Comet: an open-source MS/MS sequence database search tool.* Proteomics, 2013. **13**(1): p. 22-4.
155. Mellacheruvu, D., et al., *The CRAPome: a contaminant repository for affinity purification-mass spectrometry data.* Nat Methods, 2013. **10**(8): p. 730-6.
156. Shteynberg, D., et al., *iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates.* Mol Cell Proteomics, 2011. **10**(12): p. M111 007690.

157. Choi, H., et al., *SAINT: probabilistic scoring of affinity purification-mass spectrometry data.* Nat Methods, 2011. **8**(1): p. 70-3.
158. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.
159. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
160. Team, R.C., *R: A language and environment for statistical computing.* 2016, R Foundation for Statistical Computing, Vienna, Austria. .
161. Vizcaino, J.A., et al., *2016 update of the PRIDE database and its related tools.* Nucleic Acids Res, 2016. **44**(D1): p. D447-56.
162. Basquin, J., et al., *Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction.* Mol Cell, 2012. **48**(2): p. 207-18.
163. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.
164. Temme, C., et al., *Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation.* RNA, 2010. **16**(7): p. 1356-70.
165. Kerr, S.C., et al., *The Ccr4-Not complex interacts with the mRNA export machinery.* PLoS One, 2011. **6**(3): p. e18302.
166. Jain, S., et al., *ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure.* Cell, 2016. **164**(3): p. 487-98.
167. Guo, L., et al., *Nuclear-Import Receptors Reverse Aberrant Phase Transitions of RNA-Binding Proteins with Prion-like Domains.* Cell, 2018. **173**(3): p. 677-692 e20.
168. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10**(6): p. 845-58.
169. States, D.J. and W. Gish, *Combined use of sequence similarity and codon bias for coding region identification.* J Comput Biol, 1994. **1**(1): p. 39-50.
170. Sievers, F. and D.G. Higgins, *Clustal omega.* Curr Protoc Bioinformatics, 2014. **48**: p. 3 13 1-16.
171. Munoz, E.E., et al., *ALBA4 modulates its stage-specific interactions and specific mRNA fates during Plasmodium yoelii growth and transmission.* Mol Microbiol, 2017. **106**(2): p. 266-284.
172. Dupressoir, A., et al., *Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding.* BMC Genomics, 2001. **2**: p. 9.
173. Hart, K.J., et al., *Plasmodium male gametocyte development and transmission are critically regulated by general and transmission-specific members of the CAF1/CCR4/NOT complex.* bioRxiv, 2018.
174. Anders, S. and W. Huber, *Differential expression analysis for sequence count data.* Genome Biol, 2010. **11**(10): p. R106.
175. Talman, A.M., et al., *Proteomic analysis of the Plasmodium male gamete reveals the key role for glycolysis in flagellar motility.* Malar J, 2014. **13**: p. 315.
176. Tremp, A.Z., et al., *LCCL protein complex formation in Plasmodium is critically dependent on LAP1.* Mol Biochem Parasitol, 2017. **214**: p. 87-90.
177. Talman, A.M., et al., *PbGEST mediates malaria transmission to both mosquito and vertebrate host.* Mol Microbiol, 2011. **82**(2): p. 462-74.
178. Holmes-McNary, M.Q., A.S. Baldwin, Jr., and S.H. Zeisel, *Opposing regulation of choline deficiency-induced apoptosis by p53 and nuclear factor kappaB.* J Biol Chem, 2001. **276**(44): p. 41197-204.
179. Gissot, M., et al., *High mobility group protein HMGB2 is a critical regulator of plasmodium oocyst development.* J Biol Chem, 2008. **283**(25): p. 17030-8.

180. Ikadai, H., et al., *Transposon mutagenesis identifies genes essential for Plasmodium falciparum gametocytogenesis.* Proc Natl Acad Sci U S A, 2013. **110**(18): p. E1676-84.
181. Farber, V., et al., *Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs.* Nucleic Acids Res, 2013. **41**(2): p. 1211-22.
182. Young, J.A., et al., *The Plasmodium falciparum sexual development transcriptome: a microarray analysis using ontology-based pattern identification.* Mol Biochem Parasitol, 2005. **143**(1): p. 67-79.
183. Yeoh, L.M., et al., *Comparative transcriptomics of female and male gametocytes in Plasmodium berghei and the evolution of sex in alveolates.* BMC Genomics, 2017. **18**(1): p. 734.
184. Otto, T.D., et al., *A comprehensive evaluation of rodent malaria parasite genomes and gene expression.* BMC Biol, 2014. **12**: p. 86.
185. Lopez-Barragan, M.J., et al., *Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum.* BMC Genomics, 2011. **12**: p. 587.
186. Gaillard, H., et al., *Genome-wide analysis of factors affecting transcription elongation and DNA repair: a new role for PAF and Ccr4-not in transcription-coupled repair.* PLoS Genet, 2009. **5**(2): p. e1000364.
187. Waghray, S., et al., *Xenopus CAF1 requires NOT1-mediated interaction with 4E-T to repress translation in vivo.* RNA, 2015. **21**(7): p. 1335-45.
188. Nishimura, T., et al., *The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs.* Cell Rep, 2015. **11**(9): p. 1425-36.
189. Kamenska, A., et al., *The DDX6-4E-T interaction mediates translational repression and P-body assembly.* Nucleic Acids Res, 2016. **44**(13): p. 6318-34.
190. Warrenfeltz, S., et al., *EuPathDB: The Eukaryotic Pathogen Genomics Database Resource.* Methods Mol Biol, 2018. **1757**: p. 69-113.
191. Cherry, J.M., et al., *Saccharomyces Genome Database: the genomics resource of budding yeast.* Nucleic Acids Res, 2012. **40**(Database issue): p. D700-5.
192. Gallo, C.M., et al., *Processing bodies and germ granules are distinct RNA granules that interact in C. elegans embryos.* Dev Biol, 2008. **323**(1): p. 76-87.
193. Parkyn Schneider, M., et al., *Disrupting assembly of the inner membrane complex blocks Plasmodium falciparum sexual stage development.* PLoS Pathog, 2017. **13**(10): p. e1006659.
194. Tomas, A.M., et al., *P25 and P28 proteins of the malaria ookinete surface have multiple and partially redundant functions.* EMBO J, 2001. **20**(15): p. 3975-83.
195. Wahle, E. and G.S. Winkler, *RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes.* Biochim Biophys Acta, 2013. **1829**(6-7): p. 561-70.
196. Sun, M., et al., *Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation.* Genome Res, 2012. **22**(7): p. 1350-9.
197. Fabian, M.R., et al., *Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin.* Nat Struct Mol Biol, 2013. **20**(6): p. 735-9.
198. Mugler, C.F., et al., *ATPase activity of the DEAD-box protein Dhh1 controls processing body formation.* Elife, 2016. **5**.
199. Carroll, J.S., S.E. Munchel, and K. Weis, *The DExD/H box ATPase Dhh1 functions in translational repression, mRNA decay, and processing body dynamics.* J Cell Biol, 2011. **194**(4): p. 527-37.
200. Banaszynski, L.A., et al., *A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules.* Cell, 2006. **126**(5): p. 995-1004.
201. Fairhead, M. and M. Howarth, *Site-specific biotinylation of purified proteins using BirA.* Methods Mol Biol, 2015. **1266**: p. 171-84.
202. Roux, K.J., D.I. Kim, and B. Burke, *BioID: a screen for protein-protein interactions.* Curr Protoc Protein Sci, 2013. **74**: p. Unit 19 23.

203. Lam, S.S., et al., *Directed evolution of APEX2 for electron microscopy and proximity labeling.* Nat Methods, 2015. **12**(1): p. 51-4.
204. Bindels, D.S., et al., *mScarlet: a bright monomeric red fluorescent protein for cellular imaging.* Nat Methods, 2017. **14**(1): p. 53-56.
205. Billker, O., et al., *Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite.* Cell, 2004. **117**(4): p. 503-14.
206. Ojo, K.K., et al., *Transmission of malaria to mosquitoes blocked by bumped kinase inhibitors.* J Clin Invest, 2012. **122**(6): p. 2301-5.

Appendix A: Oligonucleotides used in this study

Construct	Description	Oligonucleotide sequence (5'-3')
	Creation of Recombinant Sequences	
pyccr4-1^{-}	3'-Fwd	ggccgcggCCCCGAAAGACTTATTTATATGCAAATATAATGAAATATTTATTACTTTC
	3'-Rev (Soe)	cgggcccgcaccggtGCGTATAATGTTTTGATTGGGGAGGGTGGTATGGATGTGTGAC
	5'-Fwd (Soe)	ccggtgcgggcccGAACAGTACATGCATTCTTGCCAATAGATATATTATACAAAAATATG
	5'-Rev	cgcggccgcGTTAAATGACAATGTTTAGGTCTAAACAAATAAGATTTCTAAGCATATGTC
pyccr4-1::gfp	3'-Fwd	ggccgcggCCCCGAAAGACTTATTTATATGCAAATATAATGAAATATTTATTACTTTC
	3'-Rev (Soe)	gtttcccaccggtcgggcccGGGGAAATCATGATAAAATATAAATTTCACATGTTACAAATAAATTTATG
	ORF Fwd (Soe)	ccccgggcccgaccggtGGGAAACATAAAGGAGGTAAATATACTATAGATGGGTGTGC
	ORF Rev	ggcctaggTAAAAATTTAAATTCAAATTTTGCTACTAATGGGAAATGGTCAGACGG
pyccr4-2^{-}	3'-Fwd	cggtaccGTGAAGATTTGGAACTGATGATATTAAAATTTGTACGCATTTCATCTG
	3'-Rev (Soe)	accccgggcccagctagcGCCTACAATGTCACAATTAAATAGCCTAAATATTATTCGTAC
	5'-Fwd (Soe)	acattgtaggcgctagctgggcccGGGGTTTAATAATACGCACATGTGTCAAAACATCCC
	5'-Rev	cgcggccgCCATATTATGTATGTAATTACTATAACCTTAGCATGCCGATATTTATTTTAA
pyccr4-3^{-}	3'-Fwd	ccggtaccGAGAGAAATAAAATATAAATAATTGCTAATATTATACAATTTATTCACATGTGTG
	3'-Rev (Soe)	gggcccgcgctagcCTAATATGCTTATCTTTCCTTTTGGCAAAAATTGTGTTAATTAATAGTAGGG
	5'-Fwd (Soe)	ggaaagataagcatattagggctagcgcgggcccGTAAGTACTGGTTCCCATATTTATTATTTCATTTCGTCG
	5'-Rev	ccgcggccgcCTCTCAATTGCATATGCATAATATATGCATATATATTTCTATATAATTTTAACC
pyccr4-4^{-}	3'-Fwd	ccggtaccCTACACTATTAAATTAGTAATTAATTCATATTTATACACAGTTTCCCC
	3'-Rev (Soe)	gggcccgcgctagcGTATTGAAAGTCTTCCAAAACGCAAAATAGTAATCTGTATGTC

	5'-Fwd (Soe)	caatacgctagcgcgggcccGTTCAGGTCTTGAATAATTCTTCTTCTTGCAAATGTTGAC
	5'-Rev	ccgcggccgcCTATAATGTTATTGGTCCTCGATATATAACTAAGACACCATACCAGTTC
pycaf1^{-}	3'-Fwd	ccggtaccGTATTTATAAATTTAAGAAATAATTATCACTCAAATAATCATAAGCACATATATAAGC
	3'-Rev (Soe)	ggggcccgcgcggccgcCGTTAATATATAAATGTAAAATAAAATGTAGCATTATAGTATAGCATGATAAACACACG
	5'-Fwd (Soe)	ggcggccgcgcgggcccCAATTGCTAAGTGAAAAGCATATACATAATACACATTCGTGC
	5'-Rev	ccctcgagGTATGAAGTTTAATTTAGTTCTTATGCATTTACGTTCG
pycaf1::gfp	3'-Fwd	ccggtaccGTATTTATAAATTTAAGAAATAATTATCACTCAAATAATCATAAGCACATATATAAGC
	3'-Rev (Soe)	ggggcccgcgcggccgcCGTTAATATATAAATGTAAAATAAAATGTAGCATTATAGTATAGCATGATAAACACACG
	ORF Fwd (Soe)	ggcggccgcgcgggcccCTACACAGATATGTTAAACCAGGGAATGAACTACAATTCG
	ORF Rev	ccgctagcAGAATCATAAAAATAATTTCTATCTTTATTTTTAATGTCATAATTTGGATAAGAATAAATGTTTTCC
pycaf1ΔC	3'-Fwd	ccggtaccATAATATCAGCAAATAATACTAACAACCACTATATACATAACAACATTAATG
	3'-Rev (Soe)	ggggcccgcgcggccgccGCCATCGCTCATGTTGTTACTATTCATACG
	5'-Fwd (Soe)	ggcggccgcgcgggcccGTTAAAGGTCATACAACTAGGTGTAACCTTTTCTAATGG
	5'-Rev	ccactagtATCATTTTTTCCGTCCCAATATTGGCTATTATTATTATTG
pynot1^{-}	3'-Fwd	ccggtaccCATATTGTTATTTTTAGAATATTTATAAGTTCGTTTTTATTATTTTTGTTG
	3'-Rev (Soe)	gggcccgcgcggccgccCACTCTATGCTAAGTATCCATATGCCAATCG
	5'-Fwd (Soe)	gcggccgcgcgggcccCAACCTAGAATTTACTTAAATTTCCATACCAATTTAATATATTG
	5'-Rev	ccctaggGATATGTAAAAATTCATATTTACAATTGTACAACGATAAATATATTATATGTA
pynot1::gfp	3'-Fwd	ccggtaccCATATTGTTATTTTTAGAATATTTATAAGTTCGTTTTTATTATTTTTGTTG
	3'-Rev (Soe)	gggcccgcgcggccgccCACTCTATGCTAAGTATCCATATGCCAATCG
	ORF Fwd (Soe)	ggcggccgcgcgggcccCAAGGAGAAATTTTAAAACGAAAAGACATAAAAGAAGG
	ORF Rev	cccctaggAACATTTCCTAAGCATGTCTGAAAAACTGACTTG
pynot1-g^{-}	3'-Fwd	ccggtaccCAACTCATACAAACTCTTCTTAGAATGTAATATATATGTACATG

	3'-Rev (Soe)	ggggcccgcgcggccgccGTATGGATAGACAAAGCGGGATTCATATTTCTC
	5'-Fwd (Soe)	ggcggccgcgcgggcccCCTAGTTTATTTATCTACAAACGTATGCCCACAATTTCAAC
	5'-Rev	cccctaggGCACATGTTTTATTGTAATTACCTCAATATGGTATGAACAATATATATGGG
pynot1-g::gfp	3'-Fwd	ccggtaccCAACTCATACAAACTCTTCTTAGAATGTAATATATATGTACATG
	3'-Rev (Soe)	ggggcccgcgcggccgccGTATGGATAGACAAAGCGGGATTCATATTTCTC
	ORF Fwd (Soe)	ccggtaccCAACTCATACAAACTCTTCTTAGAATGTAATATATATGTACATG
	ORF Rev	ggggcccgcgcggccgccGTATGGATAGACAAAGCGGGATTCATATTTCTC
TTP-Binding Domain Overexpression	Fwd	ccgctagcATGAATAACAATTTTAACATTAATCTTCAGATCGAGGATGG
	Rev	ccgctagcAGAACTAATTTCAGAGATGTGTTTTATCTTAAGTATTGTAGAAGC
	Genotyping of Transgenic Parasites	
pyccr4-1^{-}		
	5'-External	GTAACCCCATATATAAAGATAATTTCTTTCTATTAGGATTGTAATTATGTGC
	5'-Internal	GCTATTATTGTTGGTACTGCTGCTCAACTGGTTTTCATTGCC
	3'-Internal	CCAGTAAGGCCGTCTGACCATTTCCCATTAGTAGCAAAATTTG
	3'-External	GCATACAAGAGTGTTGTATATGTAAGGTGAGTAAGGTGTATATAGAAAGGGAGGG
pyccr4-1::gfp		
	5'-External	CTTCAAGCCGTGCCTGAGCCAGTACGAATATCAAGGCGC
pyccr4-2^{-}		
	5'-External	GTATAATATGCTTGGCAATGTTCATAAGAATATTCACTAGATAGATATTCTTAAATGAAC
	5'-Internal	CGACAAATATTGTCTTTCTCTATTGATATATCATTCCGAATATATGAACAGC
	3'-Internal	GCTTCATCGGATCACCTATATTTACATGCAACTTTAATTAGAAAGATAGAAGA
	3'-External	CAAATTCATGATAATCGATTATAGTAACAAATATATTGGTTGTATTCTTTTAGTC
pyccr4-3^{-}		

	5'-External	GTGTGCAATCATAGACACTCAAATTTTCTATTGTACTATCTTTAATTATTTTC
	5'-Internal	GACAACAAAATAGTTGACTAATTTTGTTGTTTTATAATTAAATAAATTTCATACTTG
	3'-Internal	CCAAAATACCCGTCAGATCATGCTCTACTAATATCAGATATATTTATTG
	3'-External	CACACTATTTTATTACTACGCACACATATATATAATATGTATATACAAATTTATG
pyccr4-4^{-}		
	5'-External	GGCCCTCTTATAATTGAATGACAACACATTATGTTATCAG
	5'-Internal	GTCGCCACGAAAAACTTTAATAAATTCATCCATATC
	3'-Internal	GATCAATATAAAGGAATAGTATCTCCTTTCAACCCGAGCG
	3'-External	CCCATGAGATAATATCAATGCGTATTGCCGAATATTTATCTAAAAATATTC
pycaf1^{-}		
	5'-External	GGGGAACTTGTTCACCATTTCTATTACATTACATATATTAGGACAC
	5'-Internal	GCCCATACATCTACGATTTTCGTTCTCTCATCC
	3'-Internal	CATTGAATGAAAATAATAATAGGAACTTACATGGATCGATATACAATG
	3'-External	GGGGAGTAAATAAGGAATAAGCATTATTCCCTTATTATAAAAGGCAG
pycaf1::gfp		
	5'-External	CGACAATTTGAATTCTCAAAATAATTATTCCAATAATGGAAATAACTATGG
pycaf1ΔC		
	5'-External	GAGAGAACGAAAATCGTAGATGTATGGGC
	3'-External	CTGATTATTAACATTCAAAGCACAACTGTTACTCATATTTATG
pfcaf1 RT-PCR		
	P1-Fwd	TTCAGTTAGGTGTAACCTTTTCC
	P1-Rev	AATGCTGCTTCATTGTGAGG
	P2-Fwd	TCCACCTCATCTTCCACACC
	P2-Rev	ATTTGCATTCGACAAACTGG
pydccr4-1		
	5'-External	CTTCAAGCCGTGCCTGAGCCAGTACGAATATCAAGGCGC
	5'-Internal	GCTATTATTGTTGGTACTGCTGCTCAACTGGTTTTCATTGCC
	3'-Internal	CCAGTAAGGCCGTCTGACCATTTCCCATTAGTAGCAAAATTTG
	3'-External	GCATACAAGAGTGTTGTATATGTAAGGTGAGTAAGGTGTATATAGAAAGGGAGGG

pynot1^{-}	5'-External	CTTGCCCCCCCAATAACAGTTATATATATGTACC
	5'-Internal	CACTCAATGATGATTTGTTTATATTTTCACTACTAGCATTTACATTAG
	3'-Internal	GACCCATTCATGTCAAATATTAAAATGGATTTATTACAAGAAATTAAATTAG
	3'-External	ACACTTGCATATAATTTTGCAAAATGGATGTGTATG
pynot1-g^{-}	5'-External	GCAATTCATAAAATGTAAAAGAACAGTGTATGCGCAATTATGAAC
	5'-Internal	GGTGATTCCATCCTCGATCTGAAGATTAATGTTAAAATTG
	3'-Internal	CACAATAGCTAGCTCATGTATGTCAAATAAAACAC
	3'-External	GGGGTGATTAATATATGTATATTTATTGGCAAGTCTATGATTG
	IP-RT PCR	
p28		
	P1-Fwd	CCAACTTGCGATAAGATATAATAATGCAAAAGTCACTG
	P1-Rev	GTGATAATAAATATCCATTTACACATGAACATACAAGAGC
lap2		
	P1-Fwd	GCAAAAACTATTTGAAAAGTTTATAAACAAATGTGTTTCTTTAGAAG
	P1-Rev	CAGTAATTCCATGAGTTATTTTGCACTGAAATTGACATTC
nek3		
	P1-Fwd	CAAGTAGGTTGTATATTATATGAATTAGTAACTTTGTCTTCTCC
	P1-Rev	CTTGCAATATTATAATTAGAAATTACTTGTAACCGATCTAACGC
ap2-o		
	P1-Fwd	GCTTTAAAGATTGCCATCTTGGCAAAGC
	P1-Rev	CATTATTATTTTCTTCTTTTTGGAAGTTAAAATTTGGGGTATAG
cith		
	P1-Fwd	GTCTTCTGTATCAACATTACCATATATAGGAAGTAAAATTTCTTTAATTTC
	P1-Rev	CATTGATACTATTGCTGGATCATCAGGTATCG
gapdh		
	P1-Fwd	GAAGCATCTGAAGGACCACTTAAGGG
	P1-Rev	GGTGATGTGGATAGCCAAATCTAAAAGACGG
eif2b		
	P1-Fwd	CAAAACCTGTAATTGTTGTATTACCATTATTTAAATTACTTTATGATCC

	P1-Rev	GCAATATTATATAATTGAAAAGAGTCAATAGGACCAATTTCTGTTATG
	Circular RT PCR	
gapdh		
	5' Rev	GATACGGCCAAATCCATTAATTCCTACTTTTGTTATTGCC
	3' Fwd (Nested)	CAAACCGTCTTTTAGATTTGGCTATCCACATCACCAAAC
	3' Fwd (Main)	GCCTTAAATGATAACTTTTTCAAAATTGTCTCATGGTATGACAATGAATGGGG
p28		
	5' Rev	CGTGTCTACAGTGACTTTTGCATTATTATATCTTATCGCAAGTTGG
	3' Fwd (Main)	GGTACAGGTAGTGGTACTGGAACACCAGCAAATAGTAG
	5' Rev for poly(A)	GTTAAAATAACTGTTGTTATGAGTATAATTTTATGAAATTGTTTTG
	3' Fwd for poly(A)	GTTTTTTTCATTATTCCTTTTAAGATTAATGACTTAAATTTAACTATATTG

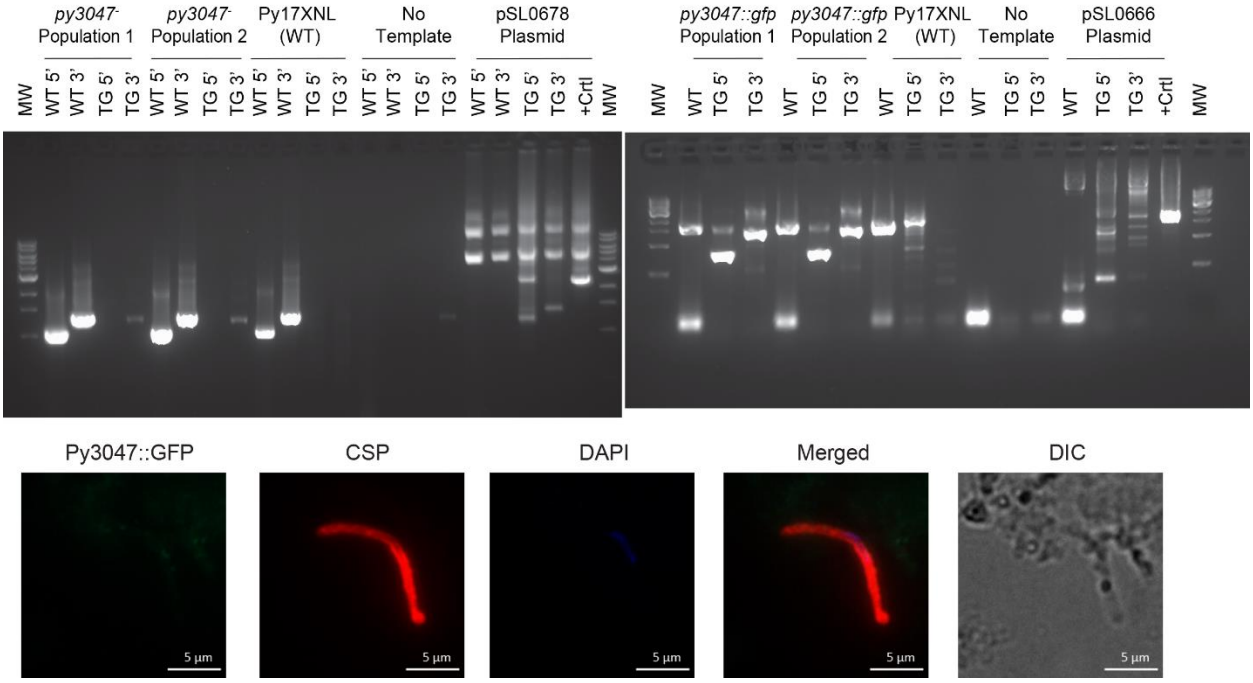
Appendix B: Additional work

Below are descriptions of work that I have participated in.

Unpublished Works

Py3047

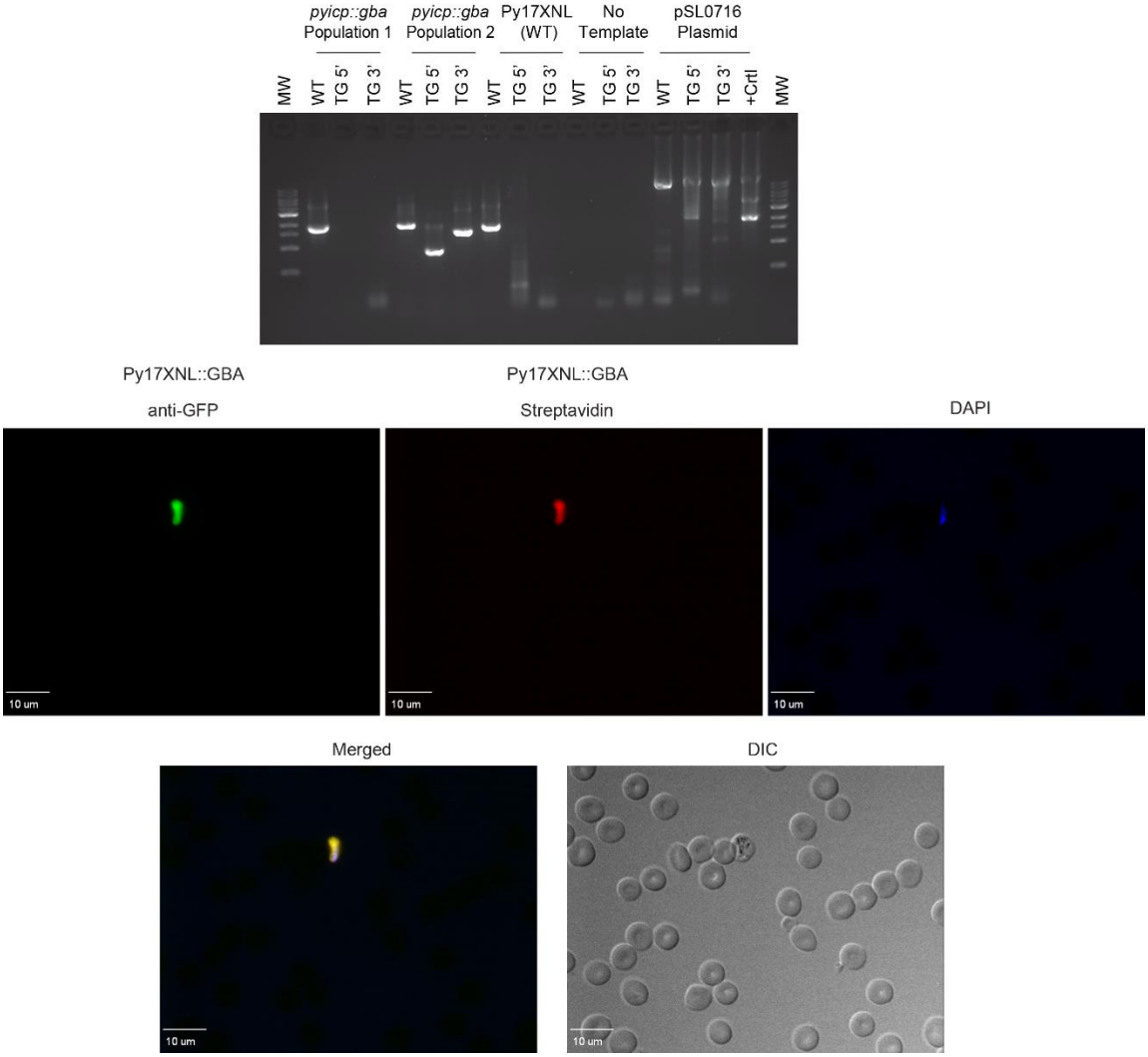
Py3047 is the gene ID of the transcript that increases the most in abundance when the *pypuf2* gene is deleted from *P. yoelii* salivary gland sporozoites. This data makes the transcript extremely interesting; however, it is not conserved between rodent and human infectious parasites. I attempted a gene deletion (pSL0678) and found that it was likely essential for asexual blood stage development. The gene was able to be successfully tagged with a C-terminal GFP (pSL0666), and was found to be expressed throughout blood stage by live fluorescence. While there is expression of PY3047 in blood stages there is none observed in sporozoites where its transcript is highly abundant. These data suggests that this transcript is being translationally repressed in the salivary gland sporozoite. Some of this work has been utilized for the translational repression in sporozoites manuscript. Future work on this protein could investigate its essential roles in asexual blood stage development as well as utilize it as a tool to study translational repression in sporozoites.



Py3047 Figure: Py3047 is likely essential for blood stage development and is not expressed as protein in salivary gland sporozoites. Top Left: A genotyping PCR demonstrating that no transgenic parasites were able to be created. Top Right: A genotyping PCR indicating that a C-terminal GFP could be integrated between the *py3047* coding sequence and it is 3' UTR. Bottom: A salivary gland sporozoite is shown with no PY3047 expression.

GBA

This project was an idea from Scott. GFP-BirA-Avitag (GBA) is a tag that could be used to purify proteins using either an anti-GFP antibody or a streptavidin-coated bead. This uses the non-promiscuous biotin ligase BirA. BirA is an *E. coli* protein that can be used to biotinylate a synthetic Avitag, which was developed to be the minimal amino acid sequence for BirA to biotinylate [201]. Theoretically, this system can be used to isolate proteins with the affinity of the biotin-streptavidin interaction in *Plasmodium* parasites. Three constructs were created, one that expresses GBA out of the p230p safe harbor locus (pSL0682), one where PyICP (pSL0716) is tagged by GBA, and one where CCR4-1 is tagged by GBA (pSL0954). Immunofluorescence of constitutively expressed GBA showed co-localization of GFP with streptavidin signal. This project was placed on hold to focus on the developing CCR4-1 story (Chapter 3).

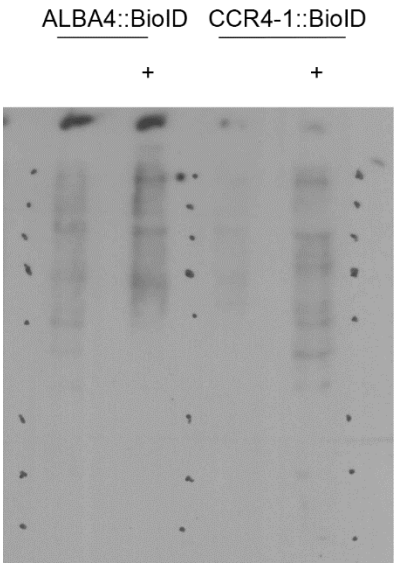


GBA Figure: GBA is expressed and can be detected by conjugated streptavidin. Top: A genotyping PCR demonstrating that *icp* was able to be tagged with GBA. Bottom: An IFA demonstrating that GBA can be detected by a streptavidin with a conjugated fluorophore and co-localizes with the GFP signal.

BioID/APEX2

When Kelly Rios began as a graduate student in the lab, her project focused on developing the BioID and APEX2 *in vivo* biotinylation approaches for *P. yoelii*. BioID utilizes a promiscuous biotin ligase, a modified BirA (BirA*) that can biotinylate lysine's and N-termini [202]. As a proof of concept, she tagged ALBA4 and I tagged CCR4-1 with BioID (pSL0953) to assist in teaching her methods that our lab commonly uses.

I additionally tagged NOT1 and NOT1-G with BioID (pSL1133, pSL1134) and APEX2 (pSL1109, pSL1110). APEX2 is a soybean ascorbic peroxidase that utilizes peroxide to function. APEX2 can biotinylate tryptophan, cysteine, tyrosine, and histidine residues using a biotin-phenol [203]. Unfortunately, neither of these techniques worked to produce labeling in either NOT1, or NOT1-G tagged parasite lines; however, they worked for ALBA4 and are being developed further by Kelly Rios.



BioID/APEX Figure: BioID labeling is increased with the addition of Biotin. Both with ALBA4::BioID and CCR4-1::BioID, an increase in labeling, is seen with the addition of 150uM Biotin overnight (Signified with "+"). Blot was probed with 1:1000 Streptavidin::HRP.

Split-Venus/S1M aptamer

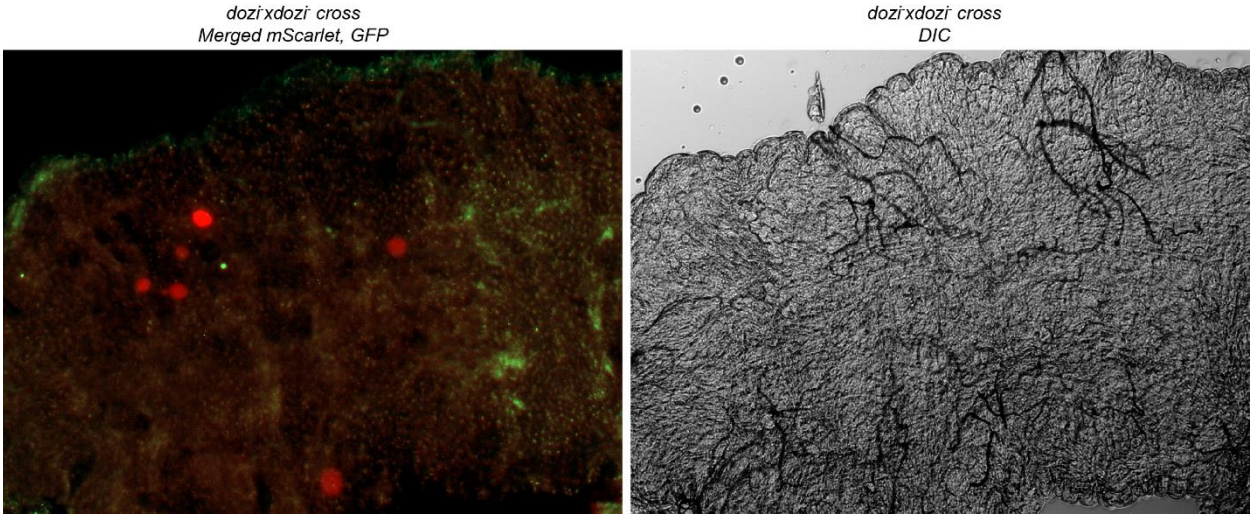
The goal of this project was to tag transcripts and use them to immunoprecipitate any proteins that were interacting with the transcript in order to identify them. Also, we could utilize this system to visualize translational repression if it was attached to a translationally repressed transcript. To this end, PY3047 was tagged with the S1M aptamer, and P25 was tagged separately with the S1M aptamer and the *ms2/pp7* hairpins with expression of a split-venus fluorophore that could bind to these hairpins. To determine if the techniques were working, cells were observed for fluorescence in the presence or absence of certain components. While initial IFA's looked promising, flow cytometry demonstrated that the technique was not working and that the controls were fluorescing as much as the experimental samples. After this point Scott and I decided that the project was going to take more time than I had to troubleshoot and I focused my efforts on the NOT1 project (Chapter 4).

mScarlet

mScarlet is an engineered red fluorescent protein that is monomeric [204]. It is important to have multiple colors to utilize when genetically manipulating parasites. The addition of mScarlet to the labs repertoire allows for its use in both traditional genetic manipulations as well as genetic cross experiments. I made all of the base plasmids for mScarlet to be used as a C-terminal tag (pSL1111), and to replace coding sequences in traditional knock-out attempts (pSL1112). This gave us another tag in addition to GFP to visualize proteins *in vivo*. I also utilized this for the *dozi* deficient parasite line.

DOZI

To perform genetic crosses, I created a line without the *dozi* gene. In *P. berghei* DOZI is essential for host-to-vector transmission and is required for female fertility. This will be useful to determine if female gametocytes are also deficient in my *not1-g* deficient lines. While the experiment worked well, it was found that parasites with *dozi* deleted from their genome can transmit to mosquitoes in *P. yoelii*; albeit, very inefficiently. I detected very few events, but the transmission was still able to occur. I also dissected salivary glands of mosquitoes infected with *dozi* deficient parasites 14 days post blood meal and saw no sporozoites. These data indicate that while the parasites can transmit from host-to-vector, they cannot produce salivary gland sporozoites.



DOZI Figure: Parasites with a gene deletion of *dozi* can transmit from host-to-vector and form oocysts. Shown is one merged image of a *dozi*$^{-}$ x *dozi*$^{-}$ cross that shows mScarlet expressing oocysts and a DIC image where oocysts can be seen.

Stage-Specific Promoters (Supervised: Laura Bowman and Logan Finger)

CRISPR/Cas9 can be utilized for a variety of purposes including manipulation of genomes and silencing or activating of genes. Small guide RNAs direct the Cas9 where to bind to the DNA strand. In order to express small guide RNAs for CRISPR in a stage-specific manner, stage-specific promoters must be identified and characterized. To do this, I led Laura Bowman and Logan Finger to characterize several stage-specific promoters for this purpose. Laura Bowman's undergraduate thesis should be referenced for the current state of this project. I plan to work with Logan to complete this project and publish it as a reference.

Immunoprecipitations using *in vitro* transcribed p25 (Supervised: Steven Griffin)

As a supplement to the S1M and split-venus project Steven Griffin created a protocol to produce both *P. yoelii* and *P. falciparum* p25 5' UTR *in vitro*. He also demonstrated that he could biotinylate this transcript. The next steps are to bind this transcript to beads and incubate it with lysate to determine what proteins may be binding to and translationally repressing it.

Published Works

A Bioinformatic Survey of RNA-binding proteins in *Plasmodium*.

PMID: 26525978

Utilizing the sequences of known RNA-binding domains, we were able to identify 189 putative RNA-binding proteins. Additionally, we reviewed the functions of the homologs of these genes in model organisms. I was responsible for identifying members of, and writing the sections on CAF1/CCR4/NOT, exosomes, and P-bodies.

5-Aminopyrazole-4-carboxamide analogues are selective inhibitors of *Plasmodium falciparum* microgametocyte exflagellation and potential malaria transmission blocking agents.

PMID: 27780638

Transmission of *Plasmodium* parasites from host-to-vector requires both male and female activatable gametocytes to productively infect the mosquito. Disruption of either sex, or the activation of these gametocytes, can prevent parasite transmission. A *Plasmodium* calcium-dependent protein kinase (CDPK4) was identified to be essential for male gametocyte activation [205]. It was then determined that bumped kinase inhibitors could inhibit activation of CDPK4 [206]. Here, compounds based on a 5-aminopyrazole-4-carboxamide scaffold were screened as inhibitors of CDPK4. We

found seven compounds with an IC50 less than 50nM. My role in this project was to test these compounds on *P. yoelii* male activation with Scott Lindner. We demonstrated these compounds have can also inhibit *P. yoelii* male gametocyte activation.

ALBA4 modulates its stage-specific interactions and specific mRNA fates during *Plasmodium yoelii* growth and transmission.

PMID: 28787542

ALBA proteins are a conserved family of RNA-binding proteins that play varying roles in RNA metabolism. We demonstrated that ALBA4, an Apicomplexan specific ALBA protein, plays roles in male gametocyte development and the synchronous development of oocyst sporozoites. I performed the first example of a crosslinking immunoprecipitation from sporozoites. I immunoprecipitated ALBA4::GFP from crosslinked sporozoites (three replicates) and then processed the data post mass spectrometry. This demonstrated that, not only was this a viable technique to determine protein-protein interactions in sporozoites, but also demonstrated that ALBA4 is interacting with some proteins during most life stages (Asexual blood stage, gametocytes, and sporozoites) and other proteins stage specifically.

Nuclear, Cytosolic, and Surface-Localized Poly (A)-Binding Proteins of *Plasmodium yoelii*.

PMID: 29359180

Poly (A)-binding proteins (PABP) bind to the poly (A) tail of mRNA and can help to regulate certain RNA regulatory processes including translational repression. In this manuscript, we demonstrated that *Plasmodium* parasites have two PABPs, one that is nuclear and one that is cytosolic. Interestingly, PABP1 was found to be on the surface of salivary gland sporozoites, and was deposited in gliding trails. This could indicate an unrecognized function of PABP. I performed the blood stage indirect immunofluorescence assays for both PABP1 and PABP2 and wrote up the section describing those results. This data demonstrated the differential localization of these two PABP's.

Puf3 participates in ribosomal biogenesis in malaria parasites.

PMID: 29487181

Puf proteins are RNA-binding proteins that play a variety of roles in RNA metabolism. Interestingly, some Puf proteins play roles in ribosome biogenesis. Puf3 in *Plasmodium* was found to be likely essential in both *P. falciparum* and *P. yoelii*. Utilizing IFAs, we determined that Puf3 localized to the nucleolus in asexual blood stage parasites, which

is consistent with the idea that this protein plays a role in ribosome biogenesis. We also demonstrated that this protein is associated with the 60S ribosome. However, we also determined that during mosquito stages this protein localizes to cytosolic granules. It is unclear what role this protein may be playing in these stages. My work included performing the experimentation in *P. yoelii* including knockout attempts and IFAs of all life cycle stages.

VITA

Education

2013-2018 The Pennsylvania State University

Ph.D. in Immunology and Infectious Disease. Advisor: Scott Lindner Ph.D.

2008-2012 Clarion University of Pennsylvania

B.S. in Molecular Biology and Biotechnology. Advisor: Douglas Smith Ph.D.

Publications

- **Hart, K.J**., Oberstaller, J., Walker, M.P., Kennedy, M.F., Padykula, I., Adams, J.H., and Lindner, S.E. (In Revision). Plasmodium transmission is critically regulated by general and transmission-specific deadenylases of the CAF1/CCR4/NOT complex. *PLoS Pathogens*
- Liang, X., **Hart, K.J.,** Dong, G., Siddiqui, F.A., Sebastian, A., Li, X., Albert, I., Miao, J., Lindner, S.E., and Cui, L. (2018). Puf3 participates in ribosomal biogenesis in malaria parasites. *Journal of Cell Science.* PMID: 29487181
- Minns, A.M., **Hart, K.J.,** Subramanian, S., Hafenstein, S., Lindner, S.E. (2018). Nuclear, Cytosolic, and Surface-Localized Poly(A)-Binding Proteins of *Plasmodium yoelii*. *mSphere.* PMID: 29359180
- Munoz, E.E., **Hart, K.J.,** Walker, M.P., Kennedy, M.F, Shipley, M.M., and Lindner, S.E. (2017). ALBA4 modulates its stage-specific interactions and specific mRNA fates during *Plasmodium yoelii* growth and transmission. *Molecular Microbiology.* PMID: 28787542
- Huang, W., Hulverson, M.A., Zhang, Z., Choi, R., **Hart, K.J.,** Kennedy, M., Vidadala, R.S., Maly, D.J., Van Voorhis, W.C., Lindner, S.E., Fan, E., Ojo, K.K. (2016). 5-Aminopyrazole-4-carboxamide analogues are selective inhibitors of *Plasmodium falciparum* microgametocyte exflagellation and potential malaria transmission blocking agents. *Bioorganic & Medicinal Chemistry Letters.* PMID: 27780638
- Reddy, B.P., Shrestha, S., **Hart, K.J**., Liang, X., Kemirembe, K., Cui, L., Lindner, S.E. (2015). A Bioinformatic Survey of RNA-binding proteins in *Plasmodium*. *BMC Genomics.* PMID: 26525978

First-Author Presentations

2018

Hart, K.J., Walker, M.P., and S.E. Lindner. 2017. NOT1-G is a Novel Member of the CAF1/CCR4/NOT Complex that is Essential for Host-to-Vector Malarial Transmission.

- **(1) Oral** Presentation at The Molecular Parasitology Meeting.
- **(1) Poster** Presentation at The Inaugural Pennsylvania Parasitology Conference (ParaCon).

2017

Hart, K.J., Walker, M.P., and S.E. Lindner. 2017. NOT1-G is a Novel Member of the CAF1/CCR4/NOT Complex that is Essential for Host-to-Vector Malarial Transmission.

- **(1) Oral** Presentation at The American Society for Tropical Medicine and Hygiene.

Hart, K.J., Oberstaller, J., Walker, M.P., Kennedy, M.F., Padykula, I., Adams, J.H., and Lindner, S.E. 2017. *Plasmodium* Transmission is Critically Regulated by CAF1/CCR4/NOT Deadenylase Assembly and Function.

- **(1) Oral** Presentation at The Molecular Parasitology Meeting.
- **(3) Poster** Presentations at The Rustbelt RNA Meeting, The Alan Magill Malaria Symposium, and World Malaria Day.

2016

Hart, K.J., Munoz, E.E., Walker, M.P., Taylor, D., Kennedy, M., and S.E. Lindner. 2016. The CCR4-1 deadenylase plays a major role regulating malarial transmission from host to vector.

- **(2) Oral** Presentations at The Rustbelt RNA Meeting and The Molecular Parasitology Meeting.
- **(2) Poster** Presentations at The Alan Magill Malaria Symposium and World Malaria Day.

2015

Hart, K.J., Kennedy, M., Walker, M.P., and S.E. Lindner. 2015. Functional Characterization of RNA Regulators in Malaria Transmission.

- **(1) Poster** Presentation at The Future of Malaria Research Meeting.

2014

Hart, K.J., Kennedy, M., and S.E. Lindner. 2014. Functional Characterization of RNA Regulators in Malaria Transmission.

- **(2) Poster** Presentations at The American Society for Tropical Medicine and Hygiene and The Molecular Parasitology Meeting.