UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Role of Autophagy in the Human Malaria Parasite, *Plasmodium falciparum*

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

Doctor of Philosophy

in

Cell, Molecular, and Developmental Biology

by

Serena Stacie Cervantes

August 2013

Dissertation Committee:
Dr. Karine Le Roch, Chairperson
Dr. Emma Wilson
Dr. Ernest Martinez

The Dissertation of Serena Stacie Cervantes is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

I would like to begin by thanking my committee chairperson and PI, Dr. Karine Le Roch, for allowing me to join, stay, and graduate from her laboratory. I doubt either one of us could have predicted the tumultuous journey we took, but I am extremely grateful for the opportunity to have been trained by her. In addition to my scientific training, I have learned to appreciate stress and cheese, and when faced with rejection to always move forward.

The lab mates I have had over the years have influenced me greatly and I thank everyone for all the help they have given me. Jacques Prudhomme without a doubt keeps the lab running and I could not have my experiments without your help. I would also like to thank Jacques for making wind and allowing me to laugh as much as possible. Dr. Evelien Bunnick thank you for editing my writing, giving me inspiration to improve on my writing, organize my file names on my computer, and sharing baking stories. The visiting postdoc and student, Dr. Elisandra Rodrigues and Nora Zeto, respectively, gave me great company during their short time here and were the two sweetest people to ever step foot into the lab. Thank you Dr. Sebastiaan Bol for reminding me to keep chasing dreams, whatever they may be and however they may change. I wish the best to former lab members Dr. Doug Chung and Dr. Nadia Ponts, and current lab members Mike Hamilton and Maggie Lu. A special thank you to all the undergrads I had the

privilege to work with over the years: Michael Cervantes, Matt Bruton, Matt Braley, Manny Torres, Sean Whitton, Aster Escalante, and Chris Conner. They all taught me patience, allowed me to practice mentoring on them, and brought a smile to my face every day they came into lab.

My dissertation committee, Dr. Emma Wilson and Dr. Ernest Martinez, were so wonderful over the years. I cannot thank them enough for all the letters of recommendations they wrote for me. Their guidance in overcoming challenges at the bench helped me become a better scientist. A special thanks to Dr. Wilson for letting me be participate in organizing the Southern California Eukaryotic Pathogen Symposium.

I would like to thank our collaborators Dr. Laurence Florens and Dr. Anita Saraf from the Stowers Institute for Medical Research for all the mass spectrometry analysis done over the past 5 years. It's been a great experience working with them. When I watched the movie Contagion I burst out laughing when Dr. Hextall said she needed two more weeks to reach an answer, because two weeks is relative and can easily turn into two years.

Another set of collaborators I would like to thank is Dr. Julia Kubanek and Dr. Paige E. Stout from the Georgia Institute of Technology. They watched me struggle writing my first two papers, and Paige reassured me it would get easier – and it did. Without this collaboration my initial drug-screening project would of not

been possible. Those first few publications may not have had a huge impact, but I would not have gotten fellowships or my postdoctoral position without them.

As an undergraduate I became interested in research I applied to programs that offered a stipend for summer research and I approached Dr. Lisa Stein for a letter of recommendation. I remember being in her office and speaking with her for about an hour. Until that moment, I did not think I would fit into any laboratory, but she smiled the whole time and was cool and collective. It was inspirational and she gave me the encouragement to pursue my dreams. That summer as an undergraduate I had the opportunity to conduct research in Dr. Morris Maduro's laboratory. While in the lab I worked with Gina Maduro, and Gina and Morris provided a nurturing environment and encouraged me to continue researching as a graduate student. Dr. Darleen DeMason wrote me letter of recommendation to graduate school and she was kind to always chat with me, even for a brief moment, whenever I saw her on campus.

Over the years I have met with many professors at UCR who have been great mentors. I was always able to ask Dr. David Lo for advice and he has given me so much insight about obtaining my research career goals. A special thank you to Dr. Leah Haimo. A few conversations with her gave me the courage to continue graduate school when things seemed so bleak. I must thank Dr. Peter Atkinson, the former director of my graduate program and current Dean of the College of Natural and Agricultural Sciences, for making sure I completed my

degree and teaching me how to become a more independent researcher. Without the support of these moonlighting mentors I would not have gotten the help I needed to finish graduate school.

The staff at UCR has been incredibly supportive. My personal oracle Tara Brown helped me grow on a personal level. Maria Franco-Aguilar believed in me when I did not believe in myself. I thank Tara and Maria for giving me multiple opportunities to succeed. Thank you Dr. David Carter for all the assistance at the microscope core facility and wonderful chats we had while acquiring images. Debbie Drake you are the best analyst that ever existed at UCR and I could not have managed all those FAUs without you. I am fortunate to have been a TA for the biology department and work with the biology laboratory prep staff, Ester Valdez, Jon Allen, Mi Kyong Kim, Xinxia Li, and Racheal Keast. Ellie Canell from the Raikel lab, I really appreciated all the conversations we had over lunch about lab management.

I would like to thank the professors that I was a teaching assistant to. Dr. John Oross taught me so much about the logistics of being an instructor. Dr. Cathy Thaler spent time with me addressing troublesome students. It was a pleasure working with Dr. Tracy Kahn, and I wish I were able to have been her TA for additional quarters. Dr. Morris Maduro taught me how to be a better instructor, to be open to new pedagogy techniques, and continue to assess myself as an instructor.

Thank you to the many graduate students who struggled with me, gave me motivation, and warned me of the great perils graduate school would have. My former TA Maneesh Manthur told me I was capable of doing research and should try to join a lab. Wendy Hung, Kathy Lin, and Randy Stout from the Maduro lab taught me the basics of being in a lab. Masha Zoudilova and John Ta reminded me that it is ok to have fun during graduate school. Kanwal Alvarez, Melissa Petreca, and Laurie Fernandez thanks for all the lovely meals we shared. Thank you to my friends from the Wilson lab Phil Nance, Danielle Worth, Clement David, and Kasia McGovern. From the shared lab space I appreciate all the chats I had with Ritu Chaudhary and Hagop Atamian. From the 2nd floor and all the time we shared in the interactive room, Sandhya Charlu, Christi Scott, and Zev Wisotsky. Thank you Kevin Keller, Tiffany Satoorian, and Ivann Martinez from the zur Nieden lab, you guys are great and always made me laugh. Thank you Hovik Gasparyan, Maiara Severo, and Olivia Sakhon for sharing dissertation-writing blues. To my UCR softball teammates who kept me sane, thank you Ryan Godfrey, Kristin Layous, Loren Alberg, Dennis Cisco-Taylor, Brittany Bannon, Ryan Rush, Philip Vieira, Christina Nicolaides, Kristina Mouzakis, Sarah Santiago, Michael Grady, Tom Lopez, and Justin Gyllen you guys are always a pleasure to hang out with. Special thanks to my all my mentees and mentoring program friends, Nolan Ung, Devon Duron-Ehnes, Cristina Flores, Abdullah Madany, Melania Ebrahimi, Kevin Rodriguez, Oscar Gonzalez, Dustin Brewton,

and Hailey Choi. To my AGEP familia Wendy Acosta, Vicente Nunez, Mario Olmedo, Rachel Behar, and Gabrielle Goodman thanks for keeping it real. At last but certainly not least, thank you Melanie Orin for appreciating me in every way possible and being such a great friend - we did it!

Lastly, I would like to thank the journals BioMed Central, ChemMedChem, and Molecular and Biochemical Parasitology for allowing me to use published material as part of my dissertation.

Dedication

I would like to dedicate this dissertation to my family. My mother and father inspired me to be persistent and they have always supported me. To my brother Michael, thank you for always being there for me in the lab, at home, and on the softball field. I cannot imagine what graduate school would have been like without you. To my sister Celina I am glad our relationship has blossomed during this process and I am going to miss our weekly hangouts. To my niece Jordyn, you brought so much joy to my life. Brandon and Ahn, thank you for being part of my family. To my pets that provided me unconditional love and comfort, thank you Coco, Dot, and Molly. My fiancé Daniel “Bucko” Montgomery, has been extremely patient and I appreciate all the compassion you showed me while writing my dissertation. Luciana Manriquez I miss you terribly, but I am grateful for all the wonderful times we had. Thank you all for encouraging me to pursue my dream.

ABSTRACT OF THE DISSERTATION

The Role of Autophagy in the Human Malaria Parasite, *Plasmodium falciparum*

by

Serena Stacie Cervantes

Doctor of Philosophy, Graduate Program in Cell, Molecular, and Developmental Biology
University of California, Riverside, August 2013
Dr. Karine G. Le Roch, Chairperson

The human malaria parasite remains a major public health burden in developing nations. Despite many years of research, the mechanisms controlling gene expression in the parasite are still poorly understood. While the *P. falciparum* genome lacks more than fifty percent of the transcription factors anticipated to regulate its 6372 genes, it encodes a large amount of genes involved in RNA metabolism and chromatin remodeling. Furthermore, preliminary data in the laboratory showed extensive nucleosome remodeling during the parasite's asexual cycle. Therefore, we hypothesized that change in chromatin structure plays an important role in controlling parasite development. To understand the role of histone post-translational modifications (PTMs) in

transcriptional regulation and histone turnover, we used a shotgun proteomic approach. A total of 246 histone PTMs were identified with 126 being novel. Parasite histones were highly acetylated and methylation marks associated with transcriptional silencing were detected at low levels. To elucidate the mechanism regulating histone turnover, we treated parasite cultures with inhibitors of two distinct pathways that degrade bulk amounts of protein; the ubiquitin-proteasome system and the autophagosome-lysosome pathway. Parasites treated with inhibitors of the autophagy pathway displayed an accumulation of histone protein. The autophagy pathway was overlooked in the parasite; thus, we investigated it at the comparative genomic, cellular, biological and genetic levels. PfATG8, an autophagosome membrane marker, was detected throughout the erythrocytic stages in the apicoplast and the cytoplasm. Proteins associated with PfAtg8 were isolated by immunoprecipitation and identified by mass spectrometry. Gene ontology enrichment showed an enrichment of proteins involved with the digestive food vacuole, the phagolysosome, and the nucleus. In summary, we determined that the autophagy pathway is multifunctional and is likely involved in vesicle traffic, apicoplast biogenesis, and protein catabolism. To further validate its role in histone turnover, we took a cellular approach and colocalized histones and PfATG8 vesicles. Collectively, our work provides key information of mechanism regulating epigenetic and its effects on gene expression in the human malaria parasite.

Table of Contents

Appendices

List of figures

List of Tables

Introduction

Malaria

Malaria is a mosquito-borne infectious disease caused by protists of the genus *Plasmodium,* an apicomplexan parasite. For millennia the malaria parasite has been transmitted throughout the world and remains a major public health issue in developing nations. The most recent World Health Organization (WHO) report estimated half the world's population was at risk of transmission, over 200 million people were infected, and approximately 660,000 malarial deaths occurred worldwide [1]. Most victims are children under the age of five and pregnant women from sub-Sahara Africa (figure 1). Back in 2000, the United Nations Millennium Declaration resolved to halt or reverse the disease by 2015, and mortality rates have decreased globally by more than 25% since then. In 2013, a child dies from malaria every minute. Populations most at risk are young children who have not developed protective immunity, pregnant women have high rates of miscarriage and maternal death, those with HIV/AIDS, international travellers from non-endemic areas, and immigrants who return back home to endemic regions. Unfortunately, funding for global malaria control has remained stagnant and without further increases in funding the international goal is unlikely to be achieved [2].

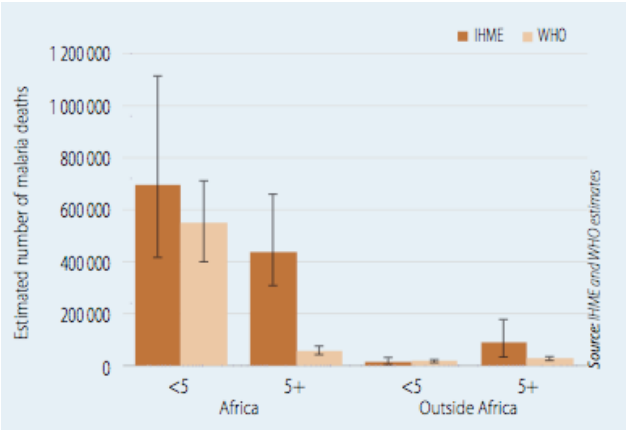


Figure 0.1 *Estimates of number of malaria deaths in 2010, by age group and geographical region* (Source: WHO World Malaria Report 2012 [1])

There are five species of Plasmodium that are transmitted to humans (*P. falciparum, P. vivax, P. malariae, P. ovale, and P. knowlesi),* with *P. falciparum* and *P. vivax* the most common infections. The most virulent and lethal species, *P. falciparum,* is responsible for up to 90% of malarial deaths and is prevalent in Africa for multiple reasons. First, *P. vivax* requires the presence of the Duffy factor on red blood cells (RBCs) to invade properly, but over 90% of sub-Saharan Africans lack the binding site. Thus, the *P. vivax* species is absent from this area. Second, the vector *Anopheles gambie* is found throughout Africa and are strongly anthrophillic, preferring almost exclusively humans to feed on [3]. This particular mosquito is an efficient vector for *P. falciparum* because it has a long life span that allows the parasite to differentiate and develop properly. Third, tropical areas with high humidity and warm temperatures are breeding grounds for *A. gambie*. Vector control in sub-Sahara Africa have had limited success, but increases in funding for insecticide-treated bed nets have reduced deaths in children by 20% and episodes of malaria by half [4, 5]. Fourth, severe malaria can be found where

there is a low-income level. Although countries, such as Oman, while unable to eliminate malaria because of their geographical location, have the resources to provide vigilant medical healthcare to treat and control the disease.

Malaria treatments

Malaria is a preventable and curable disease. Symptoms of the disease begins with fever, headache, chills and vomiting and can be easily mistaken for the flu. Due to is rapid growth rate, if not treated within 24 hours the disease can progress into a severe illness and result in death. Symptoms can escalate to severe anemia, acidification of the blood, organ failure and cerebral malaria. The most successful malaria remedies have come from natural products. Peruvian natives used the bark of cinchona trees, which lead to the purification of quinine. Chloroquine, was derived from quinine in 1934 but not readily used until the end of World War II [6]. Chloroquine treatment was cost effective and reliable, and was found to be superior to quinacrine and quinine for treatment of acute attacks of *P. vivax* [7]. The wormwood *Artemisia annua L.* has been used in traditional Chinese medicine for thousands of years. During the Vietnam War, the Chinese military created Project 523 that employed scientists to produce antimalarial drugs and to screen synthetic chemicals and recipes from traditional holistic medicine [8]. Extraction of the active ingredient, artemisinin, was optimized by

1971. In the late 90's, artemisinin-based combination therapies (ACTs) were recommended for treatment of resistant infections to prolong drug resistance to artemisinin. ACTs are currently the most effective antimalarial drug treatment [1].

Unfortunately increasing drug resistance has become a major obstacle in eradication efforts. Stories of drug resistance, identification of agene responsible for reisstance, and polymorphism found in field isolates with a geographical location echo with all commonly available antimalarials (sulfadoxine-pyrimethamine, mefloquine, quinine, artesunate, and dihydroartemisinin). Indeed, drug resistance to commonly used antimalarials can be found worldwide. The first resistant *P. falciparum* strain was detected in the late 1950s in Colombia and at the Cambodia-Thailand border [6]. By 1978 Chloroquine-resistant strains of *P. falciparum* were observed in East Africa, and resistant parasites had spread throughout tropical Africa by 1988. PfCRT, a transporter on the digestive vacuole membrane that is believed to regulate drug flux and/or pH regulation and is thought to be responsible for chloroquine resistance. Various polymorphisms in PfCRT have been found from field isolates. Resistance to artemisinins has also been detected over the past few years [9] and resistance seems to spread into Cambodia, Myanmar, Thailand and Viet Nam (figure 2) [1].

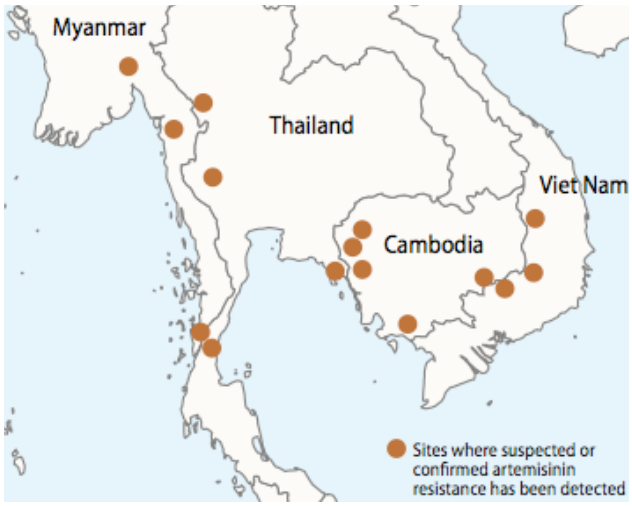


Figure 0.2 *Sites where suspected or confirmed artemisinin resistance has been detected in therapeutic efficacy studies, Mekong subregion, 2007-2012* (Source: WHO World Malaria Report 2012 [1]).

There are currently no vaccines available against malaria. A recent vaccine candidate, known as RTS,S/AS01, is in stage III clinical trials in 7 African countries. If effective, the earliest it could be used as a malaria control would be 2015. Unfortunately, recent reports show dismal efficacy of 16.8% over a 4-year period [10]. With no available vaccine and increasing resistance to common drug therapies, it is imperative to discover new antimalarial drugs. A major hurdle for antimalarial drug discovery is the limited knowledge of the mechanisms that regulate the parasite's complex life cycle (figure 2).

The malaria parasite life cycle

P. falciparum has 3 main parasitic stages that occur in the liver, bloodstream, and within the mosquito vector. A female mosquito harboring sporozoites in her salivary glands will infect a human when taking a blood meal. Hundreds of

sporozoites will be released into the bloodstream, and a few will migrate into hepatocytes where they differentiate into a schizont. The schizont will egress 7 to 10 days later and release thousands of merozoites into the bloodstream. The erythrocytic stage occurs within 48 hours, and begins when a merozoite invades a host red blood cell (RBC). Within the RBC the parasites differentiate into a ring form, grow into a trophozoite where DNA replication and global increases in transcription rates occur, and partition into daughter cells during schizogony. The blood stage is completed when up to 32 daughter merozoites egress and invade new RBCs. The health of the human host declines as parasites consume hemoglobin causing severe anemia, respiratory distress due to metabolic acidosis, or even cerebral malaria. These stresses on the host triggers parasites to undergo sexual differentiation into gametocytes, which can take up to 2 weeks to fully mature. Transmission of this disease occurs when another female mosquito takes another blood meal. Within the mosquito gametocytes enter the midgut a decrease of temperature and pH induce another transformation. Females become macrogametes, while males undergo 3 rounds of nuclear division, extrude their mitochondria and apicoplast, and become eight flagellated microgametes. Sexual reproduction occurs within the gut to form an ookinete that migrates through the gut wall and creates a cyst of the exterior of the gut. When the oocyst ruptures hundreds of sporozoites are released. These motile parasites

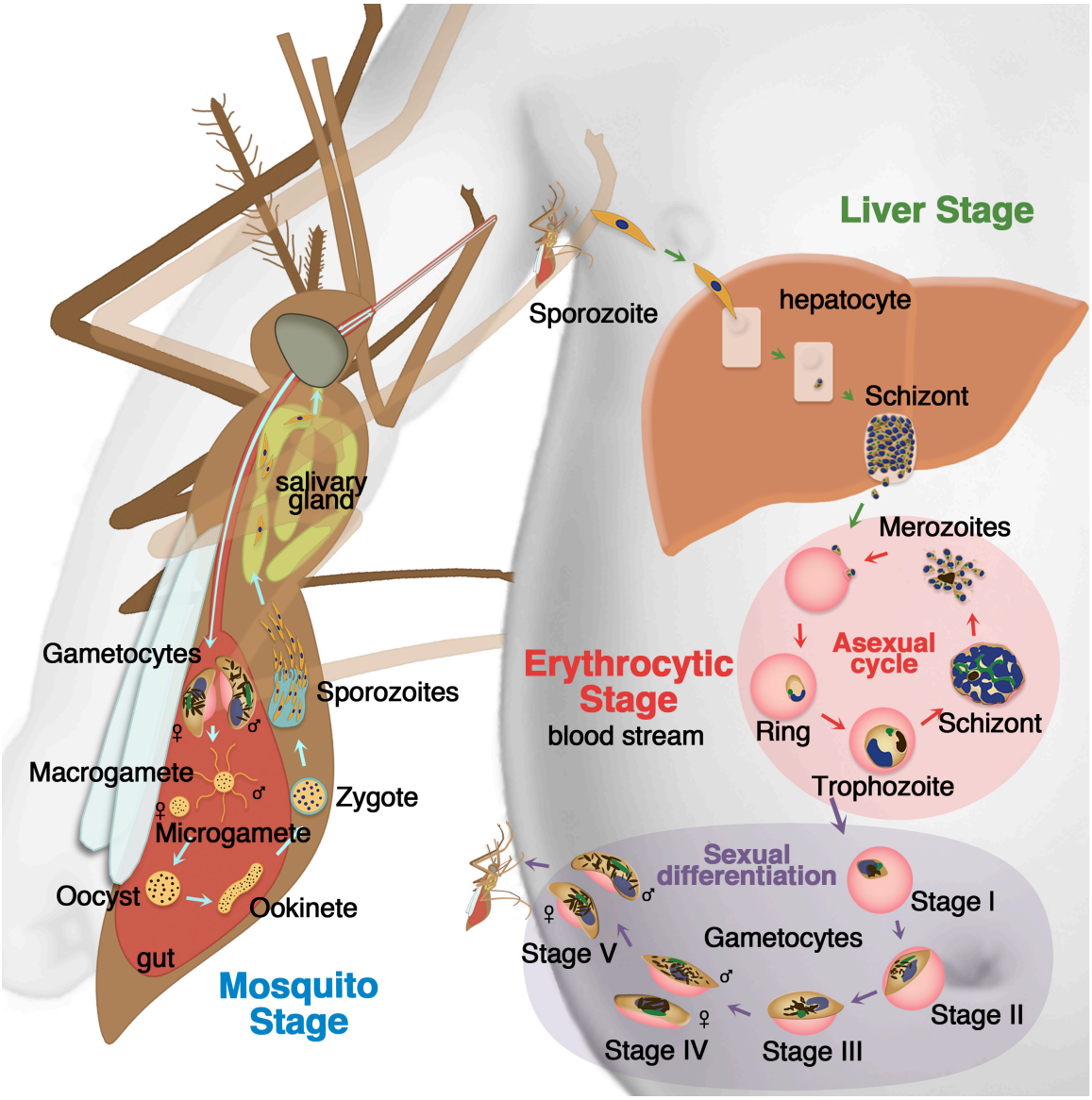


Figure 0.3: *Malaria Life Cycle*. Sporozoites enter the bloodstream from the salivary glands of an anopheles mosquito bite, and invade hepatocytes. Differentiation of sporozoites into a schizont occurs and 7 to 10 days later, merozoites enter the bloodstream and invade red blood cells. Parasites mature into ring, trophozoite, and schizont stages within 48 hours, and up to 32 new daughter cells are released as merozoites. At the ring stage, parasites can undergo gametogenesis following exposure to stress to form male and female gametocytes. Sexually mature gametocytes are taken up during a female mosquito's blood meal, and sexual reproduction occurs in the mosquito's midgut. The cycle continues when sporozoites are released into the salivary glands.

migrate into the mosquito salivary glands and wait for the mosquito to take another blood meal. By understanding the molecular processes required for parasite life cycle progression, we can develop new antimalarial therapies.

The Plasmodium falciparum Genome

In 2002, the *P. falciparum* genome was sequenced and was found to be 23 Mb consisting of 14 chromosomes that encodes approximately 5,500 genes [11]. The *P. falciparum* genome remains the most (A+T) –rich sequenced genome at 80.6% and up to 90-95% in intergenic regions [12]. The sequenced genome established a foundation for utilizing high-throughput techniques to examine transcript and protein profiles throughout the parasite's life cycle.

Characterization of transcript and protein profiles provided interesting information on the parasite biology. However, transcription regulation in the malaria parasite remains ambiguous. To date there is no defined canonical promoter and there is shortage of homologous transcription factors found in *P. falciparum* [13-17]. When the parasite genome was compared to other free-living eukaryotic genomes, it was found to encode a third of the expected transcription-associated proteins [18]. In addition, the CCCH-type zinc finger, commonly found in proteins that modulate mRNA decay and translation rates, was most abundant in the *P. falciparum* genome. When mRNA transcript and protein abundance

levels were compared throughout the erythrocytic cycle a delay between mRNA and protein accumulation was observed [19]. Parasite cultures treated with various drugs or environmental perturbations exhibit small changes in transcriptional responses compared to other eukaryotic systems [20-23]. Whether transcription profiles have significant differences between treated and untreated parasite samples remains unresolved, but cell cycle arrest or hibernation and induction of sexual differentiation is commonly observed in treated parasites. Collectively, these studies imply that gene expression may be regulated at the chromatin level and protein levels could be more precisely determined by posttranscriptional mechanisms.

Chromatin structure in Plasmodium and its role in transcriptional regulation

Chromatin consists of DNA wrapped around a core of histone proteins called nucleosomes. The structure of chromatin depends largely on nucleosome composition. There are two copies of histone proteins, H2A, H2B, H3, and H4, which assemble into an octamer, and 145-147 base pairs of DNA wrap around it to form a single nucleosome. The protein octamer is divided into four 'histone-fold' dimers, a H3-H4 and H2A-H2B histone pairs [24]. Unlike other eukaryotes, the linker histone H1 that compacts linear DNA is absent in *P. falciparum.* Nucleosome positioning is dependent on DNA sequence flexibility and histone-

induced bendability. Histones can restrict DNA transcription by distorting DNA that leads to a direct block of transcription factor binding sites. Homology with the histone-fold is well conserved, but the N-terminal tail domain can be highly variable.

Histone tails protrude away from the globular domain and can be easily assessed for posttranslational modifications (PTMs), which add an additional layer of regulating chromatin structure. These tail domains can undergo distinct covalent PTMs that include acetylation, methylation, formylation, phosphorylation, ubiquitination, hydroxylation, and crotonylation (figure 0.4).

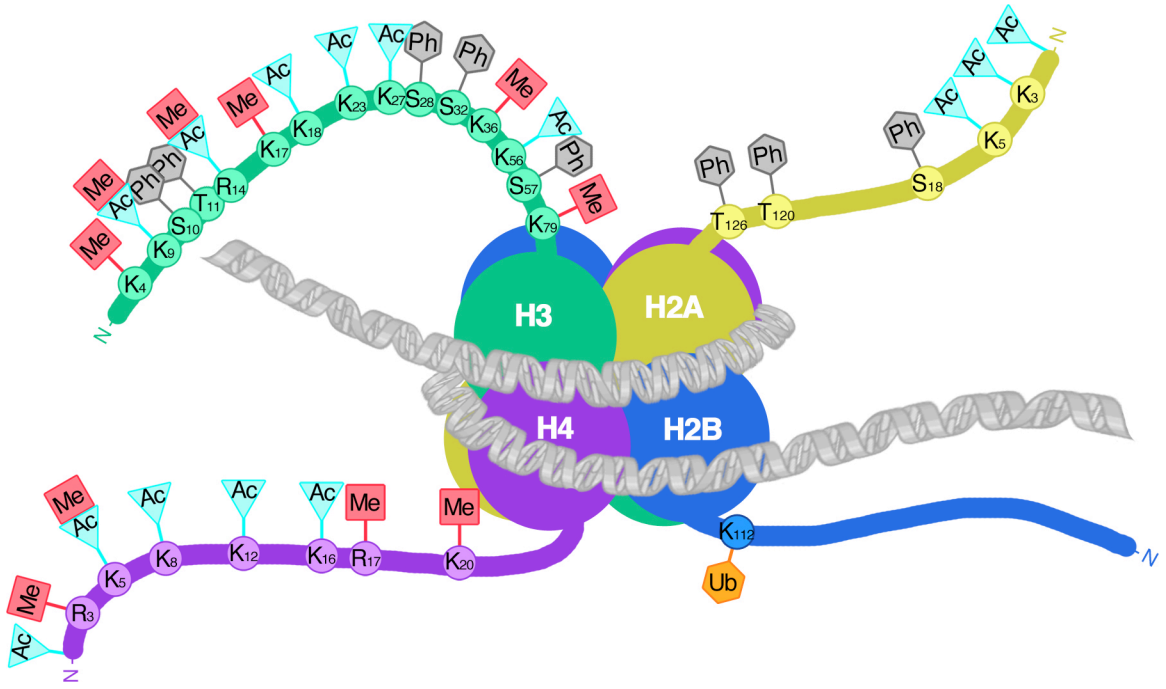


Figure 0.4. *A schematic illustration of a nucleosome in P. falciparum.* The four canonical histone proteins (H2A, H2B, H3, and H4) and the covalent PTMs are indicated on the histone tail domain (methylation, red; acetylation, light blue; phosphorylation, grey; ubiquitination, orange).

PTMs may alter chromatin structure by effecting histone-DNA and histone-histone interactions. For instance, histone tails are highly basic and acetylation decreases their affinity to DNA, which allows transcription factors to access and bind DNA. Conversely, methylation increases its basicity and hydrophobicity of histones. Methylation has a dual role and been implicated in gene silencing and heterochromatin formation, but has been found to be enriched in active promoters. Regulation of chromatin structure is believed to be important to modulate accessibility to regulatory sites and therefore play a central role in transcription regulation [25]. In addition, most PTMs are distributed in distinct patterns along DNA strands, such as, the upstream region, the promoter region, and the 5' end or 3' end of the open reading frame. The 'histone code' is a theory that histone PTMs can act sequentially or in combination to recruit effector proteins that lead to specific downstream effects.

Epigenetic regulation in the malaria parasite

Several groups have started to decipher the histone code in *P. falciparum*, but little is known about epigenetic regulation in the parasite and studies are still underway to determine all histone modifications and their functions. Mass spectrometry and immunoblotting assays have shown that acetylation occurs at a relatively steady state throughout the blood stage [26]. A histone modifier

responsible for H3K9ac lead to the identification of the histone acetyltransferase PfGCN5. Inhibition of PfGCN5 reduced H3K9ac in promoter regions and gene expression [27]. In contrast to active transcription epigenetic marks, H3K9me3 has been found on silenced genes and display dynamic changes throughout the erythrocytic cycle. *In silico* analysis and *in vitro* studies have identified potential methyltransferases and demethylases, but their function and substrate specificities need to be characterized [28]. Global histone analysis showed a high number of modifications associated with a transcriptionally active state and low levels of modifications associated with a repressive state [29]. When histone H3 PTMs were examined the inactive marker H3K4me3 and active marker H3K9ac were found to be homogenous across genes during the ring stage, but enriched at the 5' end of active genes at the schizont stage [30]. Overall, multiple studies have shown that *P. falciparum* has an extensive euchromatic epigenome and only variant surface antigen gene families are restricted to heterochromatin.

Since the late 1990's it has been speculated that epigenetic mechanisms are involved in *var* gene regulation [31]. There are approximately 60 *var* genes, but only a single *var* gene is expressed at a time, which produces a protein called PfEMP1 that is exported to the surface of infected erythrocytes. Monoallelic expression and the ability for parasites to switch expression allow parasites to evade host immune responses. Actively transcribed *var* genes and those poised for re-activation had tri- and dimethylation of H3K4 in the 5' upstream region [32].

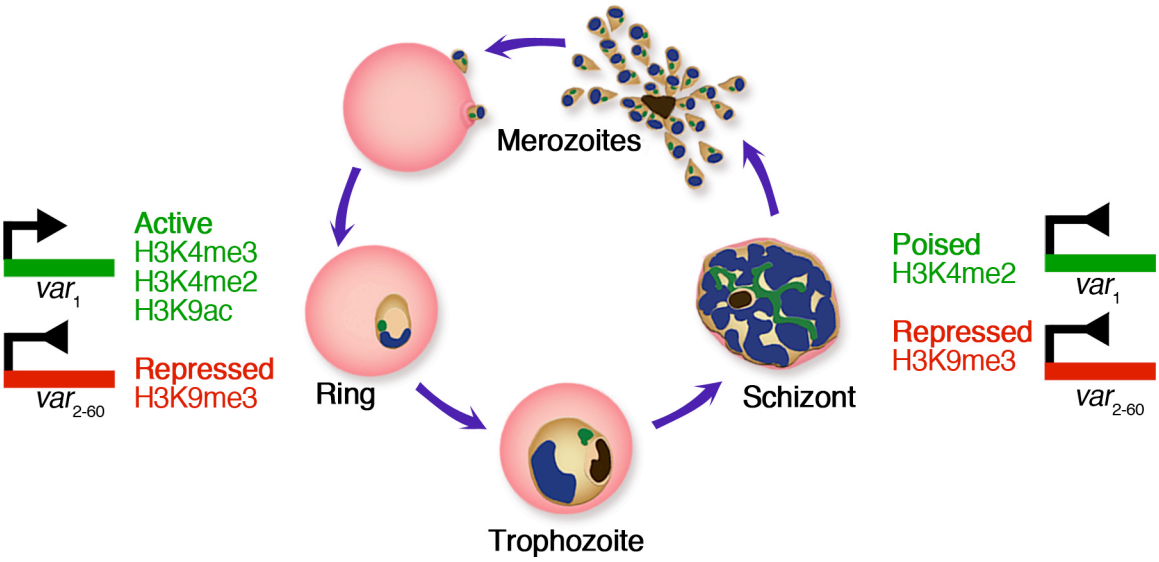


Figure 0.5. *Histone marks linked to var gene expression.* Schematic illustration of histone post-translational modifications associated with control of expression and epigenetic memory of *var* genes. Three different transcriptional states of *var* genes: Actively transcribed *var* gene in ring stages (green) is associated with H3K4me2, H3K4me3, and H3K9ac in 5'-flanking regions. In the trophozoite/schizont stage the previously active var gene is transiently silenced (poised) (green) but maintains the enrichment of H3K4me2. This association may transmit memory of the active *var* gene during cell division. Stably silenced *var* genes (5'-flanking region and exon 1) (red) are enriched in H3K9me3 throughout the asexual life cycle.

Conversely, H3K9me3 was found to stably silent *var* genes along the 5' flanking and coding region (figure 0.5). In the 5' flanking competition between H3K9 methylation and H3K9 acetylation exists. ChIP-chip showed H3K9me3 localizes with clonally variant gene families to the nuclear periphery and clusters on subtelomeric ends and some chromosome internal regions [33]. The histone deacetylase PfSir2 was found to regulate H3K9me3 and inhibition of PfSir2 disrupts monoallelic transcription of *var* genes. In addition, the heterochromatin protein 1 homolog, PfHP1, binds to tri- and dimethylated H3K9 and may participate in antigenic variation [34]. Collectively, this data implies that

expression of a single *var* gene is mutually exclusive and regulated by different epigenetic factors.

Histone depletion at the trophozoite stage

Nucleosome turnover adds an additional layer of complexity to regulating gene expression and initiation of replication. From yeast to mammalian cell lines, core histones turnover at a similar rate and histones with transcriptionally active PTMs exhibit a faster turnover rate than those with silencing PTMs [35, 36]. In *Drosophilla* cells, rapid replacement of nucleosomes has been mapped back to active gene bodies, epigenetic regulatory elements and replication origins [37]. In yeast, most nucleosome turnover occurs over promoters and small RNA genes, but also over coding regions that have high transcription rates [35]. Nucleosome depletion increases accessibility of DNA to transcription factors, but this process can erase histone modifications and maintain active or silent gene expression. In addition, DNA replication origins may be determined by chromatin structure and not by specific sequence features; therefore, nucleosome depletion facilitates initiation of replication.

We hypothesize that chromatin structure drives transcription and DNA replication in *P. falciparum.* Unlike most other eukaryotes that implement target specific chromatin remodeling complexes, the malaria parasite undergoes

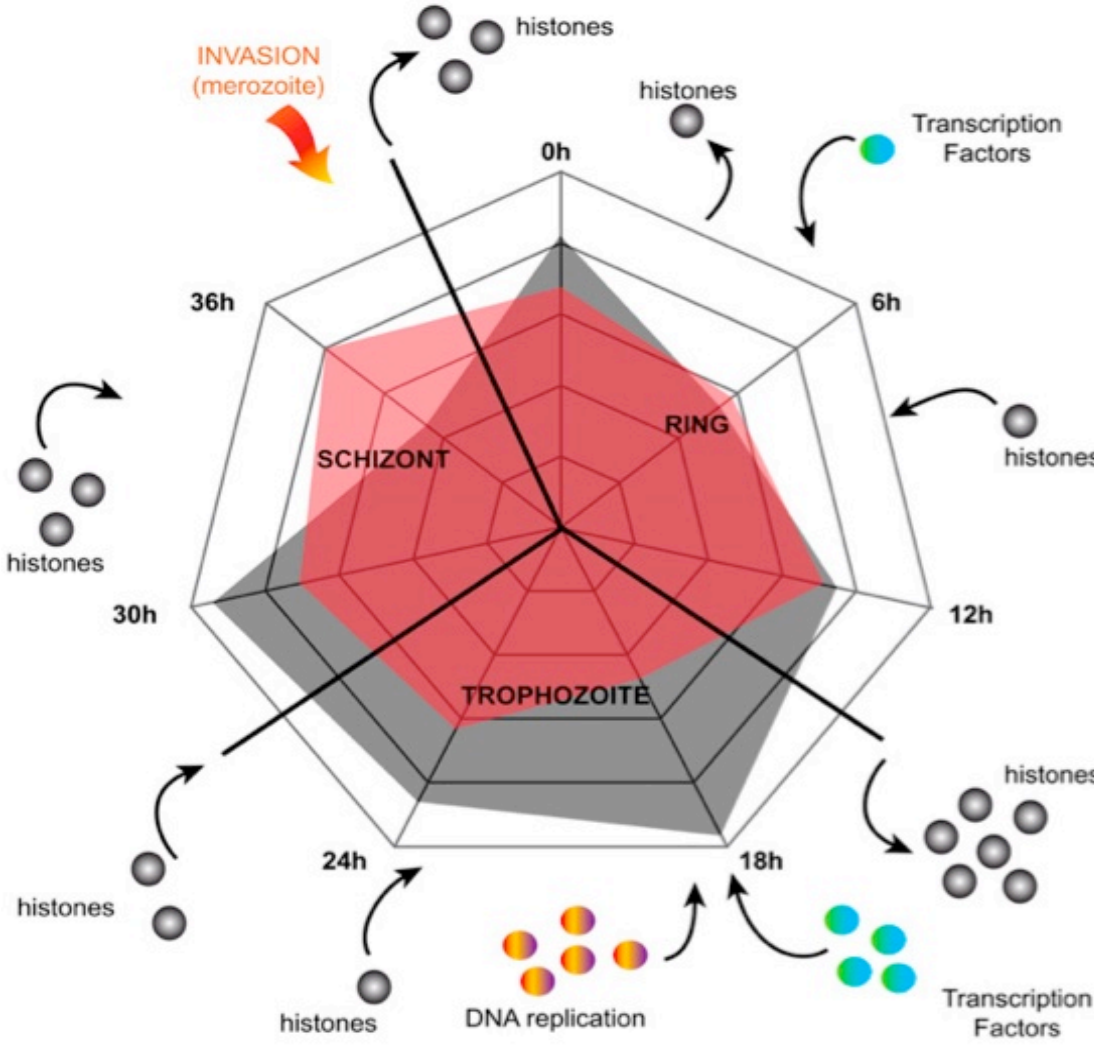


Figure 0.6. *Proposed model of chromatin remodeling throughout P. falciparum's erythrocytic cycle.* Genome-wide variations of coverage's densities are represented in two complementary assays. In red is the MNase-mediated purification of mononucleosomes to extract histone-bound DNA (MAINE) and black is the formaldehyde-assisted isolation of regulatory elements to extract protein-free DNA (FAIRE). Parasites invade host RBCs and histones are depleted and chromatin is open. At the early trophozoite stage, a significant amount of histone proteins are degraded to allow replication and transcription machinery to access DNA. After replication, DNA condenses at the schizont and parasites divide and egress as merozoites. (Source: Ponts et. al [38])

dramatic genome-wide changes to chromatin structure (figure 0.6)[38]. At 18 hours post invasion, there is a decrease in histone proteins that lead to DNA replication and an increase in transcripts. This temporally regulated event highlights the importance of histone clearance as a prerequisite for global transcription rates to increase. Previous studies showed transcript levels of histone increase throughout the cell cycle of *P. falciparum*, but there is a significant decrease in histone protein during the trophozoite stage when DNA replication is occurring [39]. Investigation of the nuclear proteome revealed a decrease in histone protein as well [40]. These observations are surprising since histone protein levels remain at a steady level throughout the cell cycle in most other eukaryotic systems [41]. Chromatin structure undergoes dynamic changes without histones and potentially gives access to transcription and replication machinery, which would subsequently lead to a global increase of transcription rates and DNA synthesis. Therefore, we hypothesize that histone protein turnover may be involved in regulating parasite life cycle progression during the blood stage.

The molecular process that regulates histone depletion remains unknown in *P. falciparum*. There are two predominant routes that lead to bulk amounts of protein clearance in eukaryotes, 1) the ubiquitin-proteasome and 2) the autophagy-lysosome pathway (figure 0.7). The ubiquitin-mediated pathway

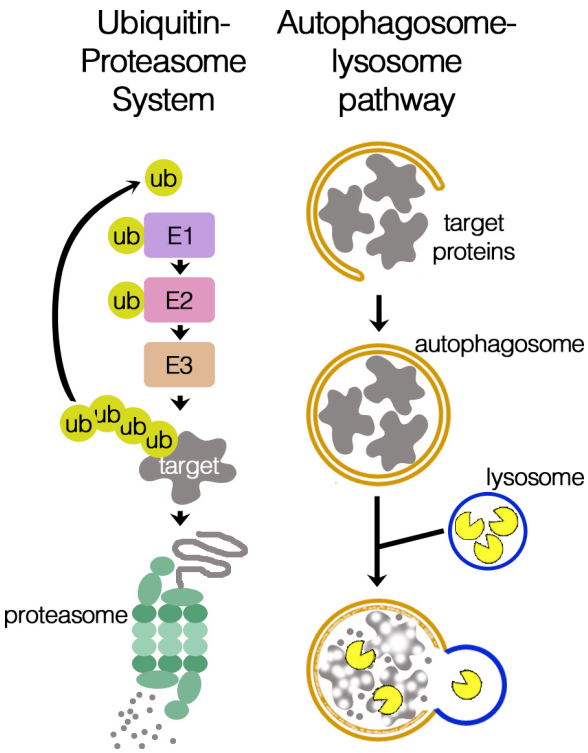


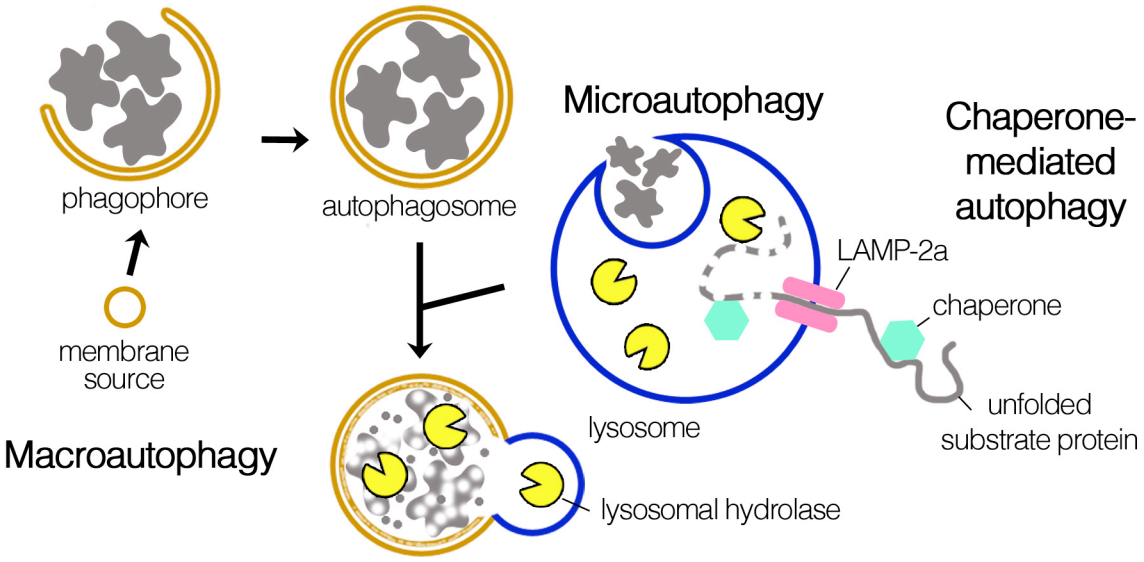
Figure 0.7. *Illustration of the ubiquitin-proteasome system (UPS) and the autophagosome-lysosome pathway.* The UPS uses ubiquitin to target proteins that will be degraded by the proteasome. The autophagosome-lysosome pathway is a membrane mediated process that engulfs material to be degraded after fusion with the lysosome.

hydrolyse peptide bonds to break proteins apart into amino acids. The ubiquitin proteasome system (UPS) typically degrades short-lived and misfolded proteins. Autophagy is a process that engulfs cellular material into vesicles that ultimately fuse with lysosomes to degrade proteins. Initially this pathway was found to be activated under nutrient limitations, but autophagosome formation occurs under numerous stimuli to maintain amino acid levels and recycle old or damaged organelles [42]. Since autophagy degrades bulk amounts of long-lived proteins and the autophagy machinery can degrade the nuclear components, we hypothesize that histone protein degradation may be regulated by autophagy. Preliminary data treating synchronized parasites with inhibitors of both the UPS and autophagy pathway suggests that the UPS does not play a role with histone protein degradation.

The autophagy pathway

The autophagosome-lysosome pathway is involved in recycling proteins and organelles. Similar to the UPS, the autophagy pathway consists of a number of sequential steps: selection and tagging of cargo, recognition and transport to proteolytic machinery, degradation of cargo, and recycling of small polypeptides or amino acids. Proteins and organelles are enveloped by a double membrane to form an autophagosome that fuses with a lysosome to degrade cargo proteins. There are three main membrane-mediated processes conserved through most eukaryotes: 1) macroautophagy (referred to as autophagy hereafter) removes unnecessary proteins and damaged organelles [43], 2) microautophagy occurs when the lysosome envelops cytoplasmic material [44], and 3) chaperone-mediated autophagy (CMA) relies on a chaperone protein to translocate individual proteins into the lysosome [45] (figure 0.8). Once thought to be a

Figure 0.8. *Schematic illustration of the three different types of autophagy.*



nonspecific molecular process that recycles proteins for survival under nutrient-limited conditions, autophagy has been shown to be involved in aging, degradation of pathogens, cell differentiation, and removal of old or damaged organelles.

Over 30 autophagy-related (Atg) genes were first discovered in *S. cerevisae* after screening for mutants unable to survive nitrogen starvation [46]. Atg genes are evolutionarily conserved in eukaryotes and have been classified into six functional units. 1) The Atg1 kinase complex (involving Atg1, Atg13, Atg17, Atg29, and Atg31) is induced by nutrient starvation and is negatively regulated by target of rapamycin (TOR) [47, 48]. 2) Atg9 is a transmembrane protein that is required for autophagosome formation and localizes to the outer membrane of the preautophagosomal structure (PAS) [49]. 3) The PI3K complex (involving Atg14, Atg6/Vps30, Vps15, and Vps34) is essential for canonical autophagosome formation [50]. Autophagy inhibitors wortmannin and 3-methyladenine (3-MA) work by inhibiting PI3K activity. 4) The Atg2-Atg18 complex completes sequestration of the autophagosome by facilitating ecruitment of Atg8-phosphatidylethanolamine (PE) [51]. 5) Analogous to the ubiquitin conjugation system, the Atg12 system involves activation of Atg12 by Atg7 (E1-like enzyme), transfer to Atg10 (E2-like enzyme), and conjugation to Atg5 (Figure 0.9)[52]. The Atg12-Atg5 conjugate then interacts with Atg16 to facilitate the lipidation of Atg8. 6) Likewise, the Atg8 conjugation system involves

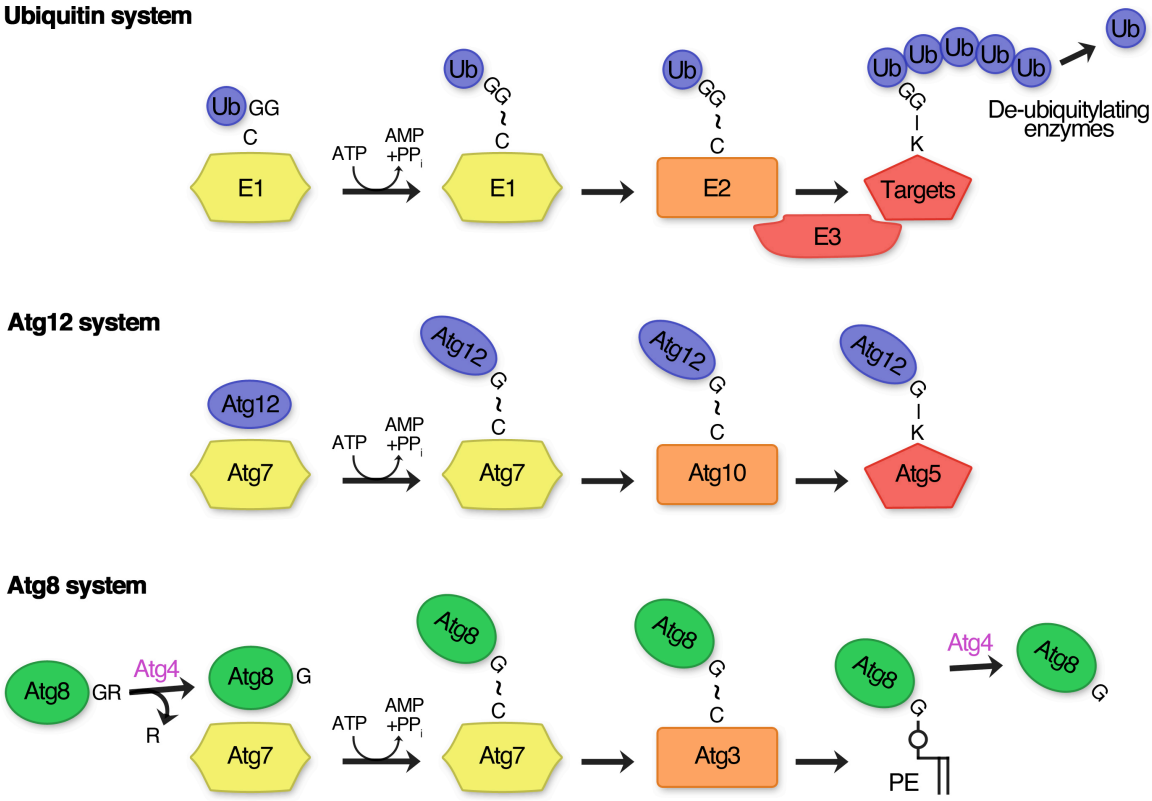


Figure 0.9. *Conjugation processes of the Atg12 and Atg8 systems in yeast.* See text for details. Atg, autophagy-related; PE, phosphatidylethanolamine; PPi, pyrophosphate; Ub, ubiquitin.

cleavage of the C-terminal of Atg8 by the cysteine protease Atg4 to expose a glycine residue, activation of Atg8 by Atg7 (E1-like enzyme), transfer to Atg3 (E2-like enzyme), and lipidation by PE [53]. Atg8-PE localizes to the developing autophagosome and is necessary for membrane expansion and completion.

Autophagy in parasitic protists

The canonical machinery of autophagy in parasitic protists has become diverse as a consequence of their adaptive life cycles. *Trypanosomatidae brucei*

is responsible for sleeping sickness in sub-Saharan Africa, and Chagas' disease in Latin America is caused by *T. cruzi*. In *T. brucei* there are three isoforms of Atg8, and *T. cruzi* has two isoforms of Atg8 and Atg4, indicating either functional redundancy and/or the ability to carry out additional function(s) [54]. In Trypanosomes, autophagy is inducted during metacyclogenesis, differentiation from the log phase to stationary phase of growth, and starvation periods. *Leishmania* have high peptidase activity during differentiation from metacyclic promastigotes to amastigotes. *L. major* contains an expansive set of 25 'Atg8' genes that undergo proteolysis by two Atg4 genes [55]. Other possible Atg gene candidates are Atg5, Atg10, and Atg12, indicating modifications and substantial loss of the canonical autophagy machinery [56]. *Toxoplasma gondii,* an Apicomplexa parasite, infects warm-blooded mammals does not have any clear orthologs of the Atg5-Atg12 conjugation system, but has a complete Atg8 conjugation system [57]. These intercellular parasites have an Atg1 ortholog and starvation has been found to induce autophagosome formation [58]. Recently, it was found that Atg3 is essential for regulating mitochondrial homeostasis during tachyzoite development [59]. Overall, comparative bioinformatics studies have revealed that Atg genes in parasitic protists have been modified, lost, or elaborated, but autophagy seems to play a role in development.

In *P. falciparum,* the autophagy-lysosome pathway has been largely overlooked. Recent studies demonstrated that under chloroquine drug pressure,

parasites die not by apoptosis, but underwent cytoplasmatic vacuolization, analogous to autophagic cell death [60]. A recent study of PfAtg8 expression implies that the parasite's Atg8 conjugation system participates in apicoplast biogenesis [61]. Similarly, in the rodent malaria *P. berghei,* the function of PbAtg8 was found to be restricted to apicoplast biology in the liver-stage [62]. In comparison, a rudimentary autophagy machinery may be involved with removing unnecessary organelles during differentiation from the motile sporozoite to a metabolically active trophozoite [63].

Summary

By discovering the epigenetic molecular mechanisms that regulate histone degradation, we can target proteins required for life cycle progression and develop new therapies. We present a comprehensive map of histone PTMs throughout the asexual cycle of *P. falciparum.* By precipitating histones and identifying PTMs by a quantitative mass spectrometry, specifically multidimensional protein identification technology (MudPIT), we detected a total of 246 histone PTMs with 126 identified as novel modifications. Low levels of methylation associated with transcriptional silencing were detected. Conversely, acetylation marks were detected a high levels and represents active transcription. In addition, histone depletion was observed and we sought to uncover the

mechanisms regulating its degradation. Surprisingly, preliminary data indicated that a rudimentary autophagy pathway was involved with histone degradation. In *P. falciparum,* the autophagy pathway has been largely overlooked. It had been suggested to participate in apicoplast biogenesis, but we determined that the autophagy pathway is multifunctional in the parasite and is involved with, but not limited to, protein catabolism, vesicle transport, and the secretory pathway. We validated the role of autophagy in histone degradation by incubating parasite cultures with inhibitors of autophagosome-lysosome fusion and observing an accumulation of histone proteins at the trophozoite stage. Therefore, we have identified an important pathway in the parasite that will lead to novel drug targets. Drug resistance remains a major problem for the field of malaria. Malaria parasites have the ability to mutate quickly, creating a challenging hurdle when the availability of effective drugs is often limited to derivations of natural compounds. By understanding the molecular mechanisms required for parasite development, we can identify new targets for therapeutic drug intervention. Collectively, this work will be the first contribution to fill in the significant gap in the current picture of epigenetic regulation of nucleosome turnover and degradation by the autophagy pathway in an important human pathogen.

References

1. WHO, U., *World Malaria Report 2012.* Geneva, Switzerland and New York, USA, 2012.

2. Snow, R.W., et al., *International funding for malaria control in relation to populations at risk of stable Plasmodium falciparum transmission.* PLoS medicine, 2008. **5**(7): p. e142.

3. Coluzzi, M., *The clay feet of the malaria giant and its African roots: hypotheses and inferences about origin, spread and control of Plasmodium falciparum.* Parassitologia, 1999. **41**(1-3): p. 277-83.

4. Lengeler, C., *Insecticide-treated nets for malaria control: real gains.* Bulletin of the World Health Organization, 2004. **82**(2): p. 84.

5. Lengeler, C., *Insecticide-treated bed nets and curtains for preventing malaria.* Cochrane database of systematic reviews, 2004(2): p. CD000363.

6. Wongsrichanalai, C., et al., *Epidemiology of drug-resistant malaria.* The Lancet infectious diseases, 2002. **2**(4): p. 209-18.

7. Whorton, C.M., T.N. Pullman, and et al., *The Chesson strain of Plasmodium vivax malaria; immunity.* The Journal of infectious diseases, 1947. **81**(1): p. 1-6.

8. Neill, U.S., *From branch to bedside: Youyou Tu is awarded the 2011 Lasker~DeBakey Clinical Medical Research Award for discovering artemisinin as a treatment for malaria.* The Journal of clinical investigation, 2011. **121**(10): p. 3768-73.

9. Dondorp, A.M., et al., *Artemisinin resistance in Plasmodium falciparum malaria.* The New England journal of medicine, 2009. **361**(5): p. 455-67.

10. Olotu, A., et al., *Four-year efficacy of RTS,S/AS01E and its interaction with malaria exposure.* The New England journal of medicine, 2013. **368**(12): p. 1111-20.

11. Gardner, M.J., et al., *Genome sequence of the human malaria parasite Plasmodium falciparum.* Nature, 2002. **419**(6906): p. 498-511.

12. Szafranski, K., et al., *Gene organization features in A/T-rich organisms.* Journal of molecular evolution, 2005. **60**(1): p. 90-8.

13. Coleman, B.I. and M.T. Duraisingh, *Transcriptional control and gene silencing in Plasmodium falciparum.* Cellular microbiology, 2008. **10**(10): p. 1935-46.

14. Gunasekera, A.M., et al., *Regulatory motifs uncovered among gene expression clusters in Plasmodium falciparum.* Molecular and biochemical parasitology, 2007. **153**(1): p. 19-30.

15. Callebaut, I., et al., *Prediction of the general transcription factors associated with RNA polymerase II in Plasmodium falciparum: conserved features and differences relative to other eukaryotes.* BMC genomics, 2005. **6**: p. 100.

16. Coulson, R.M., N. Hall, and C.A. Ouzounis, *Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum.* Genome research, 2004. **14**(8): p. 1548-54.

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

18. Nirmalan, N., P.F. Sims, and J.E. Hyde, *Quantitative proteomics of the human malaria parasite Plasmodium falciparum and its application to studies of development and inhibition.* Molecular microbiology, 2004. **52**(4): p. 1187-99.

19. Le Roch, K.G., et al., *Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle.* Genome research, 2004. **14**(11): p. 2308-18.

20. Gunasekera, A.M., et al., *Plasmodium falciparum: genome wide perturbations in transcript profiles among mixed stage cultures after chloroquine treatment.* Experimental parasitology, 2007. **117**(1): p. 87-92.

21. Young, J.A., et al., *In silico discovery of transcription regulatory elements in Plasmodium falciparum.* BMC genomics, 2008. **9**: p. 70.

22. Andrews, K.T., et al., *Comparative gene expression profiling of P. falciparum malaria parasites exposed to three different histone deacetylase inhibitors.* PloS one, 2012. **7**(2): p. e31847.
23. Mok, S., et al., *Artemisinin resistance in Plasmodium falciparum is associated with an altered temporal pattern of transcription.* BMC genomics, 2011. **12**: p. 391.

24. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

25. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription.* Cell, 2007. **128**(4): p. 707-19.

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

27. Cui, L., et al., *PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in Plasmodium falciparum.* Eukaryotic cell, 2007. **6**(7): p. 1219-27.

28. Cui, L., Q. Fan, and J. Miao, *Histone lysine methyltransferases and demethylases in Plasmodium falciparum.* International journal for parasitology, 2008. **38**(10): p. 1083-97.

29. Trelle, M.B., et al., *Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum.* Journal of proteome research, 2009. **8**(7): p. 3439-50.

30. Salcedo-Amaya, A.M., et al., *Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(24): p. 9655-60.

31. Scherf, A., et al., *Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum.* The EMBO journal, 1998. **17**(18): p. 5418-26.

32. Lopez-Rubio, J.J., et al., *5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites.* Molecular microbiology, 2007. **66**(6): p. 1296-305.

33. 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.

34. Perez-Toledo, K., et al., *Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes.* Nucleic acids research, 2009. **37**(8): p. 2596-606.

35. Dion, M.F., et al., *Dynamics of replication-independent histone turnover in budding yeast.* Science, 2007. **315**(5817): p. 1405-8.

36. Zee, B.M., et al., *Global turnover of histone post-translational modifications and variants in human cells.* Epigenetics & chromatin, 2010. **3**(1): p. 22.

37. Deal, R.B., J.G. Henikoff, and S. Henikoff, *Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones.* Science, 2010. **328**(5982): p. 1161-4.

38. Ponts N, H.E., Prudhomme J, Wick I, Eckhardt-Ludka C, Hicks GR, Hardiman G, Lonardi S, Le Roch KG., *Nucleosome landscape and control of transcription in the human malaria parasite.* Genome Research, 2010. **20**(2): p. 228-38.

39. Le Roch, K.G., et al., *Discovery of gene function by expression profiling of the malaria parasite life cycle.* Science, 2003. **301**(5639): p. 1503-8.

40. Oehring, S.C., et al., *Organellar proteomics reveals hundreds of novel nuclear proteins in the malaria parasite Plasmodium falciparum.* Genome biology, 2012. **13**(11): p. R108.

41. Corpet, A. and G. Almouzni, *Making copies of chromatin: the challenge of nucleosomal organization and epigenetic information.* Trends in cell biology, 2009. **19**(1): p. 29-41.

42. Levine, B. and D.J. Klionsky, *Development by self-digestion: molecular mechanisms and biological functions of autophagy.* Developmental cell, 2004. **6**(4): p. 463-77.

43. Mizushima, N., et al., *Autophagy fights disease through cellular self-digestion.* Nature, 2008. **451**(7182): p. 1069-75.
44. Li, W.W., J. Li, and J.K. Bao, *Microautophagy: lesser-known self-eating.* Cellular and molecular life sciences : CMLS, 2012. **69**(7): p. 1125-36.

45. Bandyopadhyay, U. and A.M. Cuervo, *Entering the lysosome through a transient gate by chaperone-mediated autophagy.* Autophagy, 2008. **4**(8): p. 1101-3.

46. Takeshige, K., et al., *Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction.* The Journal of cell biology, 1992. **119**(2): p. 301-11.

47. Reggiori, F., et al., *The Atg1-Atg13 complex regulates Atg9 and Atg23 retrieval transport from the pre-autophagosomal structure.* Developmental cell, 2004. **6**(1): p. 79-90.

48. Kabeya, Y., et al., *Characterization of the Atg17-Atg29-Atg31 complex specifically required for starvation-induced autophagy in Saccharomyces cerevisiae.* Biochem Biophys Res Commun, 2009. **389**(4): p. 612-5.

49. Reggiori, F., et al., *Atg9 cycles between mitochondria and the pre-autophagosomal structure in yeasts.* Autophagy, 2005. **1**(2): p. 101-9.

50. Itakura, E., et al., *Beclin 1 forms two distinct phosphatidylinositol 3-kinase complexes with mammalian Atg14 and UVRAG.* Molecular biology of the cell, 2008. **19**(12): p. 5360-72.

51. Obara, K., et al., *The Atg18-Atg2 complex is recruited to autophagic membranes via phosphatidylinositol 3-phosphate and exerts an essential function.* J Biol Chem, 2008. **283**(35): p. 23972-80.

52. Hanada, T., et al., *The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy.* J Biol Chem, 2007. **282**(52): p. 37298-302.

53. Suzuki, K., et al., *Hierarchy of Atg proteins in pre-autophagosomal structure organization.* Genes to cells : devoted to molecular & cellular mechanisms, 2007. **12**(2): p. 209-18.

54. Alvarez, V.E., et al., *Autophagy is involved in nutritional stress response and differentiation in Trypanosoma cruzi.* The Journal of biological chemistry, 2008. **283**(6): p. 3454-64.
55. Williams, R.A., et al., *Characterization of unusual families of ATG8-like proteins and ATG12 in the protozoan parasite Leishmania major.* Autophagy, 2009. **5**(2): p. 159-72.

56. Besteiro, S., et al., *Endosome sorting and autophagy are essential for differentiation and virulence of Leishmania major.* The Journal of biological chemistry, 2006. **281**(16): p. 11384-96.

57. Duszenko, M., et al., *Autophagy in protists.* Autophagy, 2011. **7**(2): p. 127-58.

58. Ghosh, D., et al., *Autophagy is a cell death mechanism in Toxoplasma gondii.* Cell Microbiol, 2012. **14**(4): p. 589-607.

59. Besteiro, S., et al., *Autophagy protein Atg3 is essential for maintaining mitochondrial integrity and for normal intracellular development of Toxoplasma gondii tachyzoites.* PLoS Pathog, 2011. **7**(12): p. e1002416.

60. Totino, P.R., et al., *Plasmodium falciparum: erythrocytic stages die by autophagic-like cell death under drug pressure.* Exp Parasitol, 2008. **118**(4): p. 478-86.

61. Kitamura, K., et al., *Autophagy-related Atg8 localizes to the apicoplast of the human malaria parasite Plasmodium falciparum.* PLoS One, 2012. **7**(8): p. e42977.

62. Eickel, N., et al., *Features of autophagic cell death in Plasmodium liver-stage parasites.* Autophagy, 2013. **9**(4): p. 568-80.

63. Jayabalasingham, B., N. Bano, and I. Coppens, *Metamorphosis of the malaria parasite in the liver is associated with organelle clearance.* Cell Res, 2010. **20**(9): p. 1043-59.

Chapter 1

Global patterns of histone modifications throughout the malaria parasite cell cycle

Anita Saraf[1], Serena Cervantes[2], Nadia Ponts[2], Mihaela E. Sardiu[1], Duk-Won D. Chung[2], Evelien Bunnik[2], Jacques Prudhomme[2], Zhihui Wen[1], Joseph M. Varberg[1], Michael P. Washburn[1,3], Laurence Florens[1,4]*, and Karine G. Le Roch[2, 4]*

*[1]Stowers Institute for Medical Research, 1000 E. 50th Street, Kansas City, MO 64110, USA. [2]Department of Cell Biology and Neuroscience, University of California, 900 University Avenue, Riverside, CA 92521, USA. [3]Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS 66160, USA. * These authors contributed equally to this work.*

Summary

Posttranslational modifications (PTMs) on histone tails can regulate the accessibility of DNA to transcription factors and DNA replication machinery. Thus, histone PTMs plays a role in gene regulation by acting sequentially or in combination to recruit effector proteins. Investigations of histone PTMs in P. falciparum are minimal having only examined histone H3 and a mixed population. By using a shotgun proteomic approach on histone precipitations, this study provides a quantitative assessment of histone PTMs throughout the asexual erythrocytic stage. In total, 246 histone marks were identified with 126 novel histone marks. Interestingly, PTMs associated with transcriptional silencing (methylation) were detected at very low levels and are found on genes that are involved with parasite virulence.

Abstract

While histone modifications are known to have an important role in the human malaria parasite's biology, histone modification sites and their patterns throughout the parasite cell cycle have been poorly defined. Using a comprehensive and quantitative mass spectrometry approach, we determined the global abundance of histones, histone variants, and their covalent post-transcriptional modifications throughout the parasite erythrocytic cycle. We detected a total of 246 distinct post-translational modifications on 8 *Plasmodium* histones. While about half of these were conserved in higher eukaryotes, 204 had never been detected in *Plasmodium* and 126 had never been identified in any other species. Intriguingly, we detected a very low level of modifications commonly associated with transcriptional silencing. Altogether, our data provide the most accurate and comprehensive view of dynamic changes in histone modifications throughout a eukaryotic cell cycle and suggest that the malaria parasite has a unique histone modification signature that correlates with intensive replication state and pathogenicity.

Introduction

The protozoan parasite *Plasmodium falciparum* causes the most severe form of human malaria. The World Health Organization estimates 275 million malaria cases per year, with 655,000 associated deaths. The mortality can be twice as high when including cases that are undiagnosed (Murray et al., 2012; Shetty, 2012). The parasite has a complex life cycle and can adjust its lifestyle to different hosts (Anopheles mosquitos and human) and cell types (mosquito midgut and salivary glands, as well as human hepatocytes and erythrocytes). In humans, the pathogenesis is associated with an intensive replicative state in red blood cells. During this stage, *Plasmodium* has the capacity to develop phenotypic diversity by the selection of clonally variant parasites. The exact molecular processes regulating these mechanisms at the transcriptional level are still poorly understood. Emerging evidence indicates that the chromatin structure plays a critical role (Duraisingh et al., 2005; Freitas-Junior et al., 2005; Issar et al., 2009; Lopez-Rubio et al., 2009; Ponts et al., 2011; Ponts et al., 2010; Ralph et al., 2005; Stubbs et al., 2005; Tonkin et al., 2009).

In all eukaryotic genomes DNA is packed with histone proteins to form the "chromatin". The structure of these chromatin fibers is involved in all DNA-related metabolic processes and transcriptional regulation events by modulating DNA accessibility. The chromatin unit consists of a nucleosome, *i.e.,* classically

~147 base pairs of DNA wrapped around an octamer core of histones containing two each of histones H2A, H2B, H3, and H4. Histone H1 can act as a linker to regulate the formation of the transcriptionally inactive and highly compact 30 nm chromatin fiber. Each core histone contains two functional domains: a histone-fold motif, which mediates histone/histone and histone/DNA interactions, and an N-terminal tail that projects from the core and can undergo a large number of post-translational modifications (PTMs). Known histone PTMs in model eukaryotes include acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, ADP-ribosylation, propionylation, butyrylation, formylation, citrullination, proline isomerization (Martin and Zhang, 2007; Ruthenburg et al., 2007), and the recently reported crotonylation (Tan et al., 2011). These modifications can be found in complex combinations and are primarily localized in the accessible N-terminal tails but can also occur, although less frequently, in the core regions. Variant histones, found in subsets of nucleosomes, can also provide additional complexity. The combined occurrence of histone PTMs and their variants at a particular chromosomal locus is thought to play a key role in gene regulation. Some histone PTMs have been associated with specific cellular functions such as transcriptional activation (*e.g.,* H3 and H4 lysine acetylations), transcriptional elongation (*e.g.,* H3-K36 methylation), or transcriptional repression (*e.g.,* H3-K9 methylation, H3-K27 methylation, and H4-K20 methylation). Overall the dynamic changes in these modifications in a particular cell type on individual

gene *loci* correlate with the overall genetic activity of the cell. The histone code hypothesis suggests that histone PTMs represent an epigenetic code that triggers structural chromatin changes and downstream events controlling gene expression, DNA replication and repair (Strahl and Allis, 2000). It has become clear that the combinations of histone modifications that regulate the many critical activities associated with chromatin dynamics and cell differentiation are much more complex than previously thought.

Most canonical and structural variant histones (H2A, H2A.z, H2B, H2Bv, H3, H3.3, the centromeric-specific H3.c, and H4) have been identified in the malaria parasite but no clear homolog of the linker histone H1 has been found (Gardner et al., 2002). Its absence implies that the *Plasmodium* genome exhibits a lower level of chromosome packaging, and indeed, no clear highly compact chromosomes have been observed to date. The presence of a more relaxed chromatin structure is consistent with recent observations of an atypical euchromatin structure throughout the parasite asexual erythrocytic cycle (Ponts et al., 2010; Westenberger et al., 2009). These recent data are coherent with constitutive binding of the transcription pre-initiation complex on both active and inactive promoters (Gopalakrishnan et al., 2009), as well as a ubiquitous distribution of the transcriptionally permissive histone variant H2A.z in all intergenic regions (Bartfai et al., 2010). Moreover, initial genome-wide chromatin

studies demonstrated an abnormal distribution of histone silencing marks. The silencing H3-K9me3 mark is uniquely associated with genes involved in antigenic variation within restricted subtelomeric and chromosome internal regions (Flueck et al., 2009). Additional studies confirmed the co-localization of the silencing heterochromatin protein 1 to H3-K9me3 marks, within regions of the genome that code for *Plasmodium* virulence factors (Issar et al., 2009). On the contrary, active marks such as the H3-K4me3 and H3-K9ac have a broad distribution across the genome. Finally, preliminary mass spectrometry studies identified a significant number of active acetylated histone marks in the parasite genome (Trelle et al., 2009). Collectively, these data suggest a peculiar parasite chromatin structure with silencing marks restricted to the subtelomeric regions for a tight transcriptional control of genes involved in antigenic variation. As a whole, the role of chromatin and histone PTMs in transcriptional control and pathogenicity deserves a deeper investigation.

In the past, histone PTMs in eukaryotic cells have been identified with a variety of techniques, including immunohistochemistry, chromatography, spectroscopy and mass spectrometry. The characterization of all combinatorial histone PTMs is a great analytical challenge. Antibodies can only measure known PTMs independently and are not designed to achieve a complete combinatorial data set. Furthermore, adjacent modifications can occlude the

epitope that the antibody is designed to recognize, leading to biased results (Egelhofer et al., 2011; Fuchs et al., 2011). Recent studies in the human malaria parasite characterized a few histone PTMs using chromatin immunoprecipitation (ChIP) coupled with DNA microarray hybridization or high-throughput DNA sequencing (Bartfai et al., 2010; Lopez-Rubio et al., 2009; Salcedo-Amaya et al., 2009; Scherf et al., 2008). These antibody-based techniques have only determined the specific genomic location of a restricted number of PTMs in a few parasitic stages. The broad understanding of the role of epigenetic mechanisms in controlling the parasite life cycle will require a complete characterization of all histone modifications in a combinatorial manner. Recently, the ability to determine which PTMs co-exist in a particular cell on a genome-wide scale has emerged by the development of High-Throughput Histone Code Analysis (HT-HCA) methods using mass spectrometry (reviewed in (Young et al., 2010)). While not the most accurate to conserve the combinatorial information, bottom up mass spectrometry, which sequences small peptides from enzymatic digestions, provides a sufficient throughput and quantitative analysis for biologically meaningful information. Similar to how the current high-throughput next-generation DNA sequencers can rapidly analyze genomes, HT-HCA using mass spectrometry can significantly contribute to a better understanding of the epigenome in the malaria parasite.

Here we present a complete atlas of histone PTMs detected across the parasite erythrocytic cycle using a comprehensive and quantitative Multidimensional Protein Identification Technology (MudPIT). Overall, sequence coverages, spectral counts, and relative levels of each of the eight *Plasmodium* histones were consistent across the stages considered, and reproducible between the two independent biological replicates. This dataset allowed us to comprehensively detect PTMs and to reliably quantify whether differences in modification levels existed across *Plasmodium* erythrocytic cycle. Altogether we detected 246 distinct histone post-transcriptional modifications of which 204 had never been detected in *Plasmodium* and 126 had never been identified in other species. The dynamics of histone variants and histone modifications across the seven stages of the erythrocytic cycle were further investigated.

Results

Sequence coverage, purity, and abundance of Plasmodium histones

Synchronized *P. falciparum*-infected blood cultures were harvested at seven time points corresponding to the early ring morphological stage at 0 hours post-invasion (hpi), mid-ring at 6 hpi, late ring at 12 hpi, early trophozoite at 18 hpi, late trophozoite at 24 hpi, early schizont at 30 hpi, and late schizont at 36 hpi (Table 1.1 and Figure 1.1A). Four types of digestion were performed for each of the seven stages analyzed in replicate, resulting in reproducibly high sequence coverage, ranging from 95% to 100%, for each of the seven main *Plasmodium* histones (Figure 1.1B). *Plasmodium* proteins represented up to 82% of all protein identifications (Table S1.2 and Figure 1.1D). General protein levels were similar among the different stages and between replicates (Pearson correlation; Figure 1.1E) and duplicate samples clustered more closely together than with other time points (Figure 1.1F). These tests demonstrated the overall reproducibility of the analyses, both at qualitative and quantitative levels.

Spectra mapping to peptides from *Plasmodium* histones represented between 24% (18-hpi) to 52% (36-hpi) of all spectra, demonstrating a clear enrichment in *Plasmodium* histones (Figure 1.1A). Multiple comparison analysis showed that the increased histone levels observed at early ring and late schizont

Table 1. Overview of MudPIT analysis and SEQUEST Search Settings

For each *Plasmodium falciparum* histone preparation analyzed in this study are specified:
The parasite blood stage from which proteins were acid-extracted;
Whether the cells were treated or not with the MG132 proteasome inhibitor;
The type of digestion used to generate the peptide mixture;
The number of liquid chromatography steps coupled to mass spectrometry applied to the peptide mixture;
The type of mass spectrometer used to acquire the MS/MS dataset;
The data-dependent acquisition sequence used throughout the MudPIT runs;
The software and settings used to abstract and recalibrate the MS1 precursor ion mass and extract the MS2 coordinates from Xcalibur raw files;
The peptide mass tolerance used in the SEQUEST search;
The delta-masses applied statically or differentially to various amino acids during the SEQUEST searches;
The number and type of proteins in the sequence database used to query the MS/MS dataset.

See Experimental Procedures for details.

Sample Preparation			MudPIT Analysis Settings				SEQUEST Search Settings (see details in Table S2)								
										Amino Acid Sequence Database					
Erythrocyte Stage	Proteasome inhibitor (MG132)	Digestion Type	MudPIT LC Steps	MS Instrument	Data Acquisition Settings	Digest Name	MS/MS Files Generated by	Peptide Mass Tolerance	Delta-Mass Parameters (See Supporting Table S2 for details)	Plasmodium NR Forward Amino Acid Sequences	Human NR Forward Amino Acid Sequences	NR Contaminant	NR Shuffled	NR Total	Sample Name (Supporting Tables S2-S4)
0 Hours Post-Invasion (hpi)	-	LysC + Trypsin	12	Thermo LTQ-Orbitrap	1 FT-MS @ 60K 5 IT MS/MS DE= 90s	PfHist_ER_8ug_OT-60K_Ti_1	RawDistiller D(g,6)	10 ppm	Static C: +57.02146 Da; Diff KR: +14.0157 Da; Diff MKDY: +15.9949 Da; Diff KR: +27.9949 Da; Diff KR: +28.0314 Da; Diff MKSTAV: +42.0106 Da; Diff K: +42.0470 Da; Diff M: +58.0055 Da; Diff K: +68.0261 Da; Diff STY: +79.9663 Da; Diff K: +114.0429 Da	5439 Plasmodium falciparum (PlasmoDB release 5.5)	30552 human proteins (NCBI 2008-03-04 release)	162	36153	72306	
	-	Elastase				PfHist_ER_8ug_OT-60K_Es_1									0-hpi_1
	-	LysC + GluC				PfHist_ER_8ug_OT-60K_KCEC_1									
	-	ArgC				PfHist_ER_8ug_OT-60K_RC_1									
	+	LysC + Trypsin				PfHist_ER-MG132_8ug_OT-60K_Ti_1									
	+	Elastase				PfHist_ER-MG132_8ug_OT-60K_Es_1									0-hpi_2
	+	LysC + GluC				PfHist_ER-MG132_8ug_OT-60K_KCEC_1									
	+	ArgC				PfHist_ER-MG132_8ug_OT-60K_RC_1									
6 hpi	-	LysC + Trypsin				PfHist_LR_8ug_OT-60K_Ti_1									
	-	Elastase				PfHist_LR_8ug_OT-60K_Es_1									6-hpi_1
	-	LysC + GluC				PfHist_LR_8ug_OT-60K_KCEC_1									
	+	ArgC[a]				PfHist_LR-MG132_8ug_OT-60K_RC_3[a]									
	+	LysC + Trypsin				PfHist_LR-MG132_8ug_OT-60K_Ti_1									
	+	Elastase				PfHist_LR-MG132_8ug_OT-60K_Es_1									6-hpi_2
	+	LysC + GluC				PfHist_LR-MG132_8ug_OT-60K_KCEC_1									
	+	ArgC				PfHist_LR-MG132_8ug_OT-60K_RC_1									
12 hpi	-	LysC + Trypsin				PfHist_ET_8ug_OT-60K_Ti_1									
	-	Elastase				PfHist_ET_8ug_OT-60K_Es_1									12-hpi_1
	-	LysC + GluC				PfHist_ET_8ug_OT-60K_KCEC_1									
	-	ArgC				PfHist_ET_8ug_OT-60K_RC_2									
	+	LysC + Trypsin				PfHist_ET-MG132_8ug_OT-60K_Ti_1									
	+	Elastase				PfHist_ET-MG132_8ug_OT-60K_Es_1									12-hpi_2
	+	LysC + GluC				PfHist_ET-MG132_8ug_OT-60K_KCEC_1									
	+	ArgC				PfHist_ET-MG132_8ug_OT-60K_RC_1									
18 hpi	-	LysC + Trypsin				PfHist_LT_8ug_OT-60K_Ti_1									
	-	Elastase				PfHist_LT_8ug_OT-60K_Es_1									18-hpi_1
	-	LysC + GluC				PfHist_LT_8ug_OT-60K_KCEC_2									
	-	ArgC				PfHist_LT_8ug_OT-60K_RC_1									
	+	LysC + Trypsin				PfHist_LT-MG132_8ug_OT-60K_Ti_1									
	+	Elastase				PfHist_LT-MG132_8ug_OT-60K_Es_1									18-hpi_2
	+	LysC + GluC				PfHist_LT-MG132_8ug_OT-60K_KCEC_1									
	-	ArgC[b]				PfHist_LT_8ug_OT-60K_RC_2[b]									
24 hpi	-	LysC + Trypsin				PfHist_ES_8ug_OT-60K_Ti_1									
	-	Elastase				PfHist_ES_8ug_OT-60K_Es_1									24-hpi_1
	-	LysC + GluC				PfHist_ES_8ug_OT-60K_KCEC_1									
	-	ArgC				PfHist_ES_8ug_OT-60K_RC_1									
	+	LysC + Trypsin				PfHist_ES-MG132_8ug_OT-60K_Ti_1									
	+	Elastase				PfHist_ES-MG132_8ug_OT-60K_Es_1									24-hpi_2
	+	LysC + GluC				PfHist_ES-MG132_8ug_OT-60K_KCEC_1									
	+	ArgC				PfHist_ES-MG132_8ug_OT-60K_RC_2									
30 hpi	-	LysC + Trypsin				PfHist_LS_8ug_OT-60K_Ti_1									
	-	Elastase				PfHist_LS_8ug_OT-60K_Es_1									30-hpi_1
	-	LysC + GluC				PfHist_LS_8ug_OT-60K_KCEC_1									
	-	ArgC				PfHist_LS_8ug_OT-60K_RC_1									
	+	LysC + Trypsin				PfHist_LS-MG132_8ug_OT-60K_Ti_1									
	+	Elastase				PfHist_LS-MG132_8ug_OT-60K_Es_1									30-hpi_2
	+	LysC + GluC				PfHist_LS-MG132_8ug_OT-60K_KCEC_1									
	+	ArgC				PfHist_LS-MG132_8ug_OT-60K_RC_1									
36 hpi	-	LysC + Trypsin				PfHist_MZ_8ug_OT-60K_Ti_1									
	-	Elastase				PfHist_MZ_8ug_OT-60K_Es_1									36-hpi_1
	-	LysC + GluC				PfHist_MZ_8ug_OT-60K_KCEC_1									
	-	ArgC				PfHist_MZ_8ug_OT-60K_RC_2									
	+	LysC + Trypsin				PfHist_MZ-MG132_8ug_OT-60K_Ti_1									
	+	Elastase				PfHist_MZ-MG132_8ug_OT-60K_Es_1									36-hpi_2
	+	LysC + GluC				PfHist_MZ-MG132_8ug_OT-60K_KCEC_1									
	+	ArgC				PfHist_MZ-MG132_8ug_OT-60K_RC_1									

[a] Because less protein material was obtained from the untreated 6 HPI sample, the fourth RC digest was done on the MG132-treated 6 HPI sample.
[b] Because less protein material was obtained from the MG132-treated 18 HPI sample, the fourth RC digest was done on the untreated 18 HPI sample.
Note that protein detection and levels between the untreated and MG132-treated samples were highly reproducible as calculated by Pearson (See SuppFig S1E).

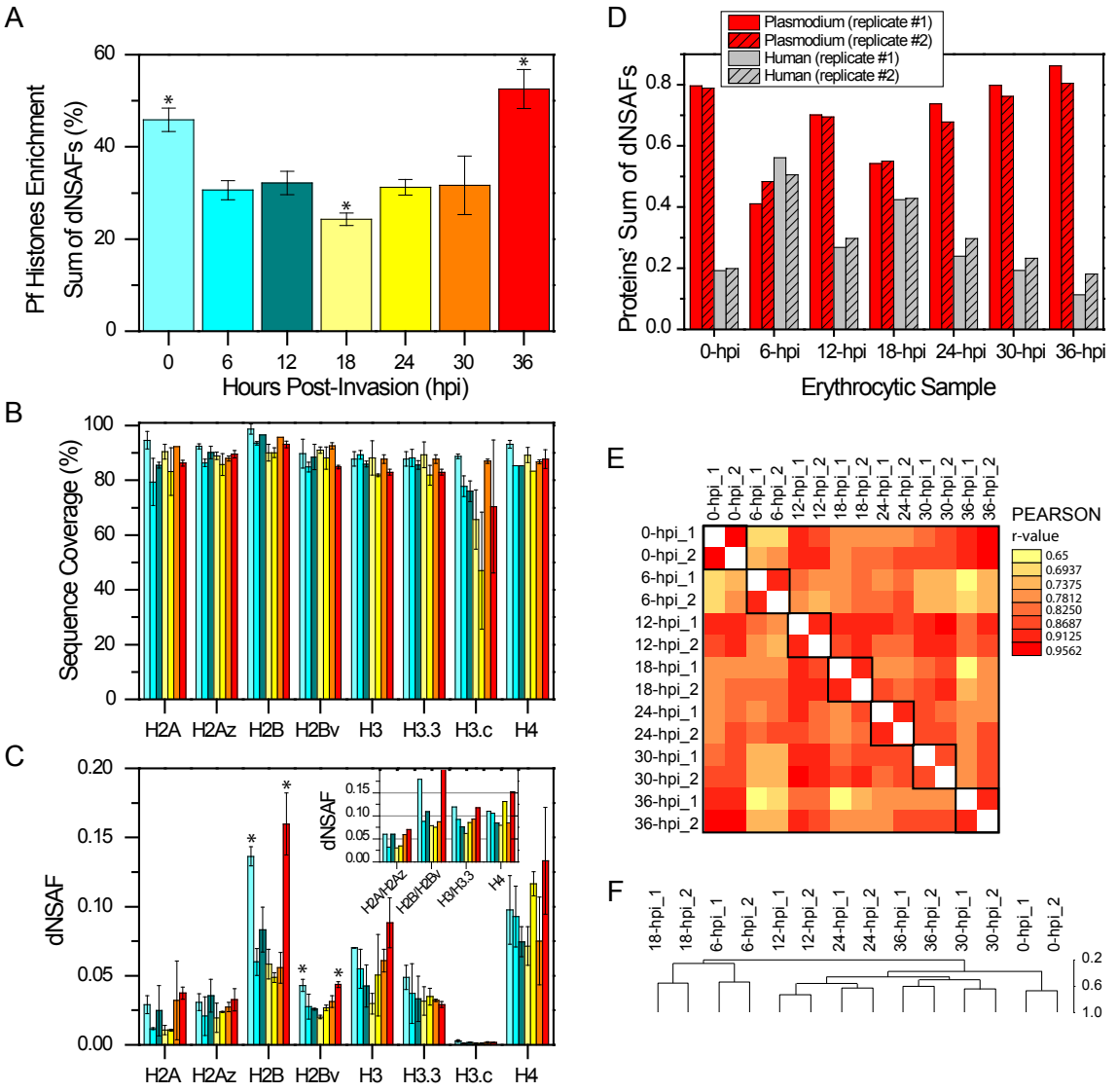


Figure 1.1: Sequence coverage, purity, abundance, and reproducibility. Synchronized *Plasmodium falciparum*-infected blood cultures were harvested every six hours from 0 hours post-invasion (hpi) to 36 hpi. The seven samples are early ring (0-hpi, light cyan bars), middle ring (6-hpi, cyan), late ring (12-hpi, dark cyan), early trophozoite (18-hpi, light yellow), late trophozoite (24-hpi, yellow), early schizont (30-hpi, orange), and late schizont (36-hpi, red). Two sets of acid-precipitated histones were independently prepared and analyzed for each time-point (Table 1.1). **A.** Enrichment of Plasmodium histones in each preparation is calculated by summing the normalized spectral counts values (Sum of dNSAF) of the eight histones (Table S1.2). To assess the variations in individual histone abundances across the intra-erythrocytic cycle, a multiple comparison analysis was performed using a bonferroni correction on histone abundances using matlab. Asterisks denote statistically significant differences in enrichment levels ($p < 0.05$) **B.** The percentage of amino acid sequence covered by peptides detected in each sample type is calculated for each *Plasmodium* histone. **C.** Relative protein levels in each sample type are estimated using dNASF values averaged between two independent biological replicates for each

stage submitted to multiple digestions (Table 1.1). Asterisks denote statistically significant differences in protein levels ($p < 0.05$). **D**. The relative content in plasmodium and contaminating human proteins detected by MudPIT analyses of each duplicate time-point was estimated by summing dNSAF values (Sum of dNSAF) for each group of proteins (Table S1.2). Summing the normalized spectral counts values (SdNSAF) of specific group of proteins (histones in Figure 1.3A, *Plasmodium* vs. human in Figure 1.1A) was used to assess their relative enrichment over all detected proteins as described in (Florens et al., 2006). **E.** Pearson correlation was calculated between all stages using the R environment function cor(). r-values obtained for each combination were plotted as a 14 x 14 symmetric matrix. Values of r close to +1 indicate a strong correlation, while values of r close to 0 indicate very poor correlation. **F.** Relative protein abundances expressed as dNSAF were used to cluster all detected proteins and each analysis (Table S1.2). Rows and columns were hierarchically clustered under the same criteria using MeV 4.8.1 (Mar et al., 2011) and HCL function with Pearson correlation as metric distance and average linkage as algorithm.

stages as well as the lower levels detected at early trophozoite stage were significantly different from those measured from other stages ($p < 0.05$, Figure 1.1A). This pattern was consistent with previously observed genome-wide variations of nucleosome-bound DNA levels (Ponts et al., 2011; Ponts et al., 2010).

Several isoforms of histones H2B and H3 are present in the *Plasmodium* genome and all were detected in our analysis. Although sequence identity exists between these isoforms, their respective abundances were calculated using the dNSAF equation (Zhang et al., 2010). H4 was the most abundant (Figure 1.1C), while H2B/H2Bv and H3/H3.3 added up to H4 levels (Figure 1.1C, insert). H3.c was the least abundant isoform, consistently with its centromere-restricted localization. Most interestingly, the transcriptionally permissive histone variant H2A.z, which has been characterized to be at least one order of magnitude less

abundant than H2A, in in human and budding yeast, was recovered at levels similar to H2A (Figure 1.1C).

Comparison of the post-translational modifications detected on Plasmodium histones against previously reported PTMs

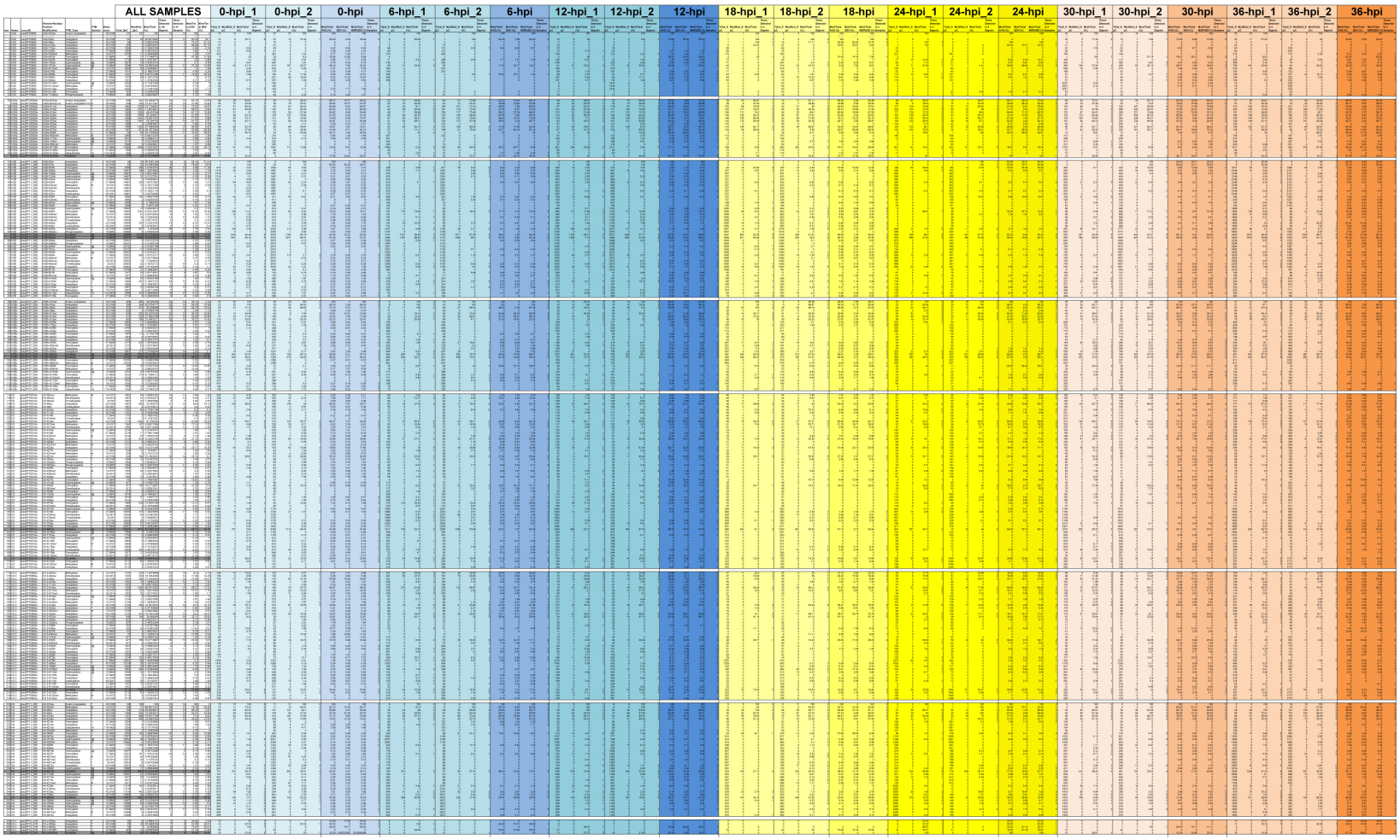
The MS/MS datasets were searched for oxidation/hydroxylation, formylation, methylation, dimethylation, trimethylation, acetylation, crotonylation, ubiquitination, phosphorylation, as well as N-terminal acetylation (Table S1.1). We identified a total of 256 modifications on the eight *Plasmodium* histones (Table 1.2 (Supplemental PDF available through dissertation publisher)). Twenty of these modifications were carried by peptides undistinguishable between the H3/H3.3 isoforms (same spectral counts, Table 1.2), and were therefore counted as one. All other "redundant" modifications on homologous residues within H3/H3.3 and H2B/H2Bv were detected on distinct peptides (different spectral counts; Table 1.2) and were hence considered independently.

Overall these 246 amino acid/PTM combinations (AA#PTMs) fell into three main categories (Table S1.3): 46 modifications had been described before in *Plasmodium* (of which 43 were previously known in other species); 74 AA#PTMs had been previously reported on evolutionary conserved residues in other organisms; and 126 were novel AA#PTMs (Figure 1.2B). Of these 126 new

PTMs, 57 were detected on residues that had been shown previously to be modified by

Table 2. Detailed spectral counts for modified and unmodified peptides mapping to Plasmodium falciparum histones measured across 7 erythrocytic stages.

In each replicate analysis, the following parameters are specified for the amino acid (AA)/post-translational modification (PTM) combination detected within each histone:
Total spectral counts for peptides bearing the amino acid residue specified in each row (column labeled "Total_SpC");
Spectral counts for peptides bearing the specified AA#PTM combination listed in each row (column labeled "Modified_SpC");
The ratio, expressed as a percentage, between modified and total spectral counts for each AA#PTM combination (column labeled "Mod/Total");
The number of independently digested samples the specified AA#PTM combination were detected in (column labeled "Times Detected in 4 Digests").
For each of the 7 sample types, average (AVG) and standard deviation (SDV) of the estimated modification levels (Mod/Total ratios) were calculated for each AA#PTM combination.
The ratio between modified and total spectral counts for each AA#PTM combination was also calculated by summing up the values measured for the 2 replicate analyses (column labeled "Mod/Total MERGED");
The number of times each AA#PTM combination was detected in the 2 biological replicate analyses is reported in the column labeled "Times Detected in 2 Samples".



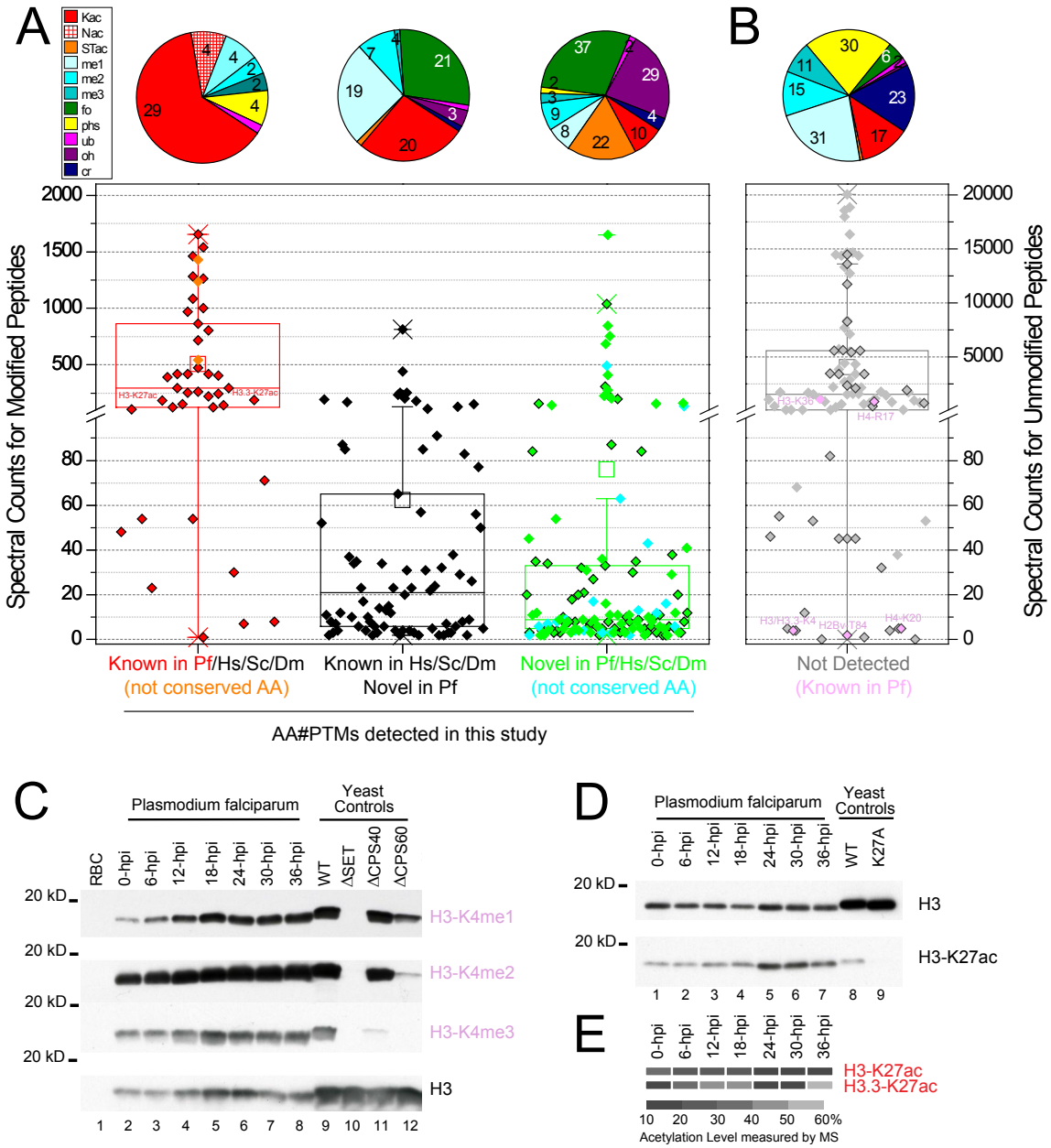


Figure 1.2: Summary statistics of detected (or not) post-translational modifications. A. *Box plot representation of spectral counts for peptides bearing the 246 distinct modifications detected in this analysis*. The summary statistics are shown for 1) PTMs previously reported in *Plasmodium*; 2) PTMs previously reported on conserved residues in human/mouse, yeast, or fly; and 3) the novel amino acid/PTM combinations uncovered in this dataset. In a box plot, the 25th and 75th percentile are represented by the upper and lower boundaries of the box, with the median being the line dissecting the box, and the mean being the small square in each box. The

5th and 95th percentiles are shown with lines attached to the box, the 'X' represents the 1st and 99th percentiles, and the stand alone '–' represents the complete range. The data points are overlaid within each box and the symbols color-coded as in Figure 1.2A. Above each category, the breakdowns of each modification type are plotted as pie charts. Pie slices with not values have 1 entry. **B.** *Summary statistics of total spectral counts for unmodified peptides bearing residues that have been previously detected with certain types of modifications but were not modified as such in our dataset*. Not detected modification sites previously reported in *Plasmodium* are plotted as labeled pink symbols, while not detected modified sites reported in human/mouse, yeast, or fly are represented by grey symbols. Grey symbols with a dark edge stand for residues that were never detected as modified in our dataset, while entirely grey symbols are for residues that were detected as modified by other types of modifications in our dataset. The box plot representation and pie chart are as described above. **C.** *Western blot analysis using antibodies directed against each of the three methylation states of S. cerevisiae H3-K4*. Histones isolated from seven *Plasmodium* erythrocytic stages and uninfected red blood cells (RBC) were separated on 18% SDS-PAGE and transferred for blotting. Whole cell extracts from wild-type yeast as well as compass-deficient mutants were analyzed in parallel to ensure proper behavior of each antibody (Takahashi et al., 2009). **D.** *Western blot analysis using an antibody directed against acetylated S. cerevisiae H3-K27*. Histone preparations isolated from seven erythrocytic stages as well as a *S. cerevisiae* strain expressing the H3-K27A mutant were processes as in **C**. In **C.** and **D.**, loading amounts in each lane were normalized using an antibody against yeast H3. **E.** *Acetylation levels on K27 in both H3 isoforms as measured by our MS analysis across the seven time points* (Table 1.2). The acetylation percentages were quantified using local spectral counts and plotted in grey scale from dark to light.

other PTMs ("known AA"), while 19 were detected on nine non conserved residues ("n.c. AA").

At the quantitative level, the peptides bearing the 46 PTMs previously reported in *Plasmodium* (Cui et al., 2008; Cui and Miao, 2010; Dastidar et al., 2013; Treeck et al., 2011; Trelle et al., 2009) were detected with large numbers of spectral counts (Figure 1.2A) and mostly consisted of acetylations and methylations on the N-terminal tails (Figure 1.4B). The peptides bearing the 74 PTMs on sites known in other species were overall less abundant, while the 126 entirely novel PTMs were detected within peptides with the lowest spectral counts with the exceptions of a few serine acetylations, formylations and

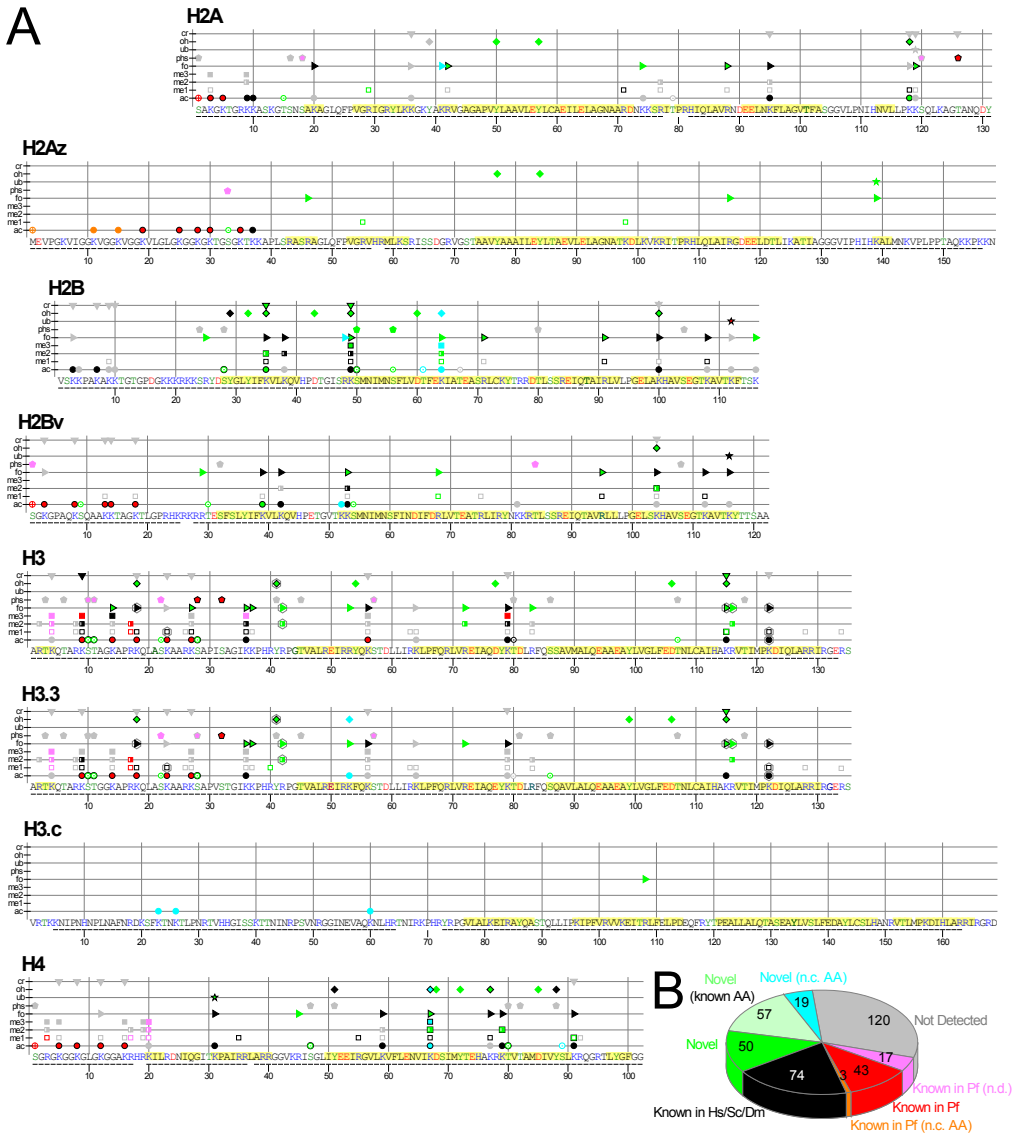


Figure 1.3: Qualitative overview of post-translational modifications along histone sequences. A. *All 256 PTMs detected on Plasmodium histones H2A, H2Az, H2B, H2Bv, H3, H3.3, H3.c, and H4 (Table 1.2).* Acetylation (circle), mono-, di-, trimethylation (open, half-filled and full square, respectively), formylation (triangle), phosphorylation (pentagon), ubiquitinations (star), hydroxylations (diamond), and crotonylations(downward triangle) are listed above the sequences from bottom to top. Underlined amino acid residues are covered by detected peptides. Amino acids predicted to be within structured domains based on sequence alignments (Figure 1.5) with histones of solved tridimensional structures are highlighted in yellow. Detected PTMs are split into three main color-coded categories: 46 PTMs previously reported in *Plasmodium falciparum* are indicated in red symbols with a black outline when these PTMs are also known in human, yeast, or fly, or orange symbols if these PTMs are on residues not conserved (n.c.) in other species (H2Az N-terminus); 74 PTMs previously reported in species other than *Plasmodium* are shown in black symbols; symbols for 116 novel AA position/PTM combinations are in bright green, with 57 of these symbols drawn with a black edge to denote that

the amino acid had been previously reported as modified by other types of PTMs; 19 novel modifications on residues not conserved in other species are colored in light cyan (Table S1.3). Twenty symbols within grey octagons in the H3 and H3.3 plots indicate that the peptides bearing these modifications are undistinguishable between the two isoforms. An additional 137 PTMs (removing isoform redundancies) that have been previously reported but not detected (n.d.) in our analysis are plotted as grey symbols. Among these, 17 modifications that have been detected on *Plasmodium* histones by mass spectrometry or using antibodies are colored in pink.
B. *Qualitative summary of the PTMs detected or not in this dataset*. Color-coded pie slices for each of the PTMs novelty categories are as defined above.

hydroxylations (Table S1.3). The majority of the novel sites were located within the histone fold domains (Figure 1.4A).

A total of 120 PTMs that had been previously reported on 84 residues conserved in *Plasmodium*, were not detected in our dataset (Table S1.3). The majority of these PTMs were methylations, phosphorylations, crotonylations, and acetylations (Figure 1.2B). Only 25% of the residues bearing these known PTMs were detected as entirely unmodified in our dataset (symbols with dark grey edges; Figure 1.2B). Some of these residues were in regions of low spectral coverage, therefore we could not conclude whether previously reported PTMs might have been present in the *Plasmodium* samples we analyzed. The majority of the "missing" PTMs were on residues that we detected as modified by another type of PTM(s) (light grey symbols; Figure 1.2B). It is then likely that these previously reported AA#PTMs were not present (or present below our detection limit) on the *Plasmodium* histones we analyzed. In addition to methylations, the two largest classes of PTMs missing from our dataset were phosphorylations and

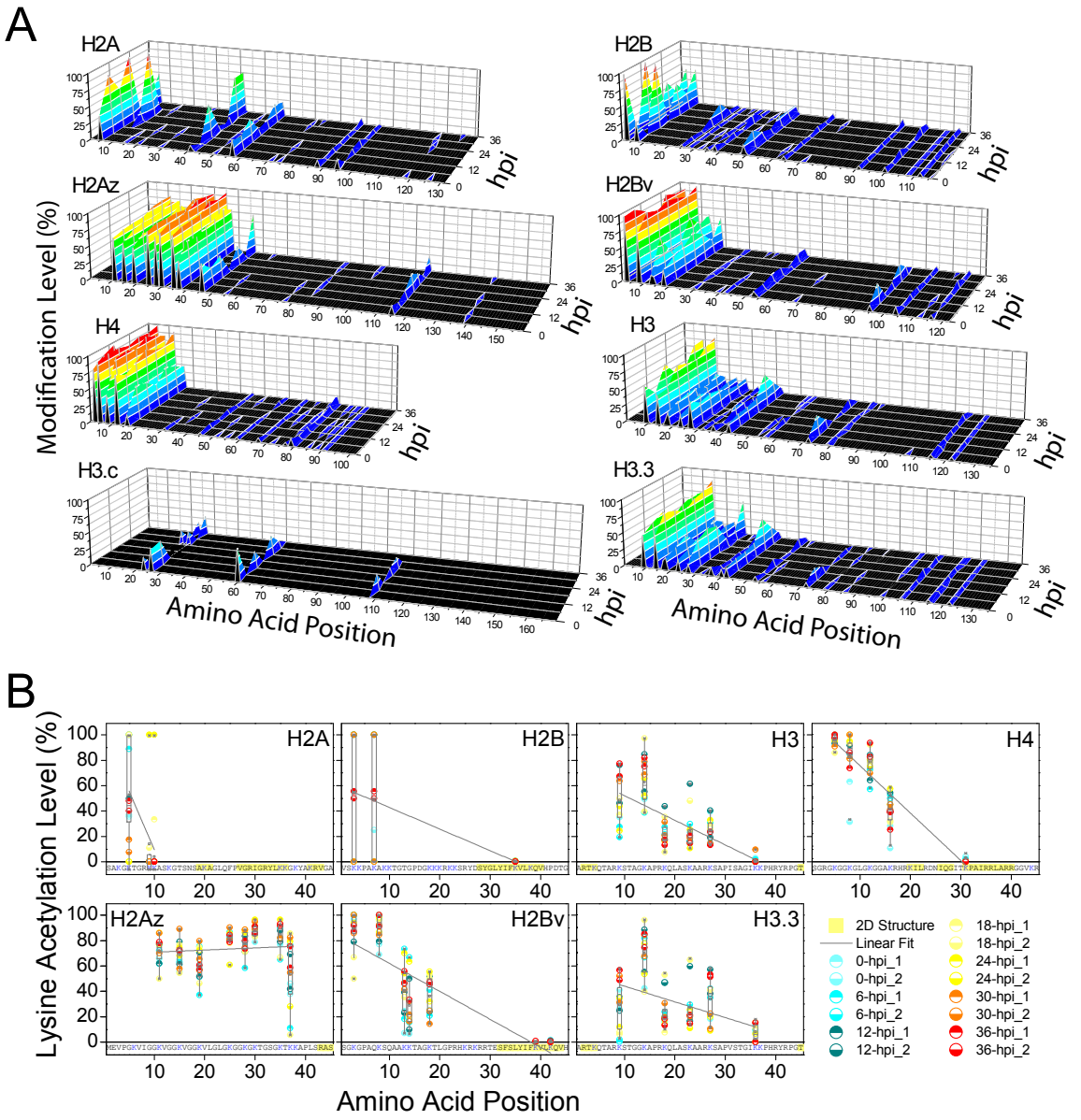


Figure 1.4: Dynamic landscape of PTM levels along histone sequences. A. *Modification levels (z-axis) for all amino acid residues bearing the 252 PTMs detected on Plasmodium histones.* Percentages of modification occupancy (Table 1.2) are plotted as a color-coded 3D-surface as a function of their positions along the sequences (x-axis) and the stages they appear in (hours post-invasion; y-axis). Note that 4 N-terminal acetylations on H2A, H2Az, H2Bv, and H4 are not plotted. When multiple types of PTMs are detected on the same residue, the modification levels are added for each of the seven time points. **B.** *Acetylation levels measured in each of the 14 analyzed samples.* Percentages of acetylation occupancy are plotted along the 45 first amino acid residues of *Plasmodium* histones H2A, H2Az, H2B, H2Bv, H3, H3.3, and H4 (Table 1.2). Box-plots are overlaid above the data points and are as described in Figure 1.3A. Linear regressions through the data points are plotted as a grey line.

crotonylations (Figure 1.2B). As we did not perform a phosphopeptide enrichment step, the low number of phosphorylated residues was not entirely surprising. However, a large number of these reported phosphorylation and crotonylation sites have been observed in yeast and human under particular cellular conditions. For example, histone phosphorylations have been linked to DNA damage (Cheung et al., 2005), mitotic chromosome condensation (Choi et al., 2005; Goto et al., 1999; Goto et al., 2002; Preuss et al., 2003), or apoptosis conditions (Ahn et al., 2005a; Ahn et al., 2005b; Cheung et al., 2003), while lysine crotonylations are enriched on sex chromosomes (Tan et al., 2011). Such cellular conditions are likely very different from the seven *Plasmodium* stages that we analyzed.

Of the 63 PTMs previously reported in *Plasmodium* (Cui et al., 2008; Cui and Miao, 2010; Dastidar et al., 2013; Treeck et al., 2011; Trelle et al., 2009), 17 were not detected in this dataset (Table S1.3). Half of these were phosphorylated sites that were detected in two recent mass spectrometry studies using phosphopeptide enrichment (Dastidar et al., 2013; Treeck et al., 2011). The H4 peptides bearing R17 were detected as unmodified with high spectral coverage (Figure 1.2B) in all of the 14 analyzed samples (Table 1.2). H3-K36 was detected as mono- and dimethylated, acetylated, and formylated on both H3 isoforms in our dataset. It is therefore likely that monomethylation of H4-R17 and trimethylation of H3-K36 were not present in the seven time points we analyzed.

Although we identified the previously reported H2A-K3ac, it was detected by a single spectrum and we hence considered this modification to be of very low abundance/confidence in our dataset (Table S1.3).

The peptides bearing H3/H3.3-K4, H4-K20, and H2Bv-T84 are in regions of very low spectral coverage (Figure 1.2B). Such low detection could be due to technical limitations in our analysis pipeline. For example due to the hydrophobic nature of the small tryptic peptide bearing H3-K4, we also had trouble detecting this H4-K4 methylation in other organisms including *S. cerevisiae* that have relatively high levels of this PTMs (Garcia et al., 2007). We therefore used highly characterized antibodies recognizing specific methylation states of H3-K4 to confirm that H3-K4me1, H3-K4me2 and H3-K4me3 are present across the erythrocytic cycle, at levels comparable to what is found in yeast cell extracts (Figure 1.2C). In support of the existence of H3-K4me3 in *Plasmodium*, trimethylation is likely carried out by the previously described PfSet1 protein (Cui et al., 2008), which shares a similar domain architecture to the MLL-like protein found in Ciliates (Smith et al., 2011). Western blot also confirmed the presence of H3-K27 acetylation in all seven time points (Figure 1.2D) and correlated well with our quantitation of H3-K27 and H3.3-K27 acetylation levels by local spectral counts in all stages analyzed (Figure 1.2E). The advantage of our MS-based dataset is that we are able to independently quantify modification levels on homologous residues present within histones isoforms.

Overview of the type of post-translational modifications detected on Plasmodium histones

N-terminal processing and acetylation — Leading methionines were post-translationally removed from histones H2A, H2Bv, and H4, similarly to other eukaryotes. The serines in the new first position were 100% N-terminally acetylated (Table S1.3). Histones H3, H3.3, H3.c, and H2B were also N-terminally processed, although acetylation of their new first residues (alanine for H3 and H3.3, and valine for H3.c and H2B) was not detected. On the contrary, the leading methionine of histone H2A.z was not processed, but was 100% N-terminally acetylated (Table S1.3). N-terminal acetylation levels did not vary across the seven time-points we analyzed (Table 1.2). The differences in the nature of the first amino acid and the acetylation state may reflect differences in histone stability *in vivo*, possibly according to an N-end rule. The exact effect of each N-terminal ending is complex and unknown. For example, serine, alanine, and valine have been fairly recently classified as stabilizing (see Tasaki and Kwon, 2007 for a review), whereas acetylation could be destabilizing (Hwang et al., 2010). N-terminal acetylation is however not considered an epigenetic mark, and hence will not be discussed any further.

Lysine acetylation — A large number of the identified PTMs consisted of well-known acetylation events (Figure 1.2A) in agreement with previously analyzed histones isolated from an asynchronous mixture of *Plasmodium*

asexual stages (Trelle et al., 2009). A total of 10 new lysine residues were found to be acetylated (Figure 1.2A) and 7 of them were on lysines not conserved across species (Tables S5-6). Three residues contributed half of the novel PTMs on non-conserved positions: H2B-K64, H3.3-K53, and H4-K67, which are all substituted for arginines in other species (including in their Plasmodium H2Bv and H3 isoforms). The presence of lysines at these positions in *Plasmodium* histones allows for a larger array of modifications (acetylations, trimethylations, and hydroxylations) that are not possible on an arginine side chains. The Four of the novel PTMs were detected on PfH3.c (K23ac, K26ac, K60ac, and R108fo). The H3.c N-terminal tail is very divergent across species (Figure 1.5) and the three acethylated lysine located within this tail in Plasmodium are not conserved. The formylated arginine H3.c-R108, located in a turn at the end of the putative second alpha helix might be conserved in human H3.cAa-K76 and fly H3.c-K163 (Table S1.3). To our knowledge, these are the only modifications ever detected on centromeric histones H3 from any species.

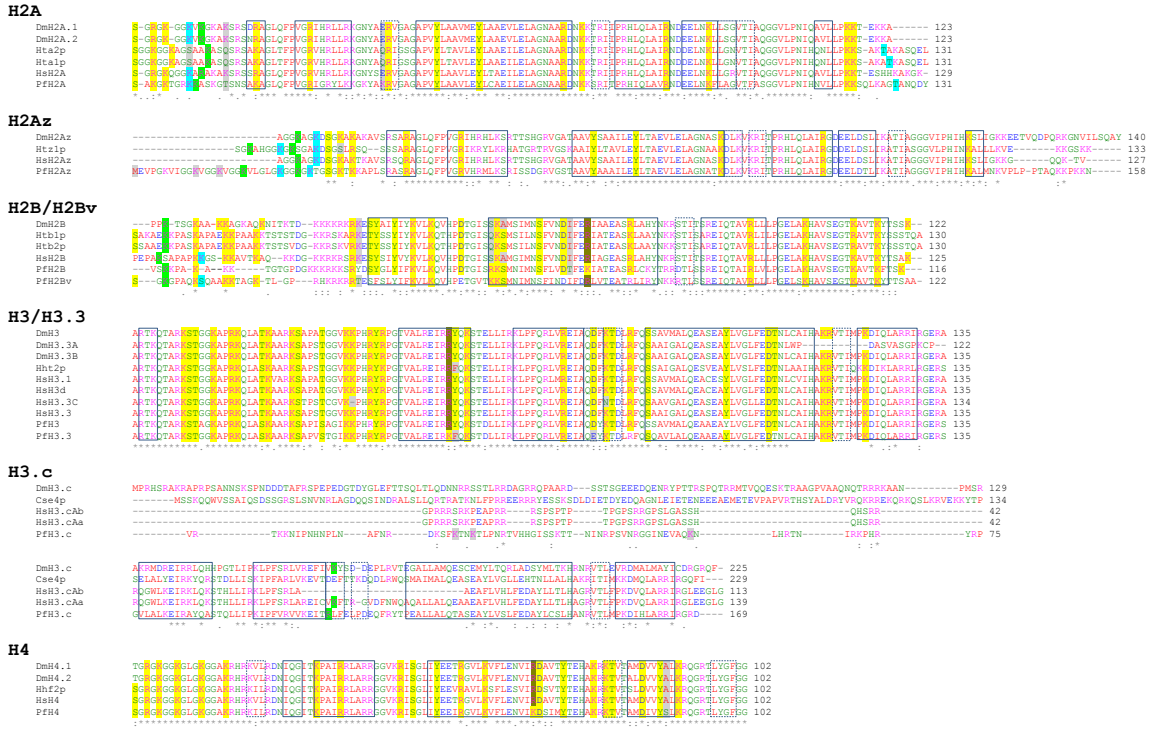


Figure 1.5: Alignments of histone sequences from *Plasmodium falciparum*, *Homo sapiens*, *Saccharomyces cerevisiae*, and *Drosophila melanogaster*. Multiple sequences alignments were mostly performed using ClustalW2 (Goujon et al., 2010; Larkin et al., 2007). However, the N-terminal sequences of the four H2Az proteins are very divergent, while the C-termini are very similar. Therefore, the alignment has been manually modified to align from the C-terminus toward the N-terminus.
Accession numbers for each amino acid sequence are:
H2A: PfH2A (psu|PFF0860c), HsH2A (NP_254280.1), ScHhta1 (NP_010511.1), ScHta2 (NP_009552.1), and DmH2A (NP_001027376.1);
H2Az: PfH2Az (psu|PFC0920w), HsH2Az (NP_002097.1), ScHhtz1 (NP_014631.1), and DmH2A.v (NP_524519.1);
H2AB/H2Bv: PfH2B (psu|PF11_0062), PfH2Bv (psu|PF07_0054), HsH2B (NP_066402.2), ScHtb1 (NP_010510.1), ScHtb2 (NP_009553.1), and DmH2B (NP_001027385.1);
H3: PfH3 (psu|PFF0510w), PfH3.3 (psu|PFF0865w), HsH3.1t (NP_003484.1), HsH3d (NP_003521.2), HsH3.3C (NP_001013721.2), HsH3.3 (NP_005315.1), ScHht2 (NP_014367.1), and DmH3 (NP_001027387.1), DmH3.3A (NP_723056.1), DmH3.3B (NP_788892.1);
H3.c: PfH3.c (psu|PF13_0185), HsH3.cAa (NP_001800.1), HsH3.cAb (NP_001035891.1), ScCse4 (NP_012875.2), and DmH3.c (NP_523730.2);
H4: PfH4 (psu|PF11_0061), HsH4 (NP_001029249.1), ScHhf (NP_014368.1), DmH4.1 (NP_001027377.1) and DmH4.2 (NP_001027382.1).
Boxed regions correspond to secondary structures based on solved tridimensional structures (solid box, alpha-helix; dotted box: turn; dashed box: strand; as described at http://www.uniprot.org). Residues detected as modified in our analysis are highlighted as well as the corresponding residues across species (grey highlights denoting non-conserved sites). These alignments were used to compare modified sites across species (Table S1.3).

Serine and threonine acetylation — Most of the novel acetylation sites were on 22 serine and threonine residues (Table S1.3). H2A was acetylated at position T15, a residue aligning with K15/K14 in human/fly H2A (Table S1.3 and Figure 1.5), which has also been shown to be acetylated (Portela and Esteller, 2010). H2A.z was acetylated on S33. Highly reproducible acetylation was detected on histone H2B residues S28, S50, S56, and T61, as well as H2Bv residues S9, T30, and S54 (Table S1.3). H2Bv-S9 aligns with K12 in human (Figure 1.5), which has been shown to be acetylated (Beck et al., 2006; Golebiowski and Kasprzak, 2005; Kim et al., 2006). Similarly H2Bv-T30 aligns with a lysine (Table S1.3) previously reported as formylated and crotonylated in other species (Tan et al., 2011; Wisniewski et al., 2008). Histone H4 was acetylated on S47, T80, and S89 (Table S1.3). Finally, H3/H3.3-S10, T11, S22, S28, as well as H3-T80, H3-T107, and H3.3-S86 were detected as acetylated multiple times. For the homologous positions in the H3 isoforms, these modifications were identified on distinct peptides. Recently, H3-T80 has been reported as acetylated in mouse H3 (Tweedie-Cullen et al., 2012).

Methylation — A large number of histone methylation marks were detected throughout the parasite erythrocytic cycle, the large majority of them being mono- and dimethylations and present at fairly low levels (Table 1.2). These methylations were mostly found within the histone cores and only histone H3 and its variant H3.3 had multiple methylation sites on their N-terminal tails

(Figure 1.3A). In addition to the classical methylations usually observed in eukaryotic cells, we identified 20 novel methylation marks (Figure 1.2A). Interestingly, two conserved arginine and lysine residues located at homologous positions within H2A and H2Az core domains (Figure 1.3A) were both found monomethylated. These residues, H2A-R29/R71 and H2Az-R55/K98 reside within the first and second putative alpha-helices of H2A and H2Az, respectively (based on H2A structure). H2B-K35 and K49, which had been shown to be me1 (Portela and Esteller, 2010) and me1/me2 (Tan et al., 2011) respectively, were further detected as K35-me2 and K49-me3 in our dataset (Table S1.3). H2B-K64 was newly identified in all three methylation states, while R68, its equivalent position in the H2Bv isoform, was monomethylated. H2Bv-K104 that has been described previously as monomethylated (Beck et al., 2006) was only detected as dimethylated in our samples. H3-K14 previously known to be me1/me2/me3 (Garcia et al., 2007; Tan et al., 2011; Tweedie-Cullen et al., 2012) was only detected as fully trimethylated. H3-R72 was dimethylated and formylated, the first report of any type of PTMs on this conserved residue. H3-K115, known to be acetylated (Lall, 2007), was newly discovered to be monomethylated, as well as formylated, hydroxylated and crotonylated (Figure 1.3A). R42 and R116 in H3/H3.3 were both newly found to be dimethylated and formylated, while H3.3 R40 was monomethylated. In our dataset, H4-K79 and H4-K91, both previously detected as acetylated and formylated (Olsen et al., 2006; Tan et al., 2011;

Wisniewski et al., 2008; Yang et al., 2011; Zhai et al., 2008), were also me2 and me1, respectively. In addition, H4-K67, which is substituted by an arginine residue known to be methylated in histones H4 from other species (Tan et al., 2011), was detected as me2/me3. Consistent identification of novel methylations on distinct peptides from isoforms and the detection of higher numbers of methyl groups on residues previously known to be methylated strengthen our confidence in such identifications.

Phosphorylation — A total of six phosphorylation sites were confidently identified in this dataset: H2A-T126, H2B-S50, H2B-S56, H3-S28, H3-S32, and H3.3-S32. The phosphorylations on H2A and H3 have been previously reported in *Plasmodium* (Treeck et al., 2011), human, and/or yeast (Table S1.3 and references therein), while H2B-S50 and S56 phosphorylations are novel despite the conservation of these residues across species (Table S1.3). Their abundance is extremely low and restricted to the beginning and the end of the erythrocytic cycle, with the exception of the conserved H3-S32phs that is present at all stages at slightly higher levels (on average 4% of all histones H3).

Ubiquitination — Ubiquitination was detected on H2B-K112, as previously reported in *Plasmodium* (Trelle et al., 2009), as well as on its homologous residue K116 in H2Bv. Ubiquitination of H2B-K112 is a well-characterized event associated with active transcription in both yeast (K123) (Tagwerker et al., 2006) and human (K120) (Pavri et al., 2006). On the other hand, the ubiquitination

marks on the conserved K31 within the H4 core domain, which had been previously reported as acetylated and formylated (Choudhary et al., 2009; Tan et al., 2011; Wisniewski et al., 2008), and K139 within the H2Az C-terminal tail are novel (Table S1.3). Abundance of these novel ubiquitylated residues was in general very low and restricted to a couple of stages. Exception are H2B-K112ub, which was present in all stages at slightly higher levels (6% of the PfH2B in average), and H2Bv-K116ub which was detected in 6 out of the 7 time points at about 5% of PfH2Bv. The roles of these rare ubiquitylation events are unknown.

Histone crotonylation — Crotonylation of lysine side chain has been recently reported on mammalian histones (Tan et al., 2011). While we were able to identify 5 crotonylated lysines in our dataset (H2B-K35, H2B-K49, H3.3-K115, H3-K115 are novel, and H3-K9 is known; Table S1.3), crotonylation constitutes one of the main classes of PTMs that was notably absent from our list (Figure 1.2B). Most of the missing crotonylations were reported on the N-terminal tail of H2B, H3 and H4 (Figure 1.3A) on lysine residues that were otherwise acetylated at high levels in the seven *Plasmodium* stages we analyzed (Figure 1.4B). Histone crotonylation has been shown to play a role in the post-meiotic stages of spermatogenesis (Tan et al., 2011), a specific epigenetic function that is not relevant to *Plasmodium* biology.

Histone hydroxylation — Hydroxylation is a reversible chemical process that attaches at least one hydroxyl group to a residue side chain. Aspartic acid and lysine may be hydroxylated on the b-C and the d-C atom of their side-chains, respectively. Tyrosine may be hydroxylated in the meta position of its aromatic ring and converted to L-3,4-dihydroxyphenylalanine (L-DOPA) by tyrosine hydroxylase. Hydroxylation is less common than other types of PTMs (Kaelin, 2005), yet hydroxylations of 13 lysine, 7 aspartate, and 9 tyrosine residues constituted the second largest class of novel PTMs in our dataset (Figure 1.2A). A total of four tyrosine hydroxylations have recently been reported in the literature on H2B-Y29, H4-Y51, H4-Y88, and H2A-Y39 (Tan et al., 2011), all but the last one being detected in our dataset as well (Table S1.3). Among the novel hydroxylated tyrosines are H2A-Y50 and H2A-Y57 and their homologous positions Y77 and Y84 in H2Az, as well as H3/H3.3-Y41. These residues are conserved and H2A-Y58 and H3-Y41 have been shown by alanine substitution scanning to be essential in *S. cerevisiae* (Nakanishi et al., 2008). While H3-Y41 has been previously reported as phosphorylated (Yu et al., 2007), this is the first modification reported on these H2A and H2Az tyrosines.

Histone formylation — While some N6-formylation of lysine has previously been associated with histones (Table S1.3 and references therein) and other nuclear proteins (Jiang et al., 2007b), lysine and arginine formylations represented the highest number of novel PTMs identified in our dataset (Figure

1.2A). Combining 21 previously reported lysine and arginine formylations with 37 novel formylated residues, this PTM represented 24% of our dataset, on par with acetylations and methylations (Table 1.3).

Positional effect of modified sites on modification levels

To estimate to what extend a particular residue was modified, we divided the local spectral counts matching a particular modification by the total spectral count for this amino acid position (Xiang et al., 2007). For each modified site, we hence obtained modification levels (Figure 1.6) that were expressed as a percentage of the total spectral count for an amino acid position (Table 1.2). A clear dichotomy in modification levels was observed for all *Plasmodium* histones analyzed: accessible N-terminal tails (Figure 1.5) were heavily modified at high levels, while most PTMs detected within the structured domains were at much lower levels of relative abundance (Figure 1.4A).

We reproducibly detected the majority of the well-known acetylation events on the lysine residues located within the N-terminal tails of H2A, H2B, H3 and H4 (Table 1.2). However, acetylation levels appeared to be lower for residues located more internally within the N-terminal tails (Figure 1.4A). To investigate this trend further, we plotted all of the quantitative data-points that were available for each acetylated lysine within N-terminal tails (Figure 1.4B).

$$AA_x\_PTM_y\_Level = \frac{AA_x\_PTM_y\_SpectralCounts}{AA_x\_Total\_SpectralCounts} x100$$

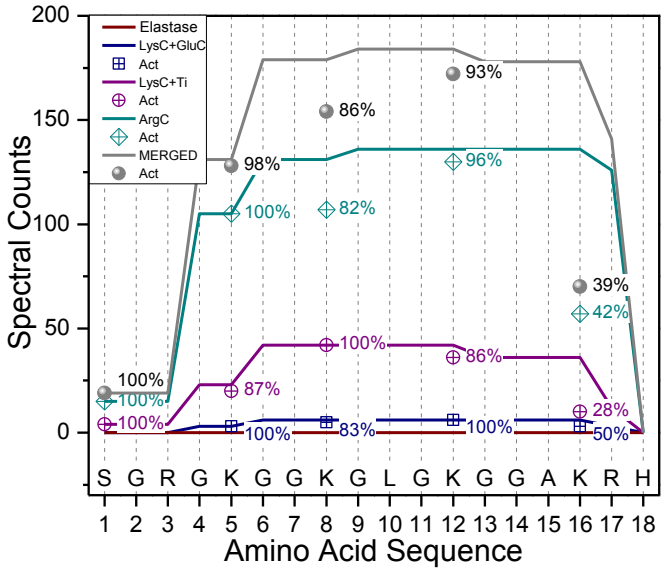


Figure 1.6: Quantification of modification levels by local spectral counts. The spectral coverage along the first 18 residues of Plasmodium histone H4 measured in the Elastase (brown line), LysC+GluC (navy line), LysC+Ti (purple line) and ArgC (dark cyan line) digested samples is plotted as an example (Plasmodium histones isolated from the 2nd biological replicate at 36-hpi). Spectral counts for acetylated residues are plotted as crossed symbols, while the percentages of spectral counts observed as acetylated vs. total spectral counts for this residue (equation above graph) are labeled next to the symbols. The final acetylated levels for histone H4 N-terminal tail in the 36-hpi_2 sample are calculated based on the spectral counts merged from all four digests (grey line, symbols, and lettering). Modification levels for all PTMs detected in this study were calculated as described here (Equation above graph) and reported in Table 1.2.

Linear decreases in lysine acetylation levels were observed for all canonical histones and were statistically significant for H4 (R^2 =0.87) and H2Bv (R^2= 0.7). H2A and H2B only had three acetylated lysines on their N-terminal tails and these residues were sporadically detected as acetylated across the 14 *Plasmodium* samples (Table 1.2), hence a lack of statistical support for the observed linear decrease. H3 and H3.3 N-terminal tails had six lysines, five of which were detected as acetylated at similar levels in all 14 samples (Table 1.2),

yet the linear regression R^2 values were low at 0.51 and 0.2, respectively. The main diversion from linearity was due to the fact that acetylation levels on K14 were consistently higher than on K9 in both isoforms. Surprisingly, lysine acetylation levels were high and stable across the long H2Az N-terminal tail, even for lysine residues close to its core domain.

Histones PTMs dynamics across the intra-erythrocytic cycle

Unsupervised hierarchical cluster analysis of the levels of 241 PTMs (excluding N-terminal acetylations) across the life cycle (Table 1.3) revealed three major groups, each exhibiting different patterns (Figure 1.7A). The first cluster (HC1) contained acetylation of lysine residues as dominating histone PTMs in all seven stages. The second cluster (HC2) contained formylation as the most predominant modification (22 sites in a group of 47 PTMs). It also contained modifications that were relatively frequently observed, yet not always present. The third cluster (HC3) contained PTMs that were only sporadically present in the seven stages at very low levels. This group was the largest and contained all possible types of PTMs, with methylations being the most represented (42 out of 158). This last cluster is potentially of great interest, since it encompasses

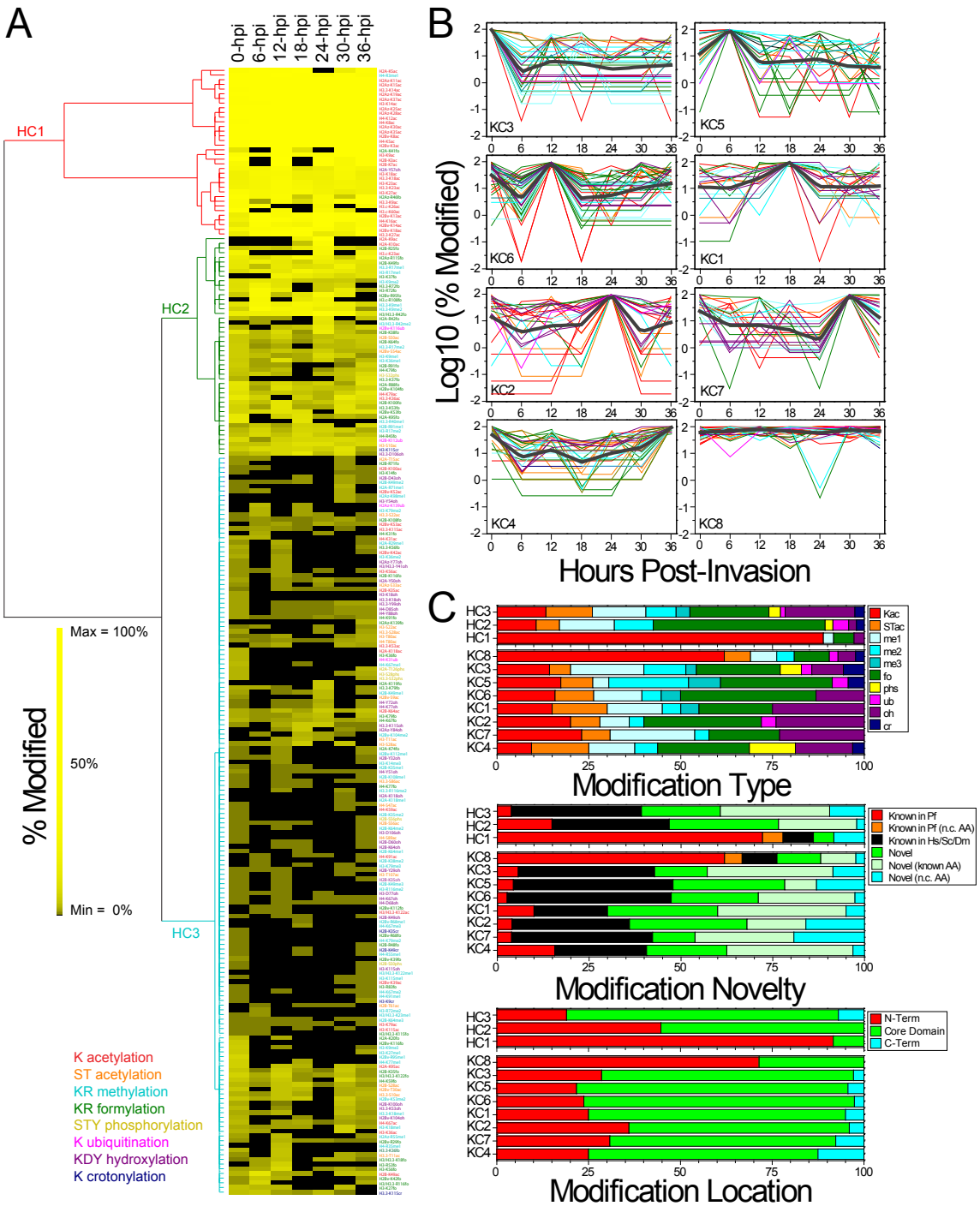

A
0-hpi
6-hpi
12-hpi
18-hpi
24-hpi
30-hpi
36-hpi
HC1
HC2
HC3
% Modified
Max = 100%
50%
Min = 0%
K acetylation
ST acetylation
KR methylation
KR formylation
STY phosphorylation
K ubiquitination
KDY hydroxylation
K crotonylation
B
Log10 (% Modified)
KC3
KC5
KC6
KC1
KC2
KC7
KC4
KC8
Hours Post-Invasion
C
HC3
HC2
HC1
KC8
KC3
KC5
KC6
KC1
KC2
KC7
KC4
Kac
STac
me1
me2
me3
fo
phs
ub
oh
cr
Modification Type
Known in Pf
Known in Pf (n.c. AA)
Known in Hs/Sc/Dm
Novel
Novel (known AA)
Novel (n.c. AA)
Modification Novelty
N-Term
Core Domain
C-Term
Modification Location

Figure 1.7: Clustering analyses of histone PTMs across seven erythrocytic stages of *P. falciparum*. A. *Unsupervised hierarchical clustering performed on the modification percentages of the 241 histone PTMs measured in each stage* (241 x 7 matrix; Table 1.3). Modifications levels (spectral counts merged between the two replicate analyses and expressed as a percentage of the modified spectral counts divided by the total spectral counts for a residue) observed in all seven time points of the *P. falciparum* life cycle (Table 1.2) were hierarchically clustered using Euclidean distance metric and WARD as a method using PermutMatrix software (Caraux and Pinloche, 2005). The tree was generated using the multiple-fragment heuristic algorithm (MF) as a seriation rule. The color intensity is proportional to the percentage of each modification with the brightest yellow indicating a high level and less intense yellow indicating residues modified at lower levels. Black symbolizes that the modification was not detected in a particular stage. The abbreviation "ac" corresponds to acetylation of lysine (K) marked in red, or serine (S), and threonine (T), both marked in orange. Mono-, di-, and trimethylations of lysines and arginines are indicated in blue. Phosphorylation and ubiquitination are indicated by "phs" in dark yellow and "ub" in pink, respectively. Hydroxylation of lysines, aspartates and tyrosines are in purple. Lysines and arginine formylation are dark green, while lysine crotonylations are written in navy. The tree lines for each of the three main hierarchical clusters are color-coded as described above to denote the type of modification that is the most represented in each cluster. **B.** *K-means clustering of histone modifications across Plasmodium seven erythrocytic stages performed on 241 PTMs*. For each AA#PTM combination, abundance values across the seven time points were normalized against the maximum measured value (maximum set at 100%; Table 1.3). We applied k-means clustering to this normalized matrix using the Hartigan-Wong algorithm (Wong, 1979) and a maximum number of iterations set at 50000. To determine the best partition, we continuously increased the number of clusters k from 3 to 20. The optimum number of k-means clusters was determined to be 8 based on the silhouette plot (see Figure 1.8). All computations were run using R environment using k-means function for the partition and daisy function to compute all the pairwise dissimilarities (Euclidean distances) between observations in the dataset for the silhouette. Log10 values of the normalized modification levels for each AA#PTM combination partitioned in one of the 8 k-means clusters are plotted as a function of hours post-invasion. Clusters are ordered based on when their relative maximum modification levels were measured (from 0 to 36-hpi), with cluster KC8 containing PTMs detected at maximum levels across all seven time points. Each line was color-coded as in **A**. Displayed values were the merge between two replicates (Figure 1.5B). Multiple curves averaging was performed for each cluster and the resulting curves are plotted in thick dark grey lines. **C.** *The relative proportions of each type of PTMs within each of hierarchical (HC) or k-means (KC) clusters plotted as stacked bars*. The summary plots are divided into three main criteria: the type of modification, the novelty status of each AA#PTM combination as defined in Figure 1.3A, and the location of each modified residue with the histone structure (N- or C-Terminal Tails and Core Domain).

histone PTMs that change over the *Plasmodium* cell cycle and therefore reflect chromatin changes in the respective transitions.

To analyze the variations in relative histone PTM abundances across the asexual cycle, we clustered abundance profiles (expressed in percentage of the maximum measured level for each AA#PTM) with the k-means algorithm. Each cluster had similar silhouette width (above 0.6) and contained similar numbers of PTMs (Figure 1.8). Each AA#PTM was plotted within each cluster as a log10 of the normalized abundance values measured in the seven time points (Figure 1.7B). Each profile within the clusters reached the maximum level in one of the seven stages (clusters KC1-7), or showed maximum levels across all time points (KC8). PTMs within each cluster were sorted based on three criteria: modification type, novelty status, and location of the modified residue within the histone structures (Figure 1.7C).

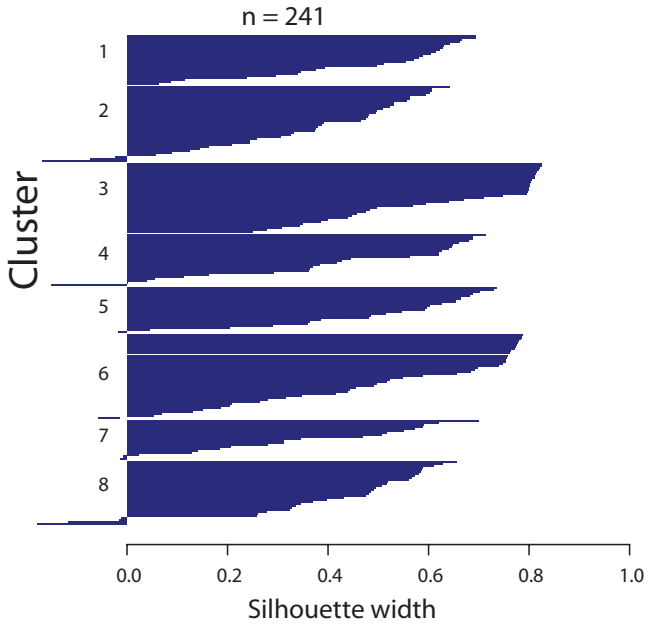


Figure 1.8: Evaluating k-means clustering. The silhouette representation after each k-means clustering is plotted to graphically evaluate how well each data point lies within its cluster. A number of clusters equals to 8 was deemed the optimal solution to this dataset (Table 1.3).

Clusters KC3, KC5, and KC6 contained PTMs with maximum relative abundances at the early, mid-, and late ring stages, respectively (Figure 1.7B). About 56% of the PTMs found in these clusters were novel and were largely located with the core domains (Figure 1.7C). Clusters KC1 and KC2 contained histone PTMs with maximum relative abundances at the early and late trophozoite stages, respectively (Figure 1.7B). These two clusters contained the most novel PTMs (about 70%). No particular enrichment in any type of PTMs was observed in any of these five clusters (Figure 1.7C). However, the six detected trimethylations fell in clusters KC3, KC5, KC6 and KC1, *i.e.* were the most abundant from early ring to early trophozoite (Figure 1.7C).

Clusters KC7 and KC4 contained PTMs with maximum relative abundances at early and late schizont stages, respectively. Most of the PTMs found in these clusters were also detected at fairly high levels at early ring stage (Figure 1.7B). Over 60% of the PTMs in these clusters were novel and located within the core domains. The six phosphorylations we identified neatly segregated into two clusters, KC3 and KC4, *i.e.* histone phosphorylations peaked in abundance at the beginning (0-hpi) and the end (36-hpi) of the erythrocytic cycle. Pfcrk-3, a kinase that phosphorylates histones, has recently been characterized in *P. falciparum* (Halbert et al., 2010) and is thought to play a crucial role in parasite proliferation.

Cluster KC8 mostly overlapped with hierarchical cluster 1 and contained the large majority of the histone acetylations previously reported in *Plasmodium* (Figure 1.7C). Abundance levels did not vary along the erythrocytic cycle. Over 75% of these PTMs were known and 71% were located on the N-terminal tails. Many PTMs detected on histone variants belonged to this group including all eight lysine acetylations on H2A.z N-terminal tail (*vs.* one for H2A; Table 1.3) and four out of five lysine acetylations detected on H2Bv N-terminal tail. Given their spatial and temporal co-localizations, these marks could be deposited by the same or closely related mechanisms. In addition, given their consistently high abundance (Table 1.3), it seems very likely that they concern a relatively large number of *loci* and regulate chromatin structure in a global manner.

Table 3: Clustering of post-translational modifications identified on Plasmodium Histones across the erythrocytic cycle

For each amino acid number / PTM type combination detected in this study, are specified:

The location of each amino acid as determined by the predicted location of N-terminal tail (NT), secondary structures forming the core domain (CD), and C-terminal tail (CT) based on known tridimensional structures (Figure S3);
Whether or not each AA#PTM combination had been previously detected in Plasmodium, human, yeast or fly on conserved or not conserved (n.c.) residues;
The total spectral counts (Total_SpC) mapping to peptides (modified or not) bearing each amino acid residue, combining all 56 MudPIT analyses of 14 histone samples;
The spectral counts matching modified peptides bearing each PTM type on this amino acid (Modified_SpC), combining all 56 MudPIT analyses of 14 histones samples;
The modification level estimated by Modified_SpC expressed as a percentage of the Total_SpC (Mod/Total %);
The number of times peptides bearing each AA#PTM combination were detected across the 56 MudPIT analyses;
The number of times peptides bearing each AA#PTM combination were detected across the 14 histone samples;
The modification level estimated by Modified_SpC expressed as a percentage of the Total_SpC merging spectral counts from both biological replicates for each time point (Mod/Total MERGED (%)) ;
The order in which each AA#PTM combination is ranked after hierarchical clustering (Figure 4A);
The hierarchical cluster (HC) in which each AA#PTM combination falls (Figure 4A);
The modification levels normalized against the highest value across the 7 time points (highest value set at 100%), i.e. input values for the k-means clustering;
The log10-transformed modification levels normalized against the highest value across the 7 time points (plotted in Figure 4B);
The k-means cluster (KC) in which each AA#PTM combination falls (Figure 4B).

Breakdown of each AA#PTM combination within clusters based on PTM types, novelty status and location within histone structure

	Kac	STac	me1	me2	me3	fo	phs	ub	oh	cr	Total
ALL	58	23	31	18	6	58	6	4	32	5	241
% of Total	24	10	13	7	2	24	2	2	13	2	

	Known in Pf	Known in Pf (n.c. AA)	Known in Hs/Sc/Dm	Novel	Novel (known AA)	Novel (n.c. AA)	Total
ALL	39	2	74	50	57	19	241
% of Total	16	1	31	21	24	8	

	NT	CD	CT	Total
ALL	84	146	11	241
% of Total	35	61	5	

Hierarchical clustering

	Kac	STac	me1	me2	me3	fo	phs	ub	oh	cr	Total
HC1	32	0	1	0	0	2	0	0	1	0	36
HC2	5	3	7	5	0	22	1	2	1	1	47
HC3	21	20	23	13	6	34	5	2	30	4	158
	58	23	31	18	6	58	6	4	32	5	241

	Known in Pf	Known in Pf (n.c. AA)	Known in Hs/Sc/Dm	Novel	Novel (known AA)	Novel (n.c. AA)	Total
HC1	26	2	3	2	0	3	36
HC2	7	0	15	14	10	1	47
HC3	6	0	56	34	47	15	158
	39	2	74	50	57	19	241

	NT	CD	CT	Total
HC1	33	3	0	36
HC2	21	26	0	47
HC3	30	117	11	158
	84	146	11	241

Figure 4C: Percent of total PTMs in each cluster

	Kac	STac	me1	me2	me3	fo	phs	ub	oh	cr
HC1	89	0	3	0	0	6	0	0	3	0
HC2	11	6	15	11	0	47	2	4	2	2
HC3	13	13	15	8	4	22	3	1	19	3

	Known in Pf	Known in Pf (n.c. AA)	Known in Hs/Sc/Dm	Novel	Novel (known AA)	Novel (n.c. AA)	Total Novel	Total Known
HC1	72	6	8	6	0	8	14	86
HC2	15	0	32	30	21	2	53	47
HC3	4	0	35	22	30	9	61	39

	NT	CD	CT
HC1	92	8	0
HC2	45	55	0
HC3	19	74	7

k-means clustering

		Kac	STac	me1	me2	me3	fo	phs	ub	oh	cr	Total
ALL 7	KC8	26	3	3	2	0	4	0	1	2	1	42
0-HPI	KC3	5	2	7	4	1	8	2	1	3	2	35
6-HPI	KC5	4	2	1	5	2	7	0	1	0	1	23
12-HPI	KC6	6	4	5	2	2	14	0	0	5	0	38
18-HPI	KC1	3	3	3	1	1	4	0	0	5	0	20
24-HPI	KC2	5	2	2	1	0	8	0	1	6	0	25
30-HPI	KC7	6	2	6	1	0	5	0	0	6	0	26
36-HPI	KC4	3	5	4	2	0	8	4	0	5	1	32
	Total	58	23	31	18	6	58	6	4	32	5	241

		Known in Pf	Known in Pf (n.c. AA)	Known in Hs/Sc/Dm	Novel	Novel (known AA)	Novel (n.c. AA)	Total
KC8		26	2	4	5	4	1	42
KC3		2	0	13	5	12	3	35
KC5		1	0	10	7	2	3	23
KC6		1	0	17	9	10	1	38
KC1		2	0	4	6	7	1	20
KC2		1	0	8	8	4	4	25
KC7		1	0	10	3	7	5	26
KC4		5	0	8	7	11	1	32
Total		39	2	74	50	57	19	241

		NT	CD	CT	Total
KC8		30	12	0	42
KC3		10	24	1	35
KC5		5	17	1	23
KC6		9	28	1	38
KC1		5	14	1	20
KC2		9	15	1	25
KC7		8	16	2	26
KC4		8	20	4	32
Total		84	146	11	241

Figure 4C: Percent of total PTMs in each cluster

		Kac	STac	me1	me2	me3	fo	phs	ub	oh	cr
ALL 7	KC8	62	7	7	5	0	10	0	2	5	2
0-HPI	KC3	14	6	20	11	3	23	6	3	9	6
6-HPI	KC5	17	9	4	22	9	30	0	4	0	4
12-HPI	KC6	16	11	13	5	5	37	0	0	13	0
18-HPI	KC1	15	15	15	5	5	20	0	0	25	0
24-HPI	KC2	20	8	8	4	0	32	0	4	24	0
30-HPI	KC7	23	8	23	4	0	19	0	0	23	0
36-HPI	KC4	9	16	13	6	0	25	13	0	16	3

		Known in Pf	Known in Pf (n.c. AA)	Known in Hs/Sc/Dm	Novel	Novel (known AA)	Novel (n.c. AA)	Total Novel	Total Known
KC8		62	5	10	12	10	2	24	76
KC3		6	0	37	14	34	9	57	43
KC5		4	0	43	30	9	13	52	48
KC6		3	0	45	24	26	3	53	47
KC1		10	0	20	30	35	5	70	30
KC2		4	0	32	32	16	16	64	36
KC7		4	0	38	12	27	19	58	42
KC4		16	0	25	22	34	3	59	41

		NT	CD	CT
KC8		71	29	0
KC3		29	69	3
KC5		22	74	4
KC6		24	74	3
KC1		25	70	5
KC2		36	60	4
KC7		31	62	8
KC4		25	63	13

				ALL SAMPLES					Mod/Total MERGED (%) Hierarchical clustering input and heat map plot (Figure 4A)									Mod/Total MERGED Normalized against maximum value in 7 time points (%) k-means input							Log10(Mod/Total MERGED) Normalized against maximum value in 7 time points (%) k-means plot (Figure 4B)							
Histone-Residue Position Modification	Type	Location	Novelty Status	Total_SpC	Modified_ SpC	Mod/Total (%)	Times Detected in 56 Digests	Times Detected in 14 Samples	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	Hierarchical Clustering output order (Fig 5A)	Hierarchical Cluster Number (Fig 5A)	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	k-means Cluster Number (Fig 5B and S6)
H2A-K10ac	Kac	NT	Known in His	204	2	0.98	2	2	0	0	0	2.94	33.33	0	0	53	HC2	0.06	0.06	0.06	8.82	100.00	0.06	0.06	-1.22	-1.22	-1.22	0.95	2.00	-1.22	-1.22	KC2
H2A-K118ac	Kac	CT	Novel (known	4355	5	0.11	3	2	1.89	0	0	0	0	0	0	138	HC3	100.00	1.06	1.06	1.06	1.06	1.06	1.06	2.00	0.02	0.02	0.02	0.02	0.02	0.02	KC3
H2A-K118me1	me1	CT	Known in His	4355	4	0.09	1	1	0	0	0	0	0	0.18	0	221	HC3	10.53	10.53	10.53	10.53	10.53	100.00	10.53	1.02	1.02	1.02	1.02	2.00	1.02	1.02	KC7
H2A-K118oh	oh	CT	Novel (known	4355	5	0.11	1	1	0	0	0	0	0	0.24	0	220	HC3	8.33	8.33	8.33	8.33	8.33	100.00	8.33	0.92	0.92	0.92	0.92	2.00	0.92	0.92	KC7
H2A-K119fo	fo	CT	Novel (known	381	2	0.52	1	1	0	0	0	0	3.33	0	0	119	HC3	0.60	0.60	0.60	0.60	100.00	0.60	0.60	-0.22	-0.22	-0.22	-0.22	2.00	-0.22	-0.22	KC2
H2A-K20fo	fo	NT	Novel (known	1609	8	0.50	2	2	2.70	0	1.28	0	0	0	0	108	HC3	100.00	0.74	47.41	0.74	0.74	0.74	0.74	2.00	-0.13	1.68	-0.13	-0.13	-0.13	-0.13	KC3
H2A-K41fo	fo	CD	Novel (n.c. A	78	17	21.79	7	7	9.09	0	50.00	7.69	0	62.50	62.50	20	HC1	14.54	0.03	80.00	12.30	0.03	100.00	100.00	1.16	-1.49	1.90	1.09	-1.49	2.00	2.00	KC7
H2A-K5ac	Kac	NT	Known in Pf	133	54	40.60	15	12	36.59	90.00	58.33	100	0	13.89	46.43	17	HC1	36.59	90.00	58.33	100.00	0.02	13.89	46.43	1.56	1.95	1.77	2.00	-1.70	1.14	1.67	KC1
H2A-K74fo	fo	CD	Novel (known	1611	3	0.19	3	3	0.31	0	0.44	0	0	0	0.44	204	HC3	70.45	4.55	100.00	4.55	4.55	4.55	100.00	1.85	0.66	2.00	0.66	0.66	0.66	2.00	KC6
H2A-K95ac	Kac	CD	Known in His	1570	31	1.97	9	7	4.49	0	2.94	1.68	0	3.09	1.38	94	HC3	100.00	0.45	65.48	37.42	0.45	68.82	30.73	2.00	-0.35	1.82	1.57	-0.35	1.84	1.49	KC6
H2A-K95fo	fo	CD	Known in His	1570	65	4.14	9	7	10.67	0	4.71	8.40	0.93	2.06	1.38	56	HC2	100.00	0.19	44.14	78.73	8.72	19.31	12.93	2.00	-0.73	1.64	1.90	0.94	1.29	1.11	KC3
H2A-K9ac	Kac	NT	Known in His	68	4	5.88	3	3	0	0	0	15.79	100	0	0	26	HC2	0.02	0.02	0.02	15.79	100.00	0.02	0.02	-1.70	-1.70	-1.70	1.20	2.00	-1.70	-1.70	KC2
H2A-R29me1	me1	CD	Novel	3916	11	0.28	3	3	0.71	0	1.37	0	0	0	0.20	140	HC3	51.82	1.46	100.00	1.46	1.46	1.46	14.60	1.71	0.16	2.00	0.16	0.16	0.16	1.16	KC6
H2A-R42fo	fo	CD	Novel (known	1719	33	1.92	9	9	2.58	1.09	0.99	0	10.31	0	2.76	73	HC2	25.02	10.57	9.60	0.19	100.00	0.19	26.77	1.40	1.02	0.98	-0.71	2.00	-0.71	1.43	KC2
H2A-R71me1	me1	CD	Known in His	3820	20	0.52	3	3	0.38	0	0	0	0	1.93	0	165	HC3	19.69	1.04	1.04	1.04	1.04	100.00	1.04	1.29	0.02	0.02	0.02	0.02	2.00	0.02	KC7
H2A-R88fo	fo	CD	Novel (known	1441	87	6.04	11	11	10.10	1.60	5.92	2.83	4.69	1.14	8.99	68	HC2	100.00	15.84	58.61	28.02	46.44	11.29	89.01	2.00	1.20	1.77	1.45	1.67	1.05	1.85	KC4
H2A-T126phs	phs	CT	Known in Pf	464	4	0.86	2	2	1.67	0	0	0	0	0	1.94	133	HC3	86.08	1.03	1.03	1.03	1.03	1.03	100.00	1.93	0.01	0.01	0.01	0.01	0.01	2.00	KC4
H2A-T15ac	STac	NT	Novel (n.c. A	1559	3	0.19	3	3	0.48	0.51	0	0	0	0.69	0	171	HC3	69.57	73.91	2.90	2.90	2.90	100.00	2.90	1.84	1.87	0.46	0.46	0.46	2.00	0.46	KC7
H2A-Y50oh	oh	CD	Novel	3975	16	0.40	9	7	0.79	0	0.44	0.31	0.92	0.32	0.14	153	HC3	85.87	2.17	47.83	33.70	100.00	34.78	15.22	1.93	0.34	1.68	1.53	2.00	1.54	1.18	KC2
H2A-Y57oh	oh	CD	Novel	3298	754	22.86	22	12	23.53	19.27	35.04	8.91	13.33	26.85	20.98	31	HC1	67.15	54.99	100.00	25.43	38.04	76.63	59.87	1.83	1.74	2.00	1.41	1.58	1.88	1.78	KC8
H2Az-K11ac	Kac	NT	Known in Pf	1653	1235	74.71	47	14	65.87	69.52	70.73	70.91	78.64	79.68	78.43	10	HC1	82.67	87.25	88.77	88.99	98.69	100.00	98.43	1.92	1.94	1.95	1.95	1.99	2.00	1.99	KC8
H2Az-K139fo	fo	CT	Novel	856	3	0.35	1	1	0	0	0	0	0	0	1.86	160	HC3	1.08	1.08	1.08	1.08	1.08	1.08	100.00	0.03	0.03	0.03	0.03	0.03	0.03	2.00	KC4
H2Az-K139ub	ub	CT	Novel	856	3	0.35	2	2	0	2.08	0	1.06	0	0	0	179	HC3	0.96	100.00	0.96	50.96	0.96	0.96	0.96	-0.02	2.00	-0.02	1.71	-0.02	-0.02	-0.02	KC5
H2Az-K15ac	Kac	NT	Known in Pf	1988	1431	71.98	48	14	60.09	70.05	72.43	71.36	73.02	78.57	71.98	11	HC1	76.48	89.16	92.19	90.82	92.94	100.00	91.61	1.88	1.95	1.96	1.96	1.97	2.00	1.96	KC8
H2Az-K19ac	Kac	NT	Known in Pf	1929	1263	65.47	37	14	54.46	66.67	63.60	67.00	68.70	75.18	58.79	13	HC1	72.44	88.68	84.60	89.12	91.38	100.00	78.20	1.86	1.95	1.93	1.95	1.96	2.00	1.89	KC8
H2Az-K25ac	Kac	NT	Known in Pf	1985	1654	83.32	40	14	81.42	85.42	83.27	82.63	79.31	87.18	81.38	8	HC1	93.39	97.98	95.52	94.78	90.97	100.00	93.35	1.97	1.99	1.98	1.98	1.96	2.00	1.97	KC8
H2Az-K28ac	Kac	NT	Known in Pf	1936	1541	79.60	37	14	74.07	73.08	78.47	82.10	87.75	84.06	76.92	7	HC1	84.41	83.28	89.42	93.56	100.00	95.79	87.66	1.93	1.92	1.95	1.97	2.00	1.98	1.94	KC8
H2Az-K30ac	Kac	NT	Known in Pf	1653	1461	88.38	41	14	80.94	86.77	85.04	86.67	95.16	93.04	88.77	3	HC1	94.95	91.18	89.37	91.08	100.00	97.77	93.28	1.93	1.96	1.95	1.96	2.00	1.99	1.97	KC8
H2Az-K35ac	Kac	NT	Known in Pf	1466	1282	87.45	39	14	79.31	83.54	84.16	84.02	95.48	89.52	92.28	4	HC1	83.06	87.49	88.14	88.00	100.00	93.76	96.65	1.92	1.94	1.95	1.94	2.00	1.97	1.99	KC8
H2Az-K37ac	Kac	NT	Known in His	1267	813	64.17	24	14	42.14	63.16	58.48	65.00	67.46	75.08	65.44	14	HC1	56.13	84.12	77.89	86.57	89.85	100.00	87.16	1.75	1.92	1.89	1.94	1.95	2.00	1.94	KC8
H2Az-K98me1	me1	CD	Novel	2096	9	0.43	2	2	0	0	0	0	0	1.50	0.38	166	HC3	1.33	1.33	1.33	1.33	1.33	100.00	25.33	0.12	0.12	0.12	0.12	0.12	2.00	1.40	KC7
H2Az-R115fo	fo	CD	Novel	12583	1648	13.10	32	14	12.94	6.49	12.04	26.64	16.39	2.87	22.05	50	HC2	48.57	24.36	45.20	100.00	61.52	10.77	82.77	1.69	1.39	1.66	2.00	1.79	1.03	1.92	KC1
H2Az-R46fo	fo	CD	Novel	143	36	25.17	12	11	42.00	25	33.33	3.70	28.57	3.70	60.00	28	HC1	70.00	41.67	55.55	6.17	47.62	6.17	100.00	1.85	1.62	1.74	0.79	1.68	0.79	2.00	KC4
H2Az-R55me1	me1	CD	Novel	2488	11	0.44	3	3	0.98	0	2.24	0	0	0	0.59	126	HC3	43.75	0.89	100.00	0.89	0.89	0.89	26.34	1.64	-0.05	2.00	-0.05	-0.05	-0.05	1.42	KC6
H2Az-S33ac	STac	NT	Novel	1466	3	0.20	3	3	0.57	0	0.50	0	0.56	0	0	152	HC3	100.00	3.51	87.72	3.51	98.25	3.51	3.51	2.00	0.55	1.94	0.55	1.99	0.55	0.55	KC2
H2Az-Y77oh	oh	CD	Novel	2613	6	0.23	3	3	0.59	0	0.79	0	0	0	0	145	HC3	74.68	2.53	100.00	2.53	2.53	2.53	2.53	1.87	0.40	2.00	0.40	0.40	0.40	0.40	KC6
H2Az-Y84oh	oh	CD	Novel	2179	7	0.32	4	4	0	0	0.45	3.66	0	0.42	0.19	182	HC3	0.55	0.55	12.30	100.00	0.55	11.48	5.19	-0.26	-0.26	1.09	2.00	-0.26	1.06	0.72	KC1
H2B-D43oh	oh	CD	Novel	3008	5	0.17	3	3	0	0	0	0.66	0	1.05	0.08	168	HC3	1.90	1.90	1.90	62.86	1.90	100.00	7.62	0.28	0.28	0.28	1.80	0.28	2.00	0.88	KC7
H2B-D60oh	oh	CD	Novel	25732	5	0.02	5	5	0.02	0	0.02	0.07	0	0	0.02	239	HC3	28.57	28.57	28.57	100.00	28.57	28.57	28.57	1.46	1.46	1.46	2.00	1.46	1.46	1.46	KC1
H2B-K100ac	Kac	CD	Known in His	3880	6	0.15	5	5	0.11	0.17	0	0	0	0.52	0.24	212	HC3	21.15	32.69	3.85	3.85	3.85	100.00	46.15	1.33	1.51	0.59	0.59	0.59	2.00	1.66	KC7
H2B-K100fo	fo	CD	Known in His	3880	201	5.18	34	13	5.90	2.37	9.79	2.06	3.16	7.29	5.13	65	HC2	60.27	24.21	100.00	21.04	32.28	74.46	52.40	1.78	1.38	2.00	1.32	1.51	1.87	1.72	KC8
H2B-K100oh	oh	CD	Novel (known	3880	35	0.90	8	7	1.53	0	0.21	0	0	2.86	1.10	103	HC3	53.50	0.70	7.34	0.70	0.70	100.00	38.46	1.73	-0.16	0.87	-0.16	-0.16	2.00	1.59	KC7
H2B-K108fo	fo	CD	Known in His	5487	22	0.40	13	10	0.27	0.66	0.36	0.33	0.55	0.80	0.16	173	HC3	33.75	82.50	45.00	41.25	68.75	100.00	20.00	1.53	1.92	1.65	1.62	1.84	2.00	1.30	KC8
H2B-K108me1	me1	CD	Known in His	5487	5	0.09	4	4	0.18	0	0.24	0	0	0	0.08	197	HC3	75.00	8.33	100.00	8.33	8.33	8.33	33.33	1.88	0.92	2.00	0.92	0.92	0.92	1.52	KC6
H2B-K112ub	ub	CD	Known in Pf	2906	153	5.26	19	13	5.96	0.48	4.55	5.62	5.56	6.46	5.12	60	HC2	92.26	7.43	70.43	87.00	86.07	100.00	79.26	1.97	0.87	1.85	1.94	1.93	2.00	1.90	KC8
H2B-K116fo	fo	CT	Novel (known	2712	8	0.29	4	4	0.70	0	0.84	0.32	0	0	0	144	HC3	83.33	2.38	100.00	38.10	2.38	2.38	2.38	1.92	0.38	2.00	1.58	0.38	0.38	0.38	KC6
H2B-K35ac	Kac	CD	Novel (known	14681	27	0.18	15	11	0.30	0	0.46	0.13	0.34	0.15	0.07	151	HC3	65.22	4.35	100.00	28.26	73.91	32.61	15.22	1.81	0.64	2.00	1.45	1.87	1.51	1.18	KC6
H2B-K35cr	cr	CD	Novel (known	14681	7	0.05	1	1	0.21	0	0	0	0	0	0	216	HC3	100.00	9.52	9.52	9.52	9.52	9.52	9.52	2.00	0.98	0.98	0.98	0.98	0.98	0.98	KC3
H2B-K35fo	fo	CD	Known in His	14681	236	1.61	25	13	3.22	0.04	3.76	0.92	0.20	3.29	0.63	93	HC3	85.64	1.06	100.00	24.47	5.32	87.50	16.76	1.93	0.03	2.00	1.39	0.73	1.94	1.22	KC6
H2B-K35me1	me1	CD	Known in His	14681	15	0.10	5	4	0.33	0	0.23	0.13	0	0	0	199	HC3	100.00	6.06	69.70	39.39	6.06	6.06	6.06	2.00	0.78	1.84	1.60	0.78	0.78	0.78	KC3
H2B-K35me2	me2	CD	Novel (known	14681	3	0.02	1	1	0.09	0	0	0	0	0	0	233	HC3	100.00	22.22	22.22	22.22	22.22	22.22	22.22	2.00	1.35	1.35	1.35	1.35	1.35	1.35	KC3
H2B-K35oh	oh	CD	Novel (known	14681	6	0.04	3	3	0.03	0	0.15	0	0	0	0.07	228	HC3	20.00	13.33	100.00	13.33	13.33	13.33	46.67	1.30	1.12	2.00	1.12	1.12	1.12	1.67	KC6
H2B-K38fo	fo	CD	Known in His	4690	181	3.86	23	12	9.03	2.23	6.30	0.52	13.82	3.13	1.10	71	HC2	65.34	16.14	45.59	3.76	100.00	22.65	7.96	1.82	1.21	1.66	0.58	2.00	1.36	0.90	KC2
H2B-K38me2	me2	CD	Known in His	4690	2	0.04	2	2	0	0.12	0	0	0	0	0.06	242	HC3	16.67	100.00	16.67	16.67	16.67	16.67	50.00	1.22	2.00	1.22	1.22	1.22	1.70	KC5
H2B-K3ac	Kac	NT	Known in His	38	23	60.53	8	8	100	0	100	0	50.00	25.00	53.85	30	HC1	100.00	0.02	100.00	0.02	50.00	25.00	53.85	2.00	-1.70	2.00	-1.70	1.70	1.40	1.73	KC6
H2B-K49ac	Kac	CD	Known in His	10605	252	2.38	13	12	1.70	9.47	1.21	0.92	0	2.46	3.12	76	HC3	17.95	100.00	12.78	9.71	0.21	25.98	32.95	1.25	2.00	1.11	0.99	-0.68	1.41	1.52	KC5
H2B-K49cr	cr	CD	Novel (known	10605	8	0.08	5	5	0.19	0	0	0.06	0	0	0.10	218	HC3	100.00	10.53	10.53	31.58	10.53	10.53	52.63	2.00	1.02	1.02	1.50	1.02	1.02	1.72	KC3
H2B-K49fo	fo	CD	Novel (known	10605	1041	9.82	20	13	10.77	15.35	10.75	4.78	13.33	10.36	9.11	48	HC2	70.16	100.00	70.03	31.14	86.84	67.49	59.35	1.85	2.00	1.85	1.49	1.94	1.83	1.77	KC8
H2B-K49me1	me1	CD	Known in His	10605	57	0.54	9	8	1.16	0	0.19	0.40	2.22	0	0.70	120	HC3	52.25	0.90	8.56	18.02	100.00	0.90	31.53	1.72	-0.05	0.93	1.26	2.00	-0.05	1.50	KC2
H2B-K49me2	me2	CD	Known in His	10605	34	0.32	13	10	0.39	0.12	0.19	0.40	0	0.91	0.20	169	HC3	42.86	13.19	20.88	43.96	2.20	100.00	21.98	1.63	1.12	1.32	1.64	0.34	2.00	1.34	KC7
H2B-K49me3	me3	CD	Novel (known	10605	5	0.05	4	4	0.08	0	0.11	0	0	0	0	227	HC3	72.73	18.18	100.00	18.18	18.18	18.18	18.18	1.86	1.26	2.00	1.26	1.26	1.26	1.26	KC6
H2B-K49oh	oh	CD	Novel (known	10605	10	0.09	5	5	0.12	0	0.04	0.23	0	0	0.10	249	HC3	52.17	8.70	17.39	100.00	8.70	43.48	1.72	0.94	1.24	2.00	0.94	0.94	1.64	KC1
H2B-K64ac	Kac	CD	Known in His	25653	492	1.92	36	14	1.39	0.68	2.35	1.15	3.78	1.48	2.84	114	HC3	36.77	17.99	62.17	30.42	100.00	39.15	75.13	1.57	1.26	1.79	1.48	2.00	1.59	1.88	KC2
H2B-K64fo	fo	CD	Novel	25653	686	2.67	35	14	3.76	1.07	2.67	0.58	3.36	2.27	3.28	86	HC2	100.00	28.46	71.01	15.43	89.36	60.37	87.23	2.00	1.45	1.85	1.19	1.95	1.78	1.94	KC8
H2B-K64me1	me1	CD	Novel	25653	2	0.01	2	2	0.02	0	0	0.03	0	0	0	238	HC3	66.67	66.67	66.67	100.00	66.67	66.67	66.67	1.82	1.82	1.82	2.00	1.82	1.82	1.82	KC8
H2B-K64me2	me2	CD	Novel	25653	10	0.04	3	3	0.13	0	0.02	0	0	0	0.02	232	HC3	100.00	15.38	15.38	15.38	15.38	15.38	15.38	2.00	1.19	1.19	1.19	1.19	1.19	1.19	KC3
H2B-K64me3	me3	CD	Novel (n.c. A	25653	43	0.17	13	10	0.26	0.34	0.12	0.14	0	0.04	0.14	187	HC3	76.47	100.00	35.29	41.18	5.88	11.76	41.18	1.88	2.00	1.55	1.61	0.77	1.07	1.61	KC5
H2B-K64oh	oh	CD	Novel (n.c. A	25653	4	0.02	3	3	0	0	0.02	0.03	0	0	0.04	240	HC3	50.00	50.00	50.00	75.00	50.00	50.00	100.00	1.70	1.70	1.70	1.88	1.70	1.70	2.00	KC8
H2B-K7ac	Kac	NT	Known in His	53	23	43.40	8	8	28.57	0	100	0	50.00	25.00	53.85	18	HC1	28.57	0.02	100.00	0.02	50.00	25.00	53.85	1.46	-1.70	2.00	-1.70	1.70	1.40	1.73	KC6
H2B-R25fo	fo	NT	Novel	8296	408	4.92	26	14	3.18	4.98	3.16	13.25	25.28	5.63	3.54	52	HC2	12.58	19.70	12.50	52.41	100.00	22.27	14.00	1.10	1.29	1.10	1.72	2.00	1.35	1.15	KC2
H2B-R48fo	fo	CD	Novel (n.c. A	2949	2	0.07	2	2	0.21	0	0	0	0	0	0.08	219	HC3	100.00	9.52	9.52	9.52	9.52	9.52	38.10	2.00	0.98	0.98	0.98	0.98	0.98	1.58	KC3
H2B-R71fo	fo	CD	Novel (known	12756	21	0.16	3	3	0.50	0	0	0	0	0.62	0	170	HC3	80.65	3.23	3.23	3.23	3.23	100.00	3.23	1.91	0.51	0.51	0.51	0.51	2.00	0.51	KC7
H2B-R91fo	fo	CD	Novel (known	1477	30	2.03	6	6	4.08	1.12	0.83	0	0	0.37	4.59	81	HC2	88.89	24.40	18.08	0.44	0.44	8.06	100.00	1.95	1.39	1.26	-0.36	-0.36	0.91	2.00	KC4
H2B-R91me1	me1	CD	Known in His	1477	31	2.10	8	8	0.96	3.37	0.83	4.82	4.12	0	5.10	57	HC2	18.82	66.08	16.27	94.51	80.78	0.39	100.00	1.27	1.82	1.21	1.98	1.91	-0.41	2.00	KC1
H2B-S28ac	STac	CD	Novel (known	14673	196	1.34	15	11	2.10	0.75	2.85	3.56	0.27	1.84	0.42	96	HC3	58.99	21.07	80.06	100.00	7.58	51.69	11.80	1.77	1.32	1.90	2.00	0.88	1.71	1.07	KC1
H2B-S50ac	STac	CD	Novel	24669	847	3.43	39	14	3.66	2.09	4.88	0.87	2.50	3.66	3.92	88	HC2	75.00	42.83	100.00	17.83	51.23	75.00	80.33	1.88	1.63	2.00	1.25	1.71	1.88	1.90	KC8

				ALL SAMPLES					Mod/Total MERGED (%) Hierarchical clustering input and heat map plot (Figure 4A)									Mod/Total MERGED Normalized against maximum value in 7 time points (%) k-means input							Log10(Mod/Total MERGED) Normalized against maximum value in 7 time points (%) k-means plot (Figure 4B)							
Histone-Residue Position Modification	Type	Location	Novelty Status	Total_SpC	Modified_SpC	Mod/Total (%)	Times Detected in 56 Digests	Times Detected in 14 Samples	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	Hierarchical Clustering output order (Fig 5A)	Hierarchical Cluster Number (Fig 5A)	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	k-means Cluster Number (Fig 5B and S6)
H2B-S50phs	phs	CD	Novel	24669	9	0.04	2	2	0.02	0	0	0	0	0	0.15	246	HC3	13.33	13.33	13.33	13.33	13.33	13.33	100.00	1.12	1.12	1.12	1.12	1.12	1.12	2.00	KC4
H2B-S56ac	STac	CD	Novel	25541	8	0.03	7	5	0.06	0.04	0.02	0.03	0	0	0.02	235	HC3	100.00	66.67	33.33	50.00	33.33	33.33	33.33	2.00	1.82	1.52	1.70	1.52	1.52	1.52	KC3
H2B-S56phs	phs	CD	Novel	25541	4	0.02	3	3	0.05	0	0	0	0	0	0.02	234	HC3	100.00	40.00	40.00	40.00	40.00	40.00	40.00	2.00	1.60	1.60	1.60	1.60	1.60	1.60	KC3
H2B-T61ac	STac	CD	Novel (n.c. A	25723	13	0.05	4	4	0	0.34	0	0.10	0.08	0.04	0	192	HC3	5.88	100.00	5.88	29.41	23.53	11.76	5.88	0.77	2.00	0.77	1.47	1.37	1.07	0.77	KC5
H2Bv-K104fo	fo	CD	Known in Hs/	2332	110	4.72	18	13	9.39	0.97	7.27	3.79	2.94	2.66	4.73	67	HC2	100.00	10.33	77.42	40.36	31.31	28.33	50.37	2.00	1.01	1.89	1.61	1.50	1.45	1.70	KC6
H2Bv-K104me2	me2	CD	Novel (know	2332	9	0.39	7	6	0	0.58	0	1.90	0.98	0.24	0	180	HC3	1.05	30.53	1.05	100.00	51.58	12.63	1.05	0.02	1.48	0.02	2.00	1.71	1.10	0.02	KC1
H2Bv-K104oh	oh	CD	Novel (know	2332	20	0.86	6	6	1.92	0	0	0.47	0	1.45	1.01	102	HC3	100.00	1.04	1.04	24.48	1.04	75.52	52.60	2.00	0.02	0.02	1.39	0.02	1.88	1.72	KC7
H2Bv-K112fo	fo	CD	Known in Hs/	3438	2	0.06	2	2	0	0.14	0.21	0	0	0	0	196	HC3	9.52	66.67	100.00	9.52	9.52	9.52	9.52	0.98	1.82	2.00	0.98	0.98	0.98	0.98	KC6
H2Bv-K112me1	me1	CD	Known in Hs/	3438	8	0.23	6	6	0.36	0	0.42	0	0	0	0.60	205	HC3	60.00	3.33	70.00	3.33	3.33	3.33	100.00	1.78	0.52	1.85	0.52	0.52	0.52	2.00	KC4
H2Bv-K116fo	fo	CD	Known in Hs/	1055	10	0.95	2	2	3.82	0	0	0	0	0	0	109	HC3	100.00	0.52	0.52	0.52	0.52	0.52	0.52	2.00	-0.28	-0.28	-0.28	-0.28	-0.28	-0.28	KC3
H2Bv-K116ub	ub	CD	Known in Hs/	1055	24	2.27	9	9	1.91	0	2.48	3.28	11.11	2.23	5.00	72	HC2	17.19	0.18	22.32	29.52	100.00	20.07	45.00	1.24	-0.74	1.35	1.47	2.00	1.30	1.65	KC2
H2Bv-K13ac	Kac	NT	Known in Pf	629	253	40.22	32	14	14.08	60.00	28.95	35.59	63.92	44.19	50.81	25	HC1	22.03	93.87	45.29	55.68	100.00	69.13	79.49	1.34	1.97	1.66	1.75	2.00	1.84	1.90	KC8
H2Bv-K14ac	Kac	NT	Known in Pf	813	220	27.06	31	14	11.56	46.15	25.61	20.59	52.58	19.28	32.93	23	HC1	21.99	87.77	48.71	39.16	100.00	36.63	62.63	1.34	1.94	1.69	1.59	2.00	1.56	1.80	KC8
H2Bv-K18ac	Kac	NT	Known in Pf	885	296	33.45	30	14	23.41	33.33	38.00	51.43	39.76	19.26	46.47	22	HC1	45.52	64.81	73.88	100.00	77.31	37.45	90.36	1.66	1.81	1.87	2.00	1.89	1.57	1.96	KC8
H2Bv-K39ac	Kac	CD	Novel (know	4760	4	0.08	3	3	0	0	0	0	0	0.10	0.38	210	HC3	5.26	5.26	5.26	5.26	5.26	26.32	100.00	0.72	0.72	0.72	0.72	0.72	1.42	2.00	KC4
H2Bv-K39fo	fo	CD	Known in Hs/	4760	3	0.06	2	1	0.42	0	0	0	0	0	0	213	HC3	100.00	4.76	4.76	4.76	4.76	4.76	4.76	2.00	0.68	0.68	0.68	0.68	0.68	0.68	KC3
H2Bv-K3ac	Kac	NT	Known in Pf	265	244	92.08	25	14	98.04	100	88.24	63.64	96.23	90.00	89.29	5	HC1	98.04	100.00	88.24	63.64	96.23	90.00	89.29	1.99	2.00	1.95	1.80	1.98	1.95	1.95	KC8
H2Bv-K42ac	Kac	CD	Known in Hs/	1679	6	0.36	3	3	1.02	0	0.53	0	0	0	0.17	143	HC3	100.00	1.96	51.96	1.96	1.96	1.96	16.67	2.00	0.29	1.72	0.29	0.29	0.29	1.22	KC3
H2Bv-K42fo	fo	CD	Known in Hs/	1679	13	0.77	4	4	0.51	3.57	1.60	0	0	0	0.52	79	HC3	14.29	100.00	44.82	0.56	0.56	0.56	14.57	1.15	2.00	1.65	-0.25	-0.25	-0.25	1.16	KC5
H2Bv-K52ac	Kac	CD	Novel (n.c. A	1425	4	0.28	4	4	0.30	0.72	0	0	0	1.53	0	164	HC3	19.61	47.06	1.31	1.31	1.31	100.00	1.31	1.29	1.67	0.12	0.12	0.12	2.00	0.12	KC7
H2Bv-K53ac	Kac	CD	Known in Hs/	3179	14	0.44	8	7	0	0.66	0.80	0.28	1.19	0.46	0.44	172	HC3	1.68	55.46	67.23	23.53	100.00	38.66	36.97	0.23	1.74	1.83	1.37	2.00	1.59	1.57	KC2
H2Bv-K53fo	fo	CD	Novel (know	3179	232	7.30	12	12	5.07	5.91	9.02	7.67	0	8.43	9.14	66	HC2	55.47	64.66	98.69	83.92	0.22	92.23	100.00	1.74	1.81	1.99	1.92	-0.66	1.96	2.00	KC8
H2Bv-K53me2	me2	CD	Known in Hs/	3179	85	2.67	12	10	3.43	1.53	3.01	3.69	0	3.87	1.47	95	HC3	88.63	39.53	77.78	95.35	0.52	100.00	37.98	1.95	1.60	1.89	1.98	-0.29	2.00	1.58	KC8
H2Bv-K8ac	Kac	NT	Known in Pf	523	472	90.25	38	14	72.73	97.44	91.78	97.83	93.41	87.01	95.00	2	HC1	74.34	99.60	93.82	100.00	95.48	88.94	97.11	1.87	2.00	1.97	2.00	1.98	1.95	1.99	KC8
H2Bv-R29fo	fo	NT	Novel	2693	11	0.41	7	7	0.58	0	1.70	0.58	0.49	0.15	0.20	125	HC3	34.12	1.18	100.00	34.12	28.82	8.82	11.76	1.53	0.07	2.00	1.53	1.46	0.95	1.07	KC6
H2Bv-R68fo	fo	CD	Novel	10220	5	0.05	1	1	0.22	0	0	0	0	0	0	215	HC3	100.00	9.09	9.09	9.09	9.09	9.09	9.09	2.00	0.96	0.96	0.96	0.96	0.96	0.96	KC3
H2Bv-R68me1	me1	CD	Novel	10220	3	0.03	1	1	0	0	0	0.31	0	0	0	250	HC3	6.45	6.45	6.45	100.00	6.45	6.45	6.45	0.81	0.81	0.81	2.00	0.81	0.81	0.81	KC1
H2Bv-R95fo	fo	CD	Novel (know	306	32	10.46	11	10	3.57	28.57	2.63	8.89	25.81	11.11	0	41	HC2	12.50	100.00	9.21	31.12	90.34	38.89	0.07	1.10	2.00	0.96	1.49	1.96	1.59	-1.15	KC5
H2Bv-R95me1	me1	CD	Known in Hs/	306	3	0.98	1	1	5.36	0	0	0	0	0	0	113	HC3	100.00	0.37	0.37	0.37	0.37	0.37	0.37	2.00	-0.43	-0.43	-0.43	-0.43	-0.43	-0.43	KC3
H2Bv-S54ac	STac	CD	Novel	9957	277	2.78	40	14	3.96	1.92	3.40	2.10	4.29	1.87	1.86	87	HC2	92.31	44.76	79.25	48.95	100.00	43.59	43.36	1.97	1.65	1.90	1.69	2.00	1.64	1.64	KC8
H2Bv-S9ac	STac	NT	Novel (n.c. A	629	3	0.48	3	2	0.70	0	0	0	2.06	0	0	121	HC3	33.98	0.97	0.97	0.97	100.00	0.97	0.97	1.53	-0.01	-0.01	-0.01	2.00	-0.01	-0.01	KC2
H2Bv-T30ac	STac	CD	Novel (n.c. A	4448	63	1.42	9	9	3.24	0	2.96	1.99	0	1.36	1.38	97	HC3	100.00	0.62	91.30	61.42	0.62	41.98	42.59	2.00	-0.21	1.96	1.79	-0.21	1.62	1.63	KC6
H2B-Y29oh	oh	CD	Known in Hs/	14673	12	0.08	8	7	0.12	0.04	0.15	0	0	0.15	0.07	224	HC3	80.00	26.67	100.00	13.33	13.33	100.00	46.67	1.90	1.43	2.00	1.12	1.12	2.00	1.67	KC6
H2B-Y32oh	oh	CD	Novel	14673	31	0.21	12	7	0.30	0.09	0.54	0	0	0	0.28	203	HC3	55.56	16.67	100.00	3.70	3.70	3.70	51.85	1.74	1.22	2.00	0.57	0.57	0.57	1.71	KC6
H3.3-D106oh	oh	CD	Novel	5082	205	4.03	13	9	1.91	2.87	5.58	10.58	2.48	7.09	0.44	63	HC2	18.05	27.13	52.74	100.00	23.44	67.01	4.16	1.26	1.43	1.72	2.00	1.37	1.83	0.62	KC1
H3.3-K15ac	Kac	CD	Known in Hs/	3086	12	0.39	8	7	0.50	1.09	0.73	0.30	0	0.12	0	174	HC3	45.87	100.00	66.97	27.52	1.83	11.01	1.83	1.66	2.00	1.83	1.44	0.26	1.04	0.26	KC5
H3.3-K115or	Gr	CD	Novel (know	3086	38	1.23	10	8	2.37	3.26	0.91	0.60	2.33	0.37	0	100	HC3	72.70	100.00	27.91	18.40	71.47	11.35	0.61	1.86	2.00	1.45	1.26	1.85	1.05	-0.21	KC5
H3.3-K115oh	oh	CD	Novel (know	3086	34	1.10	8	8	0.62	0	1.45	3.88	3.88	0.25	0.36	117	HC3	15.98	0.52	37.37	100.00	100.00	6.44	9.28	1.20	-0.29	1.57	2.00	2.00	0.81	0.97	KC1
H3.3-K14ac	Kac	NT	Known in Pf	1438	1000	69.54	40	14	48.83	76.19	69.78	59.46	69.84	72.84	82.05	12	HC1	59.51	92.86	85.05	72.47	85.12	88.78	100.00	1.77	1.97	1.93	1.86	1.93	1.95	2.00	KC8
H3.3-K18ac	Kac	NT	Known in Pf	2124	390	18.36	39	14	16.26	21.86	27.87	15.63	15.58	23.47	15.40	33	HC1	58.34	78.44	100.00	56.08	55.90	84.21	55.26	1.77	1.89	2.00	1.75	1.75	1.93	1.74	KC8
H3.3-K18me1	me1	NT	Known in Hs/	2124	23	1.08	11	9	1.46	0.55	1.64	0	0	2.53	0.56	105	HC3	57.71	21.74	64.82	0.79	0.79	100.00	22.13	1.76	1.34	1.81	-0.10	-0.10	2.00	1.35	KC7
H3.3-K18oh	oh	NT	Novel (know	2124	6	0.28	1	1	0.98	0	0	0	0	0	0	150	HC3	100.00	2.04	2.04	2.04	2.04	2.04	2.04	2.00	0.31	0.31	0.31	0.31	0.31	0.31	KC3
H3.3-K23ac	Kac	NT	Known in Pf	2120	419	19.76	33	14	18.06	22.83	29.48	23.16	17.41	20.50	17.58	35	HC1	61.26	77.44	100.00	78.56	59.06	69.54	59.63	1.79	1.89	2.00	1.90	1.77	1.84	1.78	KC8
H3.3-K27ac	Kac	NT	Known in Pf	548	145	26.46	41	14	12.87	38.46	46.77	41.18	11.72	14.29	52.33	21	HC1	24.59	73.50	89.38	78.69	22.40	27.31	100.00	1.39	1.87	1.95	1.90	1.35	1.44	2.00	KC8
H3.3-K36ac	Kac	NT	Known in Hs/	592	38	6.42	11	7	11.11	4.88	0	4.55	1.57	11.11	7.69	70	HC2	100.00	43.92	0.18	40.95	14.13	100.00	69.22	2.00	1.64	-0.74	1.61	1.15	2.00	1.84	KC7
H3.3-K36fo	fo	NT	Novel (know	592	4	0.68	2	2	1.71	0	3.13	0	0	0	0	124	HC3	54.63	0.64	100.00	0.64	0.64	0.64	0.64	1.74	-0.19	2.00	-0.19	-0.19	-0.19	-0.19	KC6
H3.3-K37fo	fo	NT	Novel (know	153	3	1.96	3	3	1.82	0	0	0	0	2.63	7.69	80	HC2	23.67	0.26	0.26	0.26	0.26	34.20	100.00	1.37	-0.58	-0.58	-0.58	-0.58	1.53	2.00	KC4
H3.3-K53ac	Kac	CD	Novel (n.c. A	1837	12	0.65	7	7	0	0.69	1.20	0	0	1.55	0.81	163	HC3	1.29	44.52	77.42	1.29	1.29	100.00	52.28	0.11	1.65	1.89	0.11	2.00	1.72	KC7
H3.3-K53fo	fo	CD	Novel	1837	54	2.94	11	11	4.90	1.03	9.64	1.45	1.39	8.53	2.83	64	HC2	50.83	10.68	100.00	15.04	14.42	88.49	29.36	1.71	1.03	2.00	1.18	1.16	1.95	1.47	KC6
H3.3-K53oh	oh	CD	Novel (n.c. A	1837	9	0.49	5	5	2.10	0	0.60	0	0	2.33	0.81	104	HC3	90.13	0.86	25.75	0.86	0.86	100.00	34.76	1.95	-0.07	1.41	-0.07	-0.07	2.00	1.54	KC7
H3.3-K56fo	fo	CD	Known in Hs/	2029	6	0.30	3	3	1.23	0	1.10	0	0	0	0.61	141	HC3	100.00	1.63	89.43	1.63	1.63	1.63	49.59	2.00	0.21	1.95	0.21	0.21	2.00	1.70	KC6
H3.3-K79fo	fo	CD	Known in Hs/	12841	129	1.00	10	7	1.52	1.30	0.16	0	3.22	0	0.47	118	HC3	47.20	40.37	4.97	0.62	100.00	0.62	14.60	1.67	1.61	0.70	-0.21	2.00	-0.21	1.16	KC2
H3.3-K9ac	Kac	NT	Known in Pf	1347	293	21.75	23	14	34.60	0.98	29.03	6.94	27.83	38.03	54.48	27	HC1	63.51	1.80	53.29	12.74	51.08	69.81	100.00	1.80	0.25	1.73	1.11	1.71	1.84	2.00	KC4
H3.3-K9me1	me1	NT	Known in Hs/	1347	233	17.30	22	14	8.86	25.00	15.32	22.22	18.26	14.08	5.52	40	HC2	35.44	100.00	61.28	88.88	73.04	56.32	22.08	1.55	2.00	1.79	1.95	1.86	1.75	1.34	KC8
H3.3-K9me2	me2	NT	Known in Hs/	1347	252	18.71	23	13	5.06	32.42	21.77	19.44	13.04	11.97	0.69	39	HC2	15.61	100.00	67.15	59.96	40.22	36.92	2.13	1.19	2.00	1.83	1.78	1.60	1.57	0.33	KC5
H3.3-R116me2	me2	CD	Novel	1437	3	0.21	2	2	0	0	0.63	0	0	0.40	0	193	HC3	3.17	3.17	100.00	3.17	3.17	63.49	3.17	0.50	0.50	2.00	0.50	0.50	1.80	0.50	KC6
H3.3-R17me1	me1	NT	Known in Pf	1553	123	7.92	31	14	10.20	2.40	11.41	7.41	8.53	17.30	7.93	47	HC2	58.96	13.87	65.95	42.83	49.31	100.00	45.84	1.77	1.14	1.82	1.63	1.69	2.00	1.66	KC7
H3.3-R17me2	me2	NT	Known in Pf	1553	30	1.93	17	10	4.28	0.55	1.34	0	2.33	2.16	3.05	85	HC2	100.00	12.85	31.31	0.47	54.44	50.47	71.26	2.00	1.11	1.50	-0.33	1.74	1.70	1.85	KC4
H3.3-R40me1	me1	NT	Novel	70	5	7.14	2	2	11.43	0	0	9.09	0	0	0	55	HC2	100.00	0.17	0.17	79.53	0.17	0.17	0.17	2.00	-0.76	-0.76	1.90	-0.76	-0.76	-0.76	KC3
H3.3-R72fo	fo	CD	Novel	1514	143	9.45	9	9	7.33	17.56	8.83	0	4.26	5.42	0	44	HC2	41.74	100.00	50.28	0.11	24.26	30.87	0.11	1.62	2.00	1.70	-0.94	1.38	1.49	-0.94	KC5
H3.3-S10ac	STac	NT	Novel (know	1423	18	1.26	12	10	1.57	0.78	2.88	1.35	0	1.32	1.92	98	HC3	54.51	26.39	100.00	46.88	0.69	45.83	66.67	1.74	1.42	2.00	1.67	-0.16	1.66	1.82	KC6
H3.3-S22ac	STac	NT	Novel	2120	8	0.38	8	7	0.16	1.63	0	1.05	0.45	0	0.37	177	HC3	9.82	100.00	1.23	64.42	27.61	1.23	22.70	0.99	2.00	0.09	1.81	1.44	0.09	1.36	KC5
H3.3-S28ac	STac	NT	Novel (know	622	3	0.48	3	3	0.79	0	1.90	0	0.74	0	1.05	159	HC3	75.24	1.90	1.90	1.90	70.48	1.90	100.00	1.88	0.28	0.28	0.28	1.85	0.28	2.00	KC4
H3.3-S32phs	phs	NT	Known in Pf	607	3	0.49	2	2	1.64	0	0	0	0	0	1.06	134	HC3	100.00	1.22	1.22	1.22	1.22	1.22	64.63	2.00	0.09	0.09	0.09	0.09	0.09	1.81	KC3
H3.3-S86ac	STac	CD	Novel (know	13319	20	0.15	9	8	0.24	0	0.32	0.05	0	0.13	0.17	198	HC3	75.00	6.25	100.00	15.63	6.25	40.63	53.13	1.88	0.80	2.00	1.19	0.80	1.61	1.73	KC6
H3.3-T11ac	STac	CD	Novel (know	1423	8	0.56	4	4	0.39	0	2.88	0	0	1.32	0.64	128	HC3	13.54	0.69	100.00	0.69	0.69	45.83	22.22	1.13	-0.16	2.00	-0.16	-0.16	1.66	1.35	KC6
H3.3-Y99oh	oh	CD	Novel	4848	29	0.60	6	6	0.92	0	0.15	0.19	0.35	0.89	0.91	157	HC3	100.00	2.17	16.30	20.65	38.04	96.74	98.91	2.00	0.34	1.21	1.31	1.58	1.99	2.00	KC4
H3.c-K23ac	Kac	NT	Novel (n.c. A	63	8	12.70	3	3	20.00	0	0		22.22	0	16.67	51	HC2	90.01	0.09	0.09	0.09	100.00	0.09	75.02	1.95	-1.05	-1.05	-1.05	2.00	-1.05	1.88	KC2
H3.c-K26ac	Kac	NT	Novel (n.c. A	63	16	25.40	7	6	33.33	36.36	0		22.22	0	27.78	29	HC1	91.67	100.00	0.00	0.06	61.11	0.06	76.40	1.96	2.00	-1.26	-1.26	1.79	-1.26	1.88	KC5
H3.c-K60ac	Kac	NT	Novel (n.c. A	54	7	12.96	4	4	50.00	0	25.00	0	28.57	11.11	0	30	HC1	100.00	0.04	50.00	0.04	57.14	22.22	0.04	2.00	-1.40	1.70	-1.40	1.76	1.35	-1.40	KC3
H3.c-R108fo	fo	CD	Novel	205	12	5.85	3	3	0	23.53	0	10.53	16.67	0	0	42	HC2	0.08	100.00	0.08	44.75	70.85	0.08	0.08	-1.07	2.00	-1.07	1.65	1.85	-1.07	-1.07	KC5
H3/H3.3-K115fo	fo	CD	Novel (know	5658	6	0.11	3	3	0.34	0.26	0	0	0	0	0	185	HC3	100.00	76.47	5.88	5.88	5.88	5.88	5.88	2.00	1.88	0.77	0.77	0.77	0.77	0.77	KC3
H3/H3.3-K122ac	Kac	CD	Known in Hs/	5712	2	0.04	2	2	0	0.14	0.27	0	0	0	0	195	HC3	7.41	51.85	100.00	7.41	7.41	7.41	7.41	0.87	1.71	2.00	0.87	0.87	0.87	0.87	KC6
H3/H3.3-K122fo	fo	CD	Known in Hs/	5712	83	1.45	21	13	4.80	1.42	4.34	0.59	1.06	2.71	0.45	90	HC3	100.00	29.58	90.42	12.29	22.08	56.46	9.38	2.00	1.47	1.96	1.08	1.34	1.75	0.97	KC6
H3/H3.3-K122me	me1	CD	Known in Hs/	5712	4	0.07	1	1	0	0	0	0	0	0	0.20	248	HC3	10.00	10.00	10.00	10.00	10.00	10.00	100.00	1.00	1.00	1.00	1.00	1.00	1.00	2.00	KC4
H3/H3.3-K18fo	fo	NT	Known in Hs/	2219	17	0.77	7	7	0.16	1.04	2.65	0	1.28	1.29	0.36	130	HC3	6.04	39.25	100.00	0.75	48.30	48.68	13.58	0.78	1.59	2.00	-0.12	1.68	1.69	1.13	KC6

			ALL SAMPLES					Mod/Total MERGED (%) Hierarchical clustering input and heat map plot (Figure 4A)									Mod/Total MERGED Normalized against maximum value in 7 time points (%) k-means input							Log10(Mod/Total MERGED) Normalized against maximum value in 7 time points (%) k-means plot (Figure 4B)							
Histone-Residue Position Modification	Type	Location	Novelty Status	Total_SpC	Modified_SpC	Mod/Total (%)	Times Detected in 56 Digests	Times Detected in 14 Samples	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	Hierarchical Clustering output order (Fig 5A)	Hierarchical Cluster Number (Fig 5A)	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	k-means Cluster Number (Fig 5B and S6)
H3/H3.3-K23me1	me1	NT	Known in Hs/	2172	2	0.09	2	2	0.16	0.53	0	0	0	0	0	190	HC3	30.19	100.00	3.77	3.77	3.77	3.77	3.77	1.48	2.00	0.58	0.58	0.58	0.58	0.58	KC5
H3/H3.3-R116fo	fo	CD	Novel	2977	12	0.40	4	4	0.41	5.48	0	0	0	0	1.19	78	HC3	7.48	100.00	0.36	0.36	0.36	0.36	21.72	0.87	2.00	-0.44	-0.44	-0.44	1.34	0.36	KC5
H3/H3.3-R42fo	fo	NT	Novel	971	158	16.27	25	14	13.98	17.11	17.95	20.83	39.71	25.00	7.58	38	HC2	35.21	43.09	45.20	52.46	100.00	62.96	19.09	1.55	1.63	1.66	1.72	2.00	1.80	1.28	KC2
H3/H3.3-R42me2	me2	NT	Novel	971	9	0.93	4	4	0	1.07	0	0	8.82	1.56	0	75	HC2	0.23	12.13	0.23	0.23	100.00	17.69	0.23	-0.64	1.08	-0.64	-0.64	2.00	1.25	-0.64	KC2
H3/H3.3-Y41oh	oh	NT	Novel (known	971	2	0.21	2	2	0.54	0	0.85	0	0	0	0	147	HC3	63.53	2.35	100.00	2.35	2.35	2.35	2.35	1.80	0.37	2.00	0.37	0.37	0.37	0.37	KC6
H3-D106oh	oh	CD	Novel	7656	2	0.03	2	2	0.06	0	0	0	0	0	0.07	236	HC3	85.71	28.57	28.57	28.57	28.57	28.57	100.00	1.93	1.46	1.46	1.46	1.46	2.00	KC4
H3-D77oh	oh	CD	Novel	16290	4	0.02	4	4	0.03	0.03	0.06	0	0.10	0	0	229	HC3	30.00	30.00	60.00	20.00	100.00	20.00	20.00	1.48	1.48	1.78	1.30	2.00	1.30	1.30	KC2
H3-K115ac	Kac	CD	Known in Hs/	5658	8	0.14	6	5	0.34	0.26	0	0.15	0	0.08	0	186	HC3	100.00	76.47	5.88	44.12	5.88	23.53	5.88	2.00	1.88	0.77	1.64	0.77	1.37	0.77	KC3
H3-K115cr	cr	CD	Novel (known	5658	309	5.46	14	10	4.31	6.39	5.51	6.05	4.58	8.30	1.29	62	HC2	51.93	76.99	66.39	72.89	55.18	100.00	15.54	1.72	1.89	1.82	1.86	1.74	2.00	1.19	KC8
H3-K115me1	me1	CD	Novel (known	5658	4	0.07	3	3	0.14	0	0	0	0	0.08	0.16	244	HC3	87.50	12.50	12.50	12.50	12.50	50.00	100.00	1.94	1.10	1.10	1.10	1.10	1.70	2.00	KC4
H3-K115oh	oh	CD	Novel (known	5658	2	0.04	2	2	0.07	0	0	0	0	0	0.16	245	HC3	43.75	12.50	12.50	12.50	12.50	12.50	100.00	1.64	1.10	1.10	1.10	1.10	1.10	2.00	KC4
H3-K14ac	Kac	NT	Known in Pf	1776	1084	61.04	46	14	43.68	45.26	69.96	65.12	59.80	72.22	78.24	15	HC1	55.83	57.85	89.42	83.23	76.43	92.31	100.00	1.75	1.76	1.95	1.92	1.88	1.97	2.00	KC8
H3-K14fo	fo	NT	Novel (known	1776	3	0.17	3	3	0.18	0	0	0	0	0.31	0.29	211	HC3	58.06	6.45	6.45	6.45	6.45	100.00	93.55	1.76	0.81	0.81	0.81	0.81	2.00	1.97	KC7
H3-K14me3	me3	NT	Known in Hs/	1776	2	0.11	2	2	0	0	0.43	0	0	0	0.29	202	HC3	4.65	4.65	100.00	4.65	4.65	4.65	67.44	0.67	0.67	2.00	0.67	0.67	1.83	KC8
H3-K18ac	Kac	NT	Known in Pf	2219	419	18.88	40	14	16.90	24.48	25.40	15.93	17.02	24.60	15.15	32	HC1	66.54	96.38	100.00	62.72	67.01	96.85	59.65	1.82	1.98	2.00	1.80	1.83	1.99	1.76	KC8
H3-K18me1	me1	NT	Known in Hs/	2219	32	1.44	13	9	1.26	0.52	1.06	0.88	0	4.53	1.09	106	HC3	27.81	11.48	23.40	19.43	0.44	100.00	24.06	1.44	1.06	1.37	1.29	-0.36	2.00	1.38	KC7
H3-K18oh	oh	NT	Novel (known	2219	7	0.32	1	1	1.11	0	0	0	0	0	0	149	HC3	100.00	1.80	1.80	1.80	1.80	1.80	1.80	2.00	0.26	0.26	0.26	0.26	0.26	0.26	KC3
H3-K23ac	Kac	NT	Known in Pf	2172	420	19.34	34	14	17.65	22.46	29.38	22.00	17.11	19.40	17.39	34	HC1	60.07	76.45	100.00	74.88	58.24	66.03	59.19	1.78	1.88	2.00	1.87	1.77	1.82	1.77	KC8
H3-K27ac	Kac	NT	Known in Pf	839	186	22.17	39	14	34.26	26.44	29.21	27.27	19.19	14.04	16.49	36	HC1	100.00	77.17	85.25	79.60	56.01	40.98	48.13	2.00	1.89	1.93	1.90	1.75	1.61	1.68	KC8
H3-K27fo	fo	NT	Novel (known	839	10	1.19	8	7	2.78	2.30	2.25	0	0	2.63	0	99	HC3	100.00	82.73	80.94	0.72	0.72	94.60	0.72	2.00	1.92	1.91	-0.14	-0.14	1.98	-0.14	KC7
H3-K27me1	me1	NT	Known in Hs/	839	6	0.72	2	2	3.70	0	0	0	0	0	1.06	111	HC3	100.00	0.54	0.54	0.54	0.54	0.54	28.65	2.00	-0.27	-0.27	-0.27	-0.27	-0.27	1.46	KC3
H3-K36ac	Kac	NT	Known in Hs/	1055	10	0.95	4	4	1.88	0.91	0	0	0	4.48	0	107	HC3	41.96	20.31	0.45	0.45	0.45	100.00	0.45	1.62	1.31	-0.35	-0.35	-0.35	2.00	-0.35	KC7
H3-K36fo	fo	NT	Novel (known	1055	3	0.28	2	2	1.88	0	0	0	0	0	0	137	HC3	100.00	1.06	1.06	1.06	1.06	1.06	1.06	2.00	0.03	0.03	0.03	0.03	0.03	0.03	KC3
H3-K36me1	me1	NT	Known in Hs/	1055	34	3.22	14	10	3.75	0.91	7.94	1.39	4.44	0.75	2.19	89	HC2	47.23	11.46	100.00	17.51	55.92	9.45	27.58	1.67	1.06	2.00	1.24	1.75	0.98	1.44	KC6
H3-K36me2	me2	NT	Known in Hs/	1055	3	0.28	2	2	1.25	0	0.79	0	0	0	0	142	HC3	100.00	1.60	63.20	1.60	1.60	1.60	1.60	2.00	0.20	1.80	0.20	0.20	0.20	0.20	KC3
H3-K37fo	fo	NT	Novel (known	142	11	7.75	6	6	0	0	17.86	12.50	9.09	6.25	8.33	48	HC2	0.11	0.11	100.00	69.99	50.90	34.99	46.64	-0.95	-0.95	2.00	1.85	1.71	1.54	1.67	KC1
H3-K56ac	Kac	CD	Known in Pf	3414	7	0.21	5	5	0.44	0.21	0.92	0	0.10	0	0.18	148	HC3	47.83	22.83	100.00	2.17	10.87	2.17	19.57	1.68	1.36	2.00	0.34	1.04	0.34	1.29	KC8
H3-K56fo	fo	CD	Known in Hs/	3414	37	1.08	12	9	0.87	1.26	5.96	0	0.51	1.19	0.71	132	HC3	14.60	21.14	100.00	0.34	8.56	19.97	11.91	1.16	1.33	2.00	-0.47	0.93	1.30	1.08	KC6
H3-K79ac	Kac	CD	Known in Hs/	16358	29	0.18	11	10	0.30	0.23	0.17	0.35	0.10	0	0.14	188	HC3	85.71	65.71	48.57	100.00	28.57	5.71	40.00	1.93	1.82	1.69	2.00	1.46	0.76	1.60	KC1
H3-K79fo	fo	CD	Known in Hs/	16358	170	1.04	20	11	1.49	0.51	0.67	1.74	2.61	0	2.24	115	HC3	57.09	19.54	25.67	66.67	100.00	0.77	85.82	1.76	1.29	1.41	1.82	2.00	-0.12	1.93	KC2
H3-K79me2	me2	CD	Known in Hs/	16358	85	0.52	4	3	0	1.90	0	0.87	0	0	0	178	HC3	1.05	100.00	1.05	45.79	1.05	1.05	1.05	0.02	2.00	0.02	1.66	0.02	0.02	0.02	KC5
H3-K79me3	me3	CD	Known in Pf	16358	8	0.05	4	4	0	0.15	0	0.09	0	0.03	0	243	HC3	13.33	100.00	13.33	60.00	13.33	20.00	13.33	1.12	2.00	1.12	1.78	1.12	1.30	1.12	KC5
H3-K9ac	Kac	NT	Known in Pf	1657	863	52.08	31	14	39.77	23.33	61.01	42.50	35.87	64.95	71.95	19	HC1	55.27	32.43	84.79	59.07	49.85	90.27	100.00	1.74	1.51	1.93	1.77	1.70	1.96	2.00	KC8
H3-K9cr	cr	NT	Known in Pf	1657	4	0.24	2	2	0.38	0	0	0	0	0	0.61	206	HC3	62.30	3.28	3.28	3.28	3.28	3.28	100.00	1.79	0.52	0.52	0.52	0.52	0.52	2.00	KC4
H3-K9me1	me1	NT	Known in Pf	1657	48	2.90	14	11	4.17	2.50	0.92	1.25	6.52	4.12	0.61	84	HC2	63.96	38.34	14.11	19.17	100.00	63.19	9.36	1.81	1.58	1.15	1.28	2.00	1.80	0.97	KC2
H3-K9me2	me2	NT	Known in Hs/	1657	91	5.49	20	12	7.77	10.00	8.26	3.75	3.26	3.78	0.91	45	HC2	77.70	100.00	82.60	37.50	32.60	37.80	9.10	1.89	2.00	1.92	1.57	1.51	1.58	0.96	KC5
H3-K9me3	me3	NT	Known in Pf	1657	23	1.39	3	3	3.98	0	0.46	0	0	0	0.30	110	HC3	100.00	0.50	11.56	0.50	0.50	0.50	7.54	2.00	-0.30	1.06	-0.30	-0.30	-0.30	0.88	KC3
H3-R116me2	me2	CD	Novel	2977	2	0.07	2	2	0.10	0	0.13	0	0	0	0	226	HC3	76.92	15.38	100.00	15.38	15.38	15.38	15.38	1.89	1.19	2.00	1.19	1.19	1.19	1.19	KC6
H3-R17me1	me1	NT	Known in Pf	1890	187	9.89	34	14	6.81	12.42	9.92	20.43	14.29	11.82	8.05	49	HC2	33.33	60.79	48.56	100.00	69.95	57.86	39.40	1.52	1.78	1.69	2.00	1.84	1.76	1.60	KC1
H3-R17me2	me2	NT	Known in Pf	1890	54	2.86	20	13	2.16	5.23	1.65	5.38	4.76	4.03	1.44	58	HC2	40.15	97.21	30.67	100.00	88.48	74.91	26.77	1.60	1.99	1.49	2.00	1.95	1.87	1.43	KC8
H3-R53fo	fo	CD	Novel	672	6	0.89	3	3	0	1.06	4.55	0	0	0	0	131	HC3	0.44	23.30	100.00	0.44	0.44	0.44	0.44	-0.36	1.37	2.00	-0.36	-0.36	-0.36	-0.36	KC6
H3-R72fo	fo	CD	Novel	1860	224	12.04	10	10	7.89	28.63	13.54	0	3.96	12.50	1.24	43	HC2	27.56	100.00	47.29	0.07	13.83	43.66	4.33	1.44	2.00	1.67	-1.16	1.14	1.64	0.64	KC5
H3-R72me2	me2	CD	Novel	1860	2	0.11	2	2	0	0.41	0	0	0	0	0	191	HC3	4.88	100.00	4.88	4.88	4.88	4.88	4.88	0.69	2.00	0.69	0.69	0.69	0.69	0.69	KC5
H3-R83fo	fo	CD	Novel (known	20057	12	0.06	4	4	0.05	0	0	0	0	0	0.35	209	HC3	14.29	5.71	5.71	5.71	5.71	5.71	100.00	1.15	0.76	0.76	0.76	0.76	0.76	2.00	KC4
H3-S10ac	STac	NT	Novel (known	1761	84	4.77	25	13	3.08	2.96	5.58	8.14	7.84	4.79	5.88	61	HC2	37.84	36.36	68.55	100.00	96.31	58.85	72.24	1.58	1.56	1.84	2.00	1.98	1.77	1.86	KC8
H3-S22ac	STac	NT	Novel	2172	8	0.37	8	6	0.32	0.53	0	0	0.44	0	0.72	158	HC3	44.44	73.61	2.78	2.78	61.11	2.78	100.00	1.65	1.87	0.44	0.44	1.79	0.44	2.00	KC4
H3-S28ac	STac	NT	Novel (known	1094	10	0.91	5	5	0.61	0.83	0.79	4.00	0	2.92	0	183	HC3	15.25	20.75	19.75	100.00	0.50	73.00	0.50	1.18	1.32	1.30	2.00	-0.30	1.86	-0.30	KC1
H3-S28phs	phs	NT	Known in Pf	1094	7	0.64	5	3	1.82	0	0	0	0	0.73	1.26	135	HC3	100.00	1.10	1.10	1.10	1.10	40.11	69.23	2.00	0.04	0.04	0.04	0.04	1.60	1.84	KC4
H3-S32phs	phs	NT	Known in Pf	1084	35	3.23	11	11	3.05	4.17	1.59	0	2.16	2.99	5.91	83	HC2	51.61	70.56	26.90	0.34	36.55	50.59	100.00	1.71	1.85	1.43	-0.47	1.56	1.70	2.00	KC4
H3-T107ac	STac	CD	Novel (known	7709	5	0.06	4	4	0.17	0	0.08	0	0	0.07	0	225	HC3	100.00	11.76	47.06	11.76	11.76	41.18	11.76	2.00	1.07	1.67	1.07	1.07	1.61	1.07	KC3
H3-T11ac	STac	NT	Novel (known	1761	5	0.28	5	5	0	0	0.86	2.33	0.98	0	0	181	HC3	0.86	0.86	36.91	100.00	42.06	0.86	0.86	-0.07	-0.07	1.57	2.00	1.62	-0.07	-0.07	KC1
H3-T80ac	STac	CD	Known in Hs/	18858	87	0.46	9	9	0.36	0.02	0.82	0	0	0.63	1.22	161	HC3	29.51	1.64	67.21	1.64	1.64	51.64	100.00	1.47	0.21	1.83	0.21	0.21	1.71	2.00	KC4
H3-Y54oh	oh	CD	Novel	3414	5	0.15	3	3	0	0	0	0	0	1.19	0.18	167	HC3	1.68	1.68	1.68	1.68	1.68	100.00	15.13	0.23	0.23	0.23	0.23	0.23	2.00	1.18	KC7
H4-D68oh	oh	CD	Novel	12666	11	0.09	9	8	0.14	0	0.05	0.08	0.20	0.07	0.07	231	HC3	70.00	10.00	25.00	40.00	100.00	35.00	35.00	1.85	1.00	1.40	1.60	2.00	1.54	1.54	KC2
H4-D85oh	oh	CD	Novel	14407	45	0.31	14	11	0.97	0	0.31	0.20	0.17	0.41	0.60	155	HC3	100.00	2.06	31.96	20.62	17.53	42.27	61.86	2.00	0.31	1.50	1.31	1.24	1.63	1.79	KC4
H4-K12ac	Kac	NT	Known in Pf	1221	969	79.36	36	14	75.94	67.14	76.30	86.61	78.08	75.69	89.72	6	HC1	84.64	74.83	85.04	96.53	87.03	84.36	100.00	1.93	1.87	1.93	1.98	1.94	1.93	2.00	KC8
H4-K16ac	Kac	NT	Known in Pf	1165	406	34.85	33	14	16.28	51.82	45.53	31.93	52.11	43.26	34.43	24	HC1	31.24	99.44	87.37	61.27	100.00	83.02	66.07	1.49	2.00	1.94	1.79	2.00	1.92	1.82	KC8
H4-K31ac	Kac	CD	Known in Hs/	2258	9	0.40	5	4	1.68	1.77	0	0	0	0	0.16	176	HC3	94.92	100.00	1.13	1.13	1.13	1.13	9.04	1.98	2.00	0.05	0.05	0.05	0.05	0.96	KC5
H4-K31fo	fo	CD	Known in Hs/	2258	7	0.31	3	3	0	1.33	0	0	0.52	0	0	175	HC3	1.50	100.00	1.50	1.50	39.10	1.50	1.50	0.18	2.00	0.18	0.18	1.59	0.18	0.18	KC5
H4-K31ub	ub	CD	Novel (known	2258	4	0.18	1	1	1.68	0	0	0	0	0	0	136	HC3	100.00	1.19	1.19	1.19	1.19	1.19	1.19	2.00	0.08	0.08	0.08	0.08	0.08	0.08	KC3
H4-K59ac	Kac	CD	Known in Hs/	7123	2	0.03	2	2	0.06	0	0	0	0	0.14	0	222	HC3	42.86	14.29	14.29	14.29	14.29	100.00	14.29	1.63	1.15	1.15	1.15	1.15	2.00	1.15	KC7
H4-K59fo	fo	CD	Known in Hs/	7123	154	2.16	14	11	3.25	1.03	3.46	0.47	0	5.42	1.34	92	HC3	59.96	19.00	63.84	8.67	0.37	100.00	24.72	1.78	1.28	1.81	0.94	-0.43	2.00	1.39	KC7
H4-K5ac	Kac	NT	Known in Pf	758	719	94.85	33	14	92.86	95.77	95.88	91.11	96.08	93.98	98.24	1	HC1	94.52	97.49	97.60	92.74	97.80	95.66	100.00	1.98	1.99	1.99	1.97	1.99	1.98	2.00	KC8
H4-K67ac	Kac	CD	Novel (n.c. A	12670	132	1.04	20	10	2.33	0	1.05	1.25	0.20	1.15	0.64	101	HC3	100.00	0.86	45.06	53.65	8.58	49.36	27.47	2.00	-0.07	1.65	1.73	0.93	1.69	1.44	KC3
H4-K67fo	fo	CD	Novel (known	12670	157	1.24	24	13	1.64	0.69	0.85	1.32	2.14	0.94	1.18	116	HC3	76.64	32.24	39.72	61.68	100.00	43.93	55.14	1.88	1.51	1.60	1.79	2.00	1.64	1.74	KC2
H4-K67me1	me1	CD	Known in Hs/	12670	56	0.44	5	3	2.33	0	0	0.37	0	0	0	139	HC3	100.00	0.86	0.86	15.88	0.86	0.86	0.86	2.00	-0.07	-0.07	1.20	-0.07	-0.07	-0.07	KC3
H4-K67me2	me2	CD	Novel (known	12670	20	0.16	5	4	0.18	0	0.15	0	0	0.07	0.31	208	HC3	58.06	6.45	48.39	6.45	6.45	22.58	100.00	1.76	0.81	1.68	0.81	0.81	1.35	2.00	KC4
H4-K67me3	me3	CD	Novel (n.c. A	12670	6	0.05	1	1	0	0	0	0.44	0	0	0	251	HC3	4.55	4.55	4.55	100.00	4.55	4.55	4.55	0.66	0.66	0.66	2.00	0.66	0.66	0.66	KC1
H4-K67oh	oh	CD	Novel (n.c. A	12670	6	0.05	5	5	0.09	0	0.05	0	0.10	0.07	0.03	230	HC3	90.00	20.00	50.00	20.00	100.00	70.00	30.00	1.95	1.30	1.70	1.30	2.00	1.85	1.48	KC2
H4-K77fo	fo	CD	Known in Hs/	8372	11	0.13	4	3	0.26	0	0.31	0.36	0	0	0	201	HC3	72.22	5.56	86.11	100.00	5.56	5.56	5.56	1.86	0.74	1.94	2.00	0.74	0.74	0.74	KC1
H4-K77me1	me1	CD	Known in Hs/	8372	77	0.92	2	2	4.38	0	0	1.22	0	0	0	112	HC3	100.00	0.46	0.46	27.85	0.46	0.46	0.46	2.00	-0.34	-0.34	1.44	-0.34	-0.34	-0.34	KC3
H4-K77oh	oh	CD	Novel (known	8372	35	0.42	9	8	0.33	0	0.08	0.12	1.73	0	0.63	122	HC3	19.08	1.16	4.62	6.94	100.00	1.16	36.42	1.28	0.06	0.67	0.84	2.00	0.06	1.56	KC2
H4-K79ac	Kac	CD	Known in Hs/	11516	442	3.84	21	13	7.65	2.55	3.73	2.76	1.95	4.53	9.79	69	HC2	78.14	26.05	38.10	28.19	19.92	46.27	100.00	1.89	1.42	1.58	1.45	1.30	1.67	2.00	KC4
H4-K79fo	fo	CD	Known in Hs/	11516	193	1.68	10	10	3.48	1.28	2.74	0	0.43	2.39	3.50	82	HC2	99.43	36.57	78.29	0.57	12.29	68.29	100.00	2.00	1.56	1.89	-0.24	1.09	1.83	2.00	KC4
H4-K79me2	me2	CD	Known in Hs/	11516	3	0.03	2	2	0.29	0	0	0	0	0	0	214	HC3	100.00	6.90	6.90	6.90	6.90	6.90	6.90	2.00	0.84	0.84	0.84	0.84	0.84	0.84	KC3
H4-K8ac	Kac	NT	Known in Pf	1059	803	75.83	36	14	42.03	90.00	92.31	85.83	96.92	90.68	82.13	9	HC1	43.37	92.86	95.24	88.56	100.00	93.56	84.74	1.64	1.97	1.98	1.95	2.00	1.97	1.93	KC8

				ALL SAMPLES					Mod/Total MERGED (%) Hierarchical clustering input and heat map plot (Figure 4A)									Mod/Total MERGED Normalized against maximum value in 7 time points (%) k-means input							Log10(Mod/Total MERGED) Normalized against maximum value in 7 time points (%) k-means plot (Figure 4B)							
Histone-Residue Position Modification	Type	Location	Novelty Status	Total_SpC	Modified_SpC	Mod/Total (%)	Times Detected in 56 Digests	Times Detected in 14 Samples	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	Hierarchical Clustering output order (Fig 5A)	Hierarchical Cluster Number (Fig 5A)	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	0-hpi	6-hpi	12-hpi	18-hpi	24-hpi	30-hpi	36-hpi	k-means Cluster Number (Fig 5B and S6)
H4-K91ac	Kac	CT	Known in Hs	14368	5	0.03	5	5	0	0.03	0	0.13	0.04	0	0.08	241	HC3	15.38	23.08	15.38	100.00	30.77	15.38	61.54	1.19	1.36	1.19	2.00	1.49	1.19	1.79	KC1
H4-K91fo	fo	CT	Known in Hs	14368	52	0.36	21	13	0.85	0.22	0.37	0.13	0.41	0.29	0.45	154	HC3	100.00	25.88	43.53	15.29	48.24	34.12	52.94	2.00	1.41	1.64	1.18	1.68	1.53	1.72	KC4
H4-K91me1	me1	CT	Novel (known	14368	8	0.06	2	2	0.26	0	0	0	0	0	0.30	207	HC3	86.67	6.67	6.67	6.67	6.67	6.67	100.00	1.94	0.82	0.82	0.82	0.82	0.82	2.00	KC4
H4-R35me1	me1	CD	Known in Hs	2209	3	0.14	1	1	0	0	2.65	0	0	0	0	127	HC3	0.75	0.75	100.00	0.75	0.75	0.75	0.75	-0.12	-0.12	2.00	-0.12	-0.12	-0.12	-0.12	KC6
H4-R3me1	me1	NT	Known in Pf	106	71	66.98	19	12	80.77	92.86	100	50.00	33.33	66.67	40.74	16	HC1	80.77	92.86	100.00	50.00	33.33	66.67	40.74	1.91	1.97	2.00	1.70	1.52	1.82	1.61	KC8
H4-R45fo	fo	CD	Novel	4529	163	3.60	20	12	3.08	6.02	3.00	6.49	3.64	1.85	2.76	58	HC2	47.46	92.76	46.22	100.00	56.09	28.51	42.53	1.68	1.97	1.66	2.00	1.75	1.45	1.63	KC1
H4-R55me1	me1	CD	Known in Hs	18480	8	0.04	3	3	0.20	0	0	0.06	0	0	0.03	217	HC3	100.00	10.00	10.00	30.00	10.00	10.00	15.00	2.00	1.00	1.00	1.48	1.00	1.00	1.18	KC3
H4-S47ac	STac	CD	Novel (know	17970	5	0.03	4	4	0.03	0	0	0	0	0.09	0.06	223	HC3	33.33	22.22	22.22	22.22	22.22	100.00	66.67	1.52	1.35	1.35	1.35	1.35	2.00	1.82	KC7
H4-S89ac	STac	CD	Novel (n.c.A	14366	2	0.01	2	2	0.07	0	0	0	0	0	0.08	237	HC3	87.50	25.00	25.00	25.00	25.00	25.00	100.00	1.94	1.40	1.40	1.40	1.40	1.40	2.00	KC4
H4-T80ac	STac	CD	Novel (know	14469	84	0.58	24	13	0.77	0.25	0.69	0.20	0.29	1.03	1.27	162	HC3	60.63	19.69	54.33	15.75	22.83	81.10	100.00	1.78	1.29	1.74	1.20	1.36	1.91	2.00	KC4
H4-Y51oh	oh	CD	Known in Hs	18534	26	0.14	12	8	0.40	0	0.20	0.06	0.03	0.18	0.13	200	HC3	100.00	5.00	50.00	15.00	7.50	45.00	32.50	2.00	0.70	1.70	1.18	0.88	1.65	1.51	KC3
H4-Y72oh	oh	CD	Novel	12237	41	0.34	10	8	0.28	0	0.05	0.24	1.41	0	0.44	123	HC3	19.86	1.42	3.55	17.02	100.00	1.42	31.21	1.30	0.15	0.55	1.23	2.00	0.15	1.49	KC2
H4-Y88oh	oh	CD	Known in Hs	14366	50	0.35	13	10	0.98	0	0.31	0.20	0.17	0.62	0.60	156	HC3	100.00	2.04	31.63	20.41	17.35	63.27	61.22	2.00	0.31	1.50	1.31	1.24	1.80	1.79	KC4

Discussion

We analyzed two independent histone preparations from seven time points of the *Plasmodium falciparum* intra-erythrocytic cycle for PTMs. Digesting each sample with multiple enzymes of varying specificities allowed us to obtain high sequence coverage for both the N-terminal tails and core domains (Figure 1.1B). After removing the redundancies within the H3 isoforms, 246 unique amino acid position/PTM combinations were detected on the eight *Plasmodium* histones (Figure 1.3A). Over half of these are novel and 85% occur on residues conserved across species, most of them located within the histone core domains (Figure 1.3A). Altogether, this demonstrates that our approach was able to dig deeper into the *Plasmodium* proteome to gain access to previously unknown and potentially critical and conserved PTMs.

Main classes of novel modifications identified on Plasmodium histones

The largest class of novel PTMs we detected is formylation (Figure 1.2A). Formylation of lysine side chains has been shown to interfere with acetylation and methylation, in particular with their signaling functions involved in oxidative and nitrosative stresses (Jiang et al., 2007b). In this context, it is interesting to note that 67% of the residues we observed as formylated are modification hot-spots, *i.e.* also detected as otherwise acetylated, methylated, or ubiquitinated (Figure 1.3A). Formylation may also play a competing role on *Plasmodium*

histones and whether formylation levels might be regulated by stress remains to be tested.

About a quarter of the novel PTMs are hydroxylations on lysines, aspartates, and tyrosines (Figure 1.2A). Proline and lysine hydroxylation have been shown to stabilize the triple helix structure of collagen (Chopra and Ananthanarayanan, 1982). Interestingly, all but four of the 29 hydroxylated sites are located within the histone globular domains (Figure 1.3A), and 22 of these are involved in alpha-helices. Since hydroxylysines could still be potentially modified on their e-N atom, it is unlikely that these hydroxylations should be *bona fide* epigenetic marks. However, the clear enrichment of these hydroxylations on residues involved in secondary structures hints at a stabilizing conformational role. Since such hydroxylations have not been systematically investigated in other species, we cannot conclude whether these hydroxylations are specific to *Plasmodium* histones or are a process common to all histones.

Acetylation on serine and threonine constitute one of the main classes of novel PTMs in our dataset (Figure 1.2A). Several serines and threonines that are usually found phosphorylated in other organisms are found exclusively acetylated in *Plasmodium* (Figure 1.3A and Table S1.3). For example, H2B-S28 aligns with S36, a known phosphorylation site in human H2B (Maile et al., 2004). Acetylations on serine and threonine are not widely known (Tweedie-Cullen et

al., 2012), but two studies have suggested that such acetylations may be a way to inhibit/prevent phosphorylation on these residues (Mittal et al., 2006; Mukherjee et al., 2006).

Activation vs. repression marks

Acetylations and monomethylations represent 37% of the PTMs we detected (Figure 1.2A and Table 1.3) and are associated with transcriptionally permissive states of chromatin. In contrast, dimethylations and trimethylations, which are usually associated with inactive states of chromatin, only constitute 7 and 2% of the detected PTMs, respectively.

Remarkably, most of the classical transcriptional repression marks were absent from the dataset or present at very low levels. For example, H3-K9me3, H3-K36me2, and H3-K27me1 were all detected at low levels during the parasite's asexual cycle and at a limited number of time points at the beginning and end of the erythrocytic cycle. The six trimethylations detected in our dataset fell in clusters KC3, KC5, KC6 and KC1 (Figure 1.7C), *i.e.* were the most abundant from early ring to early trophozoite. These observations suggest that important silencing events occur at the early trophozoite stage and could be related to the previously described deposition of H3-K9me3 in the sub-telomeric regions associated with silencing of the clonally variant "*Var*" genes (Scherf et al., 2008).

In contrast to the detection of H3-K36me3 in yeast and humans, we did not detect trimethylated H3-K36 by mass spectrometry, in agreement with (Trelle et al., 2009). A technical limitation cannot explain the absence of this modification since K36 was detected as acetylated and formylated on both H3 isoforms, and as mono- and dimethylated on H3, all of these PTMs being newly reported in *Plasmodium* (Figure 1.3A). While H3-K36me3 had previously been detected in *Plasmodium* using an antibody directed against yeast H3-K36me3 (Cui et al., 2008), a more recent study did not confirm this finding (Issar et al., 2009). Furthermore, *Plasmodium* H3 and H3.3 differ from yeast and human histones in their amino acid sequence in the few amino acids just N-terminal to K36, so it is unclear if the H3-K36me3 yeast antibody used in the initial study can recognize the *Plasmodium* me3. While *Plasmodium* possess a putative H3-K36 methylase, PfSet2 (Cui et al., 2008), which is a member of the Set2/Ash1/Nsd1-3 family of enzymes that methylates H3-K36 in other organisms, members of this family can differ in their ability to di- or trimethylate H3-K36 (Yuan et al., 2011).

Overall, the high abundance of activation marks and the low number of repressive modifications are consistent with the prevalence of euchromatin observed in the asexual cycle of the human malaria parasite (Ponts et al., 2010; Westenberger et al., 2009).

***Plasmodium* specific acetylation occupancy patterns**

Although lysine acetylations on the N-terminal tails of all histones are consistently measured at high levels across all time points (Figure 1.4A), distinct patterns of acetylation occupancy are observed. The overall trend for all *Plasmodium* canonical histones is that acetylation levels decrease linearly as the lysines become more internally located (Figure 1.4B).

The statistically supported linear decrease in acetylation levels along *Plasmodium* H4 N-terminal tail is in stark disagreement with top down work on human H4, in which acetylation occupancy is higher on K16, followed by K12, then K8/K5 (Keen et al., 2003; Pesavento et al., 2008; Smith et al., 2003; Zhang et al., 2002). In the mouse brain (Tweedie-Cullen et al., 2012), H4 acetylation also occurs most commonly on K16 (46%), K12 (27%), K8 (22%) and K5 (13%). Such occupancy numbers in human and mouse H4 support the zip model suggested in earlier MS (Zhang et al., 2002) and Edman sequencing (Thorne et al., 1990) analyses: H4 acetylation is proposed to occur from the C- to N-terminus of the N-terminal tail, while deacetylation would occur in the opposite direction (from N- to C-terminus). The reverse occupancy pattern observed in *Plasmodium* has been previously raised at the qualitative level for *Plasmodium* H4-K8 and K12 (Trelle et al., 2009) and our comprehensive quantitative data supports this contradictive pattern in *Plasmodium*.

In PfH3/H3.3, the linear decrease is not statistically supported mostly because the acetylation levels are highest on K14. This pattern of acetylation occupancy has been observed in yeast H3 (Jiang et al., 2007a). This top down analysis on yeast H3 also reported K23 as highly acetylated, on par with K14, which was not the case in *Plasmodium*. In another recent analysis (Tweedie-Cullen et al., 2012), the highest levels are also observed on H3-K14, then K23. In both yeast and mouse, H3-K9 is not as highly modified as in *Plasmodium* contributing to the most discrepancy between the three analyses.

In H2B, the acetylation occupancy reported for mouse brain is again opposite in *Plasmodium* H2Bv and H2B where the two most N-terminal lysines (K3 and K8) display the highest acetylation levels, while their homologous residues in mouse H2B (K5 and K11) only show 1% acetylation (Tweedie-Cullen et al., 2012).

No acetylation occupancy data is available for H2Az in the literature. Our analysis provides 14 data points over eight lysines (five of which conserved) located within the first 37 residues. Our data shows reproducibly high levels of acetylation across these residues, even for the ones next to the first putative alpha helix (Figure 1.4B). The surprisingly stable acetylation levels on H2Az long N-terminal tail combined with the atypical abundance of this isoform in *Plasmodium* (Figure 1.1C) might play a direct role in the dynamic condensation/relaxation of chromatin.

Altogether, our results underline a general transcriptionally permissive state of the chromatin throughout the parasite cell cycle with a restricted number of histone modifications associated with transcriptional silencing. As a whole, our data suggest that the parasite has a unique histone modification signature that correlates with an intensive replication and transcriptional activity associated with the parasite pathogenicity. Understanding the molecular mechanisms that control these atypical features may offer novel therapeutic strategies.

Experimental Procedures

Parasite culture, synchronization, and harvest

The P. falciparum strain 3D7 was maintained in human red blood cells at 5% haematocrit as described previously (Le Roch et al., 2003). Cultures were synchronized twice at ring stage with 5% sorbitol treatments performed eight hours apart (Le Roch et al., 2003). Two independent sets of samples (Table 1.1 and Figure S1) were collected every six hours for 36 hours, starting exactly 48 hours after the first sorbitol treatment, for a total of seven samples from early ring (0-hpi), middle ring (6-hpi), late ring (12-hpi), early trophozoite (18-hpi), late trophozoite (24-hpi), early schizont (30-hpi), and late schizont (36-hpi).

Histones precipitation and multiple digestions

Parasites were extracted from infected red blood cells and nuclei were released in parasite lysis buffer (Ponts et al., 2010) before histone acid extraction with 0.25 M HCl (Longhurst and Holder, 1997). Histone preparations (*ca*. 8 μg of total proteins) were independently digested using elastase, Lys-C+Glu-C, Arg-C, or Lys-C+trypsin (Florens and Washburn, 2006).

Data acquisition

Each differently digested sample was analyzed independently (Table 1.1) by Multidimensional Protein Identification Technology (MudPIT) as described previously (Florens and Washburn, 2006; Washburn et al., 2001) using an Eksigent NanoLC 2D system coupled to an LTQ-Orbitrap hybrid mass spectrometer (Thermo Scientific). Full MS spectra were recorded on the eluting peptides over a 400 to 1600 *m*/*z* range in the Orbitrap at 60K resolution, followed by fragmentation in the ion trap on the first to fifth most intense ions selected from the full MS spectrum.

Data Analysis

RAW files were extracted into .ms2 file format (McDonald et al., 2004) using RawDistiller v. 1.0 (Zhang et al., 2011). MS/MS spectra were searched using SEQUEST v.27 (rev.9) (Eng et al., 1994) with a peptide mass tolerance of 10 ppm and of +/- 0.5 amu for fragment ions. No enzyme specificity was imposed during the search against a protein database combining non-redundant *Plasmodium falciparum* and human proteins (Table 1.1). As described in (Xiang et al., 2007), 32 differential modification searches were set up (Table S1.1) in a recursive manner to query for peptides containing methylated lysines and arginines, oxidized methionines and hydroxylated lysines, aspartates and tyrosines, formylated lysines and arginines, dimethylated lysines and arginines,

trimethylated lysines, acetylated lysines, serines and threonines, crotonylated lysines, phosphorylated serines, threonines, and tyrosines, ubiquitinated lysines, as well as potentially N-terminally acetylated methionines, serines, alanines, and valines. Spectra/peptide matches were filtered using conservative criteria and compared using DTASelect/CONTRAST (Tabb et al., 2002). *NSAF v.7* (in-house written software) was used to create the final reports on all detected peptides and non-redundant proteins identified across the different runs (Table S1.2). Spectral and protein level FDRs were, on average, 0.24 ± 0.13% and 1.5 ± 0.45%, respectively (Table S1.2 (Supplemental PDF available through dissertation publisher)). To estimate relative protein levels, distributed Normalized Abundance Factors (dNSAFs) were calculated for each non-redundant protein or protein group, as described in (Zhang et al., 2010). The PTMR option in *NSAF v.7* was used to extract total and modified spectral counts for each amino acid within *Plasmodium* histones (Table 1.2) and calculate modification levels based on local spectral counts (Figure 1.6). The mass spectrometry .RAW files, .ms2 files created by RawDistiller, the .sqt files generated by SEQUEST, and the DTASelect output files for all 56 analyses will be available to download from The Stowers Institute Original Data Repository (http://srdr.stowers.org/).

Extended Experimental Procedures

Parasite harvest and histones precipitation

A first set of samples were collected every six hours for 36 hours, starting 48 hours exactly after the first sorbitol treatment, for a total of seven samples covering the intra-erythrocytic cycle from early ring stage to mature schizont. A second set of seven samples was treated with 400 nM of the proteasome inhibitor MG-132 for three hours prior to each harvest to prevent any bias related to proteasomal degradation of histone proteins (Figure S1). For each sample, cultures were washed once in PBS containing MG-132 200 nM and NEM 2 mM (MN-PBS) and parasites were extracted from infected red blood cells with ice-cold 1% acetic acid. Extracted parasites were then washed extensively with MN-PBS. Nuclei were released in parasite lysis buffer (25 mM Tris HCl pH 7.8, 1 mM EDTA, 0.25% [v/v] Igepal® CA-630, 2 mM PMSF, 20 mM NEM, and Roche complete mini EDTA-free protease inhibitor cocktail, see (Ponts et al., 2010) and washed with NaCl 0.8 M before histone acid extraction with 0.25 M HCl (Longhurst and Holder, 1997). Histones were precipitated with 25% TCA, followed by two acetone washes.

Preparation of peptide mixtures by multiple digestions

The two biological replicate sets of histone preparations, with or without prior treatment with MG-132, were independently processed by multiple digestions

(Table 1.1). The method follows that of (MacCoss et al., 2002) with modifications. TCA-precipitated protein samples were solubilized in freshly made 0.1 M Tris-HCl, pH 8.5, 8 M urea (Sigma-Aldrich), and 5 mM TCEP (Tris(2-Carboxylethyl)-Phosphine Hydrochloride, Pierce). After 30 minutes at room temperature, freshly made 0.5 M chloroacetamide (CAM, Sigma-Aldrich) was added to a final concentration of 10 mM, and the samples were left at room temperature for another 30 minutes in the dark. The denatured, reduced and alkylated protein samples were then split into aliquots of equal volume to be digested using elastase, Lys-C+Glu-C, Arg-C, or Lys-C+trypsin (*ca*. 8 μg of total proteins as determined by Bradford protein assay (Bradford, 1976)).

For elastase digestions, samples were diluted with 0.1 M Tris-HCl pH 8.5 so that the final concentration of urea was 2 M. Elastase (Calbiochem) was the added at 1:50 enzyme to protein ratio (w/w) and incubated for 8.5 hours at 37°C with constant shaking. For the endoproteinases Lys-C and Glu-C digestions, Lys-C (Roche) was first added to the samples at 1:50 enzyme to protein ratio (w/w) and incubated overnight at 37°C while shaking. Then, samples were diluted with 0.1 M Tris-HCl pH 8.5 for a final concentration 1 M urea and Glu-C, 1:20 enzyme to protein ratio (w/w), was added and incubated for over 12 hours at 25°C. For the trypsin digestions, Lys-C was first added to undiluted samples at 1:50 enzyme to protein ratio (w/w) and incubated overnight at 37°C. Samples were diluted with 0.1 M Tris-HCl pH 8.5 to a final concentration of 2 M urea, followed by the

addition of 0.5 mM $CaCl_2$. Trypsin (Promega) was added at a 1:100 enzyme to protein ratio (w/w) and incubated for over 12 hours at 37°C. For the endoproteinase Arg-C digestions, samples were diluted with 0.1 M Tris-HCl pH 8.5 to a final concentration of 4 M urea, followed by the addition of 2 mM $CaCl_2$. Then Arg-C (Roche) was added at 1:100 enzyme to protein ratio (w/w) along with the provided activation solution at 1:10 (v/v). Digestion was performed overnight at 37°C while shaking. All enzymatic digestions were finally stopped by the addition of formic acid to 5%.

Data acquisition

Each differently digested sample was analyzed independently (Table 1.1) by Multidimensional Protein Identification Technology (MudPIT) as described previously (Washburn et al., 2001), (Florens and Washburn, 2006). Peptide mixtures were pressure-loaded onto 250 μm fused silica microcapillary columns packed first with 3 cm of 5-µm Strong Cation Exchange material (Partisphere SCX, Whatman), followed by 1 cm of 5-µm C_{18} reverse phase (Aqua, Phenomenex). Loaded 250 μm columns were connected using a filtered union (UpChurch) to 100 μm fused-silica columns pulled to a 5 µm tip using a P 2000 CO_2 laser puller (Sutter Instruments) and packed with 9 cm of reverse phase material (McDonald et al., 2002). The loaded microcapillary columns were placed in-line with an Eksigent NanoLC 2D system. The four channels of the

nanoLC were filled with: 5% acetonitrile and 0.1% formic acid (channels 1A and 2A); 0.5 M ammonium acetate, 5% acetonitrile, and 0.1% formic acid (channel 1B); 80% acetonitrile and 0.1% formic acid (channel 2B). Fully automated 12 step chromatography runs were carried out with a flow rate of 500 nL/min, measured at 250 nL/min at the tip of the column. Peptides were sequentially eluted from the SCX resin to the reverse phase resin by a 10 min-long salt step of increasing concentrations (0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 100, and 100% for steps 1 through 12, respectively), followed by the organic gradient to 75% solvent. The application of a 2.5 kV distal voltage electrosprayed the eluting peptides directly into a LTQ-Orbitrap hybrid mass spectrometer (Thermo Scientific) equipped with a custom-made nano-LC electrospray ionization source. Full MS spectra were recorded on the peptides over a 400 to 1600 *m*/*z* range in the Orbitrap at 60K resolution, followed by fragmentation in the ion trap (at 35% collision energy) on the first to fifth most intense ions selected from the full MS spectrum. Dynamic exclusion was enabled for 90 s (Zhang et al., 2009), with an exclusion window of -0.03 to +1.03. Mass spectrometer scan functions and HPLC solvent gradients were controlled by the XCalibur data system (Thermo Scientific).

MS/MS dataset query

RAW files were extracted into .ms2 file format (McDonald et al., 2004) using RawDistiller v. 1.0, in-house developed software (Zhang et al., 2011). RawDistiller D(g, 6) settings were used to abstract MS1 scan profiles by Gaussian fitting and to implement dynamic offline lock mass using six background polydimethylcyclosiloxane ions as internal calibrants (Zhang et al., 2011). MS/MS spectra were first searched using SEQUEST v.27 (rev.9) (Eng et al., 1994) with a peptide mass tolerance of 10 ppm and of +/- 0.5 amu for fragment ions. No enzyme specificity was imposed during the search against a protein database combining non-redundant 5439 *Plasmodium falciparum* (PlasmoDB release 5.5) and 30552 human proteins (NCBI 2008-03-04 release), as well as 162 usual contaminants such as human keratins, IgGs, and proteolytic enzymes (Table 1.1). To estimate false discovery rates (FDR), each protein sequence was randomized (keeping the same amino acid composition and length) and the resulting "shuffled" sequences were added to the database used for the SEQUEST searches, for a total search space of 72306 amino acid sequences. To account for alkylation by CAM, +57 Da were added statically to cysteine residues for all searches (Table S1.1).

As primarily described in (Xiang et al., 2007), 32 differential modification searches were set up (Table S1.1) to query a protein database containing only the sequences of *P. falciparum* and human histones ("All Spectra *vs*. Few

Proteins" or ASFP searches, Figure 1.6). We used these combinatorial searches to query for peptides containing methylated lysines and arginines (+14.0157), oxidized methionines and hydroxylated lysines, aspartates and tyrosines (+15.9949), formylated lysines and arginines (+27.9949), dimethylated lysines and arginines (+28.0314), trimethylated lysines (+42.0470), acetylated lysines, serines and threonines (+42.0106), crotonylated lysines (+68.0261), phosphorylated serines, threonines, and tyrosines (+79.9663), ubiquitinated lysines (+114.0429), as well as potentially N-terminally acetylated methionines, serines, alanines, and valines (+42.0106). The maximum number of modified amino acids per differential modification in a peptide was limited to six. To circumvent SEQUEST limitations with heterogeneous modifications on the same type of residue within the same peptide, we used combinations of static delta-masses with differential modifications to detect K methylation marks within the same peptides as K acetylations (Table S1.1). An in-house written software, SQTMASSFixer (Zhang et al., 2011), was then used to insert the appropriate PTMs symbols after the statically-searched residues (Table S1.1 and Figure 1.6). After this round of combinatorial searches, sqt-merge (Zybailov et al., 2005) was used to combine the SEQUEST output files (.sqt files) generated from the normal search (*i.e.*, "All Spectra *vs*. All Proteins" or ASAP search, without modifications) and the 32 sets of ASFP PTMs searches described above into one set (Table S1.1 and Figure 1.6). This merging step allowed only the best matches out of *n*

differential SEQUEST queries to be ranked first based on cross-correlation scores (XCorr) and for normalized differences between XCorrs (DeltaCn) to be recalculated accordingly.

For the second round of searches, only spectra matching modified peptides were selected (-m 0 –t 0 DTASelect parameters), and their coordinates written out into smaller .ms2 files using the "-- copy" utility of DTASelect (Tabb et al., 2002). These subsetted ms2 files contained at best a few hundred MS/MS spectra and were subjected to the same 32 combinatorial searches against the complete protein database described above including shuffled sequences ("Modified Spectra *vs*. All Proteins" or MSAP searches, Table S1.1 and Figure 1.6). Again sqt-merge was used to bring together the results generated by these differential searches. This step allowed us to check that spectra matching modified peptides from *Plasmodium* and human histone sequences did not find a better match against the complete protein database . Finally, we used SQTMassFixer to correct each calculated ($M+H^+$) in the .sqt files by adding an integer multiple of 1.0033 Da (^{13}C-^{12}C) to obtain the value closest to the measured monoisotopic ($M+H^+$) as described in (Wenger et al., 2010).

Data Analysis

DTASelect v.1.9 (Tabb et al., 2002) was used to select and sort peptide/spectrum matches passing the following criteria set. Spectra/peptide matches were only

retained if they had a DeltCn of at least 0.08. Minimum XCorr values were set at 1.8 for singly-, 2.0 for doubly-, and 3.0 for triply-charged spectra for the Arg-C and trypsin-digested samples, and at 2.5 and 3.5 for +2 and +3 spectra, respectively, in the Lys-C+Glu-C and elastase-digested samples. In addition, the peptides had to be at least 7 amino acids long and comply with the specificity of the enzymes used to digest samples. All spectra matching modified peptides were visually assessed (Figures S3A-N) and given an evaluation flag (Y/M/N for yes/maybe/no) using SQTMASSFixer. The "no" matches were removed from the final data (-v 2 parameter in DTASelect).

Results from different analyses were merged and compared using CONTRAST (Tabb et al., 2002). Combining all runs, proteins had to be detected by at least three peptides and proteins that were subsets of others were removed using the parsimony option in DTASelect on the proteins detected after merging all runs. Proteins that were identified by the same set of peptides (including at least one peptide unique to such protein group to distinguish between isoforms) were grouped together, and one accession number was arbitrarily considered as representative of each protein group. *NSAF v.7* (in-house written software) was used to create the final reports on all detected peptides and non-redundant proteins identified across the different runs (Table S1.2). FDRs were calculated by NSAF *v. 7* as:

$$Spectral\_FDR = \frac{2xSHUFFLED\_SpectralCounts}{Total\_SpectralCounts} x100$$

$$Protein\_FDR = \frac{SHUFFLED\_\mathrm{Pr}oteinCounts}{Total\_\mathrm{Pr}oteinCounts} x100$$

Spectral and protein level FDRs were, on average, 0.24 ± 0.13% and 1.5 ± 0.45%, respectively (Table S1.2). To estimate relative protein levels, distributed Normalized Abundance Factors (dNSAFs) were calculated for each non-redundant protein or protein group, as described in (Zhang et al., 2010):

$$dNSAF_i = \frac{dSAF_i}{\sum_{i=1}^{N} dSAF_i}$$

$$\text{with } dSAF_i = \frac{uSpC_i + \frac{uSpC_i}{\sum_{m=1}^{M} uSpC_m} \times sSpC_{i_j}}{Length_i},$$

in which shared spectral counts (sSpC) are distributed based on spectral counts unique to each protein i (uSpC) divided by the sum of all unique spectral counts for the M protein isoforms that shared peptide *j* with protein *i*. Summing the normalized spectral counts values (SdNSAF) of specific group of proteins (histones in Figure 1.3A, plasmodium vs. human in Figure 1.1A) was used to assess their relative enrichment over all detected proteins as described in (Florens et al., 2006).

The PTMR option in *NSAF v.7* was used to extract total and modified spectral counts for each amino acid within Plasmodium histones (Table 1.2) and calculate modification levels based on local spectral counts (Xiang et al., 2007) (Figure 1.6):

$$AA_x\_PTM_y\_Level = \frac{AA_x\_PTM_y\_SpectralCounts}{AA_x\_Total\_SpectralCounts} x100$$

Sequence coverage observed for each *Plasmodium* histone across the different runs, as well as the locations of different PTMs along the sequences, was also displayed using *NSAF v. 7* PTMR option. The mass spectrometry .RAW files, .ms2 files created by RawDistiller, the .sqt files generated by SEQUEST, and the DTASelect output files for all 56 analyses will be available to download from The Stowers Institute Original Data Repository (http://srdr.stowers.org/).

Statistical analysis of protein levels

Summing the normalized spectral counts values (SdNSAF) of specific group of proteins (histones in Figure 1.1A, *Plasmodium* vs. human in Figure 1.1D) was used to assess their relative enrichment over all detected proteins as described in (Florens et al., 2006). To assess the variations in individual histone abundances across the intra-erythrocytic cycle, a multiple comparison analysis was performed using a bonferroni correction on histone abundances using matlab (Figure 1.1A and 1C). To assess the similarities in relative protein abundance (dNSAF) between all 14 samples, pairwise Pearson correlation coefficients were calculated using the R environment function *cor* (Figure 1.1E). Finally, individual proteins and independent analyses were hierarchically clustered under the same criteria using the HCL function of the MeV 4.8.1

software (Mar et al., 2011) with Pearson correlation as metric distance and average linkage as algorithm (Figure 1.1F).

Clustering of post-translational modifications across the erythrocyte cycle

Modifications levels (spectral counts merged between the two replicate analyses and expressed as a percentage of the modified spectral counts divided by the total spectral counts for a residue) observed in all seven time points of the *P. falciparum* life cycle (Table 1.2) were hierarchically clustered using Euclidean distance metric and WARD as a method using PermutMatrix software (Caraux and Pinloche, 2005). The tree was generated using the multiple-fragment heuristic algorithm (MF) as a seriation rule (Figure 1.5A). To group histone modifications based on their abundance profile across the seven time points, we first normalized each individual modification in each time point to the highest percent of modification across the seven stages (i.e. the highest value equals to 100%). We then applied k-means clustering to this normalized matrix using the Hartigan-Wong algorithm (Wong, 1979) and a maximum number of iterations set at 50000. To determine the best partition of our data, we continuously increased the number of clusters k from 3 to 20. The result showed that the optimal number of clusters was obtained when k=8 (Figure 1.5B), after carefully inspecting all the clusters and their silhouette and Hartigan indexes (Figure S6). All computations were run using R environment using k-means function for the

partition and daisy function to compute all the pairwise dissimilarities (Euclidean distances) between observations in the dataset for the silhouette.

Western blot analyses

Western blot analyses were performed using antibodies directed against each of the three methylation states of *Saccharomyces cerevisiae* H3-K4 (Figure 1.3C) and an antibody against H3-K27 acetylation (Figure 1.3D). Loading amounts in each lane were normalized using an antibody against *Saccharomyces cerevisiae* histone H3. Histone isolated from seven *Plasmodium* erythrocytic stages (treated with MG132) and uninfected red blood cells (RBC) were separated on 18% SDS-PAGE and transferred for blotting. Whole cell extract from wild-type yeast, yeast strains deficient in the H3-K4 methylating machinery, and a strain expressing the H3-K27A mutant were analyzed in parallel to ensure proper behavior of each antibody.

Table S1. Detailed Pipeline for Comprehensive Post-Translational Modifications Searches

For each of the 14 steps of the MS/MS query pipeline are specified:

the [software] used to perform each task;
the input and output file types;
the parameters entered in the params files or command lines.

STEP 1: EXTRACT AND RECALIBRATE PRECURSOR ION MASSES BY DYNAMIC OFFLINE LOCKMASS AND GENERATE MS/MS PEAK LISTS

[RawDistiller]		
input files	DOLM parameters	output files
*.RAW	D(g,6)	*.ms2

STEP 2: MATCH MS/MS SPECTRA TO PEPTIDES IN DATABASE WITH VARIOUS DIFFERENTIAL SEARCH OPTIONS

[SEQUEST]																									
input files	directory_name	database_name	ppm_peptide_mass_tolerance	isotopes	fragment_ion_tolerance	enzyme_number	diff_search_options =										max_num_differential_AA_per_mod	mass_type_parent	mass_type_fragment	static mods					output files
*.ms2	_ASAP	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	0.00	X	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASAP.sqt
	_ASFP_14KR_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	14.0157	KR	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_14KR_16M.sqt
	_ASFP_16KDY	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	0.00	X	15.9949	KDY	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_16KDY.sqt
	_ASFP_27KR_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	27.9949	KR	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_27KR_16M.sqt
	_ASFP_28KR_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	28.0314	KR	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_28KR_16M.sqt
	_ASFP_42.01AV_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	AV	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01AV_16M.sqt
	_ASFP_42.01AV_42.04K	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	AV	42.047	K	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01AV_42.04K.sqt
	_ASFP_42.01AVK_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	AVK	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01AVK_16M.sqt
	_ASFP_42.01KST_14R	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	14.0157	R	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01KST_14R.sqt
	_ASFP_42.01KST_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01KST_16M.sqt
	_ASFP_42.01KST_28R	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	28.0314	R	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01KST_28R.sqt
	_ASFP_42.01KST_58M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	58.0055	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01KST_58M.sqt
	_ASFP_42.01MK	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	MK	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01MK.sqt
	_ASFP_42.01ST_42.04K	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.0106	ST	42.047	K	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.01ST_42.04K.sqt
	_ASFP_42.04K_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	42.047	K	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_42.04K_16M.sqt
	_ASFP_68K_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	68.0261	K	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_68K_16M.sqt
	_ASFP_68K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	68.0261	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_68K_16M.sqt
	_ASFP_80STY_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	79.9663	STY	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_80STY_42.01SM.sqt
	_ASFP_114K_16M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	114.0429	K	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146								*_ASFP_114K_16M.sqt
	_ASFP_s14K_42.01K	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	27.9949	K	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157							*_ASFP_s14K_42.01K.sqt
	_ASFP_s14K_42.01K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	27.9949	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157							*_ASFP_s14K_42.01K_42.01SM.sqt
	_ASFP_s14K_68K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	54.0104	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157							*_ASFP_s14K_68K_42.01SM.sqt
	_ASFP_s14K_68K_58M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	54.0104	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157							*_ASFP_s14K_68K_58M.sqt
	_ASFP_s28K_42.01K	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	13.9792	K	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314							*_ASFP_s28K_42.01K.sqt
	_ASFP_s28K_42.01K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	13.9792	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314							*_ASFP_s28K_42.01K_42.01SM.sqt
	_ASFP_s28K_68K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	39.9947	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314							*_ASFP_s28K_68K_42.01SM.sqt
	_ASFP_s28K_68K_58M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	39.9947	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314							*_ASFP_s28K_68K_58M.sqt
	_ASFP_s42.01K_42.04K	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	0.0364	K	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106							*_ASFP_s42.01K_42.04K.sqt
	_ASFP_s42.01K_42.04K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	0.0364	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106							*_ASFP_s42.01K_42.04K_42.01SM.sqt
	_ASFP_s42.01K_68K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	26.0155	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106							*_ASFP_s42.01K_68K_42.01SM.sqt
	_ASFP_s42.01K_68K_58M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	26.0155	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106							*_ASFP_s42.01K_68K_58M.sqt
	_ASFP_s42.04K_68K_42.01SM	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	25.9791	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0470							*_ASFP_s42.04K_68K_42.01SM.sqt
	_ASFP_s42.04K_68K_58M	PfHistones_N-term-processed_wSHUFFLED.fasta	10	1 (=search isotopes)	0	0 (=No_Enzyme)	25.9791	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0470							*_ASFP_s42.04K_68K_58M.sqt

STEP 3: INSERT APPROPRIATE PTMs SYMBOLS AFTER STATICALLY-SEARCHED RESIDUES

[STQMASSFixer]		
input files	command line	output files
*_ASFP_s14K_42.01K.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(#)(14.0157)	*_ASFP_14K_42.01K.sqt
*_ASFP_s14K_42.01K_42.01SM.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(#)(14.0157)	*_ASFP_14K_42.01K_42.01SM.sqt
*_ASFP_s14K_68K_42.01SM.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(#)(14.0157)	*_ASFP_14K_68K_42.01SM.sqt
*_ASFP_s14K_68K_58M.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(#)(14.0157)	*_ASFP_14K_68K_58M.sqt
*_ASFP_s28K_42.01K.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(~)(28.0314)	*_ASFP_28K_42.01K.sqt
*_ASFP_s28K_42.01K_42.01SM.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(~)(28.0314)	*_ASFP_28K_42.01K_42.01SM.sqt
*_ASFP_s28K_68K_42.01SM.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(~)(28.0314)	*_ASFP_28K_68K_42.01SM.sqt
*_ASFP_s28K_68K_58M.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params combinestaticptms=K(~)(28.0314)	*_ASFP_28K_68K_58M.sqt
*_ASFP_s42.01K_42.04K.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params "combinestaticptms=K(!)(42.0106)	*_ASFP_42.01K_42.04K.sqt
*_ASFP_s42.01K_42.04K_42.01SM.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params "combinestaticptms=K(!)(42.0106)	*_ASFP_42.01K_42.04K_42.01SM.sqt
*_ASFP_s42.01K_68K_42.01SM.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params "combinestaticptms=K(!)(42.0106)	*_ASFP_42.01K_68K_42.01SM.sqt
*_ASFP_s42.01K_68K_58M.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params "combinestaticptms=K(!)(42.0106)	*_ASFP_42.01K_68K_58M.sqt
*_ASFP_s42.04K_68K_42.01SM.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params "combinestaticptms=K(!)(42.0470)	*_ASFP_42.04K_68K_42.01SM.sqt
*_ASFP_s42.04K_68K_58M.sqt	./SQTMassFixer sqtparamfile=/.../sequest.params "combinestaticptms=K(!)(42.0470)	*_ASFP_42.04K_68K_58M.sqt

STEP 4: MERGE ALL DIFFERENTIALLY-SEARCHED SQT FILES, RANK BASED ON XCorr AND RECALCULATE DeltaCn

[sqt-merge]		
input files	command line	output files
*_ASAP.sqt *_ASFP_14KR_16M.sqt *_ASFP_16KDY.sqt *_ASFP_27KR_16M.sqt *_ASFP_28KR_16M.sqt *_ASFP_42.01AV_16M.sqt *_ASFP_42.01AV_42.04K.sqt *_ASFP_42.01AVK_16M.sqt *_ASFP_42.01KST_14R.sqt *_ASFP_42.01KST_16M.sqt *_ASFP_42.01KST_28R.sqt *_ASFP_42.01KST_58M.sqt *_ASFP_42.01MK.sqt *_ASFP_42.01ST_42.04K.sqt *_ASFP_42.04K_16M.sqt *_ASFP_68K_16M.sqt *_ASFP_80STY_16M.sqt *_ASFP_114K_16M.sqt *_ASFP_14K_42.01K.sqt *_ASFP_14K_42.01K_42.01SM.sqt *_ASFP_14K_68K_42.01SM.sqt *_ASFP_14K_68K_58M.sqt *_ASFP_28K_42.01K.sqt *_ASFP_28K_42.01K_42.01SM.sqt *_ASFP_28K_68K_42.01SM.sqt *_ASFP_28K_68K_58M.sqt *_ASFP_42.01K_42.04K.sqt *_ASFP_42.01K_42.04K_42.01SM.sqt *_ASFP_42.01K_68K_42.01SM.sqt *_ASFP_42.01K_68K_58M.sqt *_ASFP_42.04K_68K_42.01SM.sqt *_ASFP_42.04K_68K_58M.sqt	sqt-merge --output= ../Ph4st_(ASAP,ASFP_14KR_16M,ASFP_16KDY,ASFP_27KR_16M,ASFP_28KR_16M,ASFP_42.01AV_16M,ASFP_42.01AV_42.04K,ASFP_42.01AVK_16M,ASFP_42.01KST_14R,ASFP_42.01KST_16M,ASFP_42.01KST_28R,ASFP_42.01KST_58M,ASFP_42.01MK,ASFP_42.01ST_42.04K,ASFP_42.04K_16M,ASFP_68K_16M,ASFP_68K_42.01SM,ASFP_80STY_16M,ASFP_114K_16M,ASFP_s14K_42.01K,ASFP_s14K_42.01K_42.01SM,ASFP_s14K_68K_42.01SM,ASFP_s14K_68K_58M,ASFP_s28K_42.01K,ASFP_s28K_42.01K_42.01SM,ASFP_s28K_68K_42.01SM,ASFP_s28K_68K_58M,ASFP_s42.01K_42.04K,ASFP_s42.01K_42.04K_42.01SM,ASFP_s42.01K_68K_42.01SM,ASFP_s42.01K_68K_58M,ASFP_s42.04K_68K_42.01SM,ASFP_s42.04K_68K_58M)	*_ASAP_ASFP_sqt-merged.sqt

STEP 5: SELECT SPECTRA MATCHING MODIFIED PEPTIDES

[DTASelect]			
input files	DTASelect.params	sequest.params	output files
*_ASAP_ASFP_sqt-merged.sqt	-t 0 -m 0 --copy -y 0 -p 1 -Sm 5 -2 2 -3 3 -d 0.08 -a false --mods	dm_standard # 14.0157 KR dm_standard @ 15.9949 MKDY dm_standard ~ 28.0314 RK dm_standard ! 42.0470 K dm_standard * 42.0106 MKSTAV dm_standard * 79.9663 STY dm_standard % 114.0429 K dm_standard $ 58.0055 M dm_standard ? 68.0261 K dm_standard + 27.9949 KR	Subset *.sqt Subset *.ms2 (only peak lists for spectra matching modified peptides)

STEP 6: SEARCH SPECTRA MATCHING MODIFIED PEPTIDES AGAINST COMPLETE PROTEIN DATABASE WITH VARIOUS DIFFERENTIAL SEARCH OPTIONS

[SEQUEST]																					
input files	directory_name	database_name	ppm_peptide_mass_tolerance	isotopes	fragment_ion_tolerance	enzyme_number	diff_search_options =						max_num_differential_AA_per_mod	mass_type_parent	mass_type_fragment	static mods			output files
Subset *.sqt Subset *.ms2 (only peak lists for spectra matching modified peptides)	MSAP_14KR_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	14.0157	KR	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_14KR_16M.sqt
	MSAP_14KR_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	0.00	X	15.9949	KDY	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_16KDY.sqt
	MSAP_27KR_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	27.9949	KR	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_27KR_16M.sqt
	MSAP_28KR_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	28.0314	KR	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_28KR_16M.sqt
	MSAP_42.01AV_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	AV	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01AV_16M.sqt
	MSAP_42.01AV_42.04K	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	AV	42.047	K	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01AV_42.04K.sqt
	MSAP_42.01AVK_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	AVK	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01AVK_16M.sqt
	MSAP_42.01KST_14R	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	14.0157	R	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01KST_14R.sqt
	MSAP_42.01KST_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01KST_16M.sqt
	MSAP_42.01KST_28R	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	28.0314	R	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01KST_28R.sqt
	MSAP_42.01KST_58M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	KST	58.0055	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01KST_58M.sqt
	MSAP_42.01MK	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	MK	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01MK.sqt
	MSAP_42.01ST_42.04K	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.0106	ST	42.047	K	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.01ST_42.04K.sqt
	MSAP_42.04K_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	42.047	K	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_42.04K_16M.sqt
	MSAP_68K_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	68.0261	K	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_68K_16M.sqt
	MSAP_68K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	68.0261	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_68K_42.01SM.sqt
	MSAP_80STY_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	79.9663	STY	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_80STY_16M.sqt
	MSAP_114K_16M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	114.0429	K	15.9949	M	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146			*_MSAP_114K_16M.sqt
	MSAP_s14K_42.01K	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	27.9949	K	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157		*_MSAP_s14K_42.01K.sqt
	MSAP_s14K_42.01K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	27.9949	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157		*_MSAP_s14K_42.01K_42.01SM.sqt
	MSAP_s14K_68K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	54.0104	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157		*_MSAP_s14K_68K_42.01SM.sqt
	MSAP_s14K_68K_58M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	54.0104	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 14.0157		*_MSAP_s14K_68K_58M.sqt
	MSAP_s28K_42.01K	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	13.9792	K	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314		*_MSAP_s28K_42.01K.sqt
	MSAP_s28K_42.01K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	13.9792	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314		*_MSAP_s28K_42.01K_42.01SM.sqt
	MSAP_s28K_68K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	39.9947	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314		*_MSAP_s28K_68K_42.01SM.sqt
	MSAP_s28K_68K_58M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	39.9947	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 28.0314		*_MSAP_s28K_68K_58M.sqt
	MSAP_s42.01K_42.04K	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	0.0364	K	0.00	X	0.00	X	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106		*_MSAP_s42.01K_42.04K.sqt
	MSAP_s42.01K_42.04K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	0.0364	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106		*_MSAP_s42.01K_42.04K_42.01SM.sqt
	MSAP_s42.01K_68K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	26.0155	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106		*_MSAP_s42.01K_68K_42.01SM.sqt
	MSAP_s42.01K_68K_58M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	26.0155	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0106		*_MSAP_s42.01K_68K_58M.sqt
	MSAP_s42.04K_68K_42.01SM	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	25.9791	K	0.00	X	42.0106	SM	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0470		*_MSAP_s42.04K_68K_42.01SM.sqt
	MSAP_s42.04K_68K_58M	Pf_PlasmoDB_5.5_Hs_NCBI_2008-03-04_NR_wSHUFFLED_wNIP.fasta	10	1 (+search isotopes)	0	0 (=No_Enzyme)	25.9791	K	0.00	X	58.0055	M	6	1 (=monoisotopic masses)	1 (=monoisotopic masses)	add_C_Cysteine = 57.02146	add_K_Lysine = 42.0470		*_MSAP_s42.04K_68K_58M.sqt

STEP 7: INSERT APPROPRIATE PTMs SYMBOLS AFTER STATICALLY-SEARCHED RESIDUES

[STQMASSFixer]		
input files	command line	output files
*.MSAP_s14K_42.01K.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(#)(14.0157)	*.MSAP_14K_42.01K.sqt
*.MSAP_s14K_42.01K_42.01SM.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(#)(14.0157)	*.MSAP_14K_42.01K_42.01SM.sqt
*.MSAP_s14K_68K_42.01SM.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(#)(14.0157)	*.MSAP_14K_68K_42.01SM.sqt
*.MSAP_s14K_68K_58M.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(#)(14.0157)	*.MSAP_14K_68K_58M.sqt
*.MSAP_s28K_42.01K.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(+)(28.0314)	*.MSAP_28K_42.01K.sqt
*.MSAP_s28K_42.01K_42.01SM.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(+)(28.0314)	*.MSAP_28K_42.01K_42.01SM.sqt
*.MSAP_s28K_68K_42.01SM.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(+)(28.0314)	*.MSAP_28K_68K_42.01SM.sqt
*.MSAP_s28K_68K_58M.sqt	./SQTMassFixer sdparam/file=/.../asequest.params combinestaticptms=K(+)(28.0314)	*.MSAP_28K_68K_58M.sqt
*.MSAP_s42.01K_42.04K.sqt	./SQTMassFixer sdparam/file=/.../asequest.params "combinestaticptms=K(*)(42.0106)	*.MSAP_42.01K_42.04K.sqt
*.MSAP_s42.01K_42.04K_42.01SM.sqt	./SQTMassFixer sdparam/file=/.../asequest.params "combinestaticptms=K(*)(42.0106)	*.MSAP_42.01K_42.04K_42.01SM.sqt
*.MSAP_s42.01K_68K_42.01SM.sqt	./SQTMassFixer sdparam/file=/.../asequest.params "combinestaticptms=K(*)(42.0106)	*.MSAP_42.01K_68K_42.01SM.sqt
*.MSAP_s42.01K_68K_58M.sqt	./SQTMassFixer sdparam/file=/.../asequest.params "combinestaticptms=K(*)(42.0106)	*.MSAP_42.01K_68K_58M.sqt
*.MSAP_s42.04K_68K_42.01SM.sqt	./SQTMassFixer sdparam/file=/.../asequest.params "combinestaticptms=K(*)(42.0470)	*.MSAP_42.04K_68K_42.01SM.sqt
*.MSAP_s42.04K_68K_58M.sqt	./SQTMassFixer sdparam/file=/.../asequest.params "combinestaticptms=K(*)(42.0470)	*.MSAP_42.04K_68K_58M.sqt

STEP 8: MERGE ALL DIFFERENTIALLY-SEARCHED SQT FILES, RANK BASED ON XCorr AND RECALCULATE DeltaCN

[sqt-merge]		
input files	command line	output files
*.ASAP_ASFP_sqt-merged.sqt	sqt-merge --output=./PfHist_(ASAP_ASFP_sqt-merged,MSAP_14KR_16M,MSAP_16KDY,MSAP_27KR_16M,MSAP_28KR_16M,MSAP_42.01AV_16M,MSAP_42.01AV_42.04K,MSAP_42.01AVK_16M,MSAP_42.01KST_14R,MSAP_42.01KST_16M,MSAP_42.01KST_28R,MSAP_42.01KST_58M,MSAP_42.01M,MSAP_42.04K_16M,MSAP_68K_16M,MSAP_68K_42.01SM,MSAP_80STY_16M,MSAP_114K_16M,MSAP_s14K_42.01K,MSAP_s14K_42.01K_42.01SM,MSAP_s28K_68K_42.01SM,MSAP_s14K_68K_42.01SM,MSAP_s14K_68K_58M,MSAP_s28K_42.01K,MSAP_s28K_42.01K_42.01SM,MSAP_s42.01K_42.04K,MSAP_s42.01K_68K_42.01SM,MSAP_s42.01K_68K_58M,MSAP_s42.04K_68K_42.01SM,MSAP_s42.04K_68K_58M)	*.sqt
*.MSAP_14KR_16M.sqt		
*.MSAP_16KDY.sqt		
*.MSAP_27KR_16M.sqt		
*.MSAP_28KR_16M.sqt		
*.MSAP_42.01AV_16M.sqt		
*.MSAP_42.01AV_42.04K.sqt		
*.MSAP_42.01AVK_16M.sqt		
*.MSAP_42.01KST_14R.sqt		
*.MSAP_42.01KST_16M.sqt		
*.MSAP_42.01KST_28R.sqt		
*.MSAP_42.01KST_58M.sqt		
*.MSAP_42.01MK.sqt		
*.MSAP_42.01ST_42.04K.sqt		
*.MSAP_42.04K_16M.sqt		
*.MSAP_68K_16M.sqt		
*.MSAP_80STY_16M.sqt		
*.MSAP_114K_16M.sqt		
*.MSAP_14K_42.01K.sqt		
*.MSAP_14K_42.01K_42.01SM.sqt		
*.MSAP_14K_42.01SM.sqt		
*.MSAP_14K_68K_42.01SM.sqt		
*.MSAP_14K_68K_58M.sqt		
*.MSAP_28K_42.01K.sqt		
*.MSAP_28K_42.01K_42.01SM.sqt		
*.MSAP_28K_68K_42.01SM.sqt		
*.MSAP_28K_68K_58M.sqt		
*.MSAP_42.01K_42.04K.sqt		
*.MSAP_42.01K_42.04K_42.01SM.sqt		
*.MSAP_42.01K_68K_42.01SM.sqt		
*.MSAP_42.01K_68K_58M.sqt		
*.MSAP_42.04K_68K_42.01SM.sqt		
*.MSAP_42.04K_68K_58M.sqt		

STEP 9: CORRECT CALCULATED M+H+ TO OBTAIN THE VALUE CLOSEST TO THE OBSERVED MONOISOTOPIC MASS M+H+

[STQMASSFixer]		
input files	command line	output files
*.sqt *.ms2	./SQTMassFixer sdparam/file=/.../asequest.params	*.sqt *.ms2

STEP 10: ASSEMBLE PEPTIDE INFORMATION INTO PROTEIN INFORMATION AND VISUALLY VALIDATE SPECTRA MAPPED TO MODIFIED PEPTIDES

[DTASelect]			
input files	DTASelect.params	sequest.params	output files
*.sqt *.ms2	-t 0 --copy -y 0 -p 1 -Smn 5 -2 2 -3 3 -d 0.08 -a false --mods	dm_standard # 14.0157 KR dm_standard @ 15.9949 MKDY dm_standard ~ 28.0314 RK dm_standard ! 42.0470 K dm_standard ^ 42.0106 MKSTAV dm_standard * 79.9663 STY dm_standard % 114.0429 K dm_standard $ 58.0055 M dm_standard 7 68.0261 K dm_standard + 27.9949 KR	DTASelect.bt DTASelect-mods.html

STEP 11: INSERT Y/M/N VISUAL VALIDATION FLAGS FOR SPECTRA MATCHING MODIFIED PEPTIDES

[DTASelectFixer]		
input files	command line	output files
*.sqt *.ms2 DTASelect.txt	/DTASelectFixer_Linux_GCC dtaselectfolder=/.../ spectralflagfile=flag-table.txt thread=4 allow_one_y_spectralflag_per_spectrum =1	*.sqt *.ms2 DTASelect.txt

STEP 12: SORT AND SELECT PEPTIDE SPECTRUM MATCHES PASSING CRITERIA, AND COMPARE MULTIPLE ANALYSES

[DTASelect/CONTRAST]			
input files	Contrast.params		output files
DTASelect.txt	[Criteria Sets] Mp2r2Smn7y0a -p 2 -r 2 -Smn 7 -y 0 -d 0.08 -2 2.0 -3 3.0 -a false --DB --mods -v 2 Mp3r100Smn7y0ao -p 3 -r 100 -Smn 7 -y 0 -d 0.08 -2 2.0 -3 3.0 -a false -o --DB --mods -v 2	[Explicit Mappings] 0-hpi_1 Mp2r2Smn7y0a 0-hpi_2 Mp2r2Smn7y0a 6-hpi_1 Mp2r2Smn7y0a 6-hpi_2 Mp2r2Smn7y0a 12-hpi_1 Mp2r2Smn7y0a 12-hpi_2 Mp2r2Smn7y0a 18-hpi_1 Mp2r2Smn7y0a 18-hpi_2 Mp2r2Smn7y0a 24-hpi_1 Mp2r2Smn7y0a 24-hpi_2 Mp2r2Smn7y0a 30-hpi_1 Mp2r2Smn7y0a 30-hpi_2 Mp2r2Smn7y0a 36-hpi_1 Mp2r2Smn7y0a 36-hpi_2 Mp2r2Smn7y0a ALL Mp3r100Smn7y0ao [Options] Master ALL Mp3r100Smn7y0ao	Contrast.bt Contrast.html

STEP 13: CREATE REPORTS ON ALL DETECTED PEPTIDES AND NON-REDUNDANT PROTEINS AND CALCULATE dNSAFs AND FDRs

[NSAF7]			
input files	NSAF.params		output files
Contrast.txt *.sqt *.ms2 DTASelect.txt	[Included Experiments] 0-hpi_1 Mp2r2Smn7y0a 0-hpi 0-hpi_2 Mp2r2Smn7y0a 0-hpi 6-hpi_1 Mp2r2Smn7y0a 6-hpi 6-hpi_2 Mp2r2Smn7y0a 6-hpi 12-hpi_1 Mp2r2Smn7y0a 12-hpi 12-hpi_2 Mp2r2Smn7y0a 12-hpi 18-hpi_1 Mp2r2Smn7y0a 18-hpi 18-hpi_2 Mp2r2Smn7y0a 18-hpi 24-hpi_1 Mp2r2Smn7y0a 24-hpi 24-hpi_2 Mp2r2Smn7y0a 24-hpi 30-hpi_1 Mp2r2Smn7y0a 30-hpi 30-hpi_2 Mp2r2Smn7y0a 30-hpi 36-hpi_1 Mp2r2Smn7y0a 36-hpi 36-hpi_2 Mp2r2Smn7y0a 36-hpi ALL Mp3r100Smn7y0ao TYPE_A	[NSAF Options] Correct Spc Inconsistence #Check Windows Character Compatibleness Check EXCEL Cell Text length Limit Check indistinguishable proteins in single experiment Export U_DETECTED_PEPTIDE_BY_SEQUENCE Calculate Protein Group Info #Remove SAF Normalization Step #Collect SIMR Protein Attributes NCBI TMHMM SIGNALP Process Data With Local Copy Export Details From Peptide Level Export Details From Spectral Level Annotate PTM Spectra Except PTM Symbols @ Annotate PTM Spectra With Protein Subset psu|PF11_0081 psu|PFF0510w psu|PFF0865w psu|PF13_0185 psu|PF11_0062 psu|PF07_0054 psu|PFF0860c psu|PFC0920w Annotate PTM Spectra Offline	Contrast.NSAF_v7.xls

STEP 14: EXTRACT TOTAL AND MODIFIED SPECTRAL COUNTS FOR EACH AMINO ACID WITHIN PfHISTONES AND CALCULATE MODIFICATION LEVELS

[PTMR]			
input files	PfHistones.txt	sequest.params	output files
Contrast.txt *-mods.txt PfHistones.txt *.sqt *.ms2	psu|PF11_0061 psu|PFF0510w psu|PFF0865w psu|PF13_0185 psu|PF11_0062 psu|PF07_0054 psu|PFF0860c psu|PFC0920w	dm_standard # 14.0157 KR dm_standard @ 15.9949 MKDY dm_standard ~ 28.0314 RK dm_standard ! 42.0470 K dm_standard ^ 42.0106 MKSTAV dm_standard * 79.9663 STY dm_standard % 114.0429 K dm_standard $ 58.0055 M dm_standard 7 68.0261 K dm_standard + 27.9949 KR	Contrast.PTMR.xls

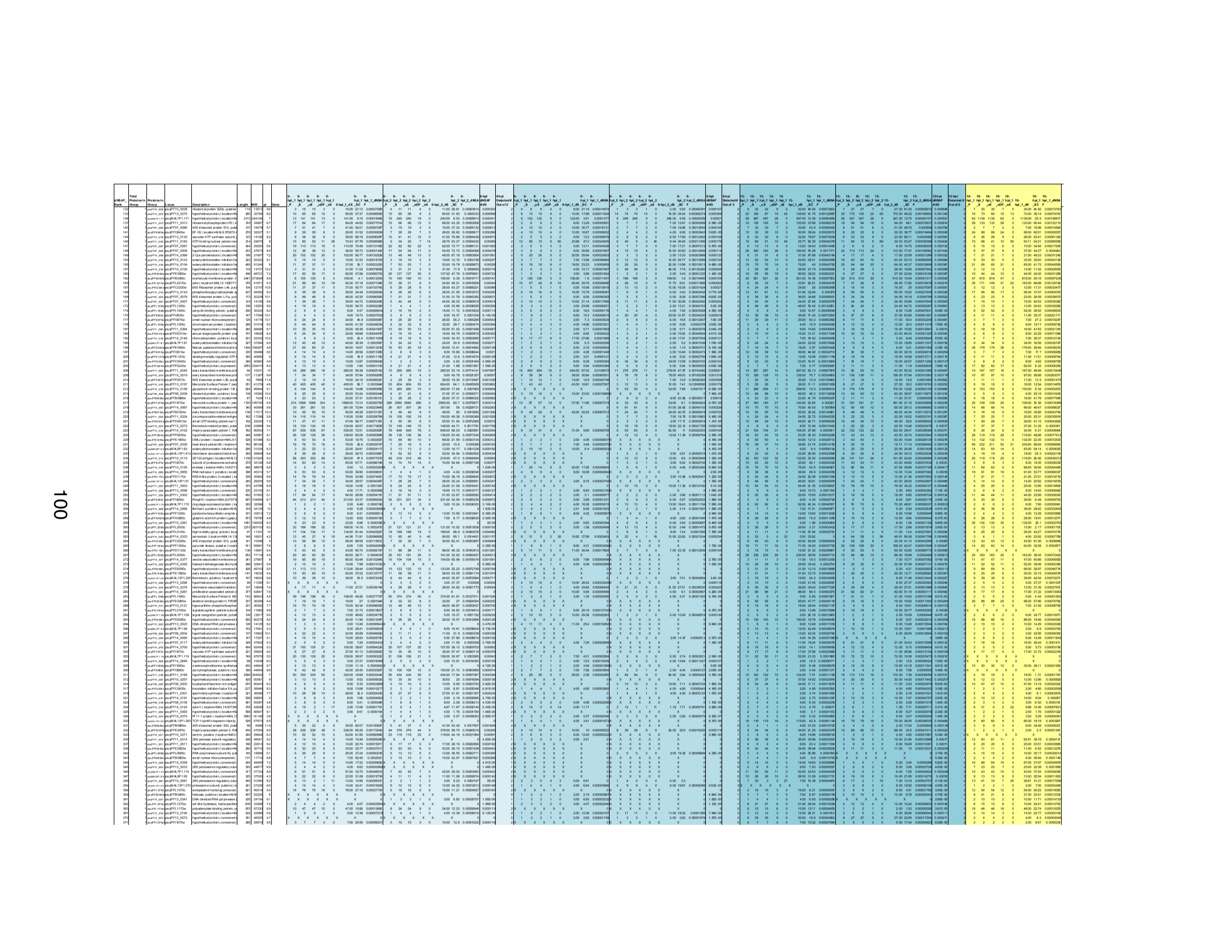
Table S2. Detailed peptide and spectral counts for proteins detected by MudPIT analyses of Plasmodium falciparum histones preparations isolated from 7 erythrocytic stages.

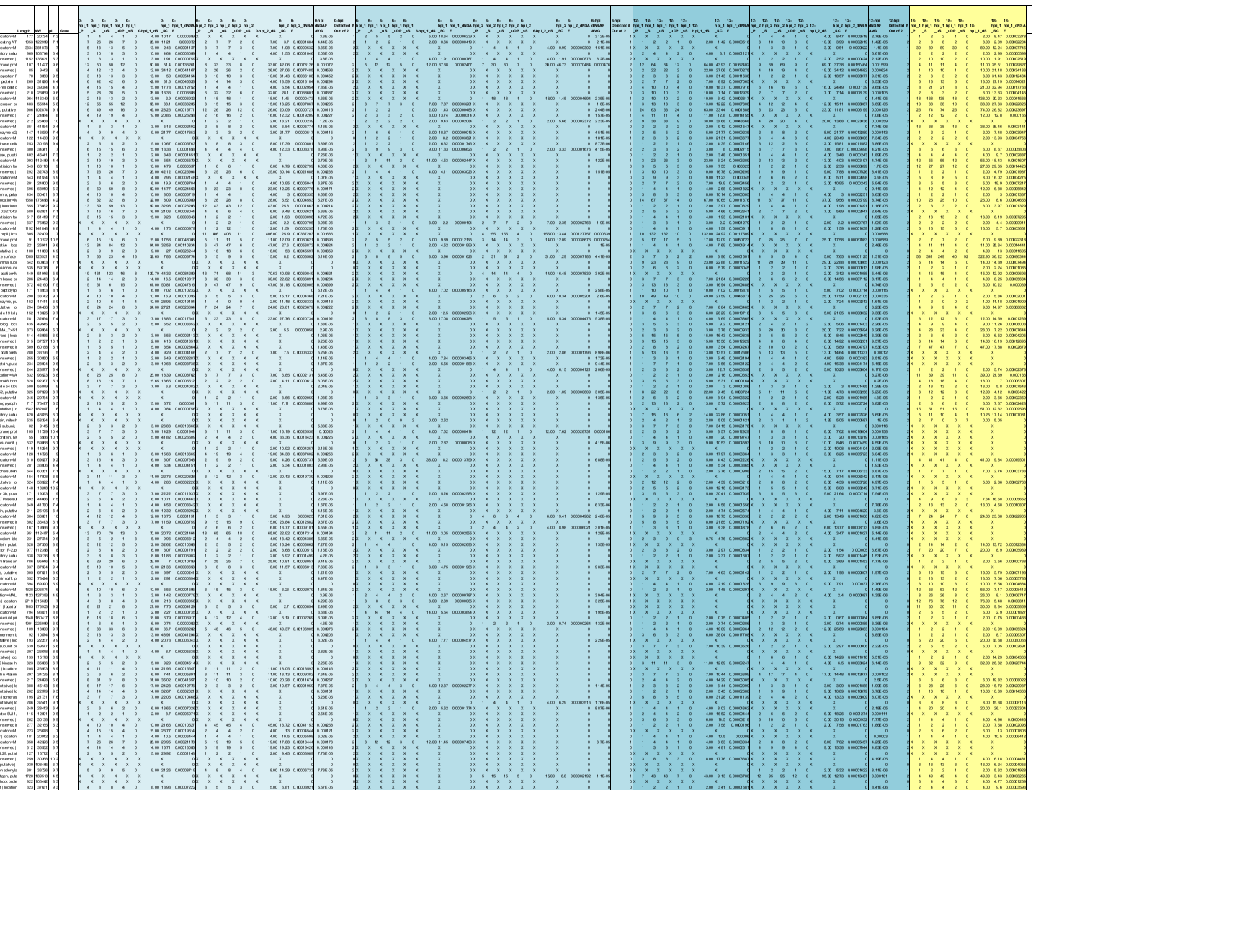
In each analysis (See details in Table S1), the following parameters are specified for each protein:

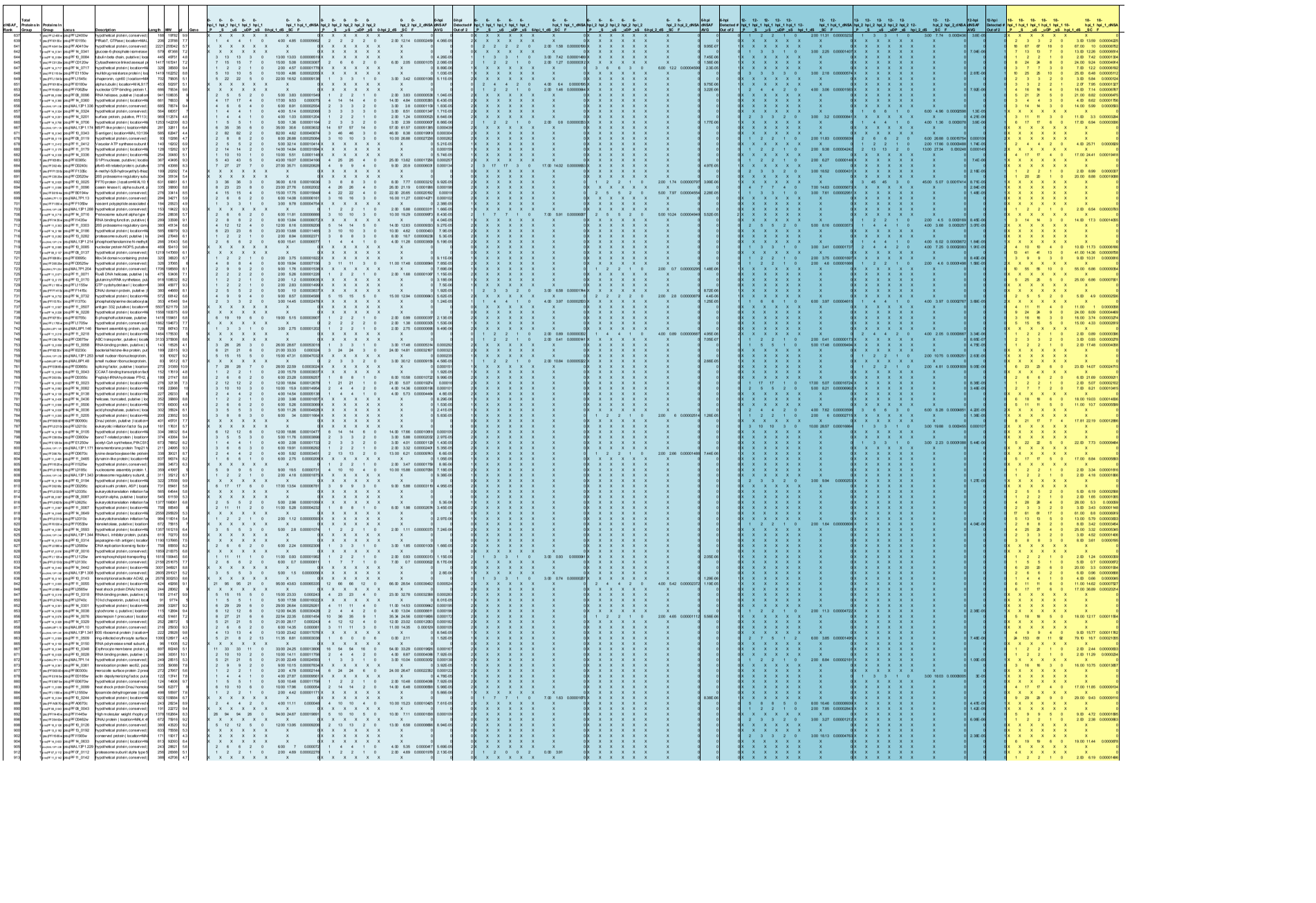
- Peptide counts (columns end with _P);
- Total spectral counts (_S);
- Spectral counts for peptides shared between protein isoforms (_sS);
- Peptide counts for peptides uniquely matching the protein (_uDP);
- Spectral counts for peptides uniquely matching the protein (_uS);
- Spectral counts distributed according to the spectral count contribution of peptides unique to each isoform (_dS);
- Sequence coverage as a percentage (_SC);
- Distributed Normalized Spectral Abundance Factors (_dNSAF), calculated based on distributed spectral counts and protein length.

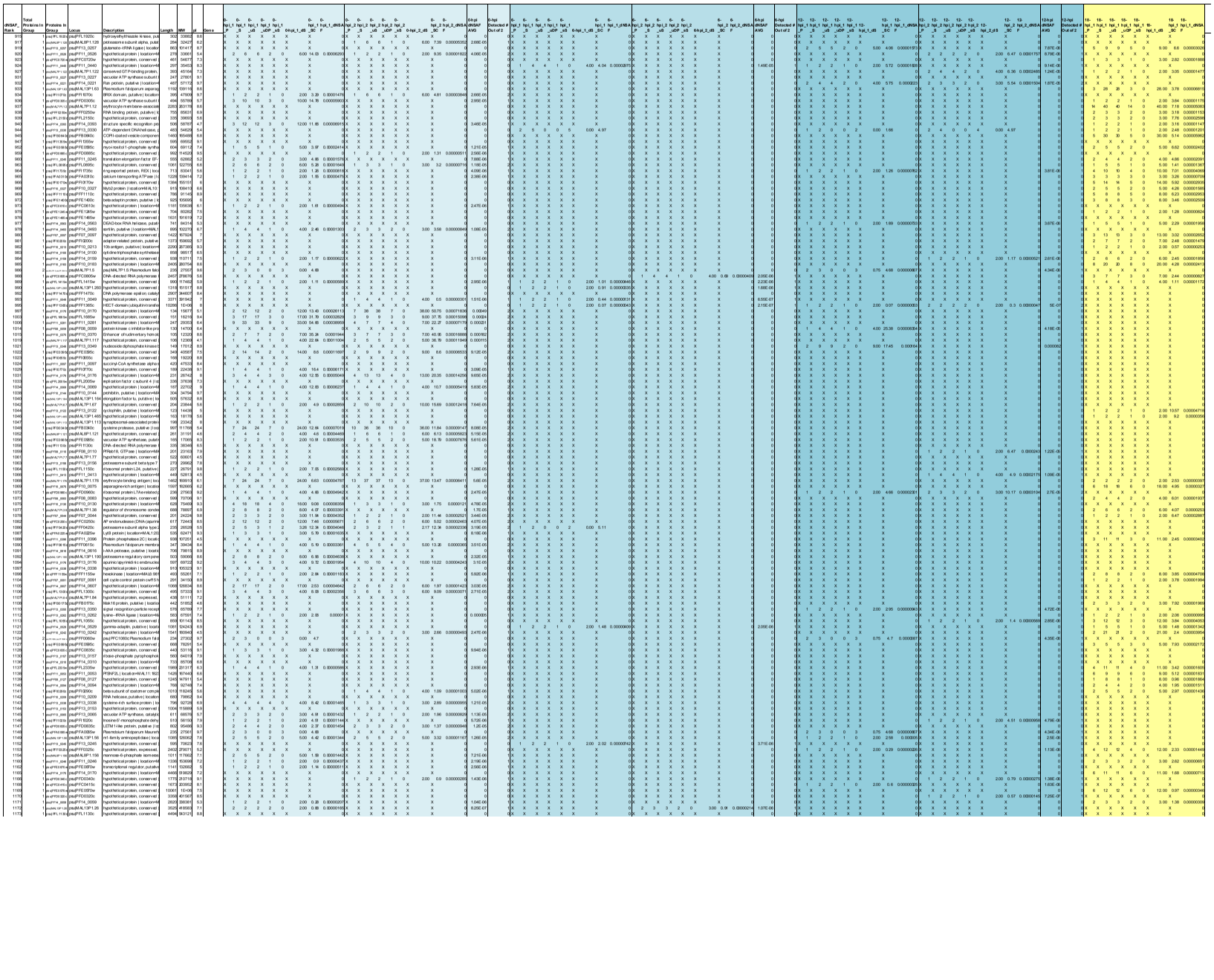
The number of times a protein is detected across replicates is listed in the "Detected # Out of" columns.
dNSAF are averaged across replicate runs ("dNSAF AVG" columns).

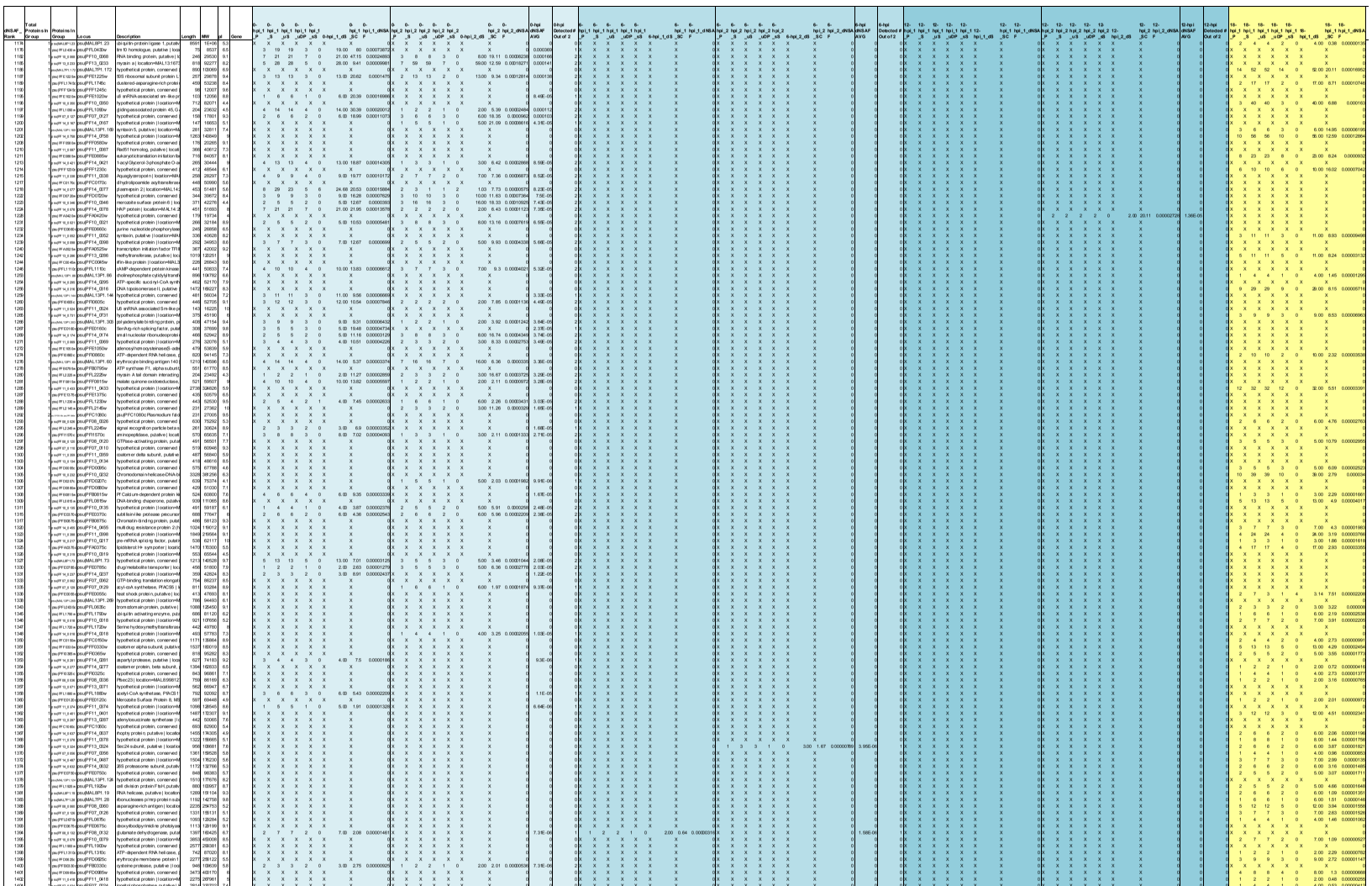
dNSAF_Rank	Total Proteins In Group	Proteins In Group	Locus	Description	Length	MW	pI	Gene	0-hpi_1_P	0-hpi_1_S	0-hpi_1_uS	0-hpi_1_uDP	0-hpi_1_sS	0-hpi_1_dS	0-hpi_1_SC	0-hpi_1_dNSAF
1	1	psu|PF11_0061	psu|PF11_0061	histone H4, putative | location=M	102	11324	11.2	H4	109	2825	2520	86	305	2803.09	94.12	0.08013517
2	1	psu|PF11_0062	psu|PF11_0062	histone H2B | location=MAL11:2	116	12994	10.3	H2B	159	5686	5369	148	317	5618.08	97.41	0.14122619
4	1	psu|PFF0510w	psu|PFF0510w	histone H3, putative | location=M	135	15315	11.1	H3	126	3614	2427	74	1187	3245.99	89.63	0.070113
6	1	psu|PFF0865w	psu|PFF0865w	histone h3 | location=MAL6:7543	135	15310	11.2	H3.3	97	2961	2099	37	862	2549.75	89.63	0.05507433
7	1	psu|PF07_0054	psu|PF07_0054	histone h2b, putative | location=	122	13632	10.8	H2Bv	98	1902	1592	89	310	1672.92	93.44	0.03998547
8	1	psu|PFC0920w	psu|PFC0920w	histone H2A variant, putative | lo	158	16453	10.6	H2Az	79	1500	1269	69	231	1437.76	91.77	0.02653469
10	1	psu|PFF0860c	psu|PFF0860c	histone h2a | location=MAL6:748	131	13991	10.3	H2A	56	1208	749	46	459	1102.08	92.37	0.02453177
89	1	psu|PF13_0185	psu|PF13_0185	histone h3, putative | location=M	169	19497	10.7	H3.c	32	193	193	32	0	193.00	89.35	0.00333009
14	1	psu|PF13_0268	psu|PF13_0268	ribosomal protein L17, putative |	203	23415	10.9		42	566	566	42	0	566.00	89.16	0.00813029
15	1	psu|PF11_0272	psu|PF11_0272	ribosomal protein S18, putative |	156	17891	10.5		30	296	284	29	12	296.00	76.28	0.0055329
16	1	psu|PFE1005w	psu|PFE1005w	40S ribosomal subunit protein S	189	22123	10.4		38	359	359	38	0	359.00	82.54	0.00553883
17	1	psu|PF13_0213	psu|PF13_0213	60S ribosomal subunit protein L6	221	25532	10.1		57	1027	1023	57	4	1023.00	83.26	0.01349799
18	1	psu|PF08_0076	psu|PF08_0076	40S ribosomal protein S16, putat	144	16285	10.3		30	213	213	30	0	213.00	77.78	0.00431323
19	1	psu|PF13_0228	psu|PF13_0228	40S ribosomal subunit protein S6	306	35384	10.4		46	556	556	46	0	556.00	61.11	0.00529833
21	1	psu|PFD1055w	psu|PFD1055w	ribosomal protein S19s, putative	170	19722	10.2		30	324	324	30	0	324.00	74.71	0.00555753
22	1	psu|PF13_0316	psu|PF13_0316	40S ribosomal protein S13 | loca	151	17250	10.4		28	255	255	28	0	255.00	81.46	0.00492435
23	1	psu|PF11_0312	psu|PF11_0312	ribosomal protein L38e | location	87	10314	10.7		18	140	140	18	0	140.00	80.46	0.00469239
24	1	psu|PF13_0014	psu|PF13_0014	40S ribosomal protein S7 homol	194	22481	9.8		24	177	177	24	0	177.00	80.93	0.00266046
25	1	psu|PFE0975c	psu|PFE0975c	40S ribosomal subunit protein S2	133	15391	10.8		23	191	191	23	0	191.00	86.47	0.00418762
26	1	psu|PF13_0304	psu|PF13_0304	elongation factor 1 alpha | locatio	443	48959	9.1		97	636	619	95	17	636.00	85.78	0.00418638
27	1	psu|PFF0885w	psu|PFF0885w	60S ribosomal protein L27a, puta	148	16723	10.5		28	157	149	25	8	157.00	79.05	0.00309331
28	1	psu|PF11_0106	psu|PF11_0106	60S Ribosomal protein L36, puta	130	15443	10.4		15	305	305	15	0	305.00	46.15	0.00684136
29	1	psu|MAL13P1.92	psu|MAL13P1.92	40S ribosomal protein S15, putat	147	17040	10.2		9	113	113	9	0	113.00	48.3	0.00224154
30	1	psu|PFE0810c	psu|PFE0810c	40S ribosomal subunit protein S1	151	16077	10.5		18	105	100	17	5	105.00	77.48	0.00202767
32	1	psu|PFC0735w	psu|PFC0735w	40S ribosomal protein S15A, put	130	14874	10		18	202	202	18	0	202.00	95.38	0.004531
33	1	psu|PFF0700c	psu|PFF0700c	60S ribosomal protein L19, putat	196	23409	11.1		29	301	293	28	8	301.00	62.24	0.00447812
34	1	psu|PFI1780w	psu|PFI1780w	hypothetical protein | location=M	383	45500	9.7		33	234	234	33	0	234.00	53.26	0.00178157
35	1	psu|PFB0830w	psu|PFB0830w	Ribosomal protein S26e, putative	107	12509	11		11	107	107	11	0	107.00	50.47	0.00291599
37	1	psu|PF14_0205	psu|PF14_0205	ribosomal protein S25, putative |	135	15741	9.9		20	117	117	20	0	117.00	60.74	0.00252719
38	1	psu|PFC1020c	psu|PFC1020c	40S ribosomal protein S3A, puta	262	30047	9.8		35	174	174	35	0	174.00	87.4	0.00193657
39	1	psu|PF08_0075	psu|PF08_0075	60S ribosomal protein L13, putat	215	25441	10.8		41	264	264	41	0	264.00	64.19	0.00358056
40	1	psu|PFC0535w	psu|PFC0535w	60S ribosomal protein L26, putat	126	14813	11.1		15	246	246	15	0	246.00	70.63	0.00569312
41	1	psu|PF14_0579	psu|PF14_0579	ribosomal protein L27, putative |	121	13795	10.3		23	243	243	23	0	243.00	80.99	0.00585607
42	1	psu|PFC0290w	psu|PFC0290w	40S ribosomal protein S23, putat	145	16130	10.8		17	94	94	17	0	94.00	73.1	0.00189036
43	1	psu|PF08_0054	psu|PF08_0054	heat shock 70 kDa protein | locat	677	73915	5.7		23	128	101	15	27	117.26	36.63	0.00050506
44	1	psu|PFC0400w	psu|PFC0400w	60S Acidic ribosomal protein P2	112	11948	4.5		19	227	227	19	0	227.00	90.18	0.00591008
45	1	psu|PFL2055w	psu|PFL2055w	40S ribosomal protein S17, putat	137	16067	10.4		8	29	29	8	0	29.00	49.64	0.00061725
46	1	psu|PF11_0260	psu|PF11_0260	ribosomal protein L35, putative |	105	12368	10.8		20	130	130	20	0	130.00	77.14	0.00361027
47	1	psu|PF11_0438	psu|PF11_0438	Ribosomal protein, putative | loca	140	16265	10.5		27	166	166	27	0	166.00	86.43	0.00345753
50	1	psu|PF13_0346	psu|PF13_0346	ubiquitin/ribosomal fusion protein	128	14617	9.9		14	148	9	1	139	148.00	46.88	0.00337161
51	1	psu|PFE0350c	psu|PFE0350c	60S ribosomal subunit protein L4	411	46212	10.5		44	207	207	44	0	207.00	59.12	0.00146864
52	1	psu|PF14_0083	psu|PF14_0083	ribosomal protein S8e, putative |	218	25051	10		23	184	177	22	7	184.00	70.18	0.0024612
53	1	psu|PFE0990w	psu|PFE0990w	hypothetical protein, conserved |	127	14940	10.5		19	128	128	19	0	128.00	81.1	0.00293895
55	1	psu|PF14_0627	psu|PF14_0627	ribosomal protein S3, putative | l	221	24668	10.2		16	72	72	16	0	72.00	63.35	0.00095001
56	1	psu|PF14_0240	psu|PF14_0240	ribosomal protein L21e, putative	161	18795	10		17	84	84	17	0	84.00	67.7	0.00152138
57	1	psu|PF14_0443	psu|PF14_0443	centrin, putative | location=MAL1	168	19311	4.7		33	293	293	33	0	293.00	88.1	0.00508562
58	1	psu|PF13_0141	psu|PF13_0141	L-lactate dehydrogenase | locatio	316	34108	7.5		16	129	129	16	0	129.00	49.68	0.00119039
62	1	psu|PF14_0231	psu|PF14_0231	ribosomal protein L7a, putative |	282	32580	10.2		54	374	374	54	0	374.00	84.4	0.0038673
64	1	psu|PF14_0598	psu|PF14_0598	glyceraldehyde-3-phosphate deh	337	36635	7.7		23	104	104	23	0	104.00	56.38	0.00089989
65	1	psu|PFB0455w	psu|PFB0455w	ribosomal L37ae protein, putative	96	10799	10.5		11	84	84	11	0	84.00	70.83	0.00255149
66	1	psu|MAL7P1.300	psu|MAL7P1.300	hypothetical protein, conserved |	53	6495	9.9		5	28	28	5	0	28.00	50.94	0.00154052
67	1	psu|PF13_0132	psu|PF13_0132	60S ribosomal protein L23a, puta	190	22093	10.3		32	147	147	32	0	147.00	74.74	0.00225605
70	1	psu|PFC0775w	psu|PFC0775w	40S ribosomal protein S11, putat	161	18760	10.2		14	40	36	13	4	40.00	52.8	0.00072447
72	1	psu|PF10_0063	psu|PF10_0063	DNA/RNA-binding protein, putati	107	11977	9.3		17	156	156	17	0	156.00	82.24	0.00425135
73	1	psu|PFE0845c	psu|PFE0845c	60S ribosomal subunit protein L8	260	28000	10.5		30	184	178	29	6	183.97	69.23	0.00206325
74	1	psu|PFI0875w	psu|PFI0875w	Heat shock protein | location=MA	652	72388	5.3		67	378	376	66	2	378.00	84.2	0.00169056
76	1	psu|PFI0190w	psu|PFI0190w	ribosomal protein L32, putative |	131	15527	11.1		14	87	85	13	2	87.00	48.85	0.00193657
77	1	psu|PFI1105w	psu|PFI1105w	Phosphoglycerate kinase | locatio	416	45427	7.8		13	42	42	13	0	42.00	38.7	0.0002944
79	1	psu|PF13_0224	psu|PF13_0224	60S ribosomal subunit protein L1	184	21746	10.6		25	116	116	25	0	116.00	89.67	0.00183834
80	1	psu|PF10_0272	psu|PF10_0272	ribosomal protein L3, putative | l	386	44221	10.2		35	186	173	33	13	186.00	64.51	0.00140511
81	1	psu|PF13_0171	psu|PF13_0171	60S ribosomal protein L23, putat	139	14993	9.9		5	14	14	5	0	14.00	48.92	0.0002937
82	1	psu|PF11_0065	psu|PF11_0065	ribosomal protein S4, putative | l	282	32354	10		25	153	153	25	0	153.00	56.38	0.00158208
83	1	psu|MAL7P1.228	psu|MAL7P1.228	Heat Shock 70 KDa Protein, (HS	660	73232	5.7		26	108	84	19	24	95.51	50.91	0.00042196
84	1	psu|MAL7P1.320	psu|MAL7P1.320	Ribosomal protein, L37e, putativ	92	10539	11.5		11	61	61	11	0	61.00	65.22	0.00193343
85	1	psu|PF13_0049	psu|PF13_0049	60S ribosomal protein L24, putat	162	19245	10.4		7	16	16	7	0	16.00	38.27	0.000288
86	1	psu|PF10_0187	psu|PF10_0187	ribosomal protein L30e, putative	108	11775	9.4		14	98	98	14	0	98.00	93.52	0.00264599
87	1	psu|PF10_0323	psu|PF10_0323	early transcribed membrane prot	355	38926	9.4		26	217	217	26	0	217.00	62.54	0.00178245
88	1	psu|MAL8P1.69	psu|MAL8P1.69	14-3-3 protein homologue, putati	256	29476	5		36	300	295	35	5	299.06	73.83	0.0034065
90	1	psu|PF13_0129	psu|PF13_0129	ribosomal protein L6 homologue,	190	21601	9.9		11	49	49	11	0	49.00	61.58	0.00075202
91	1	psu|PF14_0296	psu|PF14_0296	ribosomal protein L14, putative |	165	19297	10.2		26	169	169	26	0	169.00	71.52	0.00298668
92	1	psu|PF07_0043	psu|PF07_0043	60S ribosomal protein L34-a, put	150	17351	10.8		14	60	60	14	0	60.00	52.67	0.0011664
93	1	psu|PF11_0043	psu|PF11_0043	60S acidic ribosomal protein p1,	118	13014	4.6		11	139	139	11	0	139.00	88.98	0.00343493
94	1	psu|PFC0870w	psu|PFC0870w	elongation factor 1 (EF-1), putati	156	17706	4.6		10	62	62	10	0	62.00	63.46	0.00115892
95	1	psu|MAL13P1.209	psu|MAL13P1.209	60S ribosomal subunit porotein L	188	21617	11		15	108	108	15	0	108.00	55.85	0.00167514
98	1	psu|PFE0850c	psu|PFE0850c	60S ribosomal protein L12, putat	165	18112	9.5		17	99	99	17	0	99.00	90.3	0.00174959
99	1	psu|PF14_0678	psu|PF14_0678	exported protein 2 | location=MA	287	33412	5.3		33	343	343	33	0	343.00	57.84	0.00348496
100	1	psu|PFB0490c	psu|PFB0490c	hypothetical protein | location=M	284	32542	11.1		30	181	181	30	0	181.00	68.31	0.00185843
101	1	psu|PF11_0437	psu|PF11_0437	hypothetical protein | location=M	127	14407	10.2		15	44	44	15	0	44.00	76.38	0.00101026
102	1	psu|PF10_0086	psu|PF10_0086	adenylate kinase, putative | locat	242	27611	8.9		41	333	333	41	0	333.00	80.58	0.0040125
103	1	psu|PF14_0027	psu|PF14_0027	ribosomal S27a, putative | locatio	149	17400	9.5		11	101	99	10	2	101.00	30.87	0.00197661
105	1	psu|MAL8P1.72	psu|MAL8P1.72	high mobility group protein, putat	99	11535	10		21	120	120	21	0	120.00	77.78	0.00353453
106	1	psu|PF08_0074	psu|PF08_0074	DNA/RNA-binding protein Alba, p	248	27259	10.6		33	283	283	33	0	283.00	53.63	0.00332752
107	1	psu|PF14_0391	psu|PF14_0391	ribosomal protein L1, putative | lo	217	24805	9.8		24	126	119	22	7	126.00	81.57	0.00169315
108	1	psu|PFC0300c	psu|PFC0300c	60S ribosomal protein L7, putativ	257	30523	10.3		18	111	111	18	0	111.00	54.86	0.00125943
109	1	psu|PF11_0351	psu|PF11_0351	heat shock protein hsp70 homolo	663	73298	6.9		20	84	83	19	1	83.99	40.72	0.00036939
110	1	psu|PF14_0141	psu|PF14_0141	ribosomal protein L10, putative |	219	25216	10		11	51	51	11	0	51.00	53.88	0.00067907
111	1	psu|PF10_0043	psu|PF10_0043	ribosomal protein L13, putative |	202	23754	10.2		15	78	78	15	0	78.00	55.45	0.00112598
112	1	psu|PF14_0448	psu|PF14_0448	ribosomal protein S2, putative | l	272	29958	10		19	82	82	19	0	82.00	59.56	0.00087908
113	1	psu|PF14_0585	psu|PF14_0585	ribosomal protein S28e, putative	67	7494	10.9		1	6	6	1	0	6.00	28.36	0.00026113
114	1	psu|PF11_0250	psu|PF11_0250	high mobility group-like protein N	145	16012	5.2		9	60	60	9	0	60.00	71.72	0.00120662
115	1	psu|PF14_0230	psu|PF14_0230	Ribosomal protein family L5, puta	294	33998	9.8		12	57	57	12	0	57.00	48.3	0.00056534
118	1	psu|PF10_0372	psu|PF10_0372	Antigen UB05 | location=MAL10:	119	13712	9.4		9	72	72	9	0	72.00	40.34	0.0017643
119	1	psu|PF10_0268	psu|PF10_0268	merozoite capping protein 1 | loc	393	43930	9.7		35	194	194	35	0	194.00	45.8	0.00143944
121	1	psu|PFE0080c	psu|PFE0080c	rhoptry-associated protein 2, RA	398	46739	8.8		54	510	510	54	0	510.00	73.12	0.00373657
122	1	psu|PFE0185c	psu|PFE0185c	60S ribosomal subunit protein L3	120	14065	10.1		21	69	69	21	0	69.00	81.67	0.00167669
123	1	psu|PFA0345w	psu|PFA0345w	centrin, putative | location=MAL1	168	19599	4.9		19	89	71	13	18	89.00	86.31	0.00154478
125	1	psu|MAL8P1.52	psu|MAL8P1.52	hypothetical protein, conserved |	124	14060	9.7		9	64	64	9	0	64.00	53.23	0.00150503
126	1	psu|PF11_0266	psu|PF11_0266	small nuclear ribonucleoprotein D	126	14337	10.2		4	41	41	4	0	41.00	44.44	0.00094885
129	1	psu|PF07_0088	psu|PF07_0088	40S ribosomal protein S5, putativ	195	21863	9.6		5	10	10	5	0	10.00	27.18	0.00014954
130	1	psu|PFD0770c	psu|PFD0770c	ribosomal protein l15, putative | l	220	25923	11.3		16	77	77	16	0	77.00	53.64	0.0010206
131	1	psu|PFB0100c	psu|PFB0100c	knob associated histidine-rich pr	654	71303	9.1		21	68	68	21	0	68.00	27.52	0.00030319

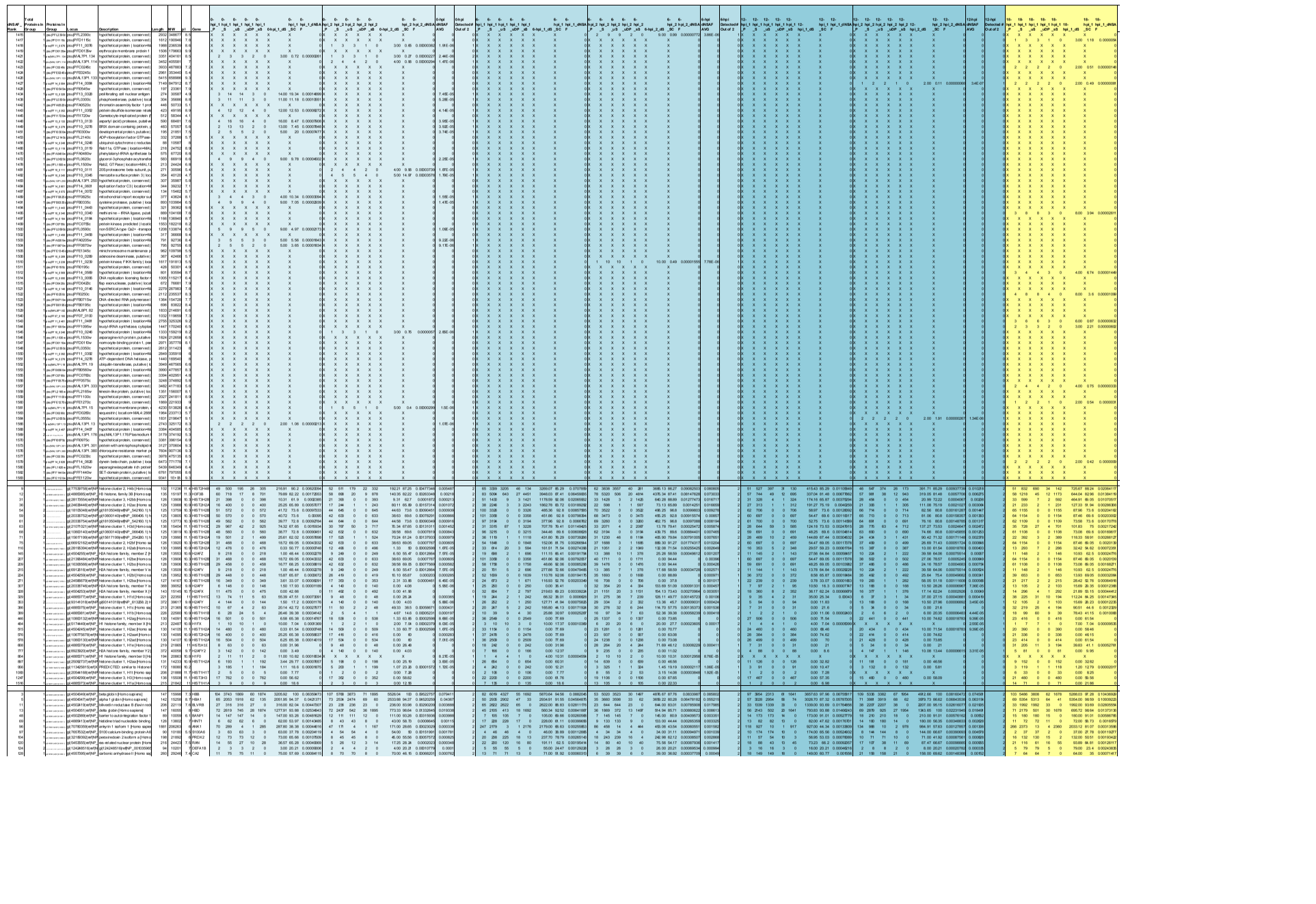


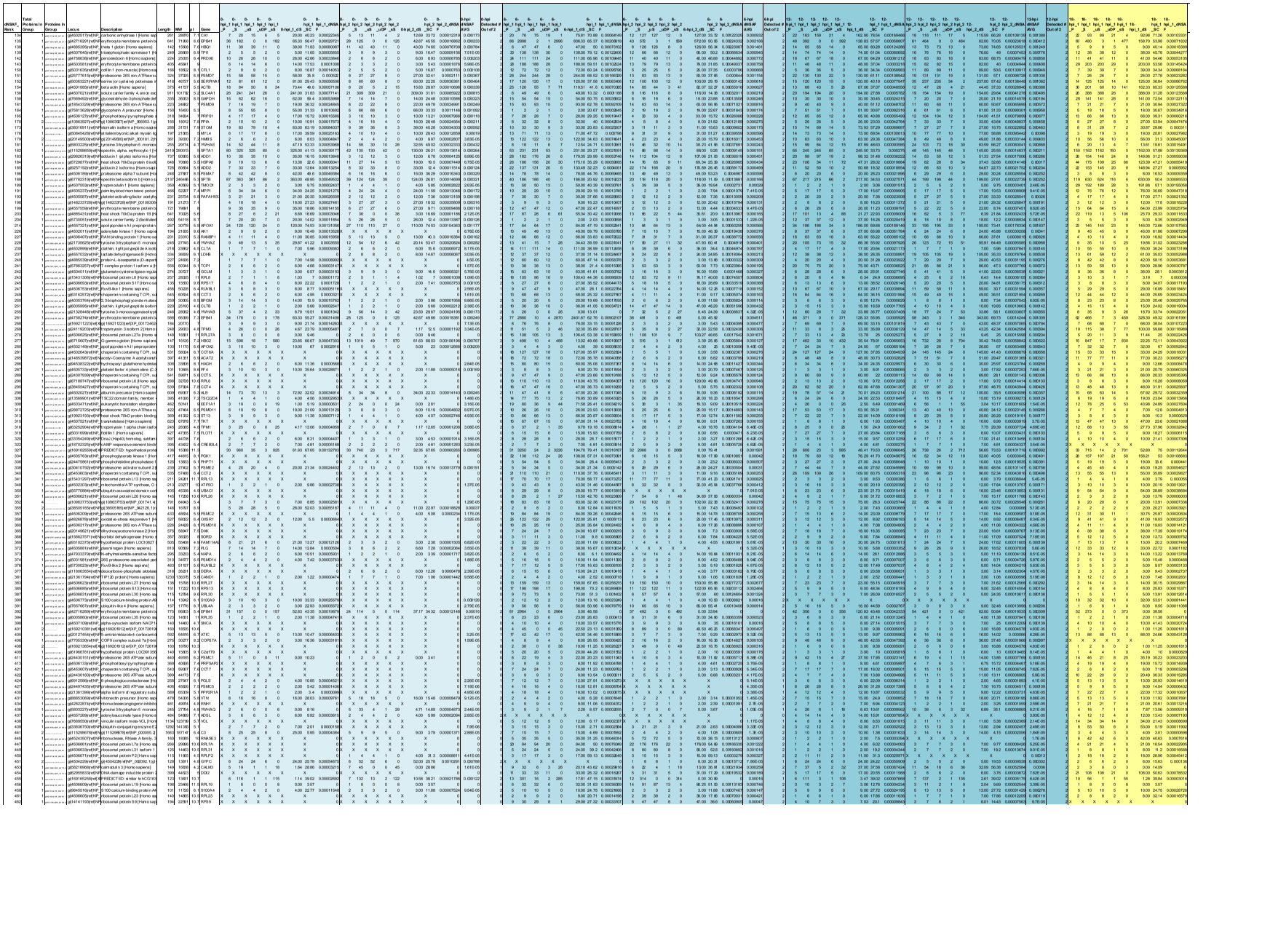


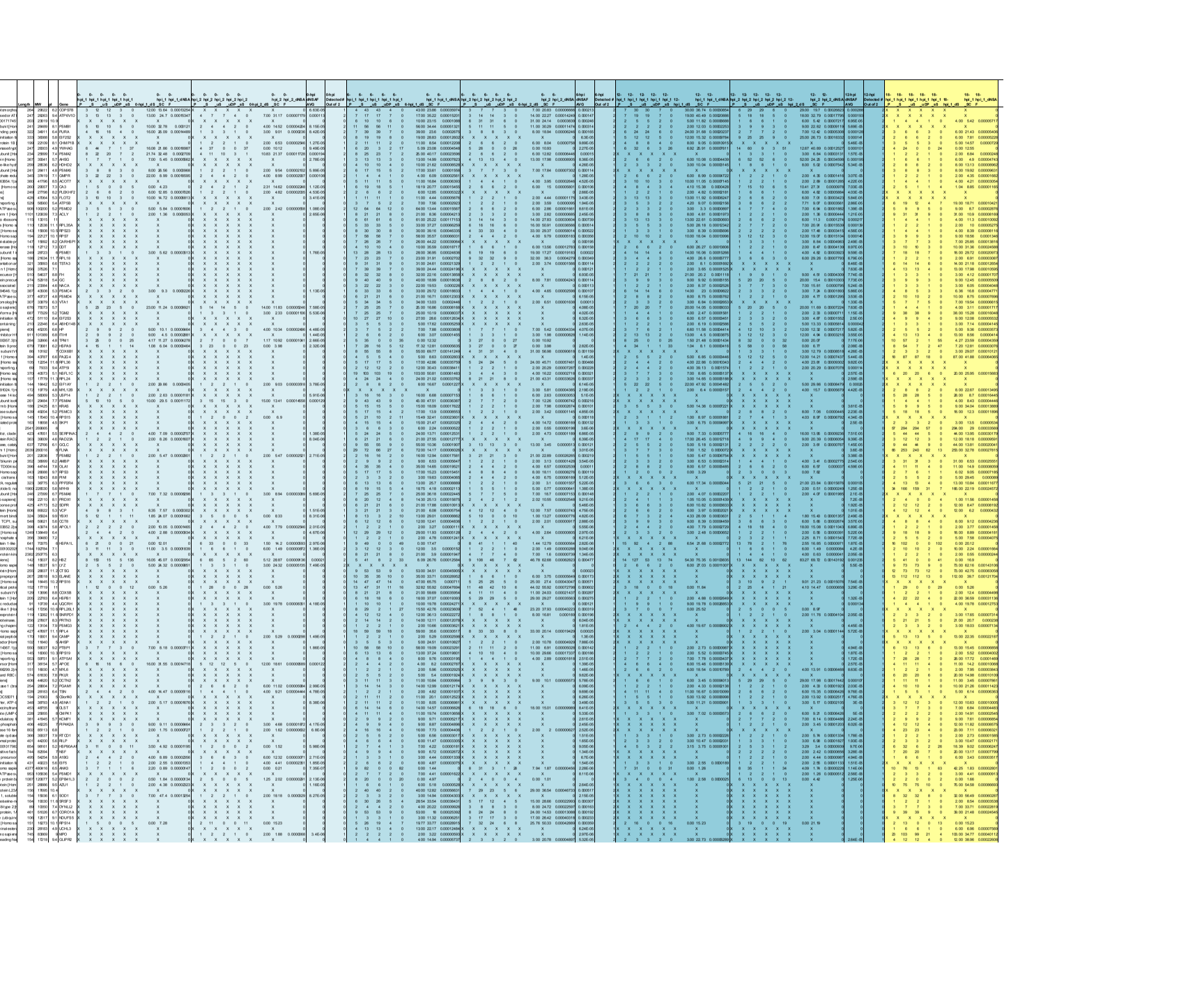




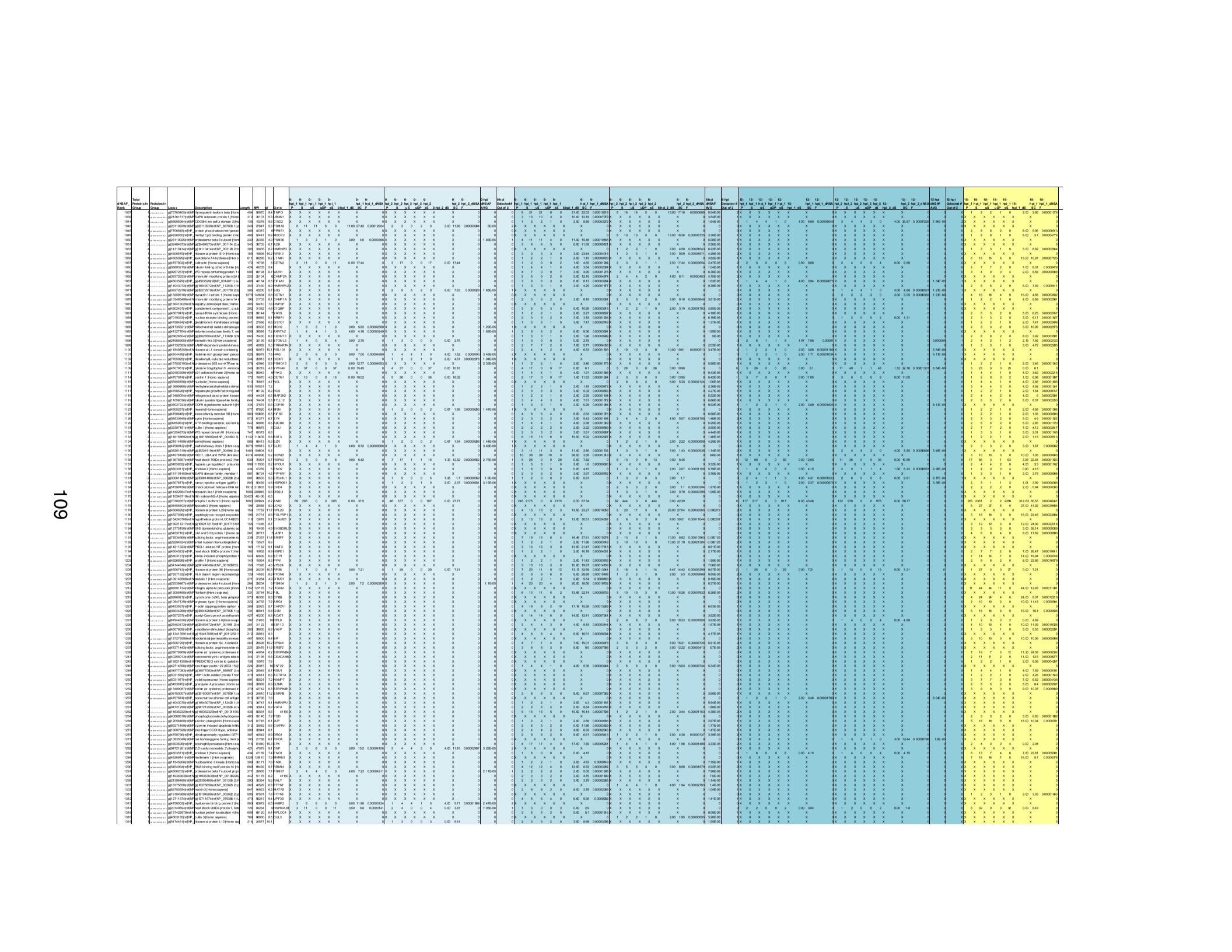


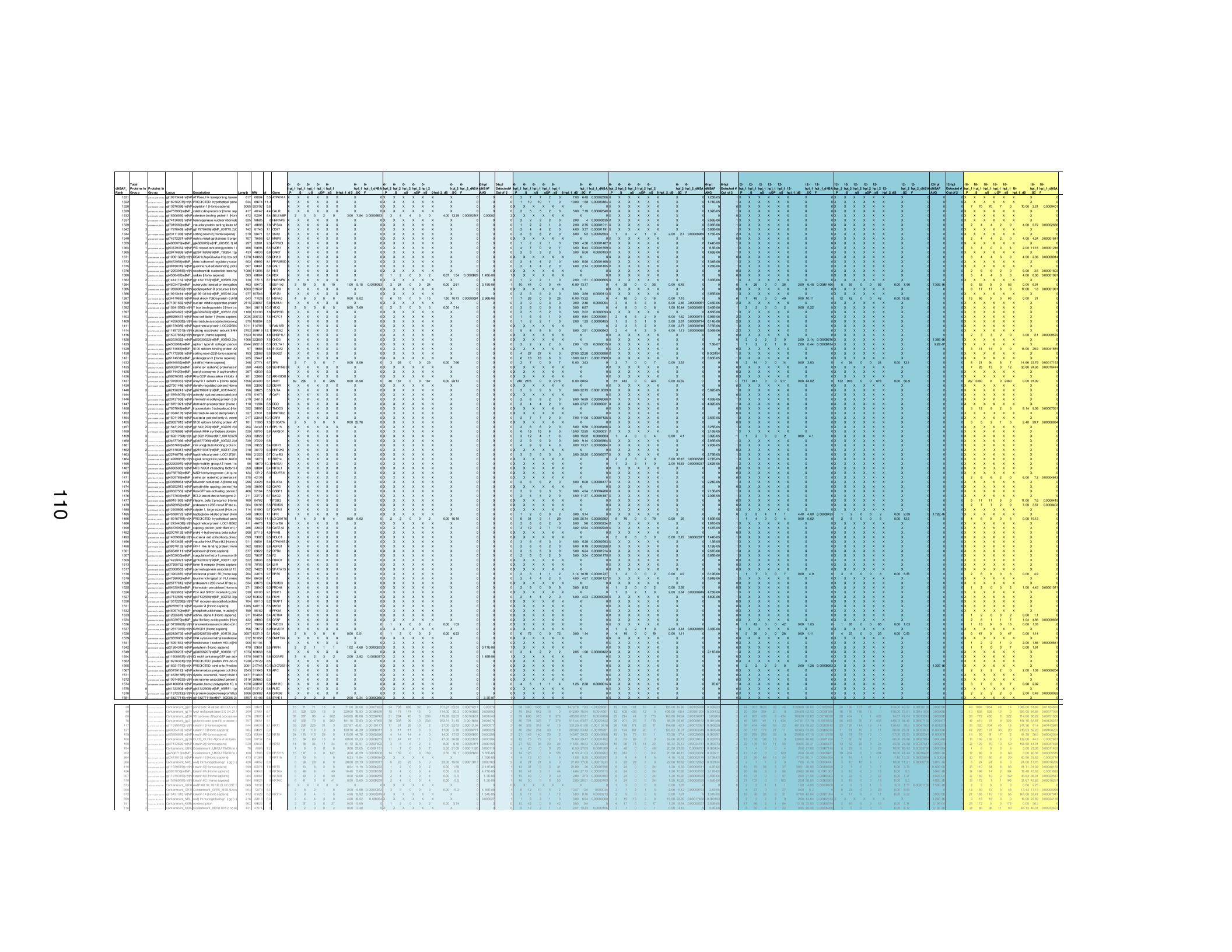


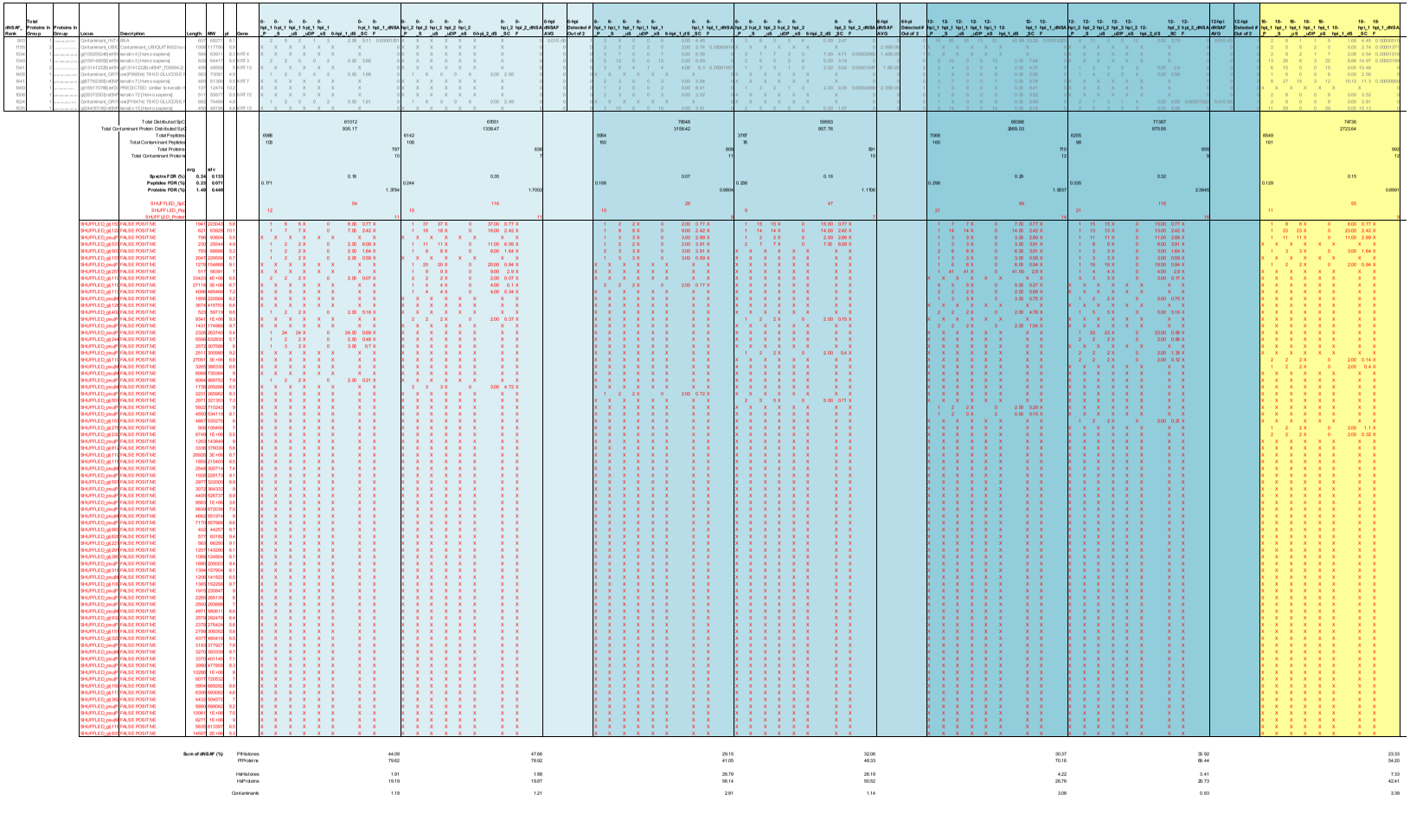


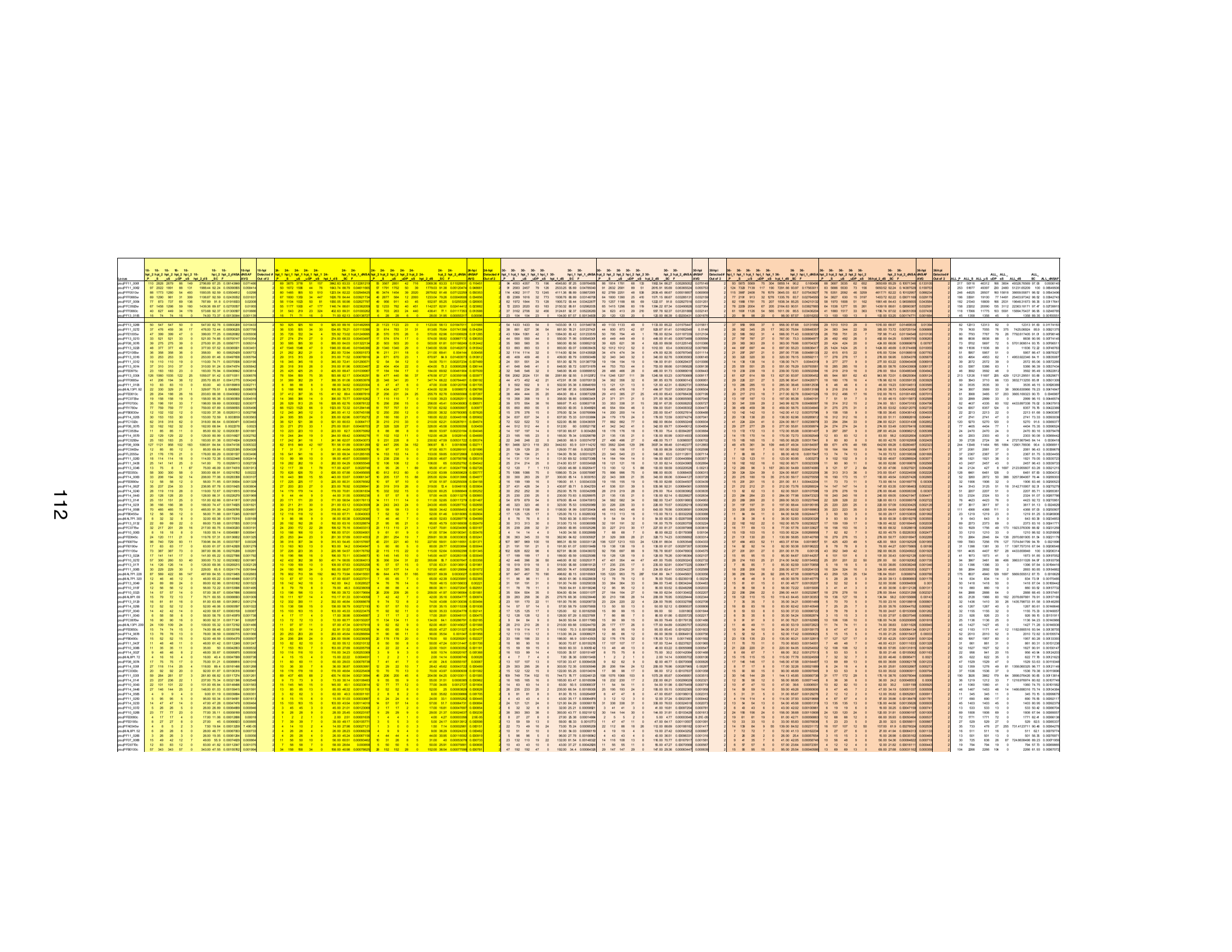


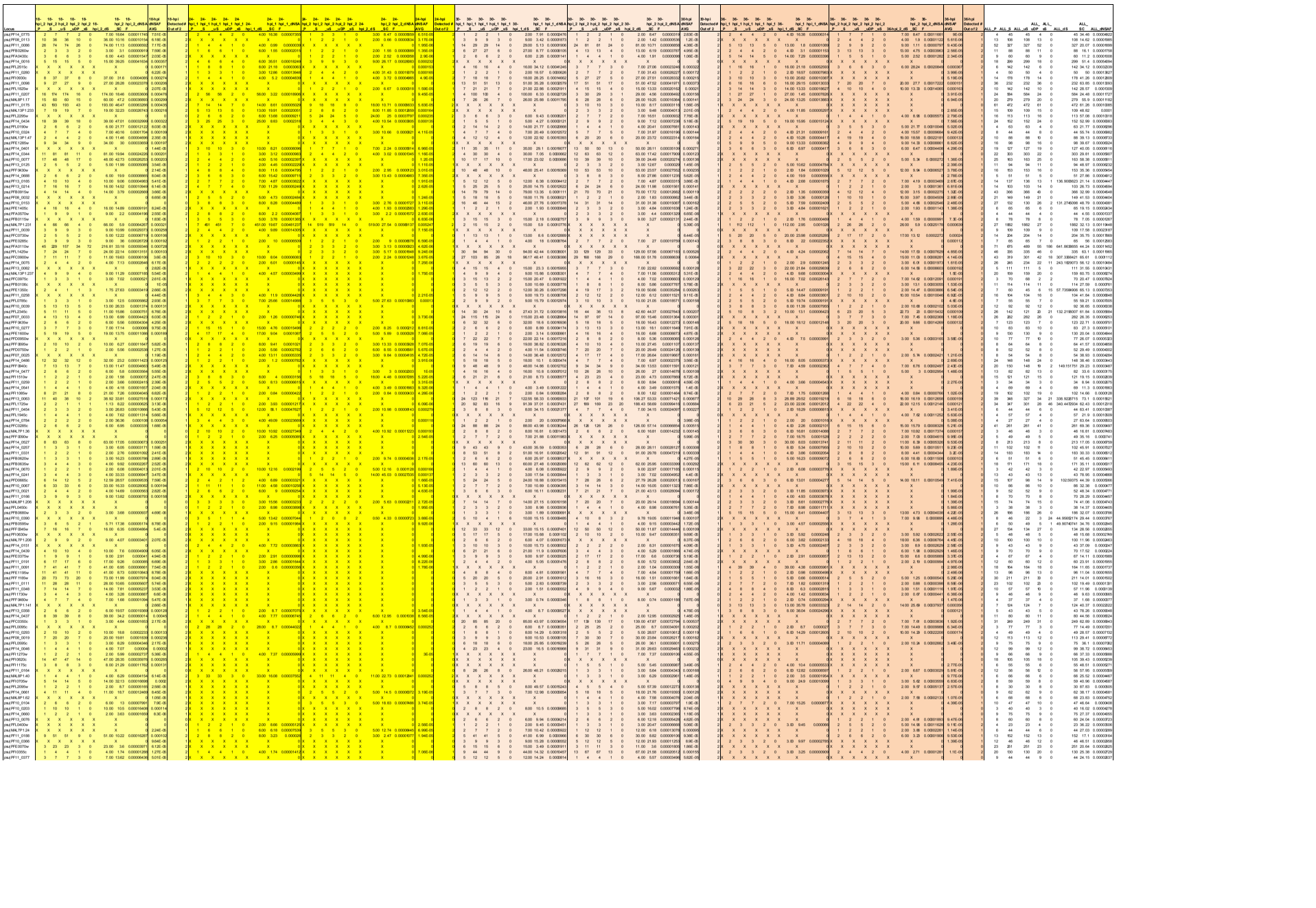
107

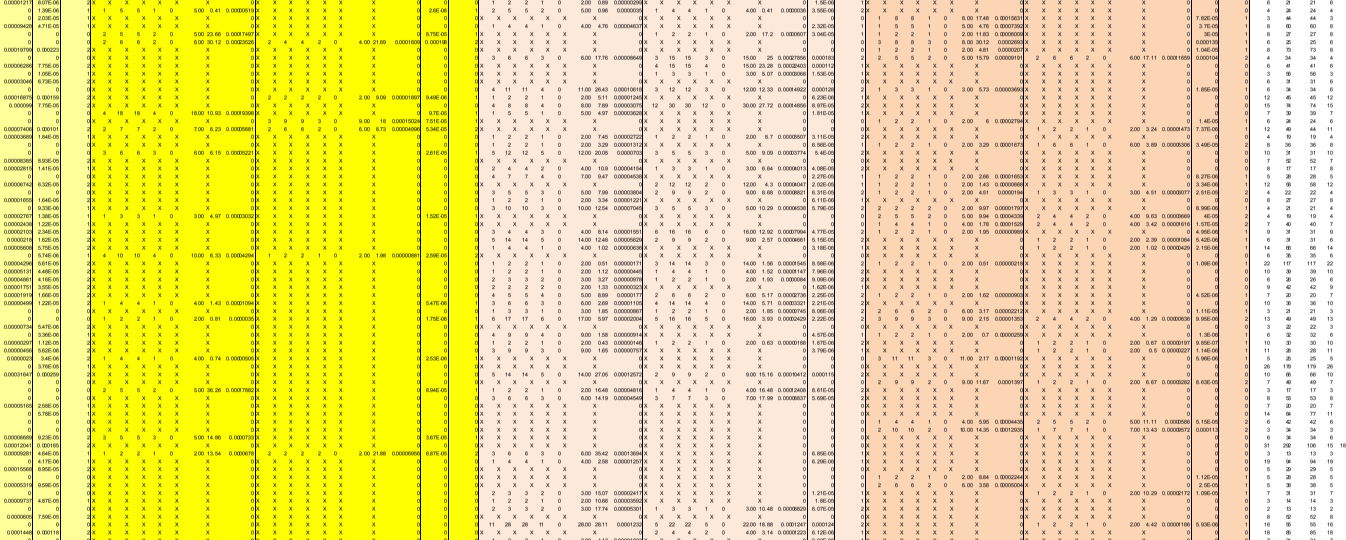


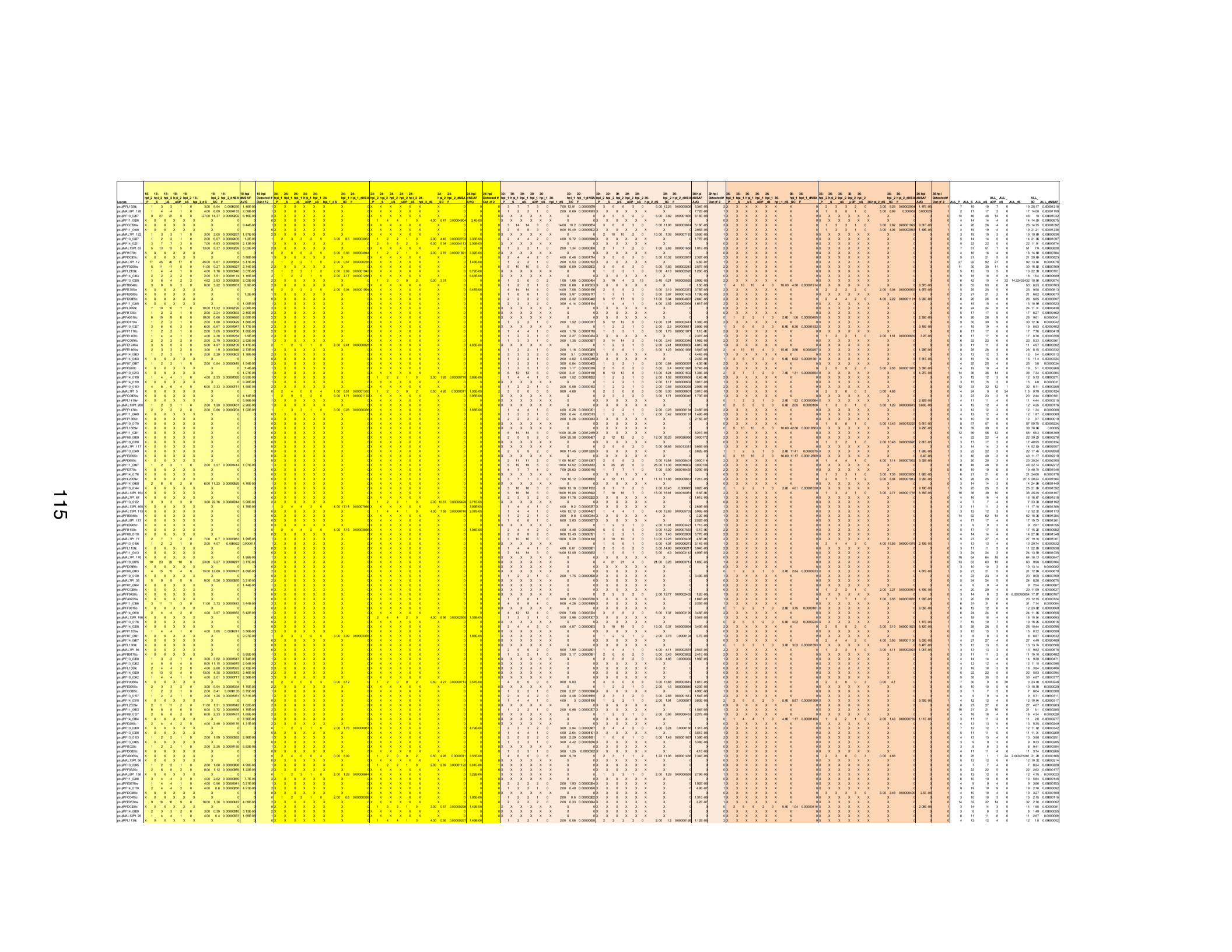


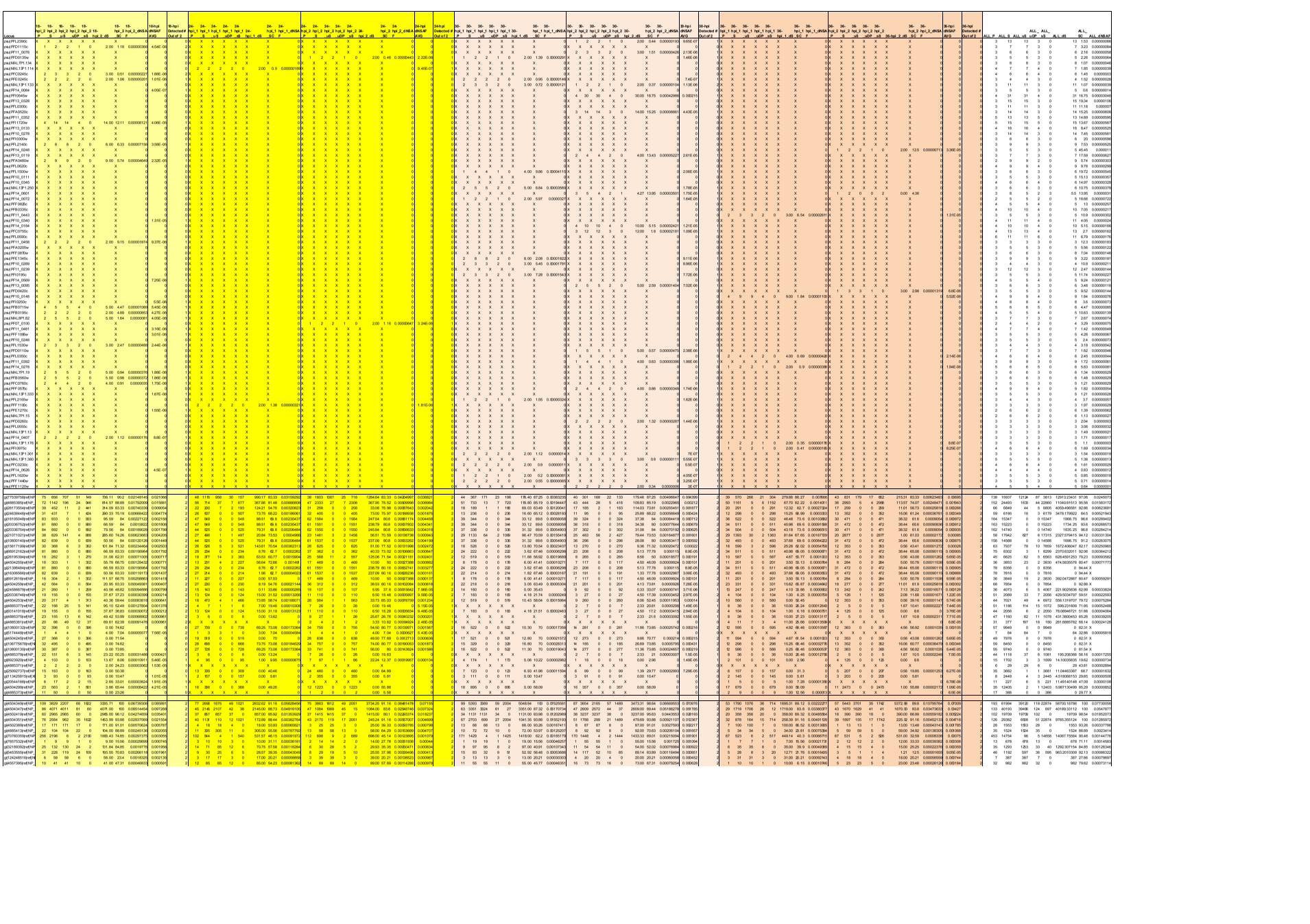


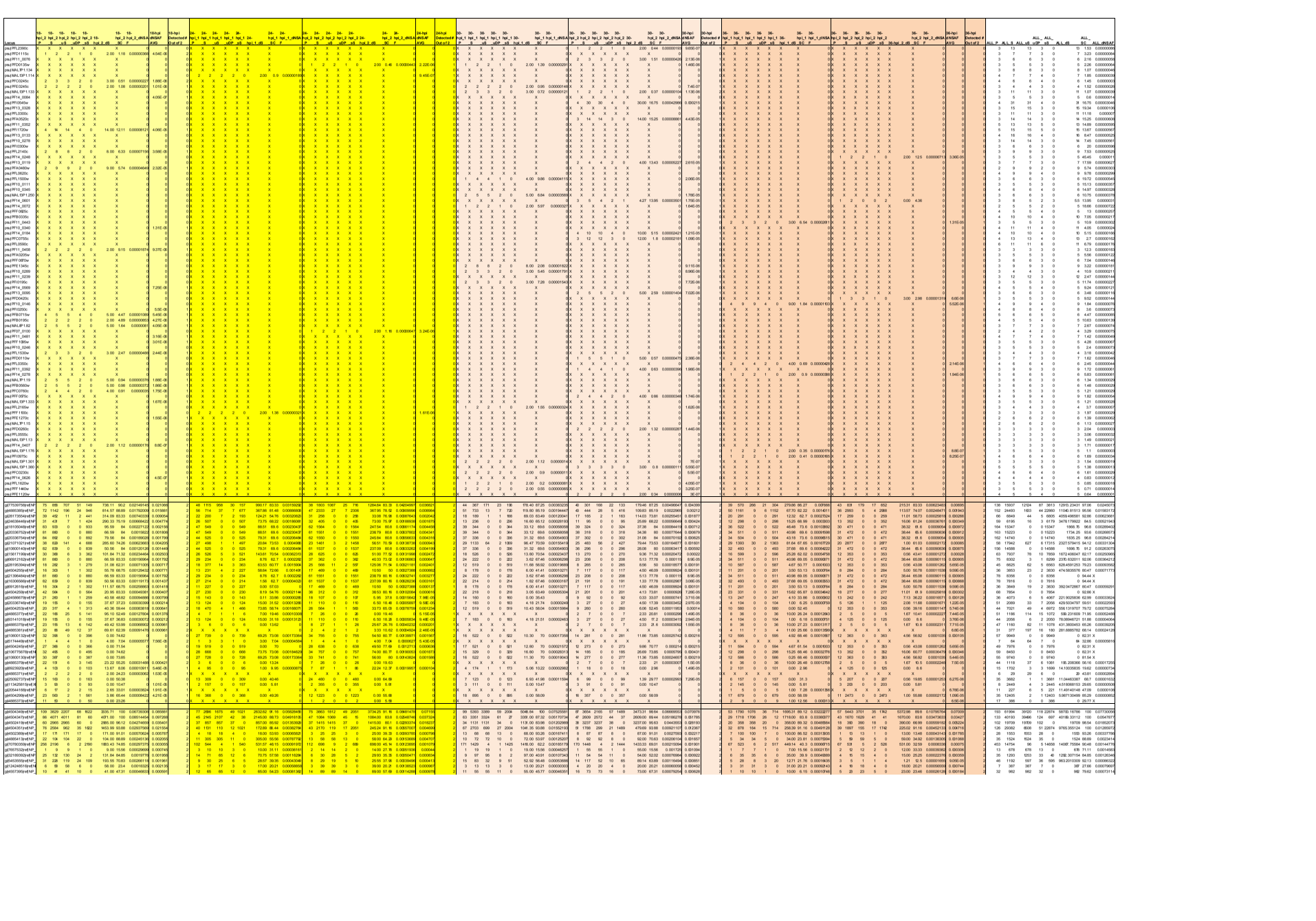


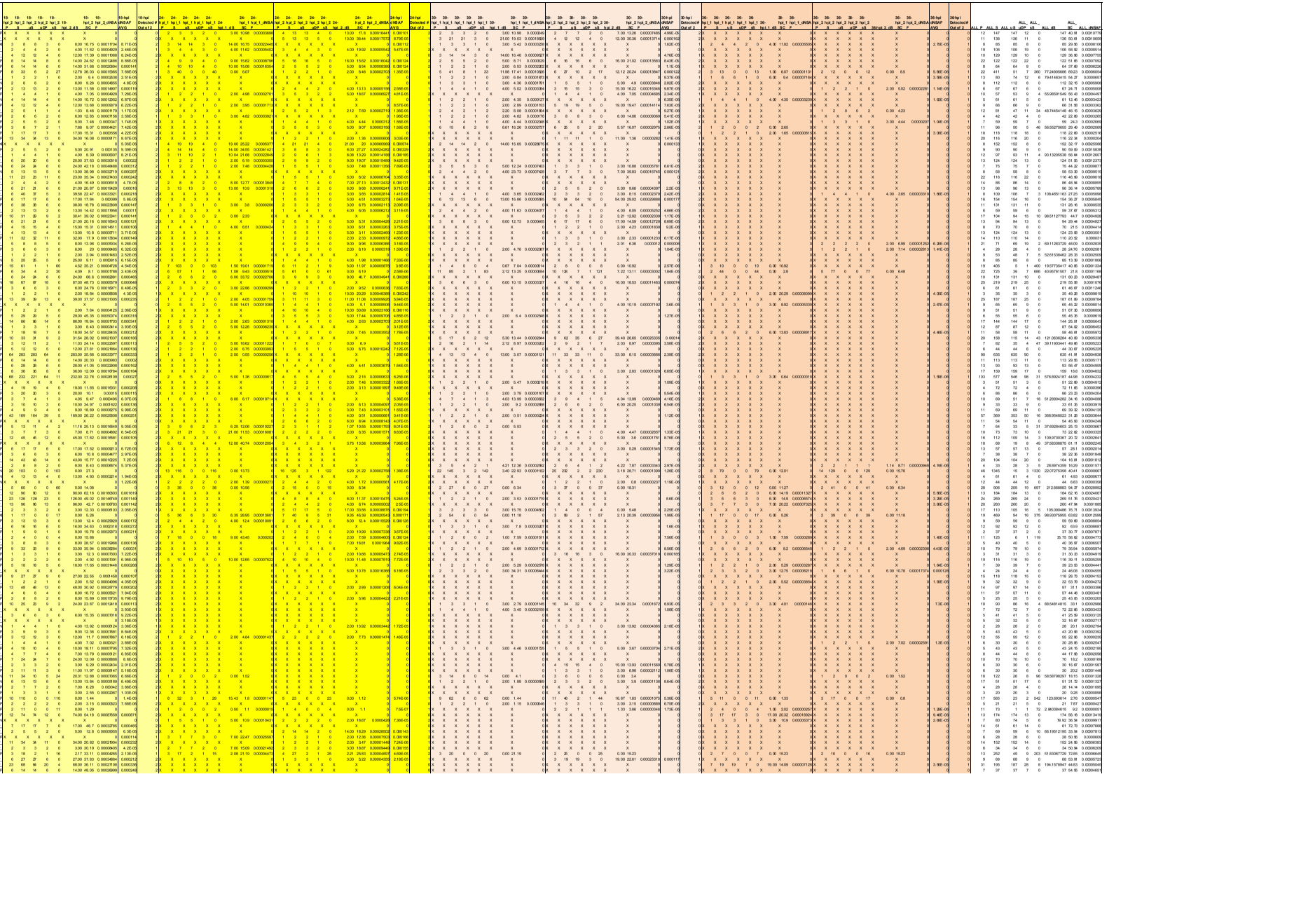


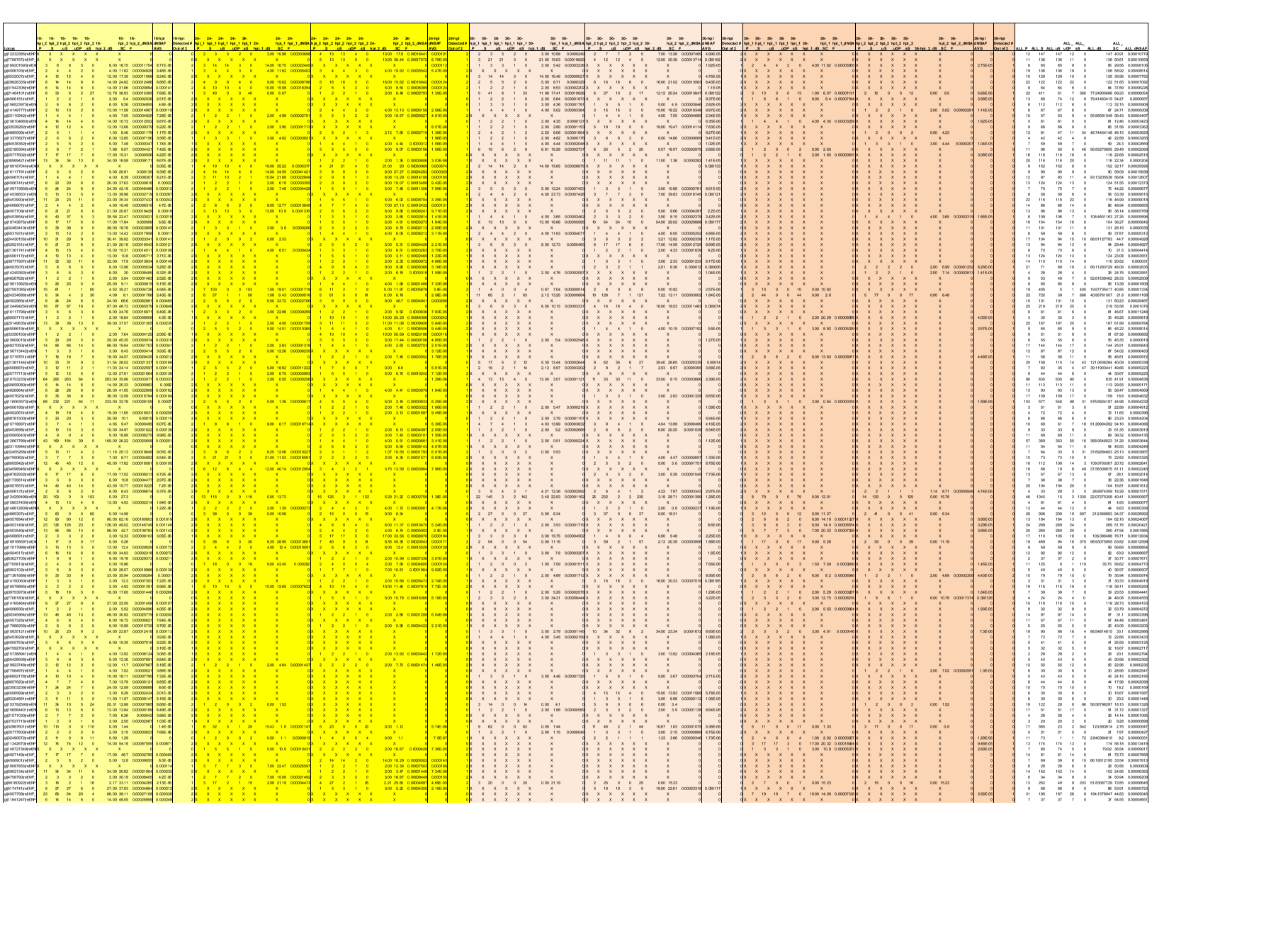


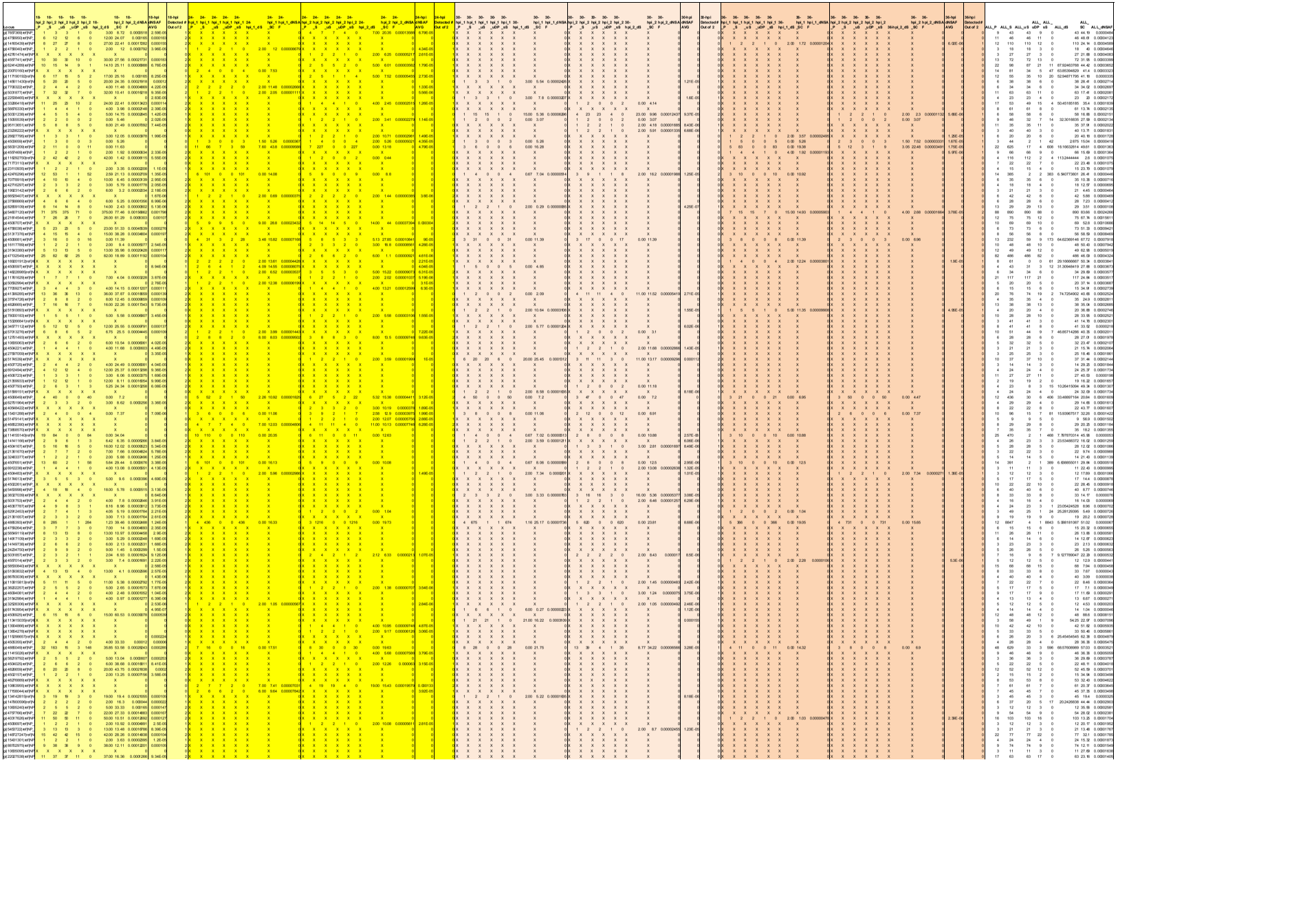


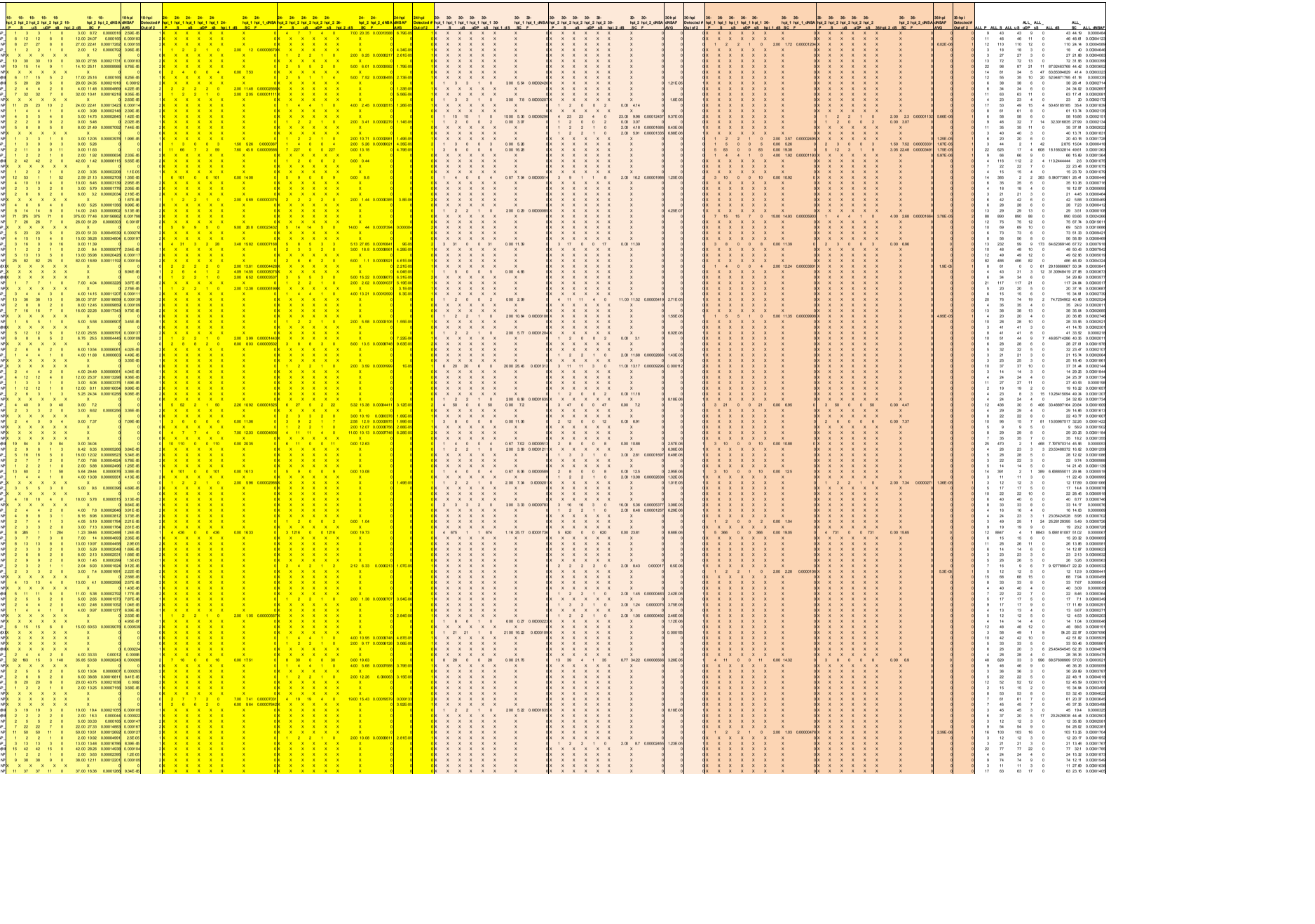




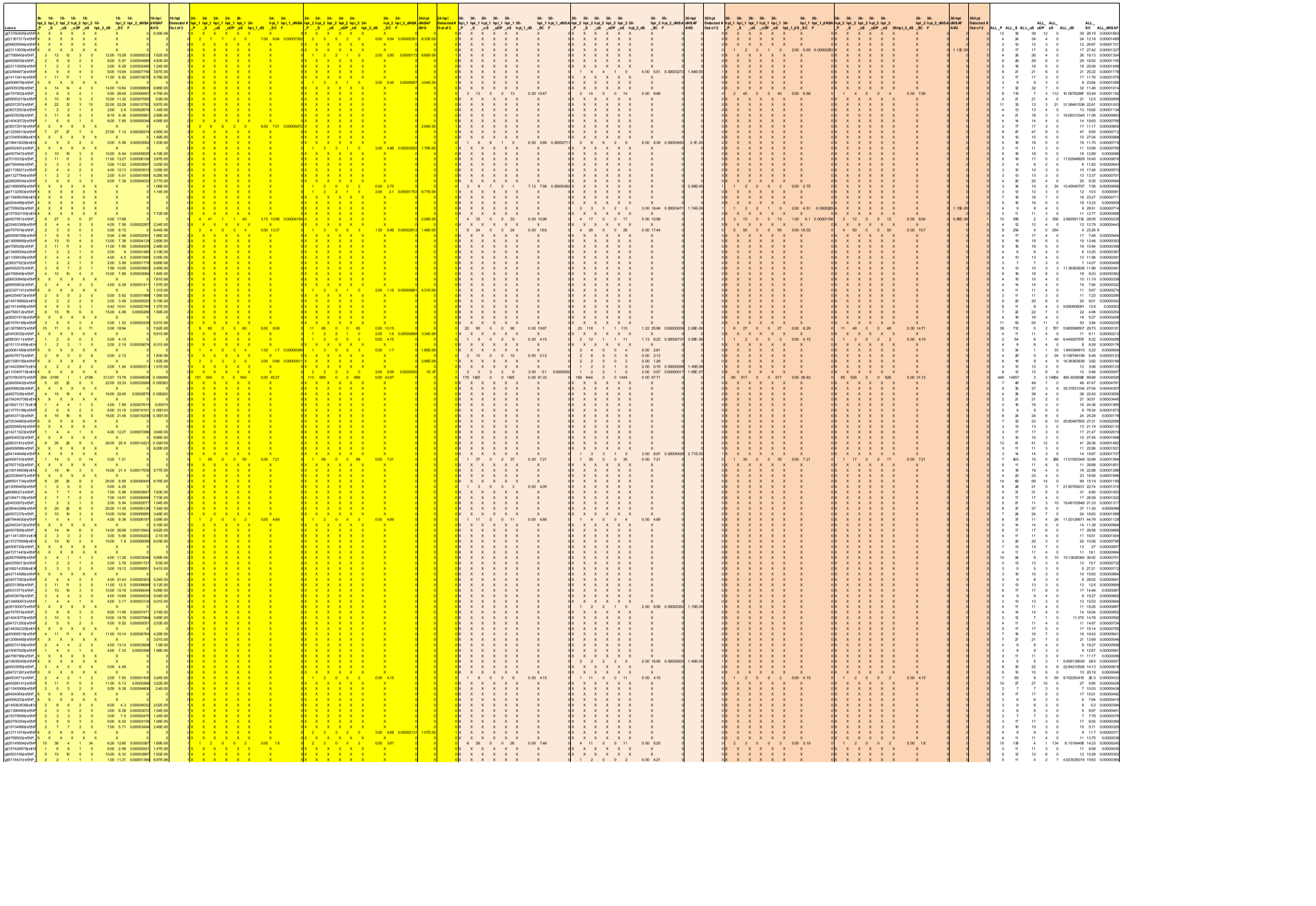


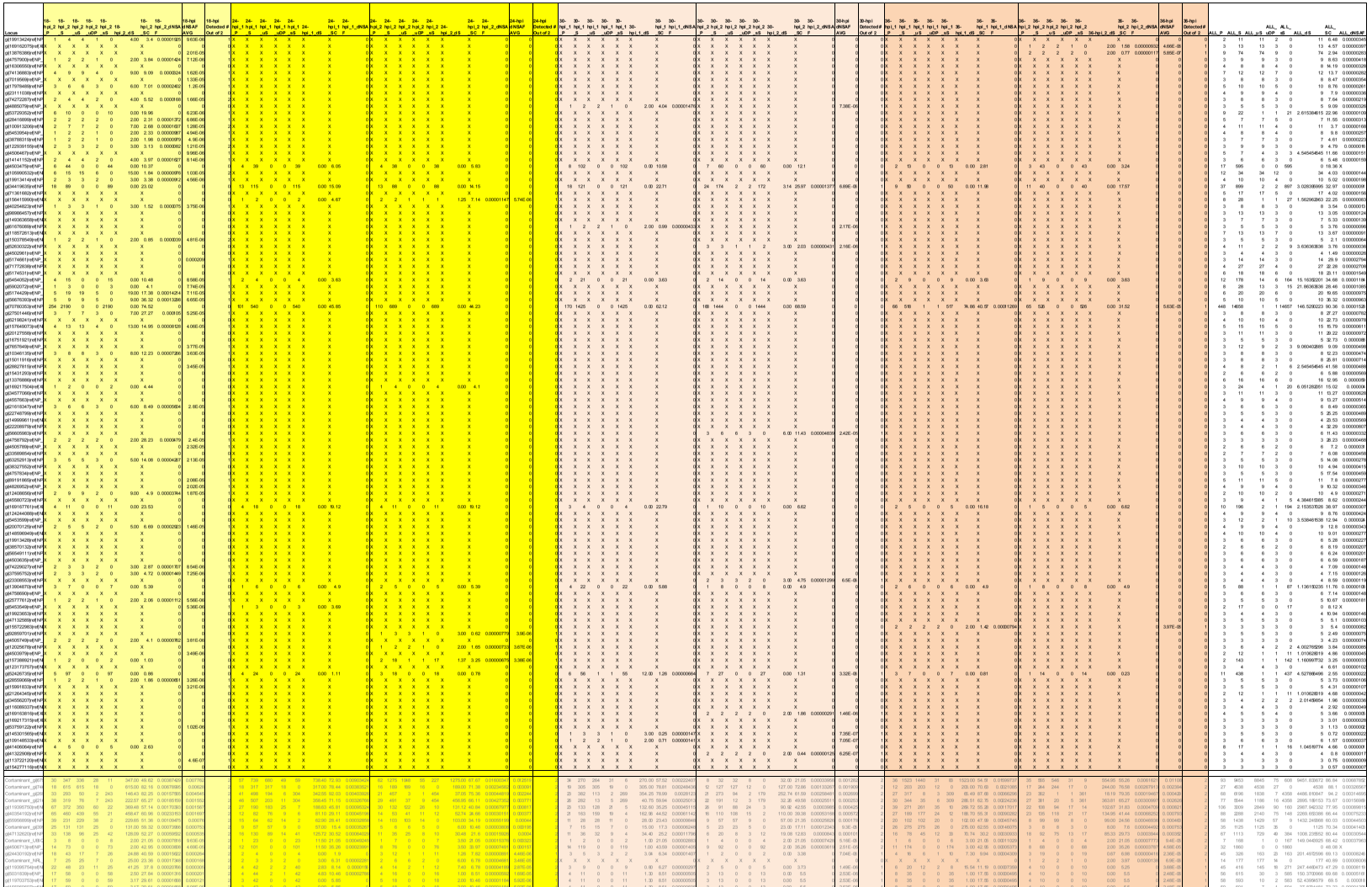






121





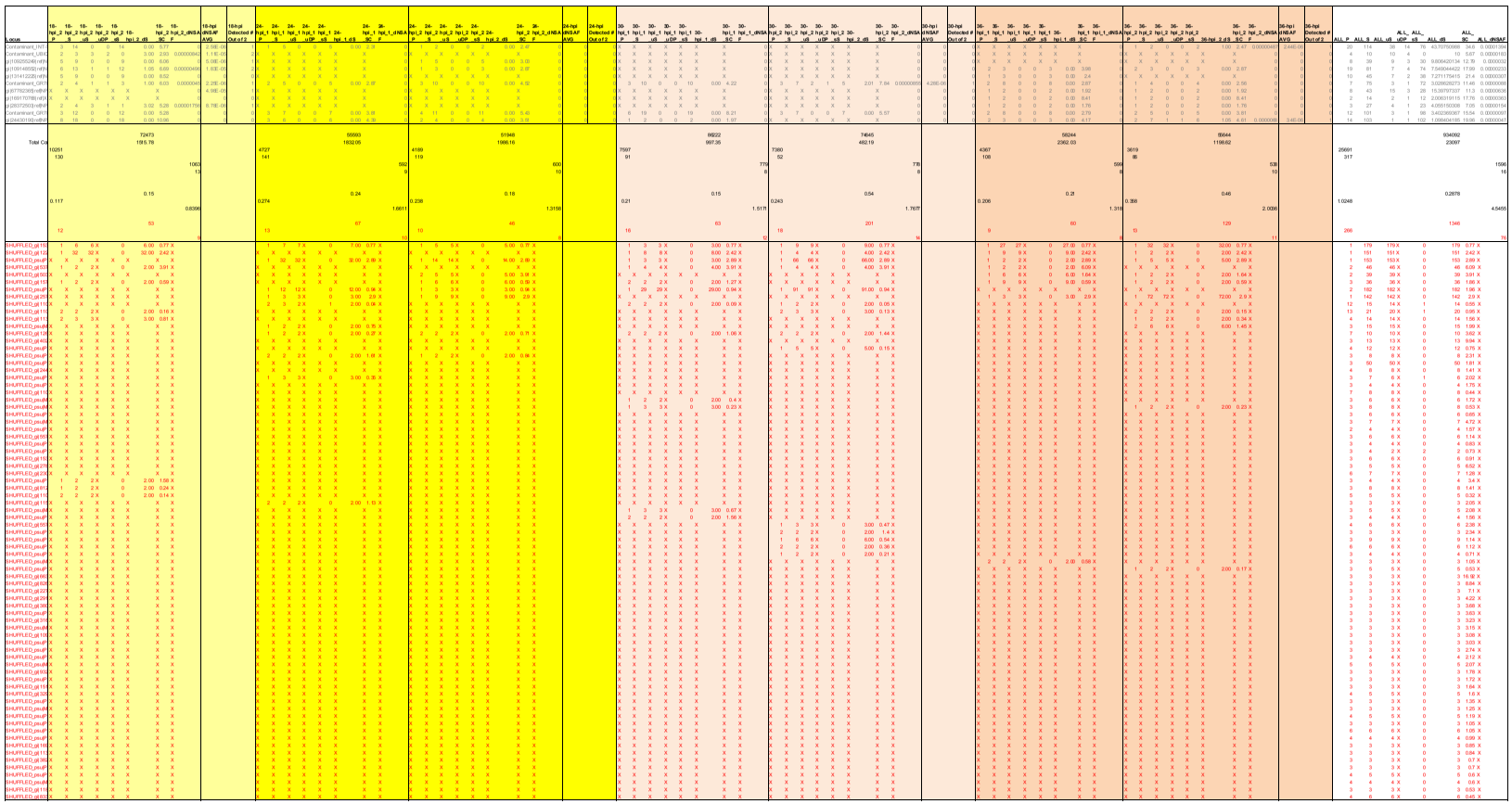


Table S3: Comparison of post-translational modifications identified on Histones from Plasmodium, Human, *Saccharomyces cerevisiae,* and Drosophila

For each amino acid number / PTM type combination detected in this study, are specified:

- The location of each amino acid as determined by the predicted location of N-terminal tail (NT), secondary structures forming the core domain (CD), and C-terminal tail (CT) based on known tridimensional structures (Figure S3);
- The total spectral counts (Total_SpC) mapping to peptides (modified or not) bearing each amino acid residue, combining all 56 MudPIT analyses of 14 histone samples;
- The spectral counts matching modified peptides bearing each PTM type on this amino acid (Modified_SpC), combining all 56 MudPIT analyses of 14 histones samples;
- The modification level estimated by Modified_SpC expressed as a percentage of the Total_SpC (Mod/Total);
- The number of times peptides bearing each AA#PTM combination were detected across the 14 histone samples;
- Whether or not (Y/N) each AA#PTM combination had been previously detected in Plasmodium (from [1] and/or [2]);
- Whether or not (Y/N) each AA#PTM combination had been previously detected in human (or mammals), yeast and/or drosophila (see References column);
- The homologous amino acids in human, yeast, and fly (sequence alignments reported in Figure S3);
- The type of modifications previously reported for this amino acid position in homolgous histones from human (or mammals), yeast and/or drosophila (see References column);
- Previously reported PTMs with check marks in columns "ac-cr" were detected in this study, while PTMs in grey cells were not.

For each amino acid previously reported as modified in Plasmodium and/or other species but not detected as modified in this study (grey rows at the bottom of each histone table), are specified:

- The total spectral counts (Total_SpC) mapping to peptides (modified or not) bearing these unmodified amino acid residues, combining all 56 MudPIT analyses of 14 histone samples;
- The number of times peptides bearing this unmodified amino acid were detected across the 14 histone samples;
- Whether or not (Y/N) each AA#PTM combination had been previously detected in Plasmodium (from [1] and/or [2]);
- Whether or not (Y/N) each AA#PTM combination had been previously detected in human (or mammals), yeast and/or drosophila (see References column);
- The homologous amino acids in human, yeast, and fly (sequence alignments reported in Figure S3);
- The type of modifications previously reported for this amino acid position in homolgous histones from human (or mammals), yeast and/or drosophila (see References column).

Abbreviations:
Nac, N-terminal acetylation; ac, acetylation; me1, monomethylation; me2, dimethylation; me3, trimethylation; fo, formylation; phs, phosphorylation; ub, ubiquitination; oh, hydroxylation; cr, crotonylation;
n.c., amino acid residue not conserved;
n.d., amino acid not detected as modified in this study.

Amino acid sequences were aligned using ClustalW2 (Figure S3):

- PfH2A (psu|PFF0860c), HsH2A (NP_254280.1), ScHhta1 (NP_010511.1), ScHta2 (NP_009552.1), and DmH2A (NP_001027376.1);
- PfH2Az (psu|PFC0920w), HsH2Az (NP_002097.1), ScHhtz1 (NP_014631.1), and Dm H2A.v (NP_524519.1);
- PfH2B (psu|PF11_0062), PfH2Bv (psu|PF07_0054), HsH2B (NP_066402.2), ScHtb1 (NP_010510.1), ScHtb2 (NP_009553.1) and Dm H2B (NP_001027385.1);
- PfH3 (psu|PFF0510w), PfH3.3 (psu|PFF0865w), HsH3.1t (NP_003484.1), HsH3d (NP_003521.2), HsH3.3C (NP_001013721.2), HsH3.3 (NP_005315.1), ScHht2 (NP_014367.1), and DmH3 (NP_001027387.1), DmH3.3A (NP_723056.1), DmH3.3B (NP_788892.1);
- PfH3.c (psu|PF13_0185), HsH3.cAa (NP_001800.1), HsH3.cAb (NP_001035891.1), ScCse4 (NP_012875.2), and DmH3.c (NP_523730.2);
- PfH4 (psu|PF11_0061), HsH4 (NP_001029249.1), ScHhf (NP_014368.1), DmH4.1 (NP_001027377.1) and DmH4.2 (NP_001027382.1).

Summary Table tallying all PTMS detected (or not) on all 8 Pf Histones (detailed tables for each Pf Histone are provided below)

Detected in this study	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
Total	155	138	246	59	23	4	31	18	6	58	6	4	32	5
Known	79	76	120	49	1	4	23	9	3	21	4	2	3	1
Novel	73	60	126	10	22	0	8	9	3	37	2	2	29	4

Previously reported as modified	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
All Known	128	113	255	60	3	4	54	23	16	22	39	4	4	26
n.d. PTM on AAs conserved in Pf	84	84	137	17	1	0	31	15	11	6	30	2	1	23
n.d. (AA detected as modified by other PTMs)	52	52	86	9	0	0	18	8	7	4	17	2	0	21
n.d. (AA n.d. as modified)	32	32	51	8	1	0	13	7	4	2	13	0	1	2
n.c. AA	15	0	21	4	1	0	4	1	1	1	6	1	0	2

Fig1B

	TOTAL	
Known in Pf	43	
Known in Pf (n.c. AA)	3	
Known in Hs/Sc/Dm	74	120
Novel in Pf/Hs/Sc/Dm	50	
Novel (AA known to be modified)	57	
Novel (n.c. AA)	19	126
Not Detected	120	
Not Detected, Known in Pf	17	137

Fig2A

	TOTAL	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
Known in Pf (and Hs/Sc/Dm)	46	29	0	4	4	2	2	0	4	1	0	0
Known in Hs/Sc/Dm (Novel in Pf)	74	20	1	0	19	7	1	21	0	1	3	1
All Known	120	49	1	4	23	9	3	21	4	2	3	1
Novel in Pf/Hs/Sc/Dm	126	10	22	0	8	9	3	37	2	2	29	4
Total	246	59	23	4	31	18	6	58	6	4	32	5

Fig2B

	TOTAL	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
Not Detected (on AAs conserved in Pf)	137	17	1	0	31	15	11	6	30	2	1	23
including n.d. PTM Known in Pf	17	0	0	0	3	2	3	0	9	0	0	0

H2A	PfHistone-AA#PTM	Location	Total_ SpC	Modified_SpC	Mod/ Total (%)	Times Detected	Known in Pf?	Known in others?	Previously reported AA#PTMs												
									Hs/Mm	Sc	Dm	ac	me1	me2	me3	fo	phs	ub	oh	cr	References
1	H2A-S1Nac	NT	127	127	100	13	Y	Y	S1	S1	S1	Nac✓					phs				[3-9]
2	H2A-K3ac	NT	133	1	0.75	1	Y	Y	R3	K4	R3	ac✓	me1		me3						[10, 11]
3	H2A-K5ac	NT	133	54	40.6	12	Y	Y	K5	K7	K5	ac✓									[4, 10, 12]
4	H2A-K9ac	NT	68	4	5.88	3	N	Y	K9	n.c.	K8	ac✓		me2	me3						[11, 13, 14]
5	H2A-K10ac	NT	204	2	0.98	2	N	Y	R11	K13	K10	ac✓									[14]
6	H2A-T15ac	NT	1559	3	0.19	3	N	N (n.c./Kac)	K15	n.c.	K14	ac									[15]
7	H2A-K20fo	NT	1609	8	0.5	2	N	N (known AA	R20	K21	R19	ac									[16]
8	H2A-R29me1	CD	3916	11	0.28	3	N	N	R29	R30	R28										
9	H2A-K41fo	CD	78	17	21.79	7	N	N (n.c.)	E41	Q42	E40										
10	H2A-R42fo	CD	1719	33	1.92	9	N	N (known AA	R42	R43	R41		me1								[17]
11	H2A-Y50oh	CD	3975	16	0.4	7	N	N	Y50	Y51	Y39										
12	H2A-Y57oh	CD	3298	754	22.86	12	N	N	Y57	Y58	Y56										
13	H2A-R71me1	CD	3820	20	0.52	3	N	Y	R71	R72	R70		me1✓								[11]
14	H2A-K74fo	CD	1611	3	0.19	3	N	N (known AA	K74	K75	K73	ac									[11]
15	H2A-R88fo	CD	1441	87	6.04	11	N	N (known AA	R88	R89	R87		me1								[17]
16	H2A-K95fo	CD	1570	65	4.14	7	N	Y	K95	K96	K94	ac✓	me1	me2		fo✓				cr	[11, 13, 18]
	H2A-K95ac	CD	1570	31	1.97	7	N	Y													
17	H2A-K118oh	CT	4355	5	0.11	1	N	N (known AA	K118	K119	K117		me1✓			fo				cr	[17, 18]
	H2A-K118me1	CT	4355	4	0.09	1	N	Y													
	H2A-K118ac	CT	4355	5	0.11	2	N	N (known AA													
18	H2A-K119fo	CT	381	2	0.52	1	N	N (known AA	K119	K120	K118	ac	me1					ub		cr	[11, 17, 19]
19	H2A-T126phs	CT	464	4	0.86	2	Y	Y	K125	T124/125	n.c.						phs✓			cr	[11, 20-22]
n.d.	H2A-K13	CD	680	0	0	7	N	Y	K13	A14	K12	ac									[15]
n.d.	H2A-S16	CD	1559	0	0	7	N	Y	S16	S17	S15						phs				[23]
n.d.	H2A-S18	CD	1559	0	0	7	Y	N	S18	S19	S17						phs				[21]
n.d.	H2A-K36	CD	45	0	0	7	N	Y	K36	R37	K35	ac				fo				cr	[17, 18]
n.d.	H2A-Y39	CD	82	0	0	11	N	Y	Y39	Y40	Y38								oh		[17]
n.d.	H2A-R77	CD	1	0	0	1	N	Y	R77	R78	R76		me1	me2							[24]
n.d.	H2A-T79	CT	0	0	0	0	N	Y	I79/T79	I80	I78	ac									[11]
n.d.	H2A-S120	CT	467	0	0	14	Y	Y	T120	S121	T119						phs				[11, 21, 22, 25]
n.c.	H2A-G99	CD	n.c.				N	Y (n.c.)	R99/K99	N100	G98		me1	me2							[15, 23]
n.c.	H2A-T126	CT	n.c.				N	Y (n.c.)	K125	K126	n.c.		me1							cr	[17]
n.c.	H2A-N128	CT	n.c.				N	Y (n.c.)	K127	S128	n.c.	ac					phs				[11, 26-28]
n.c.	H2A-D130	CT	n.c.				N	Y (n.c.)	K129	E130	n.c.	ac									[11]

H2A

Detected in this study	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
Total	19	17	22	6	1	1	3	0	0	7	1	0	3	0
Known	9	9	10	5	0	1	2	0	0	1	1	0	0	0
Novel	10	8	12	1	1	0	1	0	0	6	0	0	3	0

Previously reported as modified	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
All Known	26	22	48	14	0	1	10	4	2	3	6	1	1	6
n.d. PTM on AAs conserved in Pf	18	18	31	7	0	0	6	3	2	2	4	1	1	5
n.d. (AA detected as modified by other PTMs)	11	11	20	4	0	0	5	2	2	1	1	1	0	4
n.d. (AA n.d. as modified)	7	7	11	3	0	0	1	1	0	1	3	0	1	1
n.c. AA	4	0	7	2	0	0	2	1	0	0	1	0	0	1

H2Az	PfHistone-AA#PTM	Location	Total_ SpC	Modified_SpC	Mod/ Total (%)	Times Detected	Known in Pf?	Known in others?	Previously reported AA#PTMs												
									Hs/Mm	Sc	Dm	ac	me1	me2	me3	fo	phs	ub	oh	cr	References
1	H2Az-M1Nac	NT	545	341	62.57	14	Y	N (n.c.)	2nd residue= E not conserved			Nac✓									[1]
	H2Az-M1Nac-ox		545	204	37.43	14	N		Hs/Sc/Dm												
2	H2Az-K11ac	NT	1653	1235	74.71	14	Y	N (n.c.)	n.c.	n.c.	n.c.	ac✓									
3	H2Az-K15ac	NT	1988	1431	71.98	14	Y	N (n.c.)	n.c.	n.c.	n.c.	ac✓									
4	H2Az-K19ac	NT	1929	1263	65.47	14	Y	Y	n.c.	K3	n.c.	ac✓									[29, 30]
5	H2Az-K25ac	NT	1985	1654	83.32	14	Y	Y	n.c.	K8	n.c.	ac✓									[29, 30]
6	H2Az-K28ac	NT	1936	1541	79.6	14	Y	Y	K4	K10	K4	ac✓									[4, 13, 31, 32]
7	H2Az-K30ac	NT	1653	1461	88.38	14	Y	Y	K7	K14	K7	ac✓									[4, 13, 31, 32]
8	H2Az-S33ac	NT	1466	3	0.2	3	N (known A	N	S9	S16	S9						phs				[21]
9	H2Az-K35ac	NT	1466	1282	87.45	14	Y	Y	K11	S18	K11	ac✓									[13, 31, 32]
10	H2Az-K37ac	NT	1267	813	64.17	14	N	Y	K13	R20	K13	ac✓									[31, 32]
11	H2Az-R46fo	CD	143	36	25.17	11	N	N	R22	R27	R22										
12	H2Az-R55me1	CD	2488	11	0.44	3	N	N	R31	R36	R31										
13	H2Az-Y77oh	CD	2613	6	0.23	3	N	N	Y53	Y58	Y53										
14	H2Az-Y84oh	CD	2179	7	0.32	4	N	N	Y60	Y65	Y60										
15	H2Az-K98me1	CD	2096	9	0.43	2	N	N	K74	K79	K74										
16	H2Az-R115fo	CD	12583	1648	13.1	14	N	N	R91	R96	R91										
17	H2Az-K139fo	CT	856	3	0.35	1	N	N	K115	K120	K115										
	H2Az-K139ub	CT	856	3	0.35	2	N	N													
n.c.	H2Az-V145	CT	n.c.				N	Y (n.c.)	K121	V126	K121							ub			[33]
n.c.	H2Az-	CT	n.c.				N	Y (n.c.)	n.c.	n.c.	S137						phs				[34]

H2Az

Detected in this study	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
Total	17	14	18	8	1	1	2	0	0	3	0	1	2	0
Known	9	6	9	8	0	1	0	0	0	0	0	0	0	0
Novel	7	7	9	0	1	0	2	0	0	3	0	1	2	0

Previously reported as modified	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
All Known	12	10	12	8	0	1	0	0	0	0	2	1	0	0
n.d. PTM on AAs conserved in Pf	0	0	1	0	0	0	0	0	0	0	1	0	0	0
n.d. (AA detected as modified by other PTMs)	0	0	1	0	0	0	0	0	0	0	1	0	0	0
n.d. (AA n.d. as modified)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n.c. AA	2	0	2	0	0	0	0	0	0	0	1	1	0	0

H2B H2Bv	PfHistone-AA#PTM	Location	Total_ SpC	Modified_SpC	Mod/ Total (%)	Times Detected	Known in Pf?	Known in others?	Previously reported AA#PTMs												
									Hs/Mm	Sc	Dm	ac	me1	me2	me3	fo	phs	ub	oh	cr	References
1	H2Bv-S1Nac	NT	265	263	99.25	14	Y	N	P1	S1	P1	Nac✓					phs				[1, 21]
2	H2B-K3ac	NT	38	23	60.53	8	N	Y	K5	K6	K3	ac✓				fo				cr	[18, 31, 32, 35]
	H2Bv-K3ac	NT	265	244	92.08	14	Y	Y													
3	H2B-K7ac	NT	53	23	43.4	8	N	Y	K11	K11	K7	ac✓								cr	[12, 31, 32, 36]
	H2Bv-K8ac	NT	523	472	90.25	14	Y	Y													
4	H2Bv-S9ac	NT	629	3	0.48	2	N	N (n.c./Kac)	K12	A12	A8	ac			me3						[11, 31, 32, 35]
5	H2Bv-K13ac	NT	629	253	40.22	14	Y	Y	K15	K16	K10	ac✓	me1							cr	[31, 32, 35]
6	H2Bv-K14ac	NT	813	220	27.06	14	Y	Y	K16	K17	K11	ac✓								cr	[12, 31, 32, 36]
7	H2Bv-K18ac	NT	885	296	33.45	14	Y	Y	K20	K21	K17	ac✓	me1					ub		cr	[15, 31, 32, 35]
8	H2B-R28fo	NT	8296	408	4.92	14	N	N	R33	R36	R30										
	H2Bv-R29fo	NT	2693	11	0.41	7	N	N													
9	H2Bv-T30ac	CD	4448	63	1.42	9	N	N (n.c./Kac)	K34	K37	K31		me1			fo				cr	[11, 17, 18]
10	H2B-S28ac	CD	14673	196	1.34	11	N	N (known AA	S36	T39	S33						phs				[37]
11	H2B-Y29oh	CD	14673	12	0.08	7	N	Y	Y37	Y40	Y34						phs		oh✓		[17, 23]
12	H2B-Y32oh	CD	14673	31	0.21	7	N	N	Y40	Y43	Y37										
13	H2B-K35oh	CD	14681	6	0.04	3	N	N (known AA													
	H2B-K35fo	CD	14681	236	1.61	13	N	Y													
	H2Bv-K39fo	CD	4760	3	0.06	1	N	Y													
	H2B-K35me1	CD	14681	15	0.1	4	N	Y	K43	K46	K40		me1✓			fo✓					[15, 18]
	H2B-K35me2	CD	14681	3	0.02	1	N	N (known AA													
	H2B-K35ac	CD	14681	27	0.18	11	N	N (known AA													
	H2Bv-K39ac	CD	4760	4	0.08	3	N	N (known AA													
	H2B-K35cr	CD	14681	7	0.05	1	N	N (known AA													
14	H2B-K38fo	CD	4690	181	3.86	12	N	Y													
	H2Bv-K42fo	CD	1679	13	0.77	4	N	Y	K46	K49	K43	ac✓		me2✓		fo✓					[17, 31]
	H2B-K38me2	CD	4690	2	0.04	2	N	Y													
	H2Bv-K42ac	CD	1679	6	0.36	3	N	Y													
15	H2B-D43oh	CD	3008	5	0.17	3	N	N	D51	D54	D48										
16	H2B-R48fo	CD	2949	2	0.07	2	N	N (n.c.)	S56	Q59	S53										
	H2Bv-K52ac	CD	1425	4	0.28	4	N	N (n.c.)													
17	H2B-K49oh	CD	10605	10	0.09	5	N	N (known AA													
	H2B-K49fo	CD	10605	1041	9.82	13	N	N (known AA													
	H2Bv-K53fo	CD	3179	232	7.3	12	N	N (known AA													
	H2B-K49me1	CD	10605	57	0.54	8	N	Y													
	H2B-K49me2	CD	10605	34	0.32	10	N	Y													
	H2Bv-K53me2	CD	3179	85	2.67	10	N	Y	K57	K60	K54	ac✓	me1✓	me2✓							[11, 17]
	H2B-K49me3	CD	10605	5	0.05	4	N	N (known AA													
	H2B-K49ac	CD	10605	252	2.38	12	N	Y													
	H2Bv-K53ac	CD	3179	14	0.44	7	N	Y													
	H2B-K49cr	CD	10605	8	0.08	5	N	N (known AA													
18	H2B-S50ac	CD	24669	847	3.43	14	N	N													
	H2Bv-S54ac	CD	9957	277	2.78	14	N	N	A58	S61	A55										
	H2B-S50phs	CD	24669	9	0.04	2	N	N													
19	H2B-S56ac	CD	25541	8	0.03	5	N	N													
	H2B-S56phs	CD	25541	4	0.02	3	N	N	S64	S67	S61										
20	H2B-D60oh	CD	25732	5	0.02	5	N	N	D68	D71	D65										
21	H2B-T61ac	CD	25723	13	0.05	4	N	N (n.c.)	I69	I72	I66										
22	H2B-K64oh	CD	25653	4	0.02	3	N	N (n.c.)													
	H2B-K64fo	CD	25653	686	2.67	14	N	N													
	H2Bv-R68fo	CD	10220	5	0.05	1	N	N													
	H2B-K64me1	CD	25653	2	0.01	2	N	N													
	H2Bv-R68me1	CD	10220	3	0.03	1	N	N	R72	R75	R69										
	H2B-K64me2	CD	25653	10	0.04	3	N	N													
	H2B-K64me3	CD	25653	43	0.17	10	N	N (n.c.)													
	H2B-K64ac	CD	25653	492	1.92	14	N	N (n.c.)													
23	H2B-R71fo	CD	12756	21	0.16	3	N	N (known AA	R79	K82	R76		me1								[17]
24	H2B-R91fo	CD	1477	30	2.03	6	N	N (known AA													
	H2Bv-R95fo	CD	306	32	10.46	10	N	N (known AA	R99	R102	R96		me1✓								[17]
	H2B-R91me1	CD	1477	31	2.1	8	N	Y													
	H2Bv-R95me1	CD	306	3	0.98	1	N	Y													
25	H2B-K100oh	CD	3880	35	0.9	7	N	N (known AA													
	H2Bv-K104oh	CD	2332	20	0.86	6	N	N (known AA													
	H2B-K100fo	CD	3880	201	5.18	13	N	Y	K108	K111	K105	ac✓	me1			fo✓				cr	[11, 17, 18, 31, 32]
	H2Bv-K104fo	CD	2332	110	4.72	13	N	Y													
	H2B-K100ac	CD	3880	6	0.15	5	N	Y													
	H2Bv-K104me2	CD	2332	9	0.39	6	N	N (known AA													
26	H2B-K108fo	CD	5487	22	0.4	10	N	Y				ac									
	H2Bv-K112fo	CD	3438	2	0.06	2	N	Y	K116	R119	K113		me1✓			fo✓					[11, 17]
	H2B-K108me1	CD	5487	5	0.09	4	N	Y													
	H2Bv-K112me1	CD	3438	8	0.23	6	N	Y													
27	H2B-K112ub	CD	2906	153	5.26	13	Y	Y													
	H2Bv-K116ub	CD	1055	24	2.27	9	N	Y	K120	K123	K117	ac				fo✓		ub✓			[11, 17, 31, 38, 39]
	H2Bv-K116fo	CD	1055	10	0.95	2	N	Y													
28	H2B-K116fo	CT	2712	8	0.295	4	N	N (known AA	K125	n.c.	K122	ac									[11]
n.d.	H2B-K4	NT	53	0	0	13	N	Y	n.c.	K7	n.c.	ac									[40]
	H2Bv-G4	NT	n.c.	n.c.	n.c.	n.c.	N														
n.d.	H2B-S24	NT	8296	0	0	14	N	Y	S32	V35	K29						phs				[37]
	H2Bv-R28	NT	n.c.	n.c.	n.c.	n.c.	N														
n.d.	H2B-T67	NT	13593	0	0	14	N	Y	G75/S75	T78	A72						phs				[11]
	H2Bv-T71	NT	2368	0	0	14	N														
n.d.	H2B-R77	CD	32	0	0	10	N	Y	K85	K88	K82	ac	me1								[10, 15]
	H2Bv-K81	CD	12	0		5	N														
n.d.	H2B-T80	CD	754	0	0	14	N	N	n.c.	n.c.	n.c.						phs				[41]
	H2Bv-T84	CD	2	0	0	2	Y (n.c.)														
n.d.	H2B-S104	CD	5455	0	0	14	N	Y	S112	S115	S109	STac					phs				[11]
	H2Bv-S108	CD	3443	0	0	14	N														
n.c.	H2B-P5	NT	n.c.				N	Y (n.c.)	A7/T7	n.c.	n.c.						phs				[11]
	H2Bv-P5	NT	n.c.				N														
n.c.	H2B-	NT	n.c.				N	Y (n.c.)	S14	S10	n.c.						phs				[42-44]
	H2Bv-A11	NT	n.c.				N														
n.c.	H2B-	NT	n.c.				N	Y (n.c.)	T19	A20	A15	STac									[11]
	H2Bv-G17	NT	n.c.				N														
n.c.	H2B-G14	NT	n.c.				N	Y (n.c.)	K23	S26	K21	ac	me1								[32]
	H2Bv-G21	NT	n.c.				N														
n.c.	H2B-P15	NT	n.c.				N	Y (n.c.)	K24	V27	T22	ac									[17]
	H2Bv-P22	NT	n.c.				N														

H2B H2Bv	Detected in this study	AAs (Red	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
	Total	41	35	74	15	7	1	8	6	2	20	2	2	9	2
	Known	21	21	33	11	0	1	6	3	0	9	0	2	1	0
	Novel	20	14	41	4	7	0	2	3	2	11	2	0	8	2

Previously reported as modified	AAs (NR)	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
All Known	30	23	59	16	2	1	11	2	1	7	9	2	1	7
n.d. PTM on AAs conserved in Pf	18	18	27	6	1	0	5	0	0	1	7	1	0	6
n.d. (AA detected as modified by other PTMs)	12	12	19	4	0	0	4	0	0	1	3	1	0	6
n.d. (AA n.d. as modified)	6	6	8	2	1	0	1	0	0	0	4	0	0	0
n.c. AA	7	0	10	2	1	0	2	0	1	1	2	0	0	1

H3 H3.3	PfHistone-AA#PTM	Location	Total_ SpC	Modified_SpC	Mod/ Total (%)	Times Detected	Known in Pf?	Known in others?	Previously reported AA#PTMs												
									Hs/Mm	Sc	Dm	ac	me1	me2	me3	fo	phs	ub	oh	cr	References
1	H3-K9me1	NT	1657	48	2.9	11	Y (Ab)	Y													
	H3.3-K9me1	NT	1347	233	17.3	14	N	Y													
	H3-K9me2	NT	1657	91	5.49	12	N	Y													
	H3.3-K9me2	NT	1347	252	18.71	13	N	Y	K9			ac✓	me1✓	me2✓	me3✓					cr✓	[31, 32, 45-50]
	H3-K9me3	NT	1657	23	1.39	3	Y	Y													
	H3-K9ac	NT	1657	863	52.08	14	Y	Y													
	H3.3-K9ac	NT	1347	293	21.75	14	Y	Y													
	H3-K9cr	NT	1657	4	0.24	2	N	Y													
2	H3-S10ac	NT	1761	84	4.77	13	N	N (known AA	S10								phs				[21, 45, 50-54]
	H3.3-S10ac	NT	1423	18	1.26	10	N	N (known AA													
3	H3-T11ac	NT	1761	5	0.28	5	N	N (known AA	T11								phs				[21, 54, 55]
	H3.3-T11ac	NT	1423	8	0.56	4	N	N (known AA													
4	H3-K14fo	NT	1776	3	0.17	3	N	N (known AA													
	H3-K14me3	NT	1776	2	0.11	2	N	Y													
	H3-K14ac	NT	1776	1084	61.04	14	Y	Y	K14			ac✓	me1	me2	me3✓						[11, 17, 31, 45-47, 49, 50]
	H3.3-K14ac	NT	1438	1000	69.54	14	Y	Y													
5	H3-R17me1	NT	1890	187	9.89	14	Y	Y													
	H3.3-R17me1	NT	1553	123	7.92	14	Y	Y													
	H3-R17me2	NT	1890	54	2.86	13	Y	Y	R17				me1✓	me2✓							[49, 50, 56]
	H3.3-R17me2	NT	1553	30	1.93	10	Y	Y													
6	H3-K18oh	NT	2219	7	0.32	1	N	N (known AA													
	H3.3-K18oh	NT	2124	6	0.28	1	N	N (known AA													
	H3/H3.3-K18fo	NT	2219/2124	17	0.77	7	N	Y													
	H3-K18me1	NT	2219	32	1.44	9	N	Y	K18			ac✓	me1✓			fo✓				cr	[17, 31, 32, 46, 47]
	H3.3-K18me1	NT	2124	23	1.08	9	N	Y													
	H3-K18ac	NT	2219	419	18.88	14	Y	Y													
	H3.3-K18ac	NT	2124	390	18.36	14	Y	Y													
7	H3-S22ac	NT	2172	8	0.37	6	N	N	T22	S22	T22						phs				[21]
	H3.3-S22ac	NT	2120	8	0.38	7	N	N													
8	H3/H3.3-K23me1	NT	2172/2120	2	0.09	2	N	Y													
	H3-K23ac	NT	2172	420	19.34	14	Y	Y	K23			ac✓	me1✓			fo				cr	[17, 18, 31, 32, 46, 47, 50]
	H3.3-K23ac	NT	2120	419	19.76	14	Y	Y													
9	H3-K27fo	NT	839	10	1.19	7	N	N (known AA													
	H3-K27me1	NT	839	6	0.72	2	N	Y	K27			ac✓	me1✓	me2	me3					cr	[17, 31, 32, 45-47]
	H3-K27ac	NT	839	186	22.17	14	Y	Y													
	H3.3-K27ac	NT	548	145	26.46	14	Y	Y													
10	H3-S28ac	NT	1094	10	0.91	5	N	N (known AA													
	H3.3-S28ac	NT	622	3	0.48	3	N	N (known AA	S28								phs✓				[22, 45, 50-53, 57-59]
	H3-S28phs	NT	1094	7	0.64	3	Y	Y													
11	H3-S32phs	NT	1084	35	3.23	11	Y	Y	S31/T32								phs✓				[21, 22, 45, 47]
	H3.3-S32phs	NT	607	3	0.49	2	Y	Y													
12	H3-K36fo	NT	1055	3	0.28	2	N	N (known AA													
	H3.3-K36fo	NT	592	4	0.68	2	N	N (known AA													
	H3-K36me1	NT	1055	34	3.22	10	N	Y													
	H3-K36me2	NT	1055	3	0.28	2	N	Y	K36			ac✓	me1✓	me2✓	me3 (Ab)						[31, 32, 45-47, 60, 61]
	H3-K36ac	NT	1055	10	0.95	4	N	Y													
	H3.3-K36ac	NT	592	38	6.42	7	N	Y													
13	H3-K37fo	NT	142	11	7.75	6	N	N (known AA	K37				me1								[60]
	H3.3-K37fo	NT	153	3	1.96	3	N	N (known AA													
14	H3.3-R40me1	NT	70	5	7.14	2	N	N	R40												
15	H3/H3.3-Y41oh	NT	971	2	0.21	2	N	N (known AA	Y41								phs				[59]
16	H3/H3.3-R42fo	NT	971	158	16.27	14	N	N	R42	K42	R42										
	H3/H3.3-R42me2	NT	971	9	0.93	4	N	N													
17	H3-R53fo	CD	672	6	0.89	3	N	N													
	H3.3-K53fo	CD	1837	54	2.94	11	N	N													
	H3-K53oh	CD	1837	9	0.49	5	N	N (n.c.)	R53												
	H3.3-K53ac	CD	1837	12	0.65	7	N	N (n.c.)													
18	H3-Y54oh	CD	3414	5	0.15	3	N	N	Y54												
19	H3-K56fo	CD	3414	37	1.08	9	N	Y													
	H3.3-K56fo	CD	2029	6	0.3	3	N	Y	K56			ac✓	me1	me2	me3	fo✓				cr	[17, 46, 62-64]
	H3-K56ac	CD	3414	7	0.21	5	Y	Y													
20	H3-R72fo	CD	1860	224	12.04	10	N	N													
	H3.3-R72fo	CD	1514	143	9.45	9	N	N	R72												
	H3-R72me2	CD	1860	2	0.11	2	N	N													
21	H3-D77oh	CD	16290	4	0.02	4	N	N	D77												
22	H3-K79fo	CD	16358	170	1.04	11	N	Y													
	H3.3-K79fo	CD	12841	129	1	7	N	Y													
	H3-K79me2	CD	16358	85	0.52	3	N	Y	K79			ac✓	me1	me2✓	me3✓	fo✓				cr	[11, 17, 18, 31, 46, 47, 65]
	H3-K79me3	CD	16358	8	0.05	4	Y (Ab)	Y													
	H3-K79ac	CD	16358	29	0.18	10	N	Y													
23	H3-T80ac	CD	18858	87	0.46	9	N	Y	T80			STac✓					phs				[11, 66]
24	H3-R83fo	CD	20057	12	0.06	4	N	N (known AA	R83				me1	me2							[11]
25	H3.3-S86ac	CD	13319	20	0.15	8	N	N (known AA	S86								phs				[11]
26	H3.3-Y99oh	CD	4848	29	0.6	6	N	N	Y99												
27	H3-D106oh	CD	7656	2	0.03	2	N	N													
	H3.3-D106oh	CD	5082	205	4.03	9	N	N	D106												
28	H3-T107ac	CD	7709	5	0.06	4	N	N (known AA	T107								phs				[67]
29	H3-K115oh	CD	5658	2	0.04	2	N	N (known AA													
	H3.3-K115oh	CD	3086	34	1.1	8	N	N (known AA													
	H3/H3.3-K115fo	CD	5658/3086	6	0.11	3	N	N (known AA													
	H3-K115me1	CD	5658	4	0.07	3	N	N (known AA													
	H3-K115ac	CD	5658	8	0.14	5	N	Y													
	H3.3-K115ac	CD	3086	12	0.39	7	N	Y													
	H3-K115cr	CD	5658	309	5.46	10	N	N (known AA	K115			ac✓									[14]
	H3.3-K115cr	CD	3086	38	1.23	8	N	N (known AA													
30	H3/H3.3-R116fo	CD	2977/1437	12	0.4	4	N	N													
	H3-R116me2	CD	2977	2	0.07	2	N	N	R116												
	H3.3-R116me2	CD	1437	3	0.21	2	N	N													
31	H3/H3.3-K122fo	CD	5712	83	1.45	13	N	Y													
	H3/H3.3-K122me1	CD	5712	4	0.07	1	N	Y	K122			ac✓	me1✓			fo✓				cr	[11, 17, 18, 46, 47]
	H3/H3.3-K122ac	CD	5712	2	0.04	2	N	Y													
n.d.	H3/H3.3-R2	NT	4/4	0	0	2/2	N	Y	R2					me2							[68-70]
n.d.	H3/H3.3-T3	NT	4/4	0	0	2/2	N	Y	T3								phs				[45, 51]
n.d.	H3/H3.3-K4	NT	4/4	0	0	2/2	Y	Y	K4			ac	me1	me2	me3					cr	[17, 46, 47, 50]
n.d.	H3/H3.3-T6	NT	5/4	0	0	3/2	N	Y	T6								phs				[71]
n.d.	H3/H3.3-R8	NT	5/4	0	0	3/2	N	Y	R8				me1	me2							[72]
n.d.	H3/H3.3-R26	NT	2127/2086	0	0	14/14	N	Y	R26				me1								[14]
n.d.	H3/H3.3-T45	CD	1948/1948	0	0	14/14	N	Y	T45								phs				[15, 23]
n.d.	H3/H3.3-S57	CD	3469/2084	0	0	14/14	Y	Y	S57								phs				[21, 22, 66]
n.d.	H3/H3.3-R63	CD	3451/2025	0	0	14/14	N	Y	R63				me1								[17]
n.d.	H3/H3.3-K64	CD	55/55	0	0	11/11	N	Y	K64			ac	me1			fo					[18, 46, 47]
n.d.	H3/H3.3-T118	CD	5641/5641	0	0	14/14	N	Y	T118								phs				[73]
n.d.	H3/H3.3-R128	CD	5576/5576	0	0	14/14	N	Y	R128				me1								[74]
n.d.	H3/H3.3-R134	CD	5576/5576	0	0	14/14	N	Y	R134				me1								[11]

H3 H3.3	Detected in this study	AAs (Red	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
	Total	51	50	91	18	11	0	12	10	3	20	3	0	11	3
	Known	24	24	47	17	1	0	10	6	3	6	3	0	0	1
	Novel	27	26	44	1	10	0	2	4	0	14	0	0	11	2

Previously reported as modified	AAs (NR)	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
All Known	34	34	78	12	1	0	19	11	7	6	14	0	0	8
n.d. PTM on AAs conserved in Pf	29	29	45	2	0	0	12	7	3	2	12	0	0	7
n.d. (AA detected as modified by other PTMs)	16	16	25	0	0	0	5	4	2	1	7	0	0	6
n.d. (AA n.d. as modified)	13	13	20	2	0	0	7	3	1	1	5	0	0	1
n.c. AA	0	0	0	0	0	0	0	0	0	0	0	0	0	0

H3.c	PfHistone-AA#PTM	Location	Total_ SpC	Modified_SpC	Mod/ Total (%)	Times Detected	Known in Pf?	Known in others?	Previously reported AA#PTMs												
									Hs/Mm	Sc	Dm	ac	me1	me2	me3	fo	phs	ub	oh	cr	References
1	H3.c-K23ac	NT	63	8	12.698	3	N	N (n.c.)	n.c.	n.c.	n.c.										
2	H3.c-K26ac	NT	63	16	25.397	6	N	N (n.c.)	n.c.	n.c.	n.c.										
3	H3.c-K60ac	NT	54	7	12.963	4	N	N (n.c.)	n.c.	n.c.	n.c.										
4	H3.c-R108fo	CD	205	12	5.8537	3	N	N	K76	n.c.	K163										

H3.c	Detected in this study	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
	Total	4	1	4	3	0	0	0	0	0	1	0	0	0	0
	Known	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Novel	4	1	4	3	0	0	0	0	0	1	0	0	0	0

Previously reported as modified	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
All Known	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n.d. PTM on AAs conserved in Pf	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n.d. (AA detected as modified by other PTMs)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n.d. (AA n.d. as modified)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n.c. AA	0	0	0	0	0	0	0	0	0	0	0	0	0	0

H4	PfHistone-AA#PTM	Location	Total_ SpC	Modified_SpC	Mod/ Total (%)	Times Detected	Known in Pf?	Known in others?	Previously reported AA#PTMs												
									Hs/Mm	Sc	Dm	ac	me1	me2	me3	fo	phs	ub	oh	cr	References
1	H4-S1Nac	NT	106	106	100	14	Y	Y	S1			Nac✓					phs				[4, 73]
2	H4-R3me1	NT	106	71	66.98	12	Y	Y	R3				me1✓	me2	me3						[11, 74-76]
3	H4-K5ac	NT	758	719	94.85	14	Y	Y	K5			ac✓	me1		me3					cr	[4, 11, 13, 17, 30, 77, 78]
4	H4-K8ac	NT	1059	803	75.83	14	Y	Y	K8			ac✓								cr	[4, 13, 17, 30, 77, 78]
5	H4-K12ac	NT	1221	969	79.36	14	Y	Y	K12			ac✓	me1			fo				cr	[4, 13, 17, 30, 76, 77]
6	H4-K16ac	NT	1165	406	34.85	14	Y	Y	K16			ac✓	me1		me3					cr	[11, 13, 17, 30, 73, 77, 78]
7	H4-K31fo	CD	2258	7	0.31	3	N	Y													
	H4-K31ac	CD	2258	9	0.4	4	N	Y	K31			ac✓				fo✓					[13, 17, 18]
	H4-K31ub	CD	2258	4	0.18	1	N	N (known AA													
8	H4-R35me1	CD	2209	3	0.14	1	N	Y	R35				me1✓								[17]
9	H4-R45fo	CD	4529	163	3.6	12	N	N	R45												
10	H4-S47ac	CD	17970	5	0.03	4	N	N (known AA	S47								phs				[56, 65, 79-82]
11	H4-Y51oh	CD	18534	26	0.14	8	N	Y	Y51								phs		oh✓		[17, 83]
12	H4-R55me1	CD	18480	8	0.04	3	N	Y	R55				me1✓								[17]
13	H4-K59fo	CD	7123	154	2.16	11	N	Y	K59			ac✓	me1	me2		fo✓					[13, 17, 18, 22]
	H4-K59ac	CD	7123	2	0.03	2	N	Y													
14	H4-K67oh	CD	12670	6	0.05	5	N	N (n.c.)													
	H4-K67fo	CD	12670	157	1.24	13	N	N (known AA													
	H4-K67me1	CD	12670	56	0.44	3	N	Y	R67				me1✓								[17]
	H4-K67me2	CD	12670	20	0.16	4	N	N (known AA													
	H4-K67me3	CD	12670	6	0.05	1	N	N (n.c.)													
	H4-K67ac	CD	12670	132	1.04	10	N	N (n.c.)													
15	H4-D68oh	CD	12666	11	0.09	8	N	N	D68												
16	H4-Y72oh	CD	12237	41	0.34	8	N	N	Y72												
17	H4-K77oh	CD	8372	35	0.42	8	N	N (known AA													
	H4-K77fo	CD	8372	11	0.13	3	N	Y	K77			ac	me1✓			fo✓					[17, 18, 22]
	H4-K77me1	CD	8372	77	0.92	2	N	Y													
18	H4-K79fo	CD	11516	193	1.68	10	N	Y													
	H4-K79me2	CD	11516	3	0.03	2	N	N (known AA	K79			ac✓				fo✓					[17, 18, 84]
	H4-K79ac	CD	11516	442	3.84	13	N	Y													
19	H4-T80ac	CD	14469	84	0.58	13	N	N (known AA	T80								phs				[85]
20	H4-D85oh	CD	14407	45	0.31	11	N	N	D85												
21	H4-Y88oh	CD	14366	50	0.35	10	N	Y	Y88								phs		oh✓		[17, 83]
22	H4-S89ac	CD	14366	2	0.01	2	N	N (n.c.)	A89												
23	H4-K91fo	CT	14368	52	0.36	13	N	Y				ac✓				fo✓				cr	[17, 18, 86, 87]
	H4-K91me1	CT	14368	8	0.06	2	N	N (known AA	K91												
	H4-K91ac	CT	14368	5	0.03	5	N	Y													
n.d.	H4-R17	NT	897	0	0	14	Y	Y	R17				me1	me2							[2, 11]
n.d.	H4-R19	NT	0	0	0	0	N	Y	R19				me1	me2	me3						[11]
n.d.	H4-K20	NT	5	0	0	2	Y	Y	K20			ac	me1 (Ab)	me2 (Ab)	me3						[2, 4, 73, 88-90]
n.d.	H4-R23	NT	46	0	0	11	N	Y	R23				me1	me2	me3						[11]
n.d.	H4-T82	CD	14451	0	0	14	N	Y	T82								phs				[85]
n.d.	H4-R92	CT	11714	0	0	14	N	Y	R92				me1								[15]
n.c.	H4-V60	CD	n.c				N	Y (n.c.)	V60	S60	V60						phs				[91]
n.c.	H4-N64	CD	n.c				N	Y (n.c.)	N64	S64	N64						phs				[92]

H4	Detected in this study	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
	Total	23	21	37	9	3	1	6	2	1	7	0	1	7	0
	Known	16	16	21	8	0	1	5	0	0	5	0	0	2	0
	Novel	5	4	16	1	3	0	1	2	1	2	0	1	5	0

Previously reported as modified	AAs	Conserved	PTMs	Kac	STac	Nac	me1	me2	me3	fo	phs	ub	oh	cr
All Known	26	24	58	10	0	1	14	6	6	6	8	0	2	5
n.d. PTM on AAs conserved in Pf	19	19	33	2	0	0	8	5	6	1	6	0	0	5
n.d. (AA detected as modified by other PTMs)	13	13	21	1	0	0	4	2	3	1	5	0	0	5
n.d. (AA n.d. as modified)	6	6	12	1	0	0	4	3	3	0	1	0	0	0
n.c. AA	2	0	2	0	0	0	0	0	0	0	2	0	0	0

References

1 Trelle, M.B., et al., *Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum.* J Proteome Res, 2009. **8**(7): p. 3439-50.
2 Cui, L. and J. Miao, *Chromatin-mediated epigenetic regulation in the malaria parasite Plasmodium falciparum.* Eukaryot Cell, 2010. **9**(8): p. 1138-49.
3 Barber, C.M., et al., *The enhancement of histone H4 and H2A serine 1 phosphorylation during mitosis and S-phase is evolutionarily conserved.* Chromosoma, 2004. **112**(7): p. 360-71.
4 Bonenfant, D., et al., *Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry.* Mol Cell Proteomics, 2006. **5**(3): p. 541-52.
5 Boyne, M.T., 2nd, et al., *Precise characterization of human histones in the H2A gene family by top down mass spectrometry.* J Proteome Res, 2006. **5**(2): p. 248-53.
6 Galasinski, S.C., et al., *Global regulation of post-translational modifications on core histones.* J Biol Chem, 2002. **277**(4): p. 2579-88.
7 Hagiwara, T., Y. Hidaka, and M. Yamada, *Deimination of histone H2A and H4 at arginine 3 in HL-60 granulocytes.* Biochemistry, 2005. **44**(15): p. 5827-34.
8 Song, O.K., et al., *An Nalpha-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A.* J Biol Chem, 2003. **278**(40): p. 38109-12.
9 Zhang, Y., et al., *Phosphorylation of histone H2A inhibits transcription on chromatin templates.* J Biol Chem, 2004. **279**(21): p. 21866-72.
10 Clarke, A.S., et al., *Esa1p is an essential histone acetyltransferase required for cell cycle progression.* Mol Cell Biol, 1999. **19**(4): p. 2515-26.
11 Tweedie-Cullen, R.Y., et al., *Identification of combinatorial patterns of post-translational modifications on individual histones in the mouse brain.* PLoS One, 2012. **7**(5): p. e36980.
12 Suka, N., et al., *Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin.* Mol Cell, 2001. **8**(2): p. 473-9.
13 Choudhary, C., et al., *Lysine acetylation targets protein complexes and co-regulates major cellular functions.* Science, 2009. **325**(5942): p. 834-40.
14 Lall, S., *Primers on chromatin.* Nat Struct Mol Biol, 2007. **14**(11): p. 1110-5.
15 Portela, A. and M. Esteller, *Epigenetic modifications and human disease.* Nat Biotechnol, 2010. **28**(10): p. 1057-68.
16 Bheda, P., et al., *Biotinylation of lysine method identifies acetylated histone H3 lysine 79 in Saccharomyces cerevisiae as a substrate for Sir2.* Proc Natl Acad Sci U S A, 2012. **109**(16): p. E916-25.
17 Tan, M., et al., *Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification.* Cell, 2011. **146**(6): p. 1016-28.
18 Wisniewski, J.R., A. Zougman, and M. Mann, *Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function.* Nucleic Acids Res, 2008. **36**(2): p. 570-7.
19 Wang, H., et al., *Role of histone H2A ubiquitination in Polycomb silencing.* Nature, 2004. **431**(7010): p. 873-8.
20 Wyatt, H.R., et al., *Multiple Roles for Saccharomyces cerevisiae Histone H2A in Telomere Position Effect, Spt Phenotypes and Double-Strand-Break Repair.* Genetics, 2003. **164**(1): p. 47-64.
21 Dastidar, E.G., et al., *Comprehensive histone phosphorylation analysis and identification of pf14-3-3 protein as a histone h3 phosphorylation reader in malaria parasites.* PLoS One, 2013. **8**(1): p. e53179.
22 Treeck, M., et al., *The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries.* Cell Host Microbe, 2011. **10**(4): p. 410-9.
23 Rossetto, D., N. Avvakumov, and J. Cote, *Histone phosphorylation: A chromatin modification involved in diverse nuclear events.* Epigenetics, 2012. **7**(10): p. 1098-108.
24 Zhang, L., et al., *Identification of novel histone post-translational modifications by peptide mass fingerprinting.* Chromosoma, 2003. **112**(2): p. 77-86.
25 Aihara, H., et al., *Nucleosomal histone kinase-1 phosphorylates H2A Thr 119 during mitosis in the early Drosophila embryo.* Genes Dev, 2004. **18**(8): p. 877-88.
26 Downs, J.A., N.F. Lowndes, and S.P. Jackson, *A role for Saccharomyces cerevisiae histone H2A in DNA repair.* Nature, 2000. **408**(6815): p. 1001-4.
27 Ficarro, S.B., et al., *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae.* Nat Biotechnol, 2002. **20**(3): p. 301-5.
28 Shroff, R., et al., *Distribution and dynamics of chromatin modification induced by a defined DNA double-strand break.* Curr Biol, 2004. **14**(19): p. 1703-11.
29 Babiarz, J.E., J.E. Halley, and J. Rine, *Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomyces cerevisiae.* Genes Dev, 2006. **20**(6): p. 700-10.
30 Millar, C.B., et al., *Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast.* Genes Dev, 2006. **20**(6): p. 711-22.
31 Beck, H.C., et al., *Quantitative proteomic analysis of post-translational modifications of human histones.* Mol Cell Proteomics, 2006. **5**(7): p. 1314-25.
32 Kim, S.C., et al., *Substrate and functional diversity of lysine acetylation revealed by a proteomics survey.* Mol Cell, 2006. **23**(4): p. 607-18.
33 Jason, L.J., et al., *Histone ubiquitination: a tagging tail unfolds?* Bioessays, 2002. **24**(2): p. 166-74.
34 Madigan, J.P., H.L. Chotkowski, and R.L. Glaser, *DNA double-strand break-induced phosphorylation of Drosophila histone variant H2Av helps prevent radiation-induced apoptosis.* Nucleic Acids Res, 2002. **30**(17): p. 3698-705.
35 Golebiowski, F. and K.S. Kasprzak, *Inhibition of core histones acetylation by carcinogenic nickel(II).* Mol Cell Biochem, 2005. **279**(1-2): p. 133-9.
36 Kurdistani, S.K., S. Tavazoie, and M. Grunstein, *Mapping global histone acetylation patterns to gene expression.* Cell, 2004. **117**(6): p. 721-33.
37 Maile, T., et al., *TAF1 activates transcription by phosphorylation of serine 33 in histone H2B.* Science, 2004. **304**(5673): p. 1010-4.
38 Pavri, R., et al., *Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II.* Cell, 2006. **125**(4): p. 703-17.
39 Zhu, B., et al., *Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation.* Mol Cell, 2005. **20**(4): p. 601-11.
40 Nathan, D., et al., *Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications.* Genes Dev, 2006. **20**(8): p. 966-76.
41 Miao, J., et al., *The malaria parasite Plasmodium falciparum histones: organization, expression, and acetylation.* Gene, 2006. **369**: p. 53-65.
42 Ahn, S.H., et al., *Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in S. cerevisiae.* Cell, 2005. **120**(1): p. 25-36.
43 Ahn, S.H., et al., *H2B (Ser10) phosphorylation is induced during apoptosis and meiosis in S. cerevisiae.* Cell Cycle, 2005. **4**(6): p. 780-3.
44 Cheung, W.L., et al., *Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase.* Cell, 2003. **113**(4): p. 507-17.
45 Garcia, B.A., et al., *Modifications of human histone H3 variants during mitosis.* Biochemistry, 2005. **44**(39): p. 13202-13.
46 Garcia, B.A., et al., *Organismal differences in post-translational modifications in histones H3 and H4.* J Biol Chem, 2007. **282**(10): p. 7641-55.
47 Hake, S.B., et al., *Expression patterns and post-translational modifications associated with mammalian histone H3 variants.* J Biol Chem, 2006. **281**(1): p. 559-68.
48 Lachner, M., et al., *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins.* Nature, 2001. **410**(6824): p. 116-20.
49 Miao, F., et al., *Coactivator-associated arginine methyltransferase-1 enhances nuclear factor-kappaB-mediated gene transcription through methylation of histone H3 at arginine 17.* Mol Endocrinol, 2006. **20**(7): p. 1562-73.
50 Thomas, C.E., N.L. Kelleher, and C.A. Mizzen, *Mass spectrometric characterization of human histone H3: a bird's eye view.* J Proteome Res, 2006. **5**(2): p. 240-7.
51 Dai, J., et al., *The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment.* Genes Dev, 2005. **19**(4): p. 472-88.
52 Goto, H., et al., *Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation.* J Biol Chem, 1999. **274**(36): p. 25543-9.
53 Goto, H., et al., *Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation.* Genes Cells, 2002. **7**(1): p. 11-7.
54 Preuss, U., G. Landsberg, and K.H. Scheidtmann, *Novel mitosis-specific phosphorylation of histone H3 at Thr11 mediated by Dlk/ZIP kinase.* Nucleic Acids Res, 2003. **31**(3): p. 878-85.
55 Shimada, M., et al., *Chk1 is a histone H3 threonine 11 kinase that regulates DNA damage-induced transcriptional repression.* Cell, 2008. **132**(2): p. 221-32.
56 Ananthanarayanan, M., et al., *Ligand-dependent activation of the farnesoid X-receptor directs arginine methylation of histone H3 by CARM1.* J Biol Chem, 2004. **279**(52): p. 54348-57.
57 Choi, H.S., et al., *Phosphorylation of Ser28 in histone H3 mediated by mixed lineage kinase-like mitogen-activated protein triple kinase alpha.* J Biol Chem, 2005. **280**(14): p. 13545-53.
58 Wang, B., et al., *Evaluation of the low-specificity protease elastase for large-scale phosphoproteome analysis.* Anal Chem, 2008. **80**(24): p. 9526-33.
59 Yu, L.R., et al., *Improved titanium dioxide enrichment of phosphopeptides from HeLa cells and high confident phosphopeptide identification by cross-validation of MS/MS and MS/MS/MS spectra.* J Proteome Res, 2007. **6**(11): p. 4150-62.
60 Coon, J.J., et al., *Protein identification using sequential ion/ion reactions and tandem mass spectrometry.* Proc Natl Acad Sci U S A, 2005. **102**(27): p. 9463-8.
61 Morris, S.A., et al., *Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification.* J Biol Chem, 2007. **282**(10): p. 7632-40.
62 Masumoto, H., et al., *A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response.* Nature, 2005. **436**(7048): p. 294-8.
63 Xu, F., K. Zhang, and M. Grunstein, *Acetylation in histone H3 globular domain regulates gene expression in yeast.* Cell, 2005. **121**(3): p. 375-85.
64 Yu, Y., et al., *Histone H3 lysine 56 methylation regulates DNA replication through its interaction with PCNA.* Mol Cell, 2012. **46**(1): p. 7-17.
65 Huyen, Y., et al., *Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks.* Nature, 2004. **432**(7015): p. 406-11.
66 Vermeulen, M., et al., *Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers.* Cell, 2010. **142**(6): p. 967-80.
67 Mayya, V., et al., *Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions.* Sci Signal, 2009. **2**(84): p. ra46.
68 Guccione, E., et al., *Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive.* Nature, 2007. **449**(7164): p. 933-7.
69 Hyllus, D., et al., *PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation.* Genes Dev, 2007. **21**(24): p. 3369-80.
70 Iberg, A.N., et al., *Arginine methylation of the histone H3 tail impedes effector binding.* J Biol Chem, 2008. **283**(6): p. 3006-10.
71 Metzger, E., et al., *Phosphorylation of histone H3T6 by PKCbeta(I) controls demethylation at histone H3K4.* Nature, 2010. **464**(7289): p. 792-6.
72 Pal, S., et al., *Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes.* Mol Cell Biol, 2004. **24**(21): p. 9630-45.
73 Freitas, M.A., A.R. Sklenar, and M.R. Parthun, *Application of mass spectrometry to the identification and quantification of histone post-translational modifications.* Journal of Cellular Biochemistry, 2004. **92**(4): p. 691-700.
74 Wang, C.M., et al., *Identification of histone methylation multiplicities patterns in the brain of senescence-accelerated prone mouse 8.* Biogerontology, 2010. **11**(1): p. 87-102.
75 Pesavento, J.J., et al., *Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle.* Mol Cell Biol, 2008. **28**(1): p. 468-86.
76 Strahl, B.D., et al., *Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1.* Curr Biol, 2001. **11**(12): p. 996-1000.
77 Wang, H., et al., *Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor.* Science, 2001. **293**(5531): p. 853-7.
78 Wang, Y., et al., *Human PAD4 regulates histone arginine methylation levels via demethylimination.* Science, 2004. **306**(5694): p. 279-83.
79 O'Neill, L.P. and B.M. Turner, *Histone H4 acetylation distinguishes coding regions of the human genome from heterochromatin in a differentiation-dependent but transcription-independent manner.* EMBO J, 1995. **14**(16): p. 3946-57.
80 Turner, B.M., L.P. O'Neill, and I.M. Allan, *Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies.* FEBS Lett, 1989. **253**(1-2): p. 141-5.
81 Beranova-Giorgianni, S., et al., *Phosphoproteomic analysis of the human pituitary.* Pituitary, 2006. **9**(2): p. 109-20.
82 Dephoure, N., et al., *A quantitative atlas of mitotic phosphorylation.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10762-7.
83 Kang, B., et al., *Phosphorylation of H4 Ser 47 promotes HIRA-mediated nucleosome assembly.* Genes Dev, 2011. **25**(13): p. 1359-64.
84 Olsen, J.V., et al., *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks.* Cell, 2006. **127**(3): p. 635-48.
85 Rikova, K., et al., *Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer.* Cell, 2007. **131**(6): p. 1190-203.

86 Yang, X., et al., *HAT4, a Golgi apparatus-anchored B-type histone acetyltransferase, acetylates free histone H4 and facilitates chromatin assembly.* Mol Cell, 2011. **44**(1): p. 39-50.
87 Zhai, B., et al., *Phosphoproteome analysis of Drosophila melanogaster embryos.* J Proteome Res, 2008. **7**(4): p. 1675-82.
88 Yan, Q., et al., *BBAP monoubiquitylates histone H4 at lysine 91 and selectively modulates the DNA damage response.* Mol Cell, 2009. **36**(1): p. 110-20.
89 Ye, J., et al., *Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly.* Mol Cell, 2005. **18**(1): p. 123-30.
90 Nishioka, K., et al., *PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin.* Mol Cell, 2002. **9**(6): p. 1201-13.
91 Ong, S.E., G. Mittler, and M. Mann, *Identifying and quantifying in vivo methylation sites by heavy methyl SILAC.* Nat Methods, 2004. **1**(2): p. 119-26.
92 Yin, Y., et al., *SET8 recognizes the sequence RHRK20VLRDN within the N terminus of histone H4 and mono-methylates lysine 20.* J Biol Chem, 2005. **280**(34): p. 30025-31.
93 Albuquerque, C.P., et al., *A multidimensional chromatography technology for in-depth phosphoproteome analysis.* Mol Cell Proteomics, 2008. **7**(7): p. 1389-96.
94 Chi, A., et al., *Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2193-8.

Acknowledgments

This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID) (grant R01 AI85077-01A1 to KLR), the National Institutes of Health (fellowship F31 AI096840-01 to SC), the Human Frontier Science Program (grant LT00507/2011-L to EMB) and the Stowers Institute for Medical Research (AS, MES, WZ, JMV, MPW, LF). We thank Edwin Smith and Ali Shilatifard for the gift of the H3-K4 me1/me2/me3 and H3-K27ac antibodies and comments on our manuscript.

Author Contributions

KLR and LF conceived and designed the experiments. SC, AS, JP, and DWDC performed the experiments. MS, ZW, MPW, JV, contributed reagents, materials, and analysis tools. LF, AS, NP and KLR analyzed the data. LF and KLR wrote the paper. All authors edited and approved the manuscript.

References

Ahn, S.H., Cheung, W.L., Hsu, J.Y., Diaz, R.L., Smith, M.M., and Allis, C.D. (2005a). Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in S. cerevisiae. Cell *120*, 25-36.

Ahn, S.H., Henderson, K.A., Keeney, S., and Allis, C.D. (2005b). H2B (Ser10) phosphorylation is induced during apoptosis and meiosis in S. cerevisiae. Cell Cycle *4*, 780-783.

Bartfai, R., Hoeijmakers, W.A., Salcedo-Amaya, A.M., Smits, A.H., Janssen-Megens, E., Kaan, A., Treeck, M., Gilberger, T.W., Francoijs, K.J., and Stunnenberg, H.G. (2010). H2A.Z demarcates intergenic regions of the plasmodium falciparum epigenome that are dynamically marked by H3K9ac and H3K4me3. PLoS Pathog *6*, e1001223.

Beck, H.C., Nielsen, E.C., Matthiesen, R., Jensen, L.H., Sehested, M., Finn, P., Grauslund, M., Hansen, A.M., and Jensen, O.N. (2006). Quantitative proteomic analysis of post-translational modifications of human histones. Mol Cell Proteomics *5*, 1314-1325.

Caraux, G., and Pinloche, S. (2005). PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. Bioinformatics *21*, 1280-1281.

Cheung, W.L., Ajiro, K., Samejima, K., Kloc, M., Cheung, P., Mizzen, C.A., Beeser, A., Etkin, L.D., Chernoff, J., Earnshaw, W.C.*, et al.* (2003). Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. Cell *113*, 507-517.

Cheung, W.L., Turner, F.B., Krishnamoorthy, T., Wolner, B., Ahn, S.H., Foley, M., Dorsey, J.A., Peterson, C.L., Berger, S.L., and Allis, C.D. (2005). Phosphorylation of histone H4 serine 1 during DNA damage requires casein kinase II in S. cerevisiae. Curr Biol *15*, 656-660.

Choi, H.S., Choi, B.Y., Cho, Y.Y., Zhu, F., Bode, A.M., and Dong, Z. (2005). Phosphorylation of Ser28 in histone H3 mediated by mixed lineage kinase-like mitogen-activated protein triple kinase alpha. J Biol Chem *280*, 13545-13553.

Chopra, R.K., and Ananthanarayanan, V.S. (1982). Conformational implications of enzymatic proline hydroxylation in collagen. Proc Natl Acad Sci U S A *79*, 7180-7184.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.

Cui, L., Fan, Q., and Miao, J. (2008). Histone lysine methyltransferases and demethylases in Plasmodium falciparum. International journal for parasitology *38*, 1083-1097.

Cui, L., and Miao, J. (2010). Chromatin-mediated epigenetic regulation in the malaria parasite Plasmodium falciparum. Eukaryot Cell *9*, 1138-1149.

Dastidar, E.G., Dzeyk, K., Krijgsveld, J., Malmquist, N.A., Doerig, C., Scherf, A., and Lopez-Rubio, J.J. (2013). Comprehensive histone phosphorylation analysis and identification of pf14-3-3 protein as a histone h3 phosphorylation reader in malaria parasites. PloS one *8*, e53179.

Duraisingh, M.T., Voss, T.S., Marty, A.J., Duffy, M.F., Good, R.T., Thompson, J.K., Freitas-Junior, L.H., Scherf, A., Crabb, B.S., and Cowman, A.F. (2005). Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. Cell *121*, 13-24.

Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A.A., Cheung, M.S., Day, D.S., Gadel, S., Gorchakov, A.A*., et al.* (2011). An assessment of histone-modification antibody quality. Nature structural & molecular biology *18*, 91-93.

Eng, J., McCormack, A.L., and Yates, J.R., III (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Amer Mass Spectrom *5*, 976-989.

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

Flueck, C., Bartfai, R., Volz, J., Niederwieser, I., Salcedo-Amaya, A.M., Alako, B.T., Ehlgen, F., Ralph, S.A., Cowman, A.F., Bozdech, Z*., et al.* (2009). Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. PLoS Pathog *5*, e1000569.

Freitas-Junior, L.H., Hernandez-Rivas, R., Ralph, S.A., Montiel-Condado, D., Ruvalcaba-Salazar, O.K., Rojas-Meza, A.P., Mancio-Silva, L., Leal-Silvestre, R.J., Gontijo, A.M., Shorte, S.*, et al.* (2005). Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. Cell *121*, 25-36.

Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L., and Strahl, B.D. (2011). Influence of combinatorial histone modifications on antibody and effector protein recognition. Current biology : CB *21*, 53-58.

Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B.D., Allis, C.D.*, et al.* (2007). Organismal differences in post-translational modifications in histones H3 and H4. J Biol Chem *282*, 7641-7655.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S.*, et al.* (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. Nature *419*, 498-511.

Golebiowski, F., and Kasprzak, K.S. (2005). Inhibition of core histones acetylation by carcinogenic nickel(II). Mol Cell Biochem *279*, 133-139.

Gopalakrishnan, A.M., Nyindodo, L.A., Ross Fergus, M., and Lopez-Estrano, C. (2009). Plasmodium falciparum: Preinitiation complex occupancy of active and inactive promoters during erythrocytic stage. Exp Parasitol *121*, 46-54.

Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., Okawa, K., Iwamatsu, A., Okigaki, T., Takahashi, T.*, et al.* (1999). Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. J Biol Chem *274*, 25543-25549.

Goto, H., Yasui, Y., Nigg, E.A., and Inagaki, M. (2002). Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. Genes Cells *7*, 11-17.

Halbert, J., Ayong, L., Equinet, L., Le Roch, K., Hardy, M., Goldring, D., Reininger, L., Waters, N., Chakrabarti, D., and Doerig, C. (2010). A Plasmodium falciparum transcriptional cyclin-dependent kinase-related kinase with a crucial role in parasite proliferation associates with histone deacetylase activity. Eukaryot Cell *9*, 952-959.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*, 973-977.

Issar, N., Ralph, S.A., Mancio-Silva, L., Keeling, C., and Scherf, A. (2009). Differential sub-nuclear localisation of repressive and activating histone methyl modifications in P. falciparum. Microbes and infection / Institut Pasteur *11*, 403-407.

Jiang, L., Smith, J.N., Anderson, S.L., Ma, P., Mizzen, C.A., and Kelleher, N.L. (2007a). Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry. LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. The Journal of biological chemistry *282*, 27923-27934.

Jiang, T., Zhou, X., Taghizadeh, K., Dong, M., and Dedon, P.C. (2007b). N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. Proc Natl Acad Sci U S A *104*, 60-65.

Kaelin, W.G. (2005). Proline hydroxylation and gene expression. Annu Rev Biochem *74*, 115-128.

Keen, J.C., Yan, L., Mack, K.M., Pettit, C., Smith, D., Sharma, D., and Davidson, N.E. (2003). A novel histone deacetylase inhibitor, scriptaid, enhances expression of functional estrogen receptor alpha (ER) in ER negative human breast cancer cells in combination with 5-aza 2'-deoxycytidine. Breast cancer research and treatment *81*, 177-186.

Kim, S.C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L.*, et al.* (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol Cell *23*, 607-618.

Lall, S. (2007). Primers on chromatin. Nat Struct Mol Biol *14*, 1110-1115.

Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J.*, et al.* (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. Science *301*, 1503-1508.

Longhurst, H.J., and Holder, A.A. (1997). The histones of Plasmodium falciparum: identification, purification and a possible role in the pathology of malaria. Parasitology *114 (Pt 5)*, 413-419.

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

Maile, T., Kwoczynski, S., Katzenberger, R.J., Wassarman, D.A., and Sauer, F. (2004). TAF1 activates transcription by phosphorylation of serine 33 in histone H2B. Science *304*, 1010-1014.

Martin, C., and Zhang, Y. (2007). Mechanisms of epigenetic inheritance. Curr Opin Cell Biol *19*, 266-272.

McDonald, W.H., Tabb, D.L., Sadygov, R.G., MacCoss, M.J., Venable, J., Graumann, J., Johnson, J.R., Cociorva, D., and Yates, J.R., 3rd (2004). MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. Rapid Commun Mass Spectrom *18*, 2162-2168.

Mittal, R., Peak-Chew, S.Y., and McMahon, H.T. (2006). Acetylation of MEK2 and I kappa B kinase (IKK) activation loop residues by YopJ inhibits signaling. Proc Natl Acad Sci U S A *103*, 18574-18579.

Mukherjee, S., Keitany, G., Li, Y., Wang, Y., Ball, H.L., Goldsmith, E.J., and Orth, K. (2006). Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. Science *312*, 1211-1214.

Murray, C.J., Rosenfeld, L.C., Lim, S.S., Andrews, K.G., Foreman, K.J., Haring, D., Fullman, N., Naghavi, M., Lozano, R., and Lopez, A.D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. Lancet *379*, 413-431.

Nakanishi, S., Sanderson, B.W., Delventhal, K.M., Bradford, W.D., Staehling-Hampton, K., and Shilatifard, A. (2008). A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. Nat Struct Mol Biol *15*, 881-888.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635-648.

Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatifard, A., and Reinberg, D. (2006). Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. Cell *125*, 703-717.

Pesavento, J.J., Yang, H., Kelleher, N.L., and Mizzen, C.A. (2008). Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. Mol Cell Biol *28*, 468-486.

Ponts, N., Harris, E.Y., Lonardi, S., and Le Roch, K.G. (2011). Nucleosome occupancy at transcription start sites in the human malaria parasite: a hard-wired evolution of virulence? Infect Genet Evol *11*, 716-724.

Ponts, N., Harris, E.Y., Prudhomme, J., Wick, I., Eckhardt-Ludka, C., Hicks, G.R., Hardiman, G., Lonardi, S., and Le Roch, K.G. (2010). Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res *20*, 228-238.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nature biotechnology *28*, 1057-1068.

Preuss, U., Landsberg, G., and Scheidtmann, K.H. (2003). Novel mitosis-specific phosphorylation of histone H3 at Thr11 mediated by Dlk/ZIP kinase. Nucleic Acids Res *31*, 878-885.

Ralph, S.A., Scheidig-Benatar, C., and Scherf, A. (2005). Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. Proceedings of the National Academy of Sciences of the United States of America *102*, 5414-5419.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol *8*, 983-994.

Salcedo-Amaya, A.M., van Driel, M.A., Alako, B.T., Trelle, M.B., van den Elzen, A.M., Cohen, A.M., Janssen-Megens, E.M., van de Vegte-Bolmer, M., Selzer, R.R., Iniguez, A.L.*, et al.* (2009). Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proc Natl Acad Sci U S A *106*, 9655-9660.

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

Shetty, P. (2012). The numbers game. Nature *484*, S14-15.

Smith, C.M., Gafken, P.R., Zhang, Z., Gottschling, D.E., Smith, J.B., and Smith, D.L. (2003). Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. Analytical biochemistry *316*, 23-33.

Smith, E., Lin, C., and Shilatifard, A. (2011). The super elongation complex (SEC) and MLL in development and disease. Genes Dev *25*, 661-672.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Stubbs, J., Simpson, K.M., Triglia, T., Plouffe, D., Tonkin, C.J., Duraisingh, M.T., Maier, A.G., Winzeler, E.A., and Cowman, A.F. (2005). Molecular mechanism for switching of P. falciparum invasion pathways into human erythrocytes. Science *309*, 1384-1387.

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

Tagwerker, C., Flick, K., Cui, M., Guerrero, C., Dou, Y., Auer, B., Baldi, P., Huang, L., and Kaiser, P. (2006). A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivocross-linking. Mol Cell Proteomics *5*, 737-748.

Takahashi, Y.H., Lee, J.S., Swanson, S.K., Saraf, A., Florens, L., Washburn, M.P., Trievel, R.C., and Shilatifard, A. (2009). Regulation of H3K4 trimethylation via Cps40 (Spp1) of COMPASS is monoubiquitination independent: implication for a Phe/Tyr switch by the catalytic domain of Set1. Mol Cell Biol *29*, 3478-3486.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N.*, et al.* (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell *146*, 1016-1028.

Tasaki, T., and Kwon, Y.T. (2007). The mammalian N-end rule pathway: new insights into its components and physiological roles. Trends Biochem Sci *32*, 520-528.

Thorne, A.W., Kmiciek, D., Mitchelson, K., Sautiere, P., and Crane-Robinson, C. (1990). Patterns of histone acetylation. European journal of biochemistry / FEBS *193*, 701-713.

Tonkin, C.J., Carret, C.K., Duraisingh, M.T., Voss, T.S., Ralph, S.A., Hommel, M., Duffy, M.F., Silva, L.M., Scherf, A., Ivens, A.*, et al.* (2009). Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. PLoS biology *7*, e84.

Treeck, M., Sanders, J.L., Elias, J.E., and Boothroyd, J.C. (2011). The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries. Cell host & microbe *10*, 410-419.

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

Tweedie-Cullen, R.Y., Brunner, A.M., Grossmann, J., Mohanna, S., Sichau, D., Nanni, P., Panse, C., and Mansuy, I.M. (2012). Identification of combinatorial patterns of post-translational modifications on individual histones in the mouse brain. PloS one *7*, e36980.

Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol *19*, 242-247.

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

Wisniewski, J.R., Zougman, A., and Mann, M. (2008). Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. Nucleic Acids Res *36*, 570-577.

Wong, J.A.H.a.M.A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. The Journal of the Royal Statistical Society Series C (Applied Statistics) *28*, 100-108.

Xiang, Y., Takeo, S., Florens, L., Hughes, S.E., Huo, L.J., Gilliland, W.D., Swanson, S.K., Teeter, K., Schwartz, J.W., Washburn, M.P.*, et al.* (2007). The

inhibition of polo kinase by matrimony maintains G2 arrest in the meiotic cell cycle. PLoS Biol *5*, e323.

Yang, X., Yu, W., Shi, L., Sun, L., Liang, J., Yi, X., Li, Q., Zhang, Y., Yang, F., Han, X.*, et al.* (2011). HAT4, a Golgi apparatus-anchored B-type histone acetyltransferase, acetylates free histone H4 and facilitates chromatin assembly. Mol Cell *44*, 39-50.

Young, N.L., Dimaggio, P.A., and Garcia, B.A. (2010). The significance, development and progress of high-throughput combinatorial histone code analysis. Cell Mol Life Sci *67*, 3983-4000.

Yu, L.R., Zhu, Z., Chan, K.C., Issaq, H.J., Dimitrov, D.S., and Veenstra, T.D. (2007). Improved titanium dioxide enrichment of phosphopeptides from HeLa cells and high confident phosphopeptide identification by cross-validation of MS/MS and MS/MS/MS spectra. J Proteome Res *6*, 4150-4162.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. The Journal of biological chemistry *286*, 7983-7989.

Zhai, B., Villen, J., Beausoleil, S.A., Mintseris, J., and Gygi, S.P. (2008). Phosphoproteome analysis of Drosophila melanogaster embryos. J Proteome Res *7*, 1675-1682.

Zhang, K., Williams, K.E., Huang, L., Yau, P., Siino, J.S., Bradbury, E.M., Jones, P.R., Minch, M.J., and Burlingame, A.L. (2002). Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. Molecular & cellular proteomics : MCP *1*, 500-508.

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

Zhang, Y., Wen, Z., Washburn, M.P., and Florens, L. (2011). Improving Proteomics Mass Accuracy by Dynamic Offline Lock Mass. Anal Chem.

References for Extended Experimental Procedures

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*, 248-254.

Caraux, G., and Pinloche, S. (2005). PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. Bioinformatics *21*, 1280-1281.

Eng, J., McCormack, A.L., and Yates, J.R., III (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Amer Mass Spectrom *5*, 976-989.

Florens, L., Carozza, M.J., Swanson, S.K., Fournier, M., Coleman, M.K., Workman, J.L., and Washburn, M.P. (2006). Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods *40*, 303-311.

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

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL–EBI. Nucleic Acids Res *38*, W695-W699.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J.*, et al.* (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. Science *301*, 1503-1508.

Longhurst, H.J., and Holder, A.A. (1997). The histones of Plasmodium falciparum: identification, purification and a possible role in the pathology of malaria. Parasitology *114 (Pt 5)*, 413-419.

MacCoss, M.J., McDonald, W.H., Saraf, A., Sadygov, R., Clark, J.M., Tasto, J.J., Gould, K.L., Wolters, D., Washburn, M., Weiss, A.*, et al.* (2002). Shotgun

identification of protein modifications from protein complexes and lens tissue. Proc Natl Acad Sci U S A *99*, 7900-7905.

Mar, J.C., Wells, C.A., and Quackenbush, J. (2011). Defining an informativeness metric for clustering gene expression data. Bioinformatics *27*, 1094-1100.

McDonald, W.H., Ohi, R., Miyamoto, D.T., Mitchison, T.J., and Yates, J.R. (2002). Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LCMS/MS, 2-phase MudPIT, and 3-phase MudPIT. Int J Mass Spectrom *219*, 245-251.

McDonald, W.H., Tabb, D.L., Sadygov, R.G., MacCoss, M.J., Venable, J., Graumann, J., Johnson, J.R., Cociorva, D., and Yates, J.R., 3rd (2004). MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. Rapid Commun Mass Spectrom *18*, 2162-2168.

Ponts, N., Harris, E.Y., Prudhomme, J., Wick, I., Eckhardt-Ludka, C., Hicks, G.R., Hardiman, G., Lonardi, S., and Le Roch, K.G. (2010). Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res *20*, 228-238.

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

Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol *19*, 242-247.

Wenger, C.D., McAlister, G.C., Xia, Q., and Coon, J.J. (2010). Sub-part-per-million precursor and product mass accuracy for high-throughput proteomics on an electron transfer dissociation-enabled orbitrap mass spectrometer. Mol Cell Proteomics *9*, 754-763.

Wong, J.A.H.a.M.A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. The Journal of the Royal Statistical Society Series C (Applied Statistics) *28*, 100-108.

Xiang, Y., Takeo, S., Florens, L., Hughes, S.E., Huo, L.J., Gilliland, W.D., Swanson, S.K., Teeter, K., Schwartz, J.W., Washburn, M.P.*, et al.* (2007). The

inhibition of polo kinase by matrimony maintains G2 arrest in the meiotic cell cycle. PLoS Biol *5*, e323.

Zhang, Y., Wen, Z., Washburn, M.P., and Florens, L. (2009). Effect of dynamic exclusion duration on spectral count based quantitative proteomics. Anal Chem *81*, 6317-6326.

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

Zhang, Y., Wen, Z., Washburn, M.P., and Florens, L. (2011). Improving Proteomics Mass Accuracy by Dynamic Offline Lock Mass. Anal Chem.

Zybailov, B., Coleman, M.K., Florens, L., and Washburn, M.P. (2005). Correlation of Relative Abundance Ratios Derived from Peptide Ion Chromatograms and Spectrum Counting for Quantitative Proteomic Analysis Using Stable Isotope Labeling. Anal Chem *77*, 6218-6224.

Chapter 2

The multifunctional autophagy pathway in the human malaria parasite, *Plasmodium falciparum*

[1,2]Serena Cervantes, [2]Evelien M. Bunnik, [3]Anita Saraf, [2]Christopher M. Conner, [2]Aster Escalante, [2]Nadia Ponts, [2]Jacques Prudhomme, [3]Laurence Florens, and [2]Karine G. Le Roch*.

[1]Graduate Program in Cell, Molecular, and Developmental Biology, University of California, Riverside, CA, USA 92521, [2]Department of Cell Biology and Neuroscience, University of California, Riverside, CA, USA 92521, [3]Stowers Institute for Medical Research, Kansas City, MO, USA 64110

Summary

Malaria remains a devastating disease in developing countries where a child succumbs every minute. The human malaria parasite, *Plasmodium falciparum,* has a complex life cycle with three major stages that occur within the mosquito vector, human liver, and the blood stream. We examine the role of autophagy during the sexual and asexual blood stages. Autophagy literally means self-eating and is a process that recycles proteins or organelles when nutrients are limited. Since *P. falciparum* is an intracellular parasite, it is constantly in a nutrient-rich environment and accordingly, we found that the parasite genome does not encode the autophagy proteins involved in the detection of nutrient starvation. Instead, the autophagy pathway in the parasite is highly divergent and has taken on additional functions during development including roles in apicoplast biogenesis, vesicular trafficking, and protein degradation. This study shows the importance of the autophagy pathway for parasite survival and reveals this modified pathway as a potential target for novel antimalarial drugs.

Abstract

Autophagy is a catabolic pathway typically induced by nutrient starvation to recycle amino acids, but can also function in removing damaged organelles. In addition, this pathway plays a key role in eukaryotic development. To date, not much is known about the role of autophagy in apicomplexan parasites and more specifically in the human malaria parasite, *Plasmodium falciparum*. Comparative genomic analysis has uncovered some, but not all, orthologues of autophagy-related (Atg) genes in the malaria parasite genome. Here, using a genome-wide *in silico* analysis, we confirm that Atg genes required for vesicle expansion and completion are present, while genes involved in induction of autophagy and cargo packing are mostly absent. We subsequently focused on the molecular and cellular function of PfAtg8, an autophagosome membrane marker and key component of the autophagy pathway, throughout the parasite asexual and sexual erythrocytic stages. In this context, PfAtg8 has a distinct and atypical role in parasite development. PfAtg8 localizes in the apicoplast and vesicles throughout the cytosol, presumably autophagosomes, during parasite development. Immunofluorescence assays of PfAtg8 in apicoplast-minus parasites confirm that PfAtg8 is involved in apicoplast biogenesis. Furthermore, treatment of parasite cultures with bafilomycin and chloroquine, both lysosomotropic agents that inhibit autophagosome and lysosome fusion, resulted in dramatic morphological changes of the apicoplast and parasite death.

Furthermore, deep proteomic analysis of components associated with PfAtg8 indicated that Atg8 might be involved in ribophagy and piecemeal microautophagy of the nucleus. Collectively, our data reveals the importance and specificity of the autophagy pathway in the malaria parasite and offers potential novel therapeutic strategies.

Introduction

Malaria continues to be a major global health problem with up to a million deaths per year.[1] The human malaria parasite has a complex life cycle consisting of three major stages in the mosquito, the human liver, and the human blood. In the blood stage, merozoites invade erythrocytes and within 48 hours the parasite progresses through ring, trophozoite, and schizont stages to produce up to 32 new daughter cells that egress and reinvade new red blood cells. This asexual cycle continues until the human host is stressed. Parasites then undergo gametocytogenesis, which can require up to two weeks for full differentiation. When another female mosquito takes a blood meal, mature male and female gametocytes are taken up and undergo sexual reproduction in the mosquito midgut to complete the cycle. Throughout its life cycle, the parasite undergoes cell differentiation and metamorphosis multiple times, requiring degradation of unnecessary stage-specific proteins and organelles within host cells. In the face of increased resistance of *P. falciparum* to current antimalarials, understanding the mechanisms that control the parasite degradation pathways may open new doors for the development of novel antimalarial drugs.

In eukaryotic cells, the autophagosome-lysosome pathway is known to be involved in degradation and recycling of proteins and organelles. Similar to the ubiquitin-proteosome system that degrades short-lived and misfolded proteins[2], the autophagy pathway consists of a number of sequential steps: selection and

tagging of cargo, recognition and transport to proteolytic machinery, degradation of cargo, and recycling of small polypeptides or amino acids. Proteins and organelles targeted for degradation are enveloped by a double membrane to form an autophagosome that fuses with a lysosome to degrade cargo. There are three main membrane-mediated processes conserved through most eukaryotes: 1) macroautophagy (referred to as autophagy hereafter) removes unnecessary proteins and damaged organelles, 2) microautophagy occurs when the lysosome envelops cytoplasmic material, and 3) chaperone-mediated autophagy (CMA) relies on a chaperone protein to translocate individual proteins into the lysosome. Once thought to be a nonspecific molecular process that recycles proteins for survival under nutrient-limited conditions, autophagy has been shown to be involved in aging, degradation of pathogens, cell differentiation, and removal of old or damaged organelles.[3]

Over 30 autophagy-related (Atg) genes were first discovered in yeast,[4] and were thought to be generally conserved in eukaryotes.[5] However, preliminary data show that the canonical machinery of autophagy in parasitic protists has become diverse as a possible consequence of their adaptive life cycles.[6] Atg genes have been modified, lost or have expanded, but overall autophagy seems to play a critical role in parasite development.[7] For example, in *Toxoplasma gondii,* an apicomplexan parasite related to *Plasmodium*, orthologues of the Atg5-Atg12 conjugation system have not yet been identified. However, the Atg8

conjugation system has been detected and Atg1 and TOR orthologues have been found to mediate autophagosome formation upon starvation.[8] In addition, it was recently found that in toxoplasma, TgAtg3, an E2 conjugating enzyme for the LC3 lipidation process, is essential for regulating mitochondrial homeostasis during tachyzoite development.[9] In the murine rodent malaria parasite, *P. berghei,* autophagy is required to remove unnecessary organelles during metamorphosis of the sporozoite into merozoites in hepatocytes.[10] Genetic deletion of *P. berghei* Atg8 results in a lethal phenotype, indicating the importance of the autophagy pathway in *Plasmodium*.[11]

Despite the publication of these recent studies, not much is known about the role of autophagy in the human malaria parasite, *P. falciparum.* Genome-wide analyses using BLAST searches for either homologues or putative orthologues of known autophagy components have uncovered a rudimentary set of Atg genes. Recently, PfAtg8, marker of the autophagosomal membranes, was shown to localize in the apicoplast membrane at the late schizont and merozoite stages.[12] In the rodent malaria, *P. berghei,* the function of PbAtg8 seems restricted to apicoplast biogenesis in the liver stage.[13] The apicoplast is a plastid of secondary endosymbiotic origin that is essential to the parasite survival.[14] The role and function of PfAtg8 in the apicoplast remains ambiguous, but it has been implied that PfAtg8 may be part of ER-related organelle biogenesis, which includes apicoplast membrane biogenesis.[15] In this study, we used a combination of

genomic, molecular, cellular, and proteomic approaches to further investigate the role of Atg genes throughout the *P. falciparum* asexual and sexual blood stages. We demonstrate that the human malaria parasite has modified the canonical autophagy pathway to a rudimentary set of Atg genes that participate not only in apicoplast biogenesis, but also in additional functions of cellular development.

Results

Identification of autophagy genes in *P. falciparum*

Previous *in silico* analysis of *P. falciparum's* genome using BLAST retrieved few homologues of autophagy-related (Atg) genes.[6, 12] To uncover a complete list of Atg genes in *P. falciparum,* we conducted an HMM search for Atg protein domains. The HMM methodology has a high sensitivity for the detection of less conserved regions, and is therefore better applicable to apicomplexan parasites which seem to have modified the canonical autophagy pathway.[16]

Of the upstream autophagy pathway, none of the members of the Atg1 kinase complex functional unit were detected, even though putative Atg1 and Atg17 orthologues have been reported previously (Table 2.1).[11] Likewise, the transmembrane protein Atg9 was not identified. However, we detected potential candidates for PfAtg6 and PfAtg14, members of the class III PI3 kinase complex. In addition, a possible PfAtg2 candidate with homology in the C-terminal region was identified.[17] Altogether, this *in silico* study suggests that the human malaria parasite lacks genes critical for detection of nutrient limitation and induction of the canonical autophagy pathway, unlike higher eukaryotes and the apicomplexan, *T. gondii*.

Stage of autophagy	Functional unit	*S. cerevisiae*	*P. falciparum*
Induction	Atg1 kinase complex	Atg1	
		Atg13	
		Atg17	
		Atg29	
		Atg31	
Formation	Atg9	Atg9	
	PI3K complex	Vps15	
		Vps34	PF3D7_0515300
		Vps30/Atg6	PF3D7_1306400
		Atg14	PF3D7_1345900
Retrieval	Atg2-Atg18 complex	Atg2	PF3D7_1346400
		Atg18	PF3D7_1012900
Vesicle expansion and completion	Atg12 conjugation system	Atg5	PF3D7_1430400
		Atg7	PF3D7_1126100
		Atg10	
		Atg12	PF3D7_1470000
		Atg16	PF3D7_1250900
Vesicle breakdown	Atg8 conjugation system	Atg3	PF3D7_0905700.2
		Atg4	PF3D7_1417300
		Atg7	PF3D7_1126100
		Atg8	PF3D7_1019900
		Atg15	PF3D7_1427100

Table 2.1. *P. falciparum Atg genes identified by in silico comparative genomic studies.*

Of the downstream pathway, we identified most genes in the Atg12 conjugation system with the exception of the E2-like enzyme Atg10 (Table 2.1). In addition, all members of the Atg8 conjugation system (Atg7, Atg3 and Atg4) were detected. Finally, a homologue of the lipase PfAtg15, responsible for breakdown of intravacuolar membranes, was identified.[18] As a whole, our comparative genomics analysis has expanded the Atg gene list found in *P. falciparum* and indicates that a rudimentary autophagy pathway is present and likely functional.

PfAtg8 and GFP-PfAtg8 expression and localization during the asexual erythrocytic stage

To investigate the role of autophagy in the human malaria parasite at the molecular and cellular levels, we first explored the role of Atg8 during the erythrocytic cycle. Atg8 is the standard autophagosome marker and is essential to autophagosome formation. It remains present on the autophagosome membrane until fusion with the lysosome.[19] Using a custom PfAtg8-specific antibody, we detected Atg8 in all stages of the asexual erythrocytic cell cycle (Fig. 2.1A). In addition, PfAtg8 was found to be predominately incorporated into membranes (Fig. 2.1B).

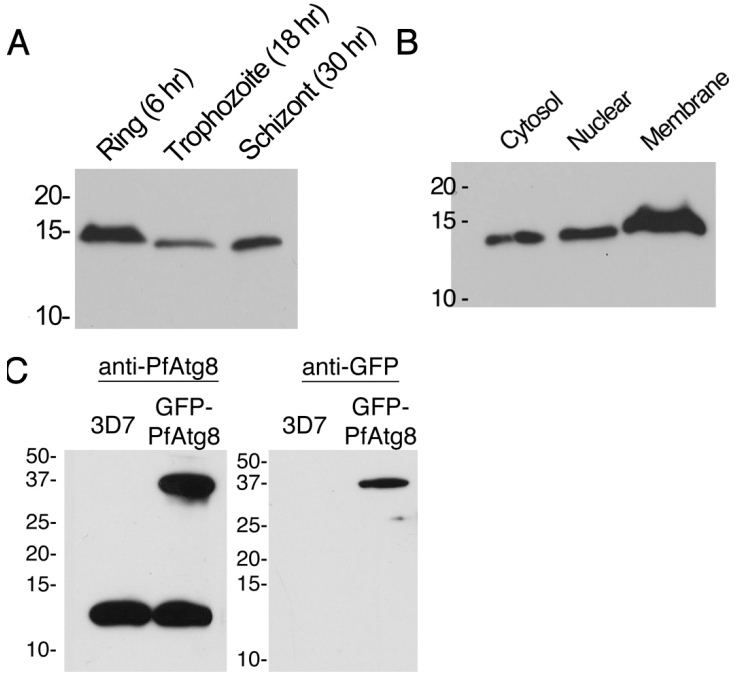


Figure 2.1. Expression of *PfAtg8 throughout the erythrocytic cycle.* A) Western blot analysis using a custom-generated anti-PfAtg8 antibody. Expression of PfAtg8 was detected at lower levels in the trophozoite stage. B) Western blot analysis showing the expression of endogenous PfAtg8 in cytosolic, nuclear, and membrane fractions of parasite lysates. PfAtg8 is predominately bound to membranes. C) Western blot analysis using high resolution acrylamide gel (16.5% Tris-Tricine gel). A faint band detected above PfAtg8-PE indicates unconjugated PfAtg8. D) Western blot analysis in a GFP-PfAtg8 transiently transfected *P. falciparum* cell line. Endogenous PfAtg8 is detected using anti-PfAtg8 at 14 kDa, while the fusion protein GFP-PfAtg8 is detected by both anti-GFP and anti-PfAtg8 antibodies at the expected molecular weight of 41 kDa.

To determine the subcellular localization of PfAtg8, synchronized 3D7 parasite cultures were immunolabeled with anti-PfAtg8 at all stages of the asexual cell cycle. PfAtg8 was found on punctate structures of 200-350 nm diameter throughout the asexual cycle (Fig. 2.2). These structures are smaller than the classical autophagosome of 300 to 900 nm diameter and larger than the typical 150 nm diameter cytoplasm-to-vacuole targeting vesicles. During the ring stage,

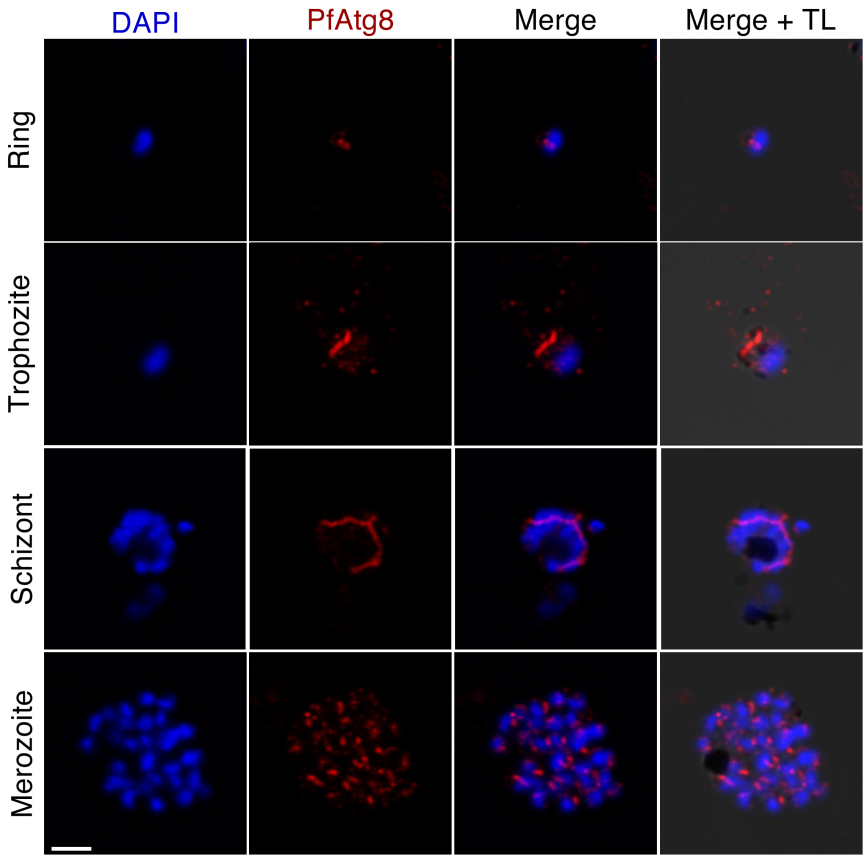


Figure 2.2. *Immunolabeling of synchronized 3D7 parasites with anti-PfAtg8.* At the ring and trophozoite stage small vesicles can be observed. During the schizont stage a large branched structure forms, but vesicles remain present in the cytosol. DAPI is used as a DNA marker in blue and scale bar indicates 2.5 μm.

small punctate structures were observed, which merged into a circular structure during the trophozoite stage (Fig. 2.2). This circular structure branched out as the parasite underwent schizogony, and finally segmented into daughter cells (Fig. 2.2). To validate our results, we generated a GFP-PfAtg8 transiently transfected cell line that correctly expressed the fusion protein (Fig. 2.1C). No apparent difference in GFP-PfAtg8 parasite growth rate or propagation was observed, although blood smears of parasites did have an enlarged digestive food vacuole phenotype at the trophozoite stage. The GFP-PfAtg8 transfected cell line

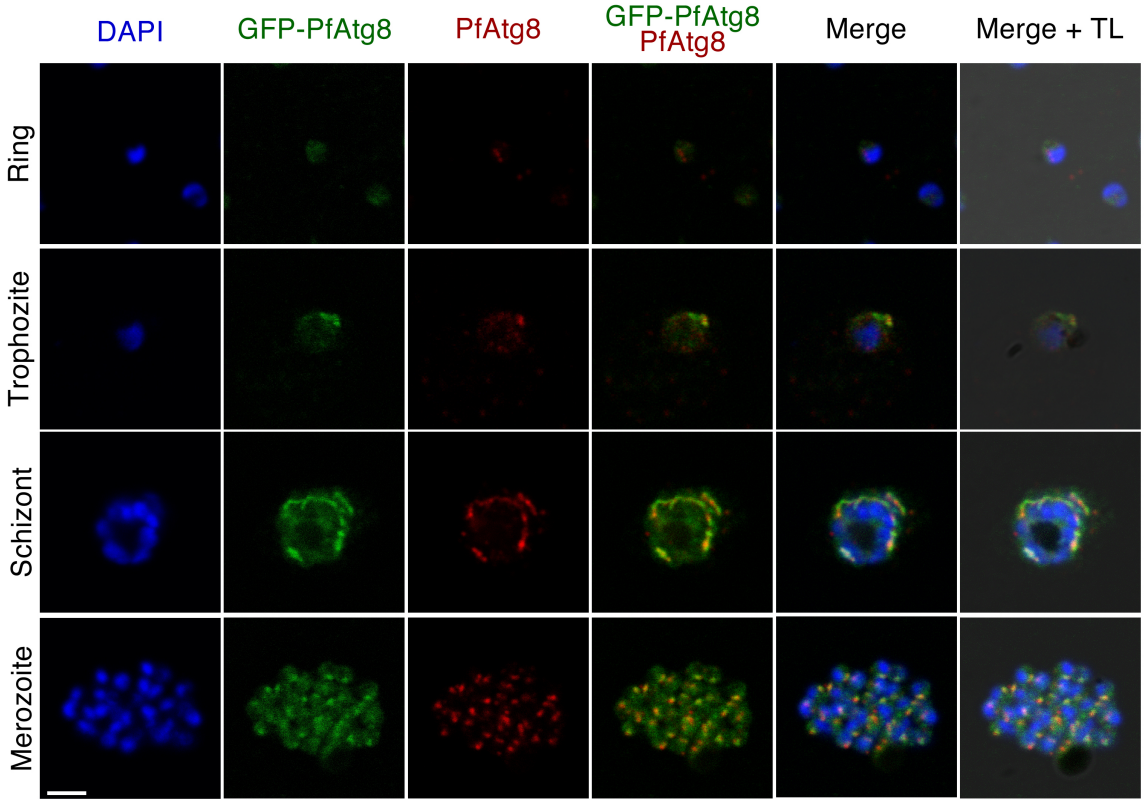


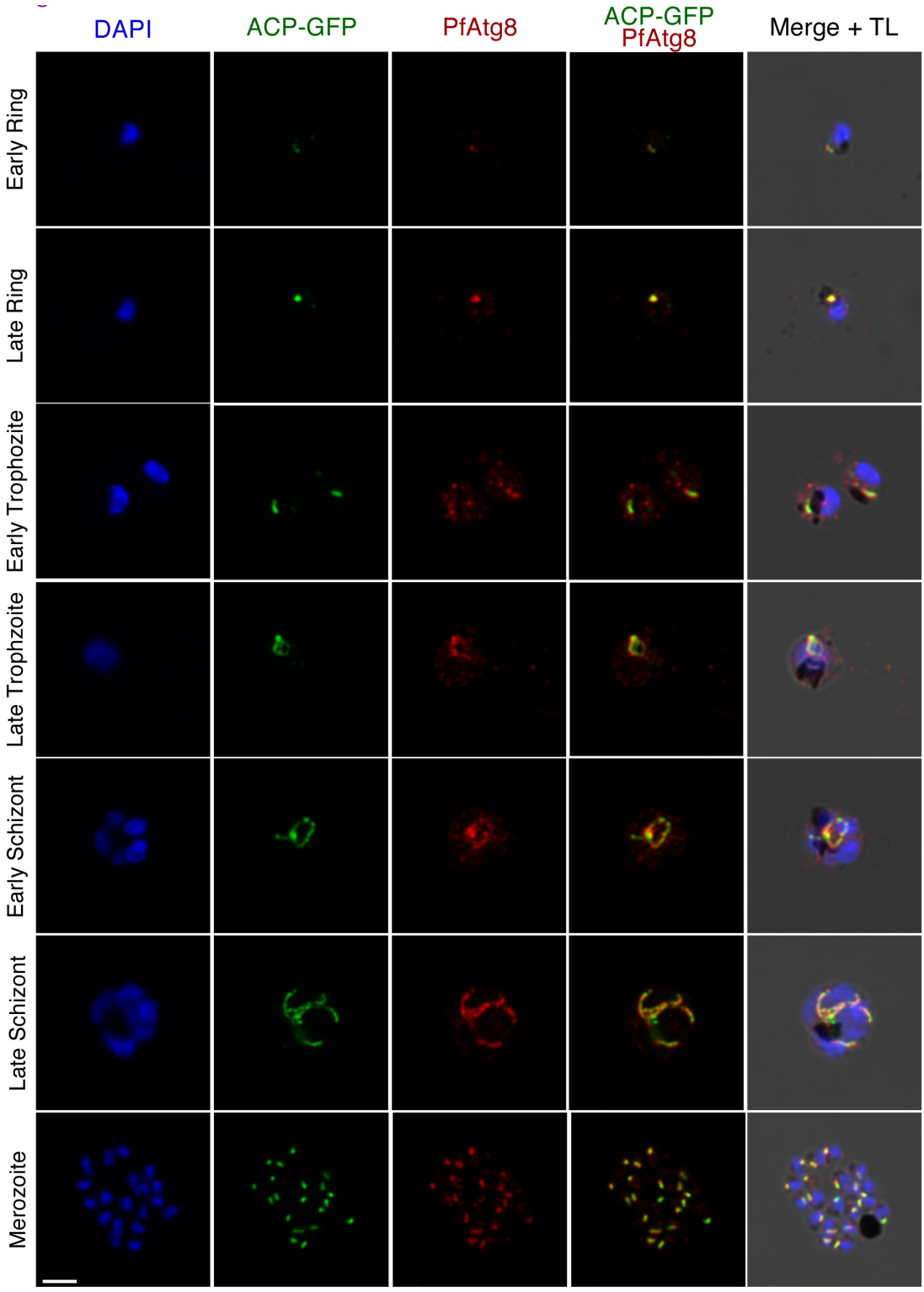
Figure 2.3. *Visualization of GFP-PfAtg8 expression by epifluorescence microscopy.* GFP-PfAtg8-labeled vesicles were observed at all stages. In addition, branching at the schizont stage was observed with a slightly diffused GFP signal throughout the cytosol. DAPI is used as a DNA marker in blue and scale bar indicates 5 μm.

displayed a similar pattern as compared to anti-PfAtg8 labeling with a diffuse GFP signal throughout the cytoplasm (Fig. 2.3). PfAtg8 was recently reported to be solely localized in the apicoplast during the merozoite stage.[12] To date, this is the first report of PfAtg8-decorated vesicles, presumably autophagosomes, observed at all stages of the asexual blood cycle.

PfAtg8 vesicles and localization with the apicoplast

The transgenic D10 strain with the apicoplast acyl carrier protein (ACP) fused to GFP was used to confirm the possible co-localization of Atg8 and the apicoplast.[20] PfAtg8 was indeed detected in close proximity to the apicoplast, suggesting a possible biological role of this protein in the biogenesis of the organelle (Fig. 2.4). However, additional small punctate structures that do not completely overlap with the apicoplast were observed throughout the asexual cycle. Quantification of vesicles revealed 1 to 2 puncta at the ring stage, while 10 to 15 puncta and 15 or more were observed at the trophozoite and the schizont stages, respectively. Moreover, the branching of PfAtg8 during the trophozoite

Figure 2.4. *PfAtg8 immunofluorescence staining throughout the erythrocytic cycle.* Synchronized *P. falciparum* D10 parasites were immunolabeled with anti-PfAtg8. D10 parasites express ACP-GFP that localizes to the apicoplast. Colocalization of ACP-GFP and PfAtg8 can be observed, although PfAtg8 appears to surround the apicoplast, indicating a possible association with the outer membrane. Additional PfAtg8 vesicles can be observed in the cytosol. DAPI is used as a DNA marker in blue. The scale bar indicates 5 μm.


DAPI
ACP-GFP
PfAtg8
ACP-GFP
PfAtg8
Merge + TL
Early Ring
Late Ring
Early Trophozite
Late Trophzoite
Early Schizont
Late Schizont
Merozoite

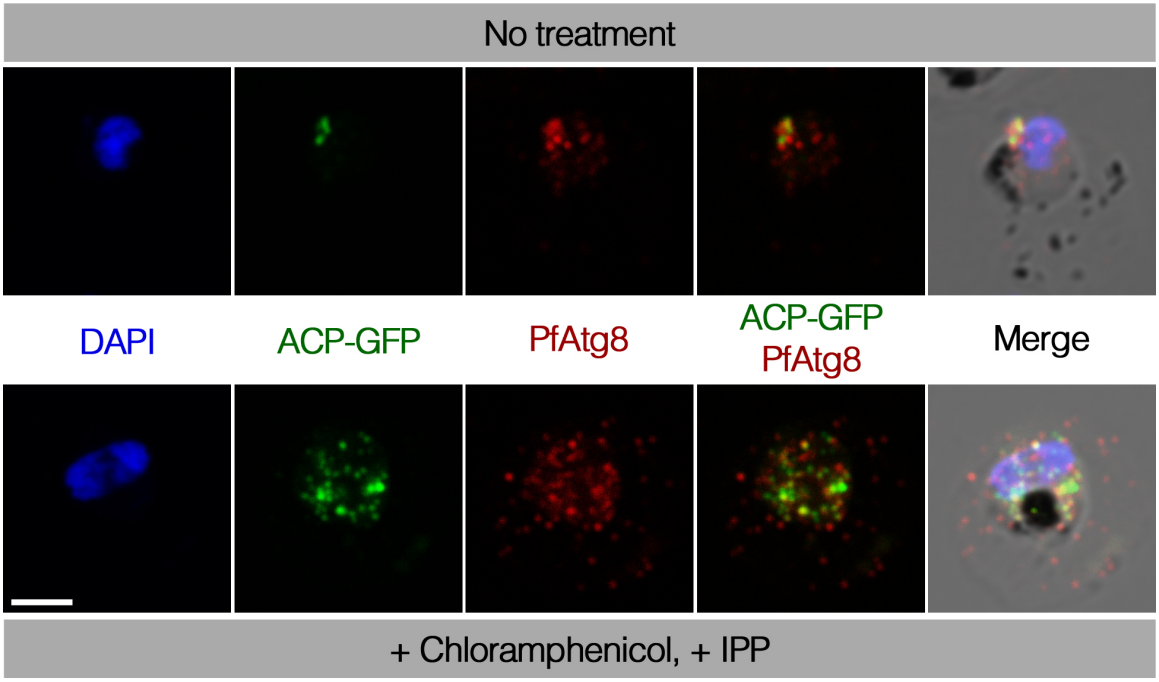


Figure 2.5. *PfAtg8 immunofluorescent staining in apicoplast-minus malaria parasites.* The integrity of the apicoplast in these parasites was lost using treatment with the antibiotic chloramphenicol. GFP-ACP parasite cultures were rescued by supplementing media with isopentenyl pyrophosphate (IPP), the essential product from the non-mevalonate isoprenoid precursor pathway in the apicoplast. The large apicoplast PfAtg8-labeled branching, as observed in untreated parasites, is disrupted into multiple foci in apicoplast-minus parasites. These foci show colocalization of GFP-ACP and PfAtg8 vesicles, although additional PfAtg8 vesicles can be observed in the cytosol. DAPI is used as a DNA marker in blue and scale bar indicates 2.5 μm.

and schizont stages covered a larger area than ACP-GFP and appeared bulbous compared to the smooth branching of the apicoplast.

Recently, it was reported that apicoplast-minus *P. falciparum* strains could be established by treatment of the parasites with antibiotics that inhibit apicoplast gene transcription and translation, followed by rescue of the parasites by isopentenyl pyrophosphate (IPP) supplementation.[21] In antibiotic-treated ACP-GFP parasites rescued with IPP, apicoplast proteins were localized in numerous

small foci spread throughout the cell with additional diffuse GFP staining, while branching of the apicoplast at the schizont stage was completely abolished. Labeling of apicoplast-minus parasites with PfAtg8 showed that PfAtg8 vesicles overlap with some, but not all, ACP-GFP vesicles (Fig. 2.5). Quantification of vesicles revealed that two-thirds of PfAtg8 vesicles did not overlap with the ACP-GFP foci.

PfAtg8 localization during gametogenesis

Next, we investigated the distribution of PfAtg8 during the sexual phase of the erythrocytic stage. In contrast to the asexual cycle which is completed within 48 hours, differentiation into gametocytes is a relatively slow process requiring 8-10 days.[22] In the NF54 strain, that readily produces gametocytes, the apicoplast ACP-GFP cell line and our newly generated GFP-PfAtg8 parasite line, PfAtg8 localized in the apicoplast adjacent to the nucleus (Fig. 2.6). In addition, gametocytes of all stages showed numerous PfAtg8-associated vesicles throughout the cytosol. During stage II–IV, PfAtg8 was also present at the apical poles (Fig. 2.6). Interestingly, the apical pole staining differed depending on the sex of gametocytes. Stage IV male gametocytes displayed a large vesicle at both poles, whereas female gametocytes had vesicles more evenly dispersed outside the nucleus (Fig. 2.6B). Once gametocytes are in the mosquito gut, females will

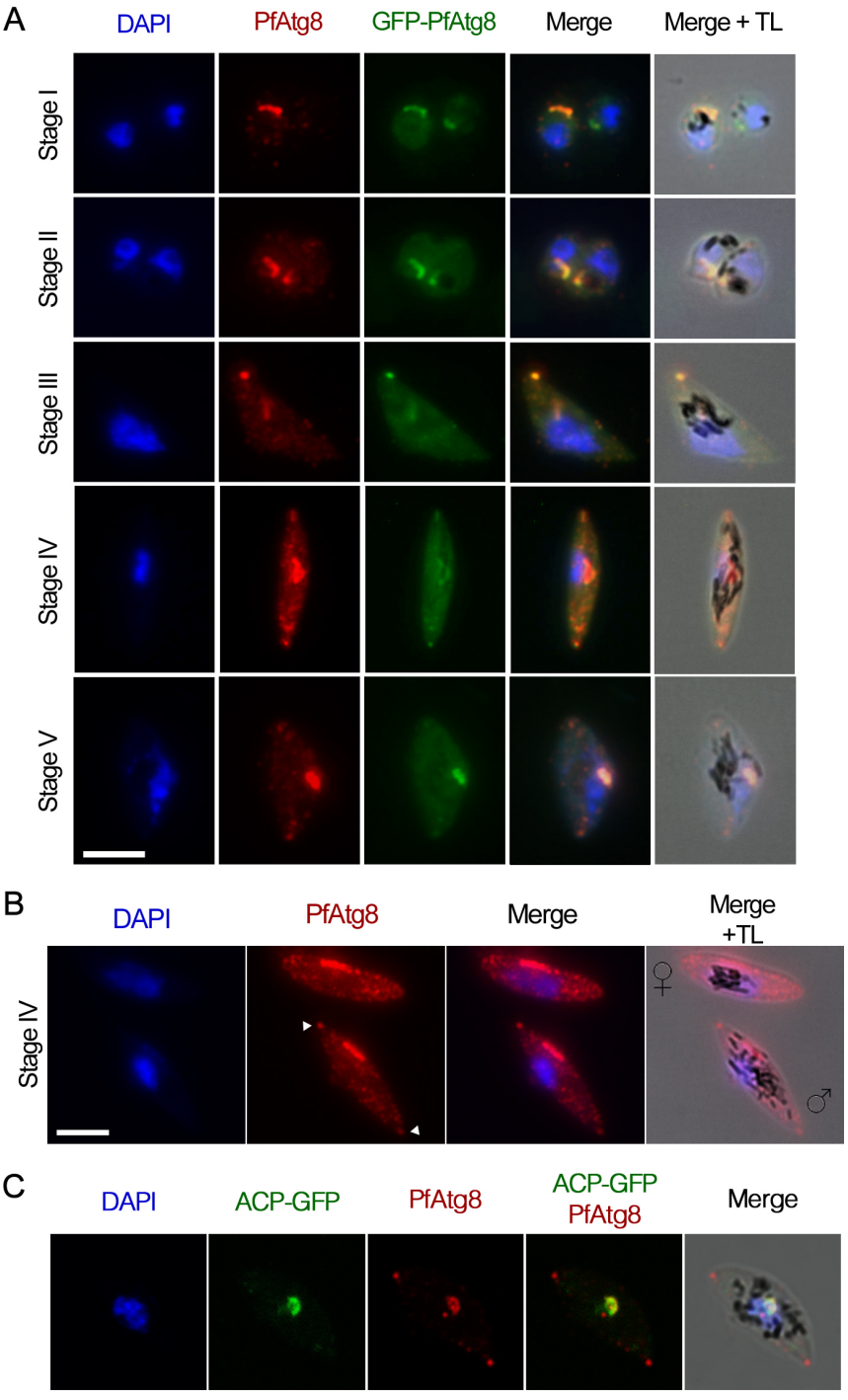


Figure 2.6. *PfAtg8 immunofluorescent staining in gametocytes.* A) Expression of PfAtg8 in GFP-PfAtg8 gametocytes. An abundance of PfAtg8-labeled autophagosomes is observed during gametocytogenesis. In stage II-IV, large vesicles are observed at the apical poles. B) Differential PfAtg8 pattern in male and female gametocytes. Female NF54 gametocytes at stage IV have rounded poles and PfAtg8 vesicles are more uniform in size and distribution in the cytosol, while male gametocytes have larger PfAtg8 vesicles at the apical poles (white arrowheads). C) Validation of PfAtg8 localization using the ACP-GFP cell line. A clear colocalization of GFP-PfAtg8 and the apicoplast can be observed. DAPI is used as a DNA marker in blue and scale bar indicates 5 μm.

produce a single macrogamete while males will produce eight microgametes. During microgametogenesis, parasites exflagellate and develop axonemes, but the apicoplast and mitochondrion remain in the original gametocyte cytoplasm.[23] A temporal and sex-dependent regulation of the autophagy pathway in male and female gametocytes is therefore not surprising.

Absence of induction of autophagy pathway by starvation in *P. falciparum*

Autophagy is typically induced by nutrient starvation.[4] However, the Atg genes involved with autophagy induction were not detected in our *in silico* genome-wide analysis (Table 2.1). To validate our *in silico* data and test whether autophagy could be induced by nutrient limitation in *P. falciparum*, parasite cultures were grown in both serum-free and minimal media devoid of amino acids and serum with minimal sugar. Ring stage parasites grown in media without human serum or albumax for 24 hours were arrested at the trophozoite stage, and normal growth was restored upon addition of complete media. A similar hibernatory state has been recently observed in isoleucine-starved parasites.[24] Parasites that were starved for 72 hours in minimal media and then given complete media produced 1% parasitemia of gametocytes when compared to control parasites. However, despite multiple attempts, incubation in minimal media for 6, 24, or 72 hours did not increase the amount of PfAtg8 autophagosomes (data not shown). These data indicate that unlike other

eukaryotic cells, autophagy in the asexual blood stage of *P. falciparum* cannot be activated under mild starving conditions, and that parasites starved for long periods of time preferably undergo gametocytogenesis.

Downstream inhibitory effects on *P. falciparum's* autophagy pathway

Traditional methods to induce and inhibit autophagy do not seem to be effective in *P. falciparum.* The lack of classical induction by starvation or drug treatments may be due to the parasite's divergent upstream signaling kinases. Indeed, treatment of parasite cultures with the TOR complex inhibitor rapamycin did not induce autophagy, most likely because the parasite lacks a homologue for TOR (data not shown). Similarly, PfAtg8 localization was previously reported to remain unchanged upon treatment of parasites with wortmannin, a PI3K inhibitor.[12] Both rapamycin and wortmannin work by targeting upstream proteins of the autophagy pathway. Bafilomycin A1, on the other hand, is a V-type H+-ATPase inhibitor that has been found to inhibit the downstream fusion events between autophagosomes and lysosomes, and has effects on the transport within the endocytic pathway.[25] Here, we observed that bafilomycin A1 is a potent inhibitor of the chloroquine-sensitive 3D7 strain, and the chloroquine-resistant strains W2 and Dd2, as well as our GFP-PfAtg8 cell line, with IC_{50}

P. falciparum strain	IC_{50} (nM)	Standard error	*R*-value	Hill slope
3D7	25.7	3.61	0.996	3.668
Dd2	22	2.73	0.998	4.401
W2	20.5	0.68	0.999	3.663
GFP-PfAtg8	23.6	0.57	0.999	3.387

Table 2.2. *IC_{50} values of bafilomycin on various P. falciparum strains.*

values ranging between 21 – 26 nM (Table 2.2). The morphological effects of bafilomycin were examined by treatment of synchronized 3D7 and GFP-PfAtg8 parasite cultures at the ring, trophozoite, and schizont stages for 6 and 12 hours. A rapid inhibition of parasite growth was observed, in particular at the mature stages (Fig. 2.7). The potent inhibitory effect of bafilomycin on parasite growth

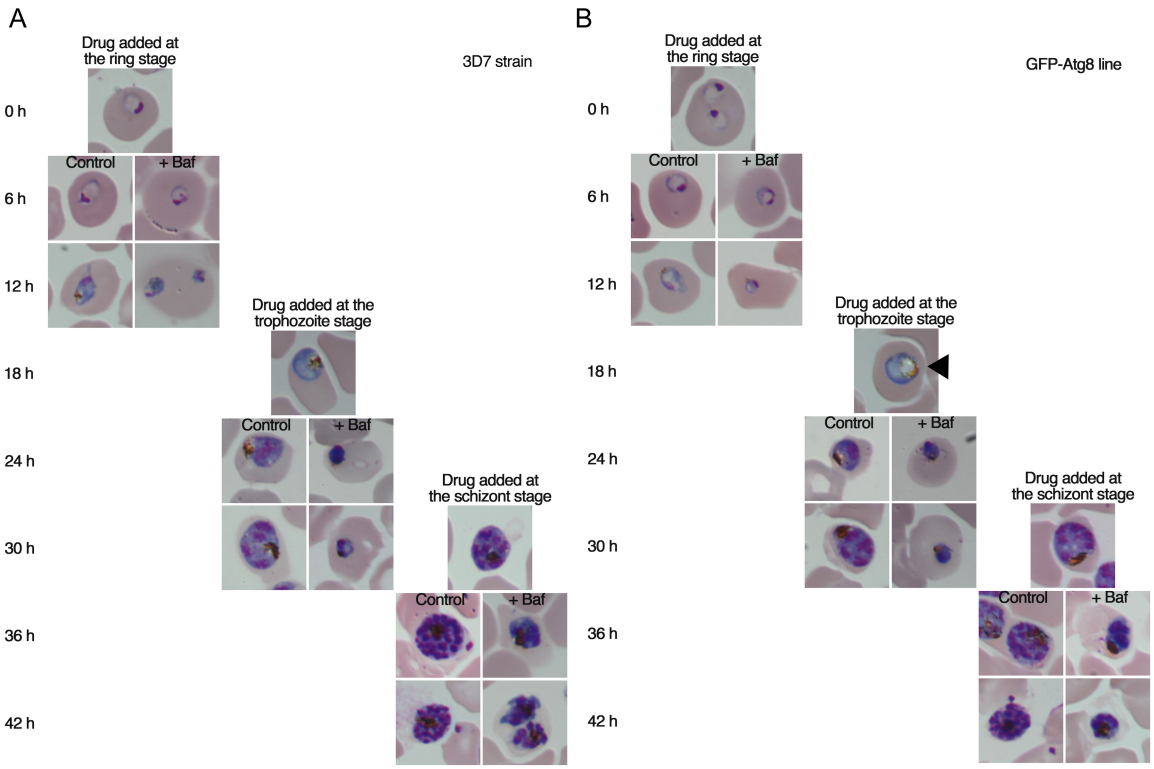


Figure 2.7. *Inhibition of GFP-PfAtg8 transfected parasites by bafilomycin treatment during the erythrocytic cycle.* A) Synchronized 3D7 parasite cultures were incubated with 75 nM bafilomycin and blood smears were taken at hours 6 and 12, followed by Giemsa staining. Inhibition of parasite growth occurred within 6 hours after the start of bafilomycin treatment. B) Synchronized GFP-Atg8 parasite cultures were incubated with 75 nM bafilomycin and blood smears were taken at hours 6 and 12, followed by Giemsa staining. Similar to what was observed for 3D7 parasites, growth of parasites was arrested within 6 hours of bafilomycin treatment. In addition, the GFP-PfAtg8 parasites have an enlarged food vacuole phenotype at the trophozoite stage (black arrowhead).

could be reversed by washing bafilomycin out of culture media within 3 hours (data not shown). After 3 hours of bafilomycin A1 incubation, early trophozoite parasites decreased in size with a slight decrease of PfAtg8-labeled vesicles, from an average of 11 punctate structures to an average of 8.5 puncta. This overall decrease of PfAtg8 staining was accompanied by dissociation from the apicoplast and an increased staining level in the cytosol (Fig. 2.8A). Western blot analysis of PfAtg8 expression after bafilomycin A1 treatment did not show a substantial difference when compared to the control (Fig. 2.8B). Similar to what we observed at the trophozoite stage, schizont stage parasites incubated with bafilomycin displayed a strong dissociation between PfAtg8 and the apicoplast and an increase in PfAtg8 vesicle staining from an average of 18 to an average of 30 puncta.

In eukaryotic cells, autophagy has also been shown to play a role in a particular form of cell death. *P. falciparum* was previously reported to undergo vacuolization after 24 hours of incubation with antimalarial compounds.[26] We examined the effects of artemisinin, currently the most effective anti-malarial agent[27], and chloroquine on PfAtg8 localization. ACP-GFP parasites treated at the trophozoite stage with a high concentration of 25 μM artemisinin did not show any significant changes in Atg8 staining pattern (Fig. 2.8A). Parasites did exhibit a delay in cell cycle progression and remained stunted at the trophozoite stage

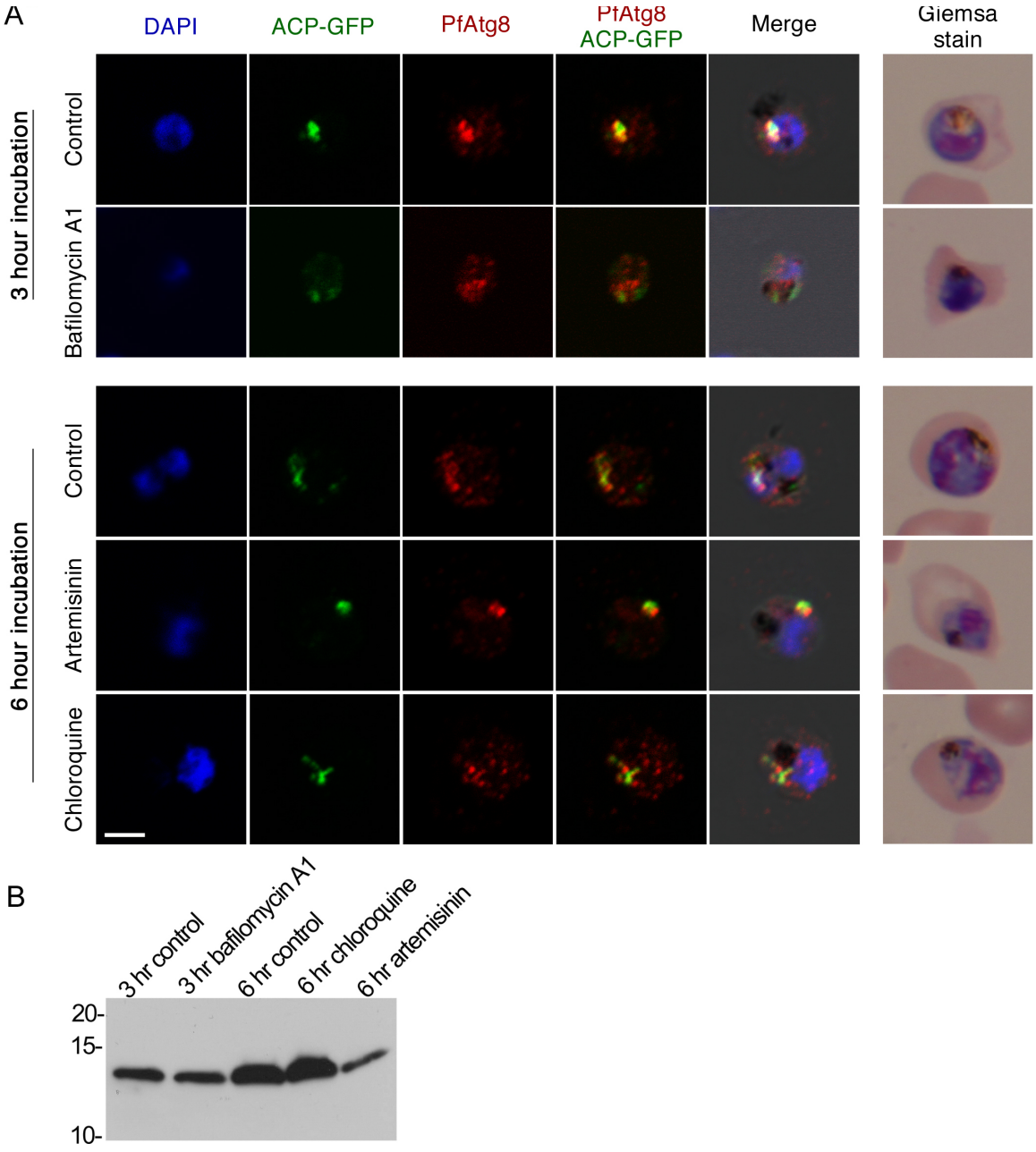


Figure 2.8. *Effect of antimalarial treatment on PfAtg8 expression and localization.* A) PfAtg8 immunofluorescent staining in 3D7 parasites treated with lysosomotropic agents and artemisinin. After 3-hour incubation with 75 nM of bafilomycin, branching of the apicoplast is disrupted and multiple PFAtg8 autophagosomes can be observed. Artemisinin treatment results in cell cycle progression arrest, and chloroquine treatment causes disruption of the apicoplast integrity. DAPI is used as a DNA marker in blue and scale bar indicates 2.5 μm. B) Western blot analysis of PfAtg8 expression levels upon bafilomycin A1, chloroquine and artemisinin treatment.

when artemisinin was added (Fig. 2.8A). On the other hand, treatment with chloroquine, which was once the most successful antimalarial drug therapy, but is also often used as a tool to study the autophagy pathway in higher eukaryotes, showed a strong phenotype. In all eukaryotic cells, chloroquine is known to increase the lysosomal pH and prevent the fusion of autophagosomes to lysosomes, and therefore block lysosomal protein degradation. Trophozoite parasites treated with 100 nM of chloroquine (IC_{80}) for 6 hours displayed fragmentation of the apicoplast with a slight increase of PfAtg8 puncta from an average of 15.5 to an average of 17.4 puncta (Fig. 2.8A). Whether this phenotype is linked to vesicle traffic inhibition or an active fragmentation process remains to be determined.

Identification of proteins associated with PfAtg8

To further confirm the role of the autophagy pathway in parasite development and to identify additional proteins involved in this process, we performed an immunoprecipitation experiment using an anti-GFP antibody on synchronized GFP-PfAtg8 cultures in combination with a shotgun mass spectrometry approach. Parasites of early trophozoite and late schizont stages were either treated or not treated with bafilomycin for 3 hours before harvesting to potentially enrich for PfAtg8-associated proteins. Proteins were identified using a state of the art method to measure protein abundance, the multidimensional

protein identification technology (MudPIT).[28] Proteins were considered positive hits when detected in at least two of the three technical replicates of at least one sample type, and when either not detected in or more than five-fold enriched compared to our two independent negative controls (a precipitation without the anti-GFP primary antibody, and a GFP-precipitation from the 3D7-HT cell line that expresses GFP in all stages of the parasite life cycle).[29] The lack of significant overlap between proteins from our precipitations and controls validates the strength and specificity of our approach. In total, 176 proteins associated with GFP-PfAtg8 passed our selection criteria (Table 2.3). As the autophagy pathway is known to regulate a large number of macromolecules in eukaryotic cells, the small number of proteins detected in our pull down experiment can be surprising. However, it is likely that our stringent cut off generated a high rate of false negatives, since proteins that precipitated non-specifically in our negative controls could still be targeted for degradation by the autophagy pathway. Evidence for enrichment of PfAtg8-associated proteins by bafilomycin treatment was observed in the schizont stage, where we detected 113 proteins in the treated sample versus 19 proteins in the untreated sample. This correlates with an increased amount of PfAtg8 vesicles that were observed when parasites were treated with bafilomycin A1. Among the proteins found to be associated with PfAtg8 were PfAtg7 (E1-like) and two undistinguishable isoforms of PfAtg3, but not PfAtg4. PfAtg2, PfAtg18, and PfAtg6 were detected in the whole dataset but

did not pass our selection criteria, while PfAtg5 was only detected in one of our controls.

					ET-nodrug_MERGED		
Proteins In Group	Description	Name	Length	MW	GFP-Atg8_ET-nodrug_Ti_Merged_dNSAF	GFP-Atg8_ET-nodrug_merged:IP-Control_merged	GFP-Atg8_ET-nodrug Detected # Out of 3
GFP_Atg8	GFP_Atg8	Atg8	317	36687	0.01624164	106.86	3
PF3D7_1126100	ThiF family protein, putative	Atg7	1316	156630	0.00019837	**infinity**	2
PF3D7_0905700.	autophagy-related protein 3, putative (ATG	Atg3	319	37476	0.00013639	**infinity**	2
PF3D7_1346400	conserved Plasmodium protein, unknown f	Atg2	5988	697066	0.00001211	0.10	2
PF3D7_1012900	conserved Plasmodium protein, unknown f	Atg18	380	43520	0.00034349	0.90	1
PF3D7_1306400	26S proteasome regulatory subunit, putati	Vps30/Atg6	393	44677	0.00014761	0.27	2
PF3D7_1430400	conserved Plasmodium protein, unknown f	Atg5	863	102230	X	0.00	0
PF3D7_1339700	conserved Plasmodium protein, unknown f	Nvj1	1936	231556	X	0.00	0
PF3D7_1363400	conserved Plasmodium protein, unknown f	Ufd3	905	107614	X	0.00	0
PF3D7_1459000	ATP-dependent RNA helicase DBP5 (DBP5)	Bre5	741	84314	0.00005872	0.91	1
PF3D7_1431700	60S ribosomal protein L14, putativeg		165	19297	0.00070318	**infinity**	3
PF3D7_1412600	deoxyhypusine synthase		496	57242	0.00026316	**infinity**	2
PF3D7_1010200	DNA2/NAM7 helicase, putative		1839	217158	0.0000552	**infinity**	2
PF3D7_0920800	inosine-5'-monophosphate dehydrogenase		510	56150	0.00036969	**infinity**	2
PF3D7_0621900	conserved Plasmodium protein, unknown function		776	92452	0.00014952	**infinity**	2
PF3D7_1426100	basic transcription factor 3b, putative		171	19393	0.00025444	**infinity**	2
PF3D7_0424600	Plasmodium exported protein (PHISTb), unknown funct		309	35962	0.00014081	**infinity**	2
PF3D7_1342400	casein kinase II beta chain (CK2beta2)		385	45285	0.00007534	**infinity**	2
PF3D7_0422300	alpha tubulin 2		450	49691	0.00006446	infinity	1
PF3D7_1305400	conserved Plasmodium protein, unknown function		609	73485	0.00004763	infinity	1
PF3D7_0810800	dihydropteroate synthetase (DHPS)		706	83374	0.00012326	infinity	1
PF3D7_1417500	pseudouridine synthase, putative		466	52942	0.00006225	infinity	1
PF3D7_1343700	kelch protein, putative		726	83665	0.00003995	infinity	1
PF3D7_0711500	regulator of chromosome condensation, putative		688	78897	0.00008432	infinity	1
PF3D7_1330400.	ER lumen protein retaining receptor 1, putative		284	32806	0.00020427	infinity	1
PF3D7_0509000	SNAP protein (soluble N-ethylmaleimide-sensitive facto		298	34646	0.00004867	infinity	1
PF3D7_1420700	surface protein, Pf113 (Pf113)		969	112574	0.00010477	**infinity**	3
PF3D7_1216900	DNA-binding chaperone, putative		939	111065	0.00004634	**infinity**	2
PF3D7_0622800	leucyl tRNA synthase		1447	170240	0.00004009	**infinity**	2
PF3D7_1464900	ATP-dependent protease la, putative		706	79815	0.00004109	**infinity**	2
PF3D7_0716300	conserved Plasmodium protein, unknown function		522	60601	0.00011113	**infinity**	2
PF3D7_0932300	M18 aspartyl aminopeptidase (M18AAP)		570	65635	0.00010178	**infinity**	2
PF3D7_1242800	rab specific GDP dissociation inhibitor (rabGDI)		459	52339	0.00009479	**infinity**	2
PF3D7_0108300	conserved Plasmodium protein, unknown function		2221	254993	0.00001959	**infinity**	2
PF3D7_1129400	RNA methyltransferase, putative		634	74369	0.00016013	**infinity**	2
PF3D7_1353100	Plasmodium exported protein, unknown function		285	32784	0.00015266	infinity	1
PF3D7_1439900	triosephosphate isomerase (TIM)		248	27935	0.00005848	infinity	1
PF3D7_1016300	glycophorin binding protein (GBP)		824	95844	0.00047522	infinity	1
PF3D7_0113900	Plasmodium exported protein (hyp8), unknown functio		243	28234	0.00029842	infinity	1
PF3D7_1419300	glutathione S-transferase (GST)		211	24789	0.00006874	infinity	1

ET-drug_MERGED			LS-nodrug_MERGED			LS-drug_MERGED			-Control_MER		Control_MERG	
GFP-Atg8_ET-drug_Ti_Merged_dNSAF	GFP-Atg8_ET-drug_merged:IP-Control_merged	GFP-Atg8_ET-drug Detected # Out of 3	GFP-Atg8_LS-nodrug_Ti_Merged_dNSAF	GFP-Atg8_LS-nodrug_merged:IP-Control_merged	GFP-Atg8_LS-nodrug Detected # Out of 3	GFP-Atg8_LS-drug_Ti_Merged_dNSAF	GFP-Atg8_LS-drug_merged:IP-Control_merged	GFP-Atg8_LS-drug Detected # Out of 3	3D7-HT-GFP-mix-unsyn_Control_Ti_Merged_dNSAF	GFP-Control Detected # Out of 2	GFP-Atg8-mix-unsyn_Control_Ti_Merged_dNSAF	IP-Control Detected # Out of 2
0.011103	73.05	3	0.0139059	91.49	3	0.00962266	63.31	3	X	0	2E-04	1
0.00006325	**infinity**	2	0.00017077	infinity	1	0.00024946	**infinity**	3	X	0	X	0
0.00018637	infinity	1	X	0	0	0.00017152	**infinity**	2	X	0	X	0
0.00000993	0.080	2	0.00001732	0.14	2	0.00002513	0.20	2	5E-05	2	1E-04	2
0.00012517	0.329	1	0.00009098	0.24	1	0.00025198	0.66	3	X	0	4E-04	1
0.00015128	0.274	2	0.00105569	1.91	2	0.00059171	1.07	2	X	0	6E-04	2
X	0.000	0	X	0.00	0	X	0.00	0	X	0	6E-05	1
X	0.000	0	0.00001786	infinity	1	0.00005652	infinity	3	X	0	X	0
X	0.000	0	X	0	0	0.00004534	infinity	1	X	0	X	0
0.00009628	1.477	2	0.00004666	0.72	1	0.00033228	5.11	3	X	0	7E-05	1
0.00050445	**infinity**	2	0.00041908	**infinity**	2	0.00058032	**infinity**	2	X	0	X	0
0.00040754	**infinity**	3	0.00003485	infinity	1	0.00016547	**infinity**	3	X	0	X	0
0.00005173	**infinity**	2	0.0000376	**infinity**	2	0.00002975	**infinity**	2	X	0	X	0
0.00002332	infinity	1	0.0000339	infinity	1	0.00013411	**infinity**	3	X	0	X	0
0.00007662	infinity	1	0.00002228	infinity	1	0.00007051	**infinity**	2	X	0	X	0
0.00020861	infinity	1	0.00010109	infinity	1	0.00007999	infinity	1	X	0	X	0
0.00007696	infinity	1	0.00016783	infinity	1	0.00017707	infinity	1	X	0	X	0
0.00006177	infinity	1	0.0001347	infinity	1	0.00007106	infinity	1	X	0	X	0
0.00022302	**infinity**	2	0.00007683	**infinity**	2	0.00006079	**infinity**	3	X	1	X	0
0.00011715	**infinity**	2	0.00017031	infinity	1	0.00020215	**infinity**	3	X	0	X	0
0.00008421	**infinity**	2	0.00004897	infinity	1	0.00017438	**infinity**	2	X	0	X	0
0.00022965	**infinity**	2	0.0000371	infinity	1	0.00008806	infinity	1	X	0	X	0
0.00004914	**infinity**	2	0.00007143	infinity	1	0.00001884	infinity	1	X	0	X	0
0.00001728	infinity	1	0.00002513	infinity	1	0.00009941	**infinity**	3	X	0	X	0
0.00004187	infinity	1	0.00024348	infinity	1	0.00019266	**infinity**	2	X	0	X	0
0.0000798	infinity	1	0.00005801	infinity	1	0.0000918	**infinity**	2	X	0	X	0
0.00006136	**infinity**	2	X	0	0	0.00004235	**infinity**	2	X	0	X	0
0.00002533	**infinity**	2	X	0	0	0.0000437	**infinity**	2	X	0	X	0
0.00002465	**infinity**	2	X	0	0	0.00002836	**infinity**	2	X	0	X	0
0.00005053	**infinity**	2	X	0	0	0.00001938	infinity	1	X	0	X	0
0.00002278	infinity	1	X	0	0	0.00028825	**infinity**	2	X	0	X	0
0.00006258	infinity	1	X	0	0	0.00009599	**infinity**	2	X	0	X	0
0.00002591	infinity	1	X	0	0	0.00029801	**infinity**	2	X	0	X	0
0.00000535	infinity	1	X	0	0	0.00001232	**infinity**	2	X	0	X	0
0.00001876	infinity	1	X	0	0	0.00002158	infinity	1	X	0	X	0
0.00062583	**infinity**	3	X	0	0	0.00014399	**infinity**	2	X	0	X	0
0.00014384	**infinity**	3	X	0	0	0.00011031	**infinity**	2	X	0	X	0
0.0003319	**infinity**	2	X	0	0	0.00029881	**infinity**	2	X	0	X	0
0.0002936	**infinity**	2	X	0	0	0.00005629	infinity	1	X	0	X	0
0.00011271	**infinity**	2	X	0	0	0.00032414	infinity	1	X	0	X	0

PF3D7_1349800	nucleolar preribosomal associated cytoplasmic ATPase,	439	52158	0.00016518	**infinity**	3
PF3D7_0936000	ring-exported protein 2 (REX2)	94	10771	0.00077144	**infinity**	2
PF3D7_0612900	nucleolar GTP-binding protein 1, putative	686	79534	0.00021142	**infinity**	2
PF3D7_0830500	tryptophan/threonine-rich antigen (TryThrA)	675	80443	0.00008594	**infinity**	2
PF3D7_1451900	small subunit rRNA processing factor, putative	1029	120667	0.00007047	**infinity**	2
PF3D7_1454200	conserved Plasmodium protein, unknown function	1404	168932	0.00007231	**infinity**	2
PF3D7_0509400	RNA polymerase I (RNAPI)	2914	340678	0.00006968	**infinity**	2
PF3D7_0936800	Plasmodium exported protein (PHISTc), unknown functi	383	45500	0.00037867	**infinity**	2
PF3D7_1477500	Plasmodium exported protein (PHISTb), unknown funct	500	60553	0.00017404	**infinity**	2
PF3D7_0424500	serine/threonine protein kinase, FIKK family (FIKK4.1)	622	74443	0.0001399	**infinity**	2
PF3D7_0823900	dicarboxylate/tricarboxylate carrier (DTC)	318	35387	0.00013682	**infinity**	2
PF3D7_1319300	N2,N2-dimethylguanosine tRNA methyltransferase, put	804	94817	0.00012627	**infinity**	2
PF3D7_1218300	adaptor protein subunit, putative	621	72901	0.00009342	**infinity**	2
PF3D7_0520100	protein phosphatase, putative	706	83405	0.00008217	**infinity**	2
PF3D7_1410200	cytidine triphosphate synthetase	858	98517	0.00006761	**infinity**	2
PF3D7_1033100	S-adenosylmethionine decarboxylase/ornithine decarbo	1434	168169	0.00006068	**infinity**	2
PF3D7_1333600	U3 snoRNA-associated small subunit rRNA processing a	1335	159337	0.00004345	**infinity**	2
PF3D7_1457700	large ribosomal subunit nuclear export factor, putative	1478	177666	0.00002944	**infinity**	2
PF3D7_1241200	conserved Plasmodium protein, unknown function	2309	276530	0.00001884	**infinity**	2
PF3D7_1207100	small subunit rRNA processing factor, putative	840	100169	0.00003453	infinity	1
PF3D7_0811500	histone-arginine methyltransferase, putative (CARM1)	912	108427	0.0000159	infinity	1
PF3D7_1134700	DNA-directed RNA polymerase 1, subunit 2, putative	1517	175531	0.00001912	infinity	1
PF3D7_1206600	DNA-directed RNA polymerase III subunit, putative	1450	167292	0.00001	infinity	1
PF3D7_1329500	conserved Plasmodium protein, unknown function	1844	218437	0.00000787	infinity	1
PF3D7_1442100	conserved Plasmodium protein, unknown function	109	12610	0.00066528	**infinity**	3
PF3D7_1434800	conserved protein, unknown function	252	28872	0.00057552	**infinity**	3
PF3D7_0923900	RNA binding protein, putative	202	22979	0.00064618	**infinity**	2
PF3D7_1027300	peroxiredoxin (nPrx)	393	43930	0.00014761	**infinity**	2
PF3D7_0413900	ubiquitin carboxyl-terminal hydrolase a, putative	853	98471	0.00005101	**infinity**	2
PF3D7_0110700	chromatin assembly factor 1 protein WD40 domain, put	446	50733	0.00009755	**infinity**	2
PF3D7_1447900	multidrug resistance protein 2 (heavy metal transport f	1024	119012	0.00014163	**infinity**	2
PF3D7_1126000	threonyl-tRNA synthetase,Threonine--tRNA ligase	1013	119546	0.00011454	**infinity**	2
PF3D7_1244100	conserved Plasmodium protein, unknown function	1296	154902	0.00002238	**infinity**	2
PF3D7_0201800	RESA-like protein with PHIST and DnaJ domains	421	48283	0.00034449	**infinity**	3
PF3D7_0410600	conserved Plasmodium protein, unknown function	326	37065	0.00013346	**infinity**	2
PF3D7_1441100	conserved Plasmodium protein, unknown function	227	26306	0.00012778	**infinity**	2
PF3D7_0621700	conserved Plasmodium protein, unknown function	1616	192026	0.00011667	**infinity**	2
PF3D7_1217200	snoRNA-associated small subunit rRNA processing prote	1119	131112	0.0000648	**infinity**	2
PF3D7_1236400	conserved Plasmodium protein, unknown function	605	73070	0.00009589	**infinity**	2
PF3D7_1405800	large subunit rRNA processing protein, putative	966	114239	0.00009008	**infinity**	2
PF3D7_0409200	40S ribosomal processing protein, putative	482	56661	0.00006018	**infinity**	2
PF3D7_1422400	conserved Plasmodium protein, unknown function	1630	195366	0.00005339	**infinity**	2
PF3D7_0925700	histone deacetylase (HDAC1)	449	51376	0.00064602	**11.96**	2
PF3D7_0813000	conserved Plasmodium protein, unknown function	599	70706	0.00043582	**10.90**	2
PF3D7_1341300	60S ribosomal protein L18-2, putative	187	21454	0.00139602	**10.82**	3
PF3D7_0401800	Plasmodium exported protein (PHISTb), unknown funct	560	60270	0.00036258	**8.44**	2
PF3D7_1230500	WD-repeat protein, putative	502	59270	0.00037558	**7.83**	2
PF3D7_1456800	V-type H( )-translocating pyrophosphatase, putative	717	76417	0.00046523	**6.94**	2
PF3D7_1105000	histone H4 (H4)	102	11324	0.00156406	**6.63**	3
PF3D7_1227100	DNA helicase 60 (DH60)	742	87020	0.00060592	**6.25**	3
PF3D7_1409600	conserved Plasmodium protein, unknown function	768	92748	0.00018884	**6.10**	2
PF3D7_1120100	phosphoglycerate mutase, putative (PGM1)	250	28770	0.00174037	**6.02**	3
PF3D7_1347700	ethanolamine-phosphate cytidylyltransferase, putative	573	66040	0.0002278	**5.43**	2
PF3D7_0521700	DEAD/DEAH box ATP-dependent RNA helicase, putative	841	97338	0.00015521	**5.34**	2
PF3D7_1323400	60S ribosomal protein L23a, putative	190	22093	0.00061066	4.81	3
PF3D7_1461300	40S ribosomal protein S28e, putative	67	7494	0.00064939	0.90	2

0.00005417	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00075898	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00006933	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00008808	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00003467	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00005928	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00005305	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00006209	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00009513	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00001912	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00003739	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00002958	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00001915	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00001684	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00002772	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00004146	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00000891	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00000805	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00000515	infinity	1	X	0	0	X	0	0	X	0	X	0
0.00002831	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00009127	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00003919	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.0000164	**infinity**	2	X	0	0	X	0	0	X	0	X	0
0.00001934	**infinity**	2	X	0	0	X	0	0	X	0	X	0
X	0.000	0	0.00047579	infinity	1	0.00050198	**infinity**	2	X	0	X	0
X	0.000	0	0.0001372	infinity	1	0.00027141	**infinity**	2	X	0	X	0
X	0.000	0	0.00042789	**infinity**	2	0.00013543	infinity	1	X	0	X	0
X	0.000	0	0.00013196	infinity	1	0.00031326	**infinity**	3	X	0	X	0
X	0.000	0	0.00010133	infinity	1	0.00016036	**infinity**	3	X	0	X	0
X	0.000	0	0.00003876	infinity	1	0.00006134	**infinity**	2	X	0	X	0
X	0.000	0	X	0	0	0.00004007	**infinity**	2	X	0	X	0
X	0.000	0	X	0	0	0.00008102	**infinity**	2	X	0	X	0
X	0.000	0	X	0	0	0.00007388	**infinity**	2	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
X	0.000	0	X	0	0	X	0	0	X	0	X	0
0.00039724	**7.352**	3	0.0001155	2.15	1	0.00042651	**7.907407**	2	X	0	5E-05	1
0.00019851	4.975	3	0.00011544	2.88	1	0.00029687	**7.425**	2	X	0	4E-05	1
0.00076304	**5.915**	2	0.00083199	**6.45**	2	0.00087779	**6.806202**	3	X	0	1E-04	1
0.00014863	3.465	2	0.00003087	0.72	1	0.00007328	1.697674	2	X	0	4E-05	1
0.00023687	4.938	2	0.00003444	0.71	1	0.00013624	2.833333	2	X	0	5E-05	1
0.00018242	2.716	2	0.00004822	0.72	1	0.00024801	3.701493	3	X	0	7E-05	1
X	0.000	0	X	0.00	0	0.00026821	1.135593	1	X	0	2E-04	1
0.00004808	0.495	1	0.00006989	0.72	1	0.00018435	1.896907	2	X	0	1E-04	1
0.00012386	4.000	1	0.00013505	4.35	1	0.00014249	4.580645	3	X	0	3E-05	1
0.00099882	3.457	3	0.00110636	3.83	3	0.0028452	**9.844291**	3	X	0	3E-04	1
0.00020752	4.952	2	0.00024135	5.74	1	0.00021485	**5.119048**	3	X	0	4E-05	1
0.00002828	0.966	1	X	0.00	0	0.00009759	3.37931	3	X	0	3E-05	1
0.00068841	**5.417**	2	0.00027295	2.15	2	0.00064795	**5.102362**	3	X	0	1E-04	1
0.00017747	0.246	1	0.00464425	**6.46**	3	0.00102081	1.420028	3	X	0	7E-04	1

PF3D7_0306800	T-complex protein beta subunit, putative	533	59199	0.00057142	3.15	2
PF3D7_0627500	4-methyl-5(B-hydroxyethyl)-thiazol monophosphate bio	189	20292	0.00046042	0.52	2
PF3D7_0205900	proteasome 26S regulatory subunit, putative	966	108358	0.00007507	1.50	2
PF3D7_1349200	glutamate--tRNA ligase, putative	863	101417	0.00003361	1.21	2
PF3D7_0302900	exportin 1, putative	1254	147938	0.00013879	3.66	2
PF3D7_1466300	26S proteasome regulatory subunit, putative	1172	132766	0.00008662	2.12	1
PF3D7_1012400	hypoxanthine-guanine phosphoribosyltransferase (HGP	231	26362	0.00056506	2.70	2
PF3D7_1360800	falcilysin (FLN)	1193	138862	0.00004863	1.23	1
PF3D7_0511800	myo-inositol 1-phosphate synthase, putative	604	69112	0.00021611	2.70	2
PF3D7_1104200	chromatin remodeling protein (SNF2L)	1426	167440	0.00005085	3.00	1
PF3D7_1311800	M1-family alanyl aminopeptidase (M1AAP)	1085	126062	0.00034754	2.25	2
PF3D7_1225800	ubiquitin-activating enzyme e1, putative (Uba1)	1140	131778	0.00002544	1.19	1
PF3D7_1434300	Hsp70/Hsp90 organizing protein (HOP)	564	66057	0.00007714	0.91	1
PF3D7_1361000	arginine methyltransferase 5, putative (PRMT5)	724	85552	0.00016026	2.39	2
PF3D7_1115300	cysteine proteinase falcipain 2b (FP2B)	482	55803	0.00081241	**5.41**	3
PF3D7_1461900	valine-tRNA ligase, putative	1090	128070	0.00007983	infinity	1
PF3D7_0601200	Pfmc-2TM Maurer's cleft two transmembrane protein (	234	27302	0.00031875	infinity	1
PF3D7_1357900	pyrroline carboxylate reductase	262	28379	0.00027678	infinity	1
PF3D7_0304200	EH (Eps15 homology) protein (PAST1)	533	62303	0.00002721	infinity	1
PF3D7_0622200	radical SAM protein, putative	860	99459	0.00001686	infinity	1
PF3D7_0511500	RNA pseudouridylate synthase, putative	10061	1187631	0.00000288	infinity	1
PF3D7_1206800	conserved Plasmodium protein, unknown function	727	86315	0.0000399	infinity	1
PF3D7_1407500	tRNA m2G10 methyltransferase subunit, putative	125	14234	0.00011602	infinity	1
PF3D7_0932800	conserved Plasmodium protein, unknown function	1342	159114	0.00004323	infinity	1
PF3D7_1304500	small heat shock protein, putative	211	25195	0.00020621	infinity	1
PF3D7_0527100	ubiquitin-conjugating enzyme E2 N, putative (UBC13)	152	17411	0.00009542	infinity	1
PF3D7_0801800	mannose-6-phosphate isomerase, putative	1011	117662	0.00001435	infinity	1
PF3D7_0727800	cation transporting ATPase, putative	1918	225075	0.00001512	infinity	1
PF3D7_1037700	enhancer of rudimentary homolog, putative (ERH)	105	12320	0.00013812	infinity	1
PF3D7_0810200	ABC1 family, putative (ABCK1)	940	111728	0.00006172	infinity	1
PF3D7_1012600	GMP synthetase (GMPS)	555	63932	0.00005226	infinity	1
PF3D7_1127000	protein phosphatase, putative	287	34461	0.00005053	infinity	1
PF3D7_1243800	microtubule associated katanin, putative	370	42733	0.0000392	infinity	1
PF3D7_0903400	DEAD/DEAH box helicase, putative	2536	300924	0.00004003	infinity	1
PF3D7_1412800	N-myristoyltransferase (NMT)	410	47970	0.00003537	infinity	1
PF3D7_0625400	conserved Plasmodium protein, unknown function	412	48544	X	0.00	0
PF3D7_0807900	tyrosyl-tRNA synthetase, putative	373	43453	X	0.00	0
PF3D7_0815200	importin beta, putative	877	99964	X	0.00	0
PF3D7_0107000	centrin-1 (CEN1)	168	19599	X	0.00	0
PF3D7_0617200	conserved Plasmodium protein, unknown function	446	52981	X	0.00	0
PF3D7_0906500	arginase	411	46494	X	0.00	0
PF3D7_1109900	apicoplast ribosomal protein L36e precursor, putative	112	12787	X	0.00	0
PF3D7_0109800	phenylalanyl-tRNA synthetase, putative	575	67720	X	0.00	0
PF3D7_1434100	queuine tRNA-ribosyltransferase, putative	407	47829	X	0.00	0
PF3D7_0724100	conserved Plasmodium protein, unknown function	1003	124126	X	0.00	0
PF3D7_1138500	protein phosphatase 2c	924	105404	X	0.00	0
PF3D7_1367700	alanyl-tRNA synthetase,Alanine--tRNA ligase	1408	165135	X	0.00	0
PF3D7_1209200	U6 snRNA-associated Sm-like protein LSm7, putative (LS	108	11996	X	0.00	0
PF3D7_1205600	conserved Plasmodium protein, unknown function	333	39018	X	0.00	0
PF3D7_1436000	glucose-6-phosphate isomerase (GPI)	579	67368	X	0.00	0
PF3D7_1340600	RNA lariat debranching enzyme, putative (DBR1)	575	67555	X	0.00	0
PF3D7_0205400	conserved Plasmodium protein, unknown function	457	54351	X	0.00	0
PF3D7_0720400	ferrodoxin reductase-like protein	642	72722	X	0.00	0
PF3D7_0513600	deoxyribodipyrimidine photolyase (photoreactivating e	1113	129199	X	0.00	0
PF3D7_1335100	merozoite surface protein 7 (MSP7)	351	41276	X	0.00	0
PF3D7_1224500	chromatin assembly protein (ASF1), putative	273	31547	X	0.00	0

0.00033464	1.851	3	0.00094057	**5.20**	2	0.00089824	4.961326	3	X	0	2E-04	1
0.00050331	0.563	3	0.00457327	**5.12**	3	0.00101325	1.134378	3	X	0	9E-04	1
0.00003693	0.740	2	0.00025054	**5.02**	2	0.00024073	4.82	3	X	0	5E-05	1
0.00005511	1.964	2	0.00002003	0.71	1	0.00036456	**13.03571**	3	X	0	3E-05	1
0.00011379	3.000	2	0.00015164	4.00	1	0.00034906	**9.184211**	3	X	0	4E-05	1
0.00007102	1.732	2	0.00004425	1.07	1	0.00033847	**8.243902**	3	X	0	4E-05	1
0.00046327	2.215	2	0.00014967	0.72	1	0.00165804	**7.933014**	3	X	0	2E-04	2
0.00004984	1.250	1	0.00002898	0.73	1	0.00029811	**7.45**	3	X	0	4E-05	1
0.00007875	0.988	2	0.00005724	0.71	2	0.00058882	**7.3625**	3	X	0	8E-05	1
0.00001668	1.000	2	0.00003637	2.12	1	0.00010552	**6.235294**	3	X	0	2E-05	1
0.00019727	1.271	2	0.00007966	0.52	1	0.00089511	**5.774194**	3	X	0	2E-04	1
0.00001043	0.476	1	0.00001516	0.71	1	0.00010799	**5.142857**	3	X	0	2E-05	1
0.00023191	2.729	2	0.0000613	0.72	1	0.00043656	**5.141176**	2	X	0	9E-05	1
0.00018066	2.701	3	0.00009551	1.43	1	0.00034008	**5.074627**	3	X	0	7E-05	1
0.0005674	3.780	3	0.00025106	1.67	3	0.00045407	3.026667	3	1E-04	1	1E-04	1
0.00007636	infinity	1	X	0.00	0	0.00016314	**infinity**	3	X	0	X	0
0.00020326	infinity	1	X	0.00	0	0.00029228	**infinity**	2	X	0	X	0
0.00009077	infinity	1	X	0.00	0	0.00020884	**infinity**	2	X	0	X	0
0.00011155	infinity	1	X	0.00	0	0.00015398	**infinity**	2	X	0	X	0
0.00001383	infinity	1	X	0.00	0	0.00003181	**infinity**	2	X	0	X	0
0.00000118	infinity	1	X	0.00	0	0.00000408	**infinity**	2	X	0	X	0
X	0.000	0	0.00011889	**infinity**	3	0.00022579	**infinity**	3	X	0	X	0
X	0.000	0	0.00055318	**infinity**	3	0.00043772	**infinity**	2	X	0	X	0
X	0.000	0	0.00003864	infinity	1	0.00018347	**infinity**	3	X	0	X	0
X	0.000	0	0.00049157	infinity	1	0.00019449	**infinity**	2	X	0	X	0
X	0.000	0	0.00011373	infinity	1	0.00026998	**infinity**	2	X	0	X	0
X	0.000	0	0.0000171	infinity	1	0.00004059	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00002853	**infinity**	3	X	0	X	0
X	0.000	0	X	0.00	0	0.00065137	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00004366	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00009859	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00023831	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00011091	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00001618	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00006673	**infinity**	2	X	0	X	0
0.00020203	**infinity**	3	0.00004196	infinity	1	0.0000332	infinity	1	X	0	X	0
0.00009564	**infinity**	2	0.00018538	infinity	1	0.00011002	**infinity**	2	X	0	X	0
0.00008135	**infinity**	2	0.00003942	infinity	1	0.00007799	**infinity**	2	X	0	X	0
0.00007078	infinity	1	0.0001029	infinity	1	0.00024427	**infinity**	2	X	0	X	0
0.00005332	infinity	1	0.00003876	infinity	1	0.00015335	**infinity**	2	X	0	X	0
0.00011572	**infinity**	2	0.00004206	infinity	1	X	0	0	X	0	X	0
0.00053084	**infinity**	2	X	0.00	0	0.00061066	**infinity**	3	X	0	X	0
0.00004136	**infinity**	2	X	0.00	0	0.00009516	infinity	1	X	0	X	0
0.00005843	**infinity**	2	X	0.00	0	X	0	0	X	0	X	0
0.00001186	0.500	1	0.00015512	**6.46**	3	0.00004091	1.708333	1	X	0	2E-05	1
0.00003861	1.500	1	0.00003742	1.42	1	0.00014804	**5.692308**	3	X	0	3E-05	1
0.00000845	infinity	1	X	0.00	0	0.00003886	**infinity**	3	X	0	X	0
0.0003303	infinity	1	X	0.00	0	0.00025331	**infinity**	2	X	0	X	0
0.00003571	infinity	1	X	0.00	0	0.00016431	**infinity**	2	X	0	X	0
0.00008215	infinity	1	X	0.00	0	0.00016537	**infinity**	2	X	0	X	0
0.00004136	infinity	1	X	0.00	0	0.00004758	**infinity**	2	X	0	X	0
0.00005204	infinity	1	X	0.00	0	0.0000898	**infinity**	2	X	0	X	0
0.00001852	infinity	1	X	0.00	0	0.00008523	**infinity**	2	X	0	X	0
0.00001068	infinity	1	X	0.00	0	0.00002458	**infinity**	2	X	0	X	0
X	0.000	0	0.000197	**infinity**	2	0.00011691	infinity	1	X	0	X	0
X	0.000	0	0.00012664	**infinity**	2	0.00020042	infinity	1	X	0	X	0

PF3D7_0207700	serine repeat antigen 4 (SERA4)	962	108681	X	0.00	0
PF3D7_1346300	DNA/RNA-binding protein Alba 2 (ALBA2)	211	24984	X	0.00	0
PF3D7_1236900	conserved Plasmodium protein, unknown function	372	43712	X	0.00	0
PF3D7_1316600	choline-phosphate cytidylyltransferase (CCT)	896	104782	X	0.00	0
PF3D7_0803200	filament assembling protein, putative	728	86743	X	0.00	0
PF3D7_0212900	Leu/Phe-tRNA protein transferase, putative	365	42940	X	0.00	0
PF3D7_1356600	regulator of chromosome condensation, putative	480	54595	X	0.00	0
PF3D7_1323500	plasmepsin V (PMV)	590	68481	X	0.00	0
PF3D7_1220100	pre-mRNA splicing factor, putative	618	72536	X	0.00	0
PF3D7_0106000	conserved Plasmodium protein, unknown function	922	109280	X	0.00	0
PF3D7_1013600	conserved Plasmodium protein, unknown function	2238	269271	X	0.00	0
PF3D7_1027700.	centrin-3 (CEN3)	179	20913	X	0.00	0
PF3D7_0207500	serine repeat antigen 6 (SERA6)	1031	118836	X	0.00	0
PF3D7_1135400	conserved Plasmodium protein, unknown function	245	28496	X	0.00	0
PF3D7_1107400	Rad51 homolog (RAD51)	350	38592	X	0.00	0
PF3D7_0511000	translationally controlled tumor protein homolog, putat	171	19979	X	0.00	0
PF3D7_1349600	conserved Plasmodium protein, unknown function	307	35987	X	0.00	0
PF3D7_1127100	deoxyuridine 5'-triphosphate nucleotidohydrolase (dUT	173	19573	X	0.00	0
PF3D7_1113400	ubiquitin domain containing protein	388	42706	X	0.00	0
PF3D7_1348500	TBC domain protein, putative	347	41207	X	0.00	0
PF3D7_1237000	ubiquitin-activating enzyme, putative (Uba2)	686	81120	X	0.00	0
PF3D7_1454700	6-phosphogluconate dehydrogenase, decarboxylating, p	468	52994	X	0.00	0
PF3D7_0413700	lysine decarboxylase-like protein, putative	338	39021	X	0.00	0
PF3D7_0111500	UMP-CMP kinase, putative	371	44245	X	0.00	0
PF3D7_1033700	bromodomain protein, putative	488	55938	X	0.00	0
PF3D7_0217900	conserved Plasmodium protein, unknown function	519	60598	X	0.00	0
PF3D7_1466100	protein serine/threonine phosphatase	889	100775	X	0.00	0
PF3D7_0509100	chromosome condensation protein, putative	1708	201666	X	0.00	0
PF3D7_0705300	origin recognition complex subunit 2, putative (ORC2)	825	98010	X	0.00	0

X	0.000	0	0.00001797	infinity	1	0.00008532	**infinity**	3	X	0	X	0
X	0.000	0	0.00016386	infinity	1	0.00012966	**infinity**	2	X	0	X	0
X	0.000	0	0.00009294	infinity	1	0.00018386	**infinity**	2	X	0	X	0
X	0.000	0	0.00005788	infinity	1	0.0000916	**infinity**	2	X	0	X	0
X	0.000	0	0.00004749	infinity	1	0.00013153	**infinity**	2	X	0	X	0
X	0.000	0	0.00004736	infinity	1	0.00011243	**infinity**	2	X	0	X	0
X	0.000	0	0.00003601	infinity	1	0.000057	**infinity**	2	X	0	X	0
X	0.000	0	0.0000293	infinity	1	0.00009274	**infinity**	2	X	0	X	0
X	0.000	0	0.00002797	infinity	1	0.00004427	**infinity**	2	X	0	X	0
X	0.000	0	0.00001875	infinity	1	0.00002967	**infinity**	2	X	0	X	0
X	0.000	0	0.00000772	infinity	1	0.00001834	**infinity**	2	X	0	X	0
X	0.000	0	0.0003863	**infinity**	2	X	0	0	X	0	X	0
X	0.000	0	0.00021797	**9.48**	2	0.00018575	**8.086957**	3	X	0	2E-05	1
X	0.000	0	0.00028224	2.88	1	0.00050249	**5.122449**	2	X	0	1E-04	1
X	0.000	0	0.00014817	2.14	1	0.00035174	**5.101449**	2	X	0	7E-05	1
X	0.000	0	X	0.00	0	0.00063995	**infinity**	3	X	0	X	0
X	0.000	0	X	0.00	0	0.00035645	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00031627	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00014102	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00011826	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00011964	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00008768	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00008094	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00007374	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00005606	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00005271	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00004616	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00004004	**infinity**	2	X	0	X	0
X	0.000	0	X	0.00	0	0.00003316	**infinity**	2	X	0	X	0

Table 2.3. *Proteins identified by MudPIT analyses of GFP-PfAtg8 immunoprecipitations.* In each analysis, the following parameters are specified for each protein: peptide counts (column headers end with _P); total spectral counts (_S); spectral counts for peptides shared between protein isoforms (_sS); peptide counts for peptides uniquely matching the protein (_uDP); spectral counts for peptides uniquely matching the protein (_uS); spectral counts distributed according to the spectral count contribution of peptides unique to each isoform (_dS); sequence coverage as a percentage (_SC); distributed Normalized Spectral Abundance Factors (_dNSAF), calculated based on distributed spectral counts and protein length. The number of times a protein is detected across replicates is listed in the "Detected # Out of" columns.

Subsequently, a gene ontology (GO) analysis was performed to find functional enrichments among the proteins associated with PfAtg8. A relatively large overlap of significantly enriched GO terms was observed between the untreated and bafilomycin-treated early trophozoite samples, as well as between the treated early trophozoite and the treated late schizont samples (Fig. 2.9A).

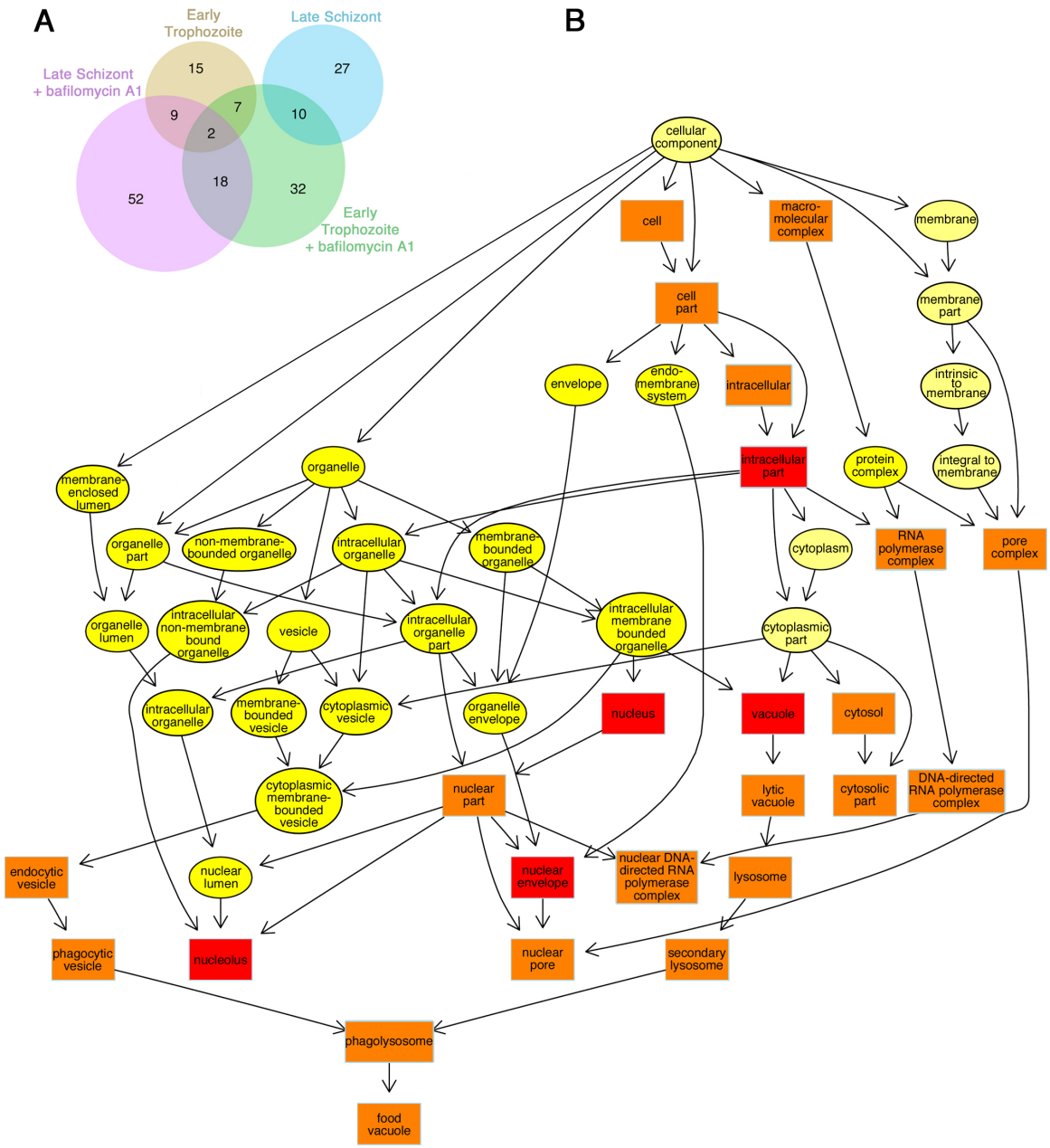


Figure 2.9. *Gene ontology (GO) enrichment analysis of PfAtg8-associated proteins.* A) Venn diagram of enriched GO terms among PfAtg8-associated proteins detected in mass spectrometry samples. The number of significantly enriched GO terms is shown for proteins associated with Atg8 in early trophozoite (ET) or late schizont (LS) stages, either untreated or treated with bafilomycin. B) Visualization of the network structure of significantly enriched GO terms. In total, 24 GO terms were enriched with a p-value < 0.01. These terms are shown in the diagram as rectangles. All interconnecting GO terms are shown in ovals. GO terms with lower p-values are labeled in red, while GO terms with higher p-values are labeled increasingly more yellow.

Cellular Component	P-value	Gene ontology term
GO:0044428	4.50399E-05	nuclear part
GO:0005634	0.000101356	nucleus
GO:0044424	0.000378797	intracellular part
GO:0005635	0.000443458	nuclear envelope
GO:0005730	0.000830602	nucleolus
GO:0032991	0.001189583	macromolecular complex
GO:0005829	0.001546748	cytosol
GO:0005622	0.001669865	intracellular
GO:0005773	0.002253456	vacuole
GO:0005643	0.0023142	nuclear pore
GO:0046930	0.0023142	pore complex
GO:0043234	0.002508582	protein complex
GO:0000428	0.003028735	DNA-directed RNA polymerase complex
GO:0030880	0.003028735	RNA polymerase complex
GO:0055029	0.003028735	nuclear DNA-directed RNA polymerase complex
GO:0005767	0.003864872	secondary lysosome
GO:0020020	0.003864872	food vacuole
GO:0032010	0.003864872	phagolysosome
GO:0045335	0.003864872	phagocytic vesicle
GO:0044446	0.004779569	intracellular organelle part
GO:0030139	0.004828474	endocytic vesicle
GO:0044422	0.005011077	organelle part
GO:0000323	0.005924728	lytic vacuole
GO:0005764	0.005924728	lysosome

Table 2.4. *Enrichment of gene ontology (GO) terms of proteins associated with PfAtg8.* Proteins associated with PfAtg8, as identified by MudPIT, were analyzed for enrichment of GO terms from the three GO domains (i.e. cellular component, biological process or molecular function).

Among all 176 proteins that were found to be associated with PfAtg8, significantly enriched GO terms included those involved with the digestive food vacuole, the phagolysosome, and the nucleus (Fig. 2.9B; Table 2.4).

In line with the fact that more than half of *P. falciparum*'s proteome is incompletely characterized, approximately one third of the proteins identified to be associated with PfAtg8 were hypothetical proteins. These hypothetical proteins were BLASTed against the yeast genome to find evidence for their function. We found homology of PF3D7_1339700 to the yeast protein Nvj1 that is present at the formation of nucleus-vacuole junctions.[30-32] In addition, we

detected homology of PF3D7_1459000 to yeast Bre5, which is associated with the process of ribophagy.[33, 34] Further validation will be required to determine if these homologous proteins are indeed involved in selective autophagy pathways and whether they have a functional role in the human malaria parasite. Collectively, this data indicates that a functional Atg8 conjugating system is detected at the proteomic level and highlights the multiple facets of this pathway in the parasite's infectious cycle progression.

Discussion

The data presented in this study provide a basic understanding of the role of autophagy in the human malaria parasite. Our *in silico*, cellular, and molecular approaches demonstrate that the canonical autophagy pathway has become modified in *P. falciparum*. Similar to *T. gondii*, *P. falciparum* encodes homologues for only a subset of yeast Atg proteins.[6, 7] Upstream kinases involved in detection of nutrient limitation and induction of autophagy are absent. *P. falciparum* is an intracellular parasite residing inside the red blood cell where hemoglobin, its primary nutritional source, is abundant. The parasite may have lost genes required for the nutrient-dependent induction of autophagy, as seen in other systems.[35] This was further confirmed by our observation that autophagy cannot be induced by mild nutrient starvation and specific inhibitors of key upstream kinase proteins in the pathway. Our results are in agreement with a recent study that showed that *P. falciparum* enters a hibernatory state in isoleucine free media.[24]

On the contrary, most downstream proteins involved with autophagy were detected in *P. falciparum* and appear to be functional. PfAtg genes involved with autophagosome formation, retrieval, and vesicle breakdown were clearly identified. From the 12 Atg genes uncovered from our HMM search, our proteomic analysis identified 7 PfAtg proteins, although PfAtg2, PfAtg5, PfAtg6, and PfAtg18 were detected at low abundance levels. Proteins of the Atg8

conjugation system (PfAtg7, PfAtg8, and two isoforms of PfAtg3) were significantly enriched in the GFP-Atg8 pull-downs, with the exception of PfAtg4. The role of PfAtg4 in priming PfAtg8 for PE lipidation remains uncertain. PfAtg8 was primarily found conjugated to PE and proteolytic cleavage by PfAtg4 may not be required for lipidation since PfAtg8's C-terminal residue is a glycine. Further investigation will be required to determine the exact role of PfAtg4 in recycling PfAtg8 from membranes.

In addition to the detection of PfAtg8 localizing with the apicoplast[12, 13] and its possible involvement in apicoplast biogenesis, we report for the first time the presence of PfAtg8-coated vesicles throughout the entire asexual and sexual erythrocytic cycle. Over the years, protein trafficking to the apicoplast has remained ambiguous[15, 20, 36, 37], and the inability to isolate the apicoplast from *P. falciparum* has been a limiting factor to understanding the biogenesis, trafficking, and segregation of the apicoplast. Apicoplast proteins derived from nuclear-encoded genes are known to utilize a bipartite N-terminal signal peptide to enter the secretory pathway and a transit peptide for apicoplast import.[20, 36] The two outer membranes of the apicoplast are ER-related.[38] Similarly, the preautophagosomal structure may also derive from the ER membrane.[39-41] It is therefore possible that these two pathways could interact with each other. In apicoplast-minus parasites, PfAtg8 localized with vesicles that contain apicoplast-targeted proteins, which suggests an alternative form of bulk protein

trafficking to the apicoplast. Further examination into PfAtg8 cargo proteins will be necessary to determine the role of PfAtg8 and the autophagy machinery's involvement with apicoplast biogenesis.

Interestingly, known inhibitors of the autophagy pathway and parasite development seem to interfere with apicoplast biogenesis. The branch-like structure of the apicoplast observed during the schizont stage was destroyed into multiple foci in both apicoplast-minus parasites and cultures treated with lysosomotropic agents. Both chloroquine and bafilomycin A1 are (V)-ATPase inhibitors that cause vacuolar deacidifcation and inhibit fusion of autophagosomes to lysosomes. Bafilomycin A1 has also been reported to block transport within the endocytic pathway[42], but this is the first report of bafilomycin blocking transport to a plastid organelle. Chloroquine is an antimalarial that is known to accumulate in the digestive food vacuole and inhibit the formation of hemozoin, the byproduct of hemoglobin catabolism.[43] In human fibroblasts, chloroquine is also known to raise intralysosomal pH and disrupt the intracellular pathway for newly synthesized acid hydrolases and uptake by pinocytosis.[44] Reports of cytoplasmatic vacuolization, similar to autophagic-like cell death, have been observed in parasites after chloroquine treatment, but whether or not unicellular protists undergo apoptosis remains ambiguous.[26] Our lysosomotropic drug-treated parasites displayed an accumulation of PfAtg8 vesicles and apicoplast fragmentation, which could be interpreted as either cytoplasmatic

vacuolization or an inhibition of vesicle fusion. In comparison, incubation with the antimalarial artemisinin did not induce autophagic-like cell death or vacuolization. Further investigations will be needed to determine whether these lysosomotropic agents inhibit fusion to the apicoplast or cause apicoplast fragmentation by vacuolization.

While the asexual stage of the erythrocytic cycle does not last for more than 48 hours, the differentiation into gametocytes can take up to two weeks.[45] During this time, the parasite resides within the same red blood cell and will need to recycle its proteins in order to survive. The abundance and atypical localization of PfAtg8 in gametocytes suggests that autophagy indeed plays an important role in protein turnover and parasite survival. It has been previously reported that the autophagy pathway degrades unnecessary organelles, such as micronemes, during sporozoite metamorphosis into merozoites in the liver stage.[10] Our data shows that localization of PfAtg8 remains on the apicoplast, but there is an increase in autophagosomes during gametocytogenesis. Considering the apparent role of PfAtg8 in gametocyte development, the observation that harsh nutrient starvation induced gametocytogenesis within 48 hours without reinvasion of a new erythrocyte may be indicative of a specialized role of autophagy in cell differentiation. Possibly, a delayed or modified form of autophagy induction could trigger differentiation into the sexual stage in *P. falciparum*. Further studies into

this process are required to determine the developmental cues necessary for gametocytogenesis.

Conclusion

Collectively, this study presents the critical role of the modified autophagy pathway in the human malaria parasite. PfAtg8 has evolved to participate in multiple functions including, but not limited to, cell differentiation, protein turnover, apicoplast biogenesis, and partitioning into daughter cells. While macroautophagy is the most well studied type of autophagy, our proteomic approach validates the importance of this pathway, but suggests that ribophagy and piecemeal microautophagy of the nucleus may also occur in *P. falciparum*, although additional experiments will be necessary to validate these findings. A detailed understanding of the molecular mechanisms regulating this pathway throughout the human malaria parasite life cycle will not only increase our knowledge of parasite biology, but will most likely uncover novel antimalarial strategies.

Materials and Methods

P. falciparum parasite culture

The following reagents were obtained through the MR4 as part of the BEI Resources Repository, NIAID, NIH: *Plasmodium falciparum* 3D7, MRA-102, deposited by DJ Carucci; Dd2, MRA-156, deposited by TE Wellems; D10 ACP(leader)-GFP, MRA-568, deposited by AF Cowman; W2, MRA-157, deposited by DE Kyle; 3D7HT-GFP, MRA-1029, deposited by A Talman and R Sinden; and NF54, MRA-1000, deposited by M Dowler, Walter Reed Army Institute of Research. *P. falciparum* strains were cultured in human O^+ erythrocytes as previously described.[46] The D10 ACP-GFP transfectant strain was supplemented with 100 nM pyrimethamine.[47]

HMM search of Autophagy related Pfam domains

Functional Pfam domains found in Atg genes were examined in 11 genomes: *P. falciparum* (release version 8.1), *P. vivax* (release version 8.1), *P. knowlesi* (release version 8.1), *P. yoelii* (release version 8.1), *P. chabaudi* (release version 8.1), *P. berghei* (release version 8.0), *Toxoplasma gondii* (release version 7.1), *Cryptosporidium hominis* (release version 4.5), *C. parvum* (release version 4.5),

Saccharomyces cerevisiae, and *Homo sapiens.* The Pfamm HMM profiles were downloaded at http://pfam.janelia.org from their Pfam_ls HMM library version 26.0. Additional information regarding Pfam accession numbers are given in Table 2.5. HMM searches were performed, as previously described, in a series of threshold E-values, increasing stringency from E-value <0.5 to E-value <0.05 (data not shown).[16] An E-value of 0.05 gave the most consistent results when compared to previously published results.

Atg genes	Pfam accession number	Domain
ATG1	PF00069.19	Pkinase
ATG10	PF03987	autophagy_act_c
ATG11	PF10377.3	
ATG12	PF04110.7	AGP12
ATG13	PF10033.3	
ATG14	PF10186.3	
ATG15	PF01764.1	Lipase_3
ATG16	PF08614.5	
ATG17	PF04108.6	APG17
ATG18	PF00400.26	WD40
ATG19	PF12744.1	
ATG2	PF09333.5	ATG C
ATG20	PF00787.18	PX
ATG20	PF09325.4	Vps5
ATG21	PF00400.26	WD40
ATG22	PF11700.2	
ATG24	PF00787.18	PX
ATG24	PF09325.4	Vps5
ATG25	PF11559.2	ADIP
ATG26	PF00169.23	PH
ATG26	PF00201.12	UDPGT
ATG26	PF02893.14	GRAM
ATG26	PF03033.14	glyco_transf_28
ATG27	PF09451.4	
ATG3	PF03986.7	Nterm
ATG3	PF03987.9	autophagy_act_c
ATG3	PF10381.3	Cterm
ATG31	PF09795.3	
ATG4	PF03416.13	peptidase_C54
ATG5	PF04106.6	APG5
ATG6	PF04111.6	APG6 (Vps30)
ATG7	PF00899.15	ThiF
ATG8	PF02991.10	
ATG9	PF04109.10	ATG9
mTOR	PF00454.21	PI3_P14_kinase
mTOR	PF02259.17	FAT
mTOR	PF08771.5	Rapamycin_bind
Vac8	pf00514.17	ARM

Table 2.5. *Pfam HMM profiles for autophagy related gene HMM search.* Pfam accession numbers and domains that correspond to Atg genes are listed.

Generation of PfAtg8 antibody

PfAtg8 was amplified from cDNA using primers forward: CAA TCTAGA CCCGGG AATAATTTTGTTTAACTTTAAGAAG and reverse: CAA GGATCC CCATGG GG GTATAGGGGACATACCGCTGC, bearing the restriction enzyme sites XbaI/XmaI and BamHI/NcoI, respectively (underlined). The fragment was inserted into the pGS21a plasmid vector using the XbaI and BamHI restriction sites. PfAtg8-His_{6x} recombinant protein was expressed in arctic bacterial cells, pulled down using Ni-NTA agarose beads (Qiagen), and then purified and concentrated by 10 and 30 kDa size exclusion centrifugation (Millipore). Anti-PfAtg8 antisera were raised in rabbits and were then purified using protein A/G (Thermo Fisher Scientific).

Cloning of P.falciparum GFP-fused Atg8

The DNA corresponding to *P. falciparum* Atg8 orthologue (PF10_0193) was obtained by PCR from cDNA using primers forward: TCCC CCCGGG ATGCCATCGCTTAAAGACG and reverse: ACGC GTCGAC CTGCAG TTATCCTAGACAACTCTCACAAC, bearing the restriction enzyme sites XmaI

and SalI/PstI, respectively (underlined). The fragment was cloned into the pCC-1 vector using the XmaI and SalI restriction sites. GFP was then amplified from the pARL/GFP vector using the primers forward: CCTG CTCGAG ATGAGTAAAGGAGAAGAACT and reverse: TCCC CCCGGG TTGTATAGTTCATCCATGCC, bearing the restriction enzyme sites XhoI and XmaI, respectively (underlined). The fragment was cloned into the N-terminal of the pCC-1/PfAtg8 vector to make pCC-1/GFP-PfAtg8. Using the primers forward: CAGG CTCGAG ACTAGT ATGAGTAAAGGAGAAGAACT and reverse: GGTC ACCGGT CCCGGG GTCGAC TTATCCTAGACAACTCTCAC, bearing the restriction enzyme sites XhoI/SpeI and AgeI/XmaI/SalI, respectively (underlined), the GFP-PfAtg8 fragment was amplified and inserted into the pARL vector using XhoI and AgeI restriction sites. The plasmid pARL/GFP-PfAtg8 was then used for transient expression of N-terminally GFP-tagged Atg8 in *P. falciparum*.

Generation of GFP-PfAtg8 transient expressing P. falciparum cell lines

Parasite transfection was carried out as previously described.[48] 3D7 parasite cultures were synchronized with 5% sorbitol and ring stage parasites were used for electroporation with 40 to 50 μg of plasmid DNA. Each electroporation was performed in a 0.2 cm cuvette (Bio-rad), with time constants ranging from 11-14 ms. The following day mature parasites were separated by magnetic purification and uninfected RBCs were electroporated with 40 to 50 μg of plasmid DNA. All

electroporations were pooled and 50 μl of fresh RBC was added to the culture. The transfected cultures were supplemented with 5 nM WR99210 for drug selection.[47]

Subcellular fractionation

Parasite subcellular fractions were performed as previously described.[49] Parasites were extracted from red blood cells using 0.15% saponin. Parasite pellets were incubated with cytoplasmic lysis buffer (0.65% Igepal, 10 mM Tris HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 2 mM AEBSF, 1X Roche protease inhibitor tablet) for 5 minutes followed by centrifugation at 2,000 rpm for 10 minutes. Samples were washed with PBS and were subsequently incubated with nuclear lysis buffer (10 mM HEPES pH 7.9, 0.1 mM EGTA, 0.1 mM EDTA, 1.5 mM MgCl2, 420 mM NaCl, 0.5 mM DTT, 25% glycerol, 2 mM PMSF, 1x Roche protease inhibitor tablet) for 20 minutes, followed by centrifugation at 6,000 rpm for 10 minutes. Samples were washed with PBS, incubated with solubilization lysis buffer (2% triton x-100, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM EGTA, 2 mM PMSF, 1x Roche protease inhibitor tablet) for 30 minutes and were then pelleted at 13,000 rpm for 10 minutes.[50] Protein concentration was determined by the Bradford method[51] and fractions were analyzed by western blot.

Western blot analysis

Parasite cultures were incubated in lysis buffer (1% Triton x-100, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM EGTA, 2 mM PMSF, 1x Roche protease inhibitor tablet) for 30 min on ice and then mechanically lysed by needle and syringe. Samples were then centrifuged at 10,000 rpm for 15 minutes and the supernatant was removed. Equal amounts of protein, 10 μg, were used for immunoblotting. Regular protein separations were performed using 16.5% Tris-Tricine SDS-PAGE gels (Bio-rad). Primary antibodies anti-PfAtg8 (Thermo Fisher Scientific) and anti-GFP (Abcam, ab290) were diluted 1:1,500 and 1:3,500, respectively. Secondary antibody goat anti-rabbit IgG HRP conjugate (Bio-Rad) was incubated at a dilution of 1:25,000.

Fluorescence microscopy

P. falciparum parasites were permeabilized with 0.015% saponin, fixed with 4% paraformaldehyde, and blocked with IFA buffer (2% BSA, 0.05% Tween-20, 100 mM glycine, 3 mM EDTA, 150 mM NaCl, 1x PBS). Cells were incubated with anti-PfAtg8 antibody (Thermo Fisher Scientific; 1:400) followed by anti-rabbit DyLight 550 (abcam ab98489; 1:500). DAPI (Invitrogen) was included in a wash at a final concentration of 5 μg/ml. Images were acquired using the Olympus BX40 epifluorescence microscope or the Leica SP5 confocal microscope.

Generation of apicoplast-minus P. falciparum

Following a similar protocol described by Yeh and DeRisi[21], parasite cultures were synchronized using 5% sorbitol and ring-stage parasites at 3% parasitemia were grown in 6-well plates (Costar #3506). The antibiotic chloramphenicol (Sigma) was added to a final concentration of 100 μM and incubated for 48 hours. After parasites reinvaded red blood cells, isopentenyl pyrophosphate (IPP; Isoprenoids LC) was added to a final concentration of 200 μM to rescue apicoplast-minus parasites. Parasites were grown for an additional generation before sampling cultures for immunofluorescence assays.

P. falciparum starvation

P. falciparum cutures were synchronized by 5% sorbitol. Early and late stage parasites were grown for 6, 24, and 72 hours in serum-free complete media or in starvation media (0.043 mg/ml gentamicin, 38.5 mM HEPES, 0.18% sodium bicarbonate, 0.003 mM NaOH in RPMI 1640 (Corning)).

SYBR Green I-based fluorescence assay

Bafilomycin was diluted to 10 μM and serially diluted by thirds into clear bottom 96-well plates (Costar #3904). *P. falciparum* cultures were added at a 2.5% hematocrit with a 1% parasitemia to 96-well plates and incubated for 72 hours at 37ºC. Plates were frozen at -80ºC overnight. Lysis buffer (20 mM Tris-HCl, 5 mM

EDTA, 0.008% saponin, 0.08% Triton X-100, and 0.2μl/ml SYBR Green I dye (Invitrogen) was added to thawed plates and incubated at 37ºC. The plates were read using the Spectra Max Gemini EM reader (Molecular Devices). Data were then analyzed with SoftMax Pro v5 (Molecular Devices Software, Inc.).[52]

Immunoprecipitation

Parasites were extracted from red blood cells using 0.15% saponin. To enrich for Atg8 vesicles, samples were incubated with 75 nM of bafilomycin for 3 hours before harvest. Approximately 1 ml of parasite pellets was homogenized by needle and syringe in a total volume of 3.5 ml of lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 1 mM PMSF, 1% Triton x-100, 5μM E-64, Roche complete EDTA-free protease inhibitor cocktail). Lysates were centrifuged at 13,000 rpm for 10 minutes and the supernatant was transferred to a clean tube. The supernatant was precleared with PureProteome protein G magnetic beads (Millipore) for 1.5 hours. Precleared beads were removed and either anti-PfAtg8 or anti-GFP (Abcam #ab290) antibodies were incubated for 6 hours at 4 μg/ml before fresh beads were added for 30 minutes. Our custom anti-PfAtg8 antibody was found to not efficiently pull-down endogenous PfAtg8 (data not shown). Resin-containing immune complexes were washed with wash buffers A (1% NP40, 1mM EDTA, 1%BSA in PBS), B (buffer A plus 0.5 M NaCl), C (1% NP40, 1mM EDTA in PBS), and D (1 mM EDTA in PBS). Proteins were first

eluted with 0.1 M glycine, pH 2.8 and neutralized with 100 mM Tris-HCl, pH 8.0, followed by a second elution using the same solution, a third elution using 10% 1,4-dioxane, and a fourth elution with elution buffer (50 mM Tris-HCl pH 6.8, 100 mM DTT, 2% SDS) for 40 min at 65º C. Eluates were pooled and precipitated with 20% trichloroacetic acid (TCA). The resulting pellet was washed once with 10% TCA and four times with cold acetone.

MudPIT analysis of immunoprecipitated proteins

TCA-precipitated proteins were solubilized in 100 mM Tris-HCl, pH 8.5, 8M Urea, 5mM Tris(2-Carboxylethyl)-Phosphine Hydrochloride before chloroacetamide was added to a final concentration of 10 mM. Endoproteinase Lys-C was used at 1:100 w/w to digest proteins at overnight at 37ºC. Samples were then brought to a final concentration of 2 M urea with 100 mM Tris-HCl, pH 8.5 and 2 mM $CaCl_2$ before adding Trypsin at 1:100 w/w for a second overnight digestion at 37ºC. Formic acid was added to a final concentration of 5% before samples were loaded onto a split-triple-phase fused-silica microcapillary column[53], and placed in-line with an quaternary agilent 1100/1200 series HPLC (ThermoScientific) LTQ ion trap mass spectrometer equipped with a custom-made nanospray ionization source. Full MS spectra were recorded on the peptides over a 400 to 1,600 m/z range, followed by 5 tandem mass (MS/MS) events sequentially generated in a data-dependent manner on the first to fifth most intense ions selected from the

full MS spectrum (at 35% collision energy).[54] The dynamic exclusion was enabled for 120 sec. A total of 3 technical replicates were acquired for each of the four types of samples analyzed (early trophozoite with and without bafilomycin treatment and late schizont with or without bafilomycin treatment). GFP-immunoprecipitations from two batches of asynchronized wild-type 3D7 cells and two batches of asynchronized 3D7-HT-GFP cells were also analyzed as negative controls.

MudPIT data analysis

The MS/MS datasets were searched using SEQUEST against a protein database consisting of 5538 *P. falciparum* (PlasmoDB rerelease 9.1) and 34521 *H. sapiens* (NCBI 2012-08-27 release) proteins, complemented with 177 sequences from usual contaminants (human keratins, IgGs, proteolytic enzymes). To estimate false positive discovery rates (FDR), each sequence was randomized (keeping amino acid composition and length the same) and the resulting "shuffled" sequences were added to the "normal" DB (doubling its size) and searched at the same time. The GFP-PfAtg8 sequence was added to this database. To limit the number of false positive "shuffled" peptides passing the selection criteria, the filtering criteria were strict, notably enforcing that both peptide ends match the enzyme used for the digestion. The average protein FDR was 2.6%, peptide FDR was 1.06%, and the spectral FDR is 0.40%.[55]

Gene Ontology Analysis

Enrichment of gene ontology (GO) terms within each sample was analyzed using the software package 'topGO' (written in R and maintained by the BioConductor project).[56] For each GO domain (i.e. cellular component, biological process or molecular function), we compared the proteins identified by MudPIT to the full proteome of *P. falciparum* using the 'classic' algorithm in combination with a Fisher's exact test. In addition, GO enrichment was analyzed among all Atg8-associated proteins identified by MudPIT, irrespective of time point or drug treatment. Gene ontology structures were visualized for GO terms with a *p*-value < 0.01 using the R software package 'Rgraphviz'.[57] To illustrate overlaps in GO terms between the different samples, a Venn diagram was drawn including all GO terms with a *p*-value < 0.05 using the R software package 'venneuler'.[58]

List of Abbreviations

Atg: autophagy related gene

PE: phosphatidylethanolamine

PMN: piecemeal microautophagy of the nucleus

PAS: preautophagosomal structure

RBC: red blood cell

HMM: hidden Markov model

MudPIT: multidimensional protein identification technology

GO: gene ontology

Acknowledgments

This work was supported by the National Institutes of Health (grant R01 AI85077-01A1 to KLR and fellowship F31 AI096840-01 to SC) as well as the Human Frontier Science Program (grant LT00507/2011-L to EMB).

References

1. WHO U. World Malaria Report 2012. Geneva, Switzerland and New York, USA 2012.

2. Tsukada M, Ohsumi Y. Isolation and characterization of autophagy-defective mutants of Saccharomyces cerevisiae. FEBS Lett 1993; 333:169-74.

3. Mizushima N, Levine B, Cuervo AM, Klionsky DJ. Autophagy fights disease through cellular self-digestion. Nature 2008; 451:1069-75.

4. Klionsky DJ, Abdalla FC, Abeliovich H, Abraham RT, Acevedo-Arozena A, Adeli K, et al. Guidelines for the use and interpretation of assays for monitoring autophagy. Autophagy 2012; 8:445-544.

5. Nakatogawa H, Suzuki K, Kamada Y, Ohsumi Y. Dynamics and diversity in autophagy mechanisms: lessons from yeast. Nat Rev Mol Cell Biol 2009; 10:458-67.

6. Duszenko M, Ginger ML, Brennand A, Gualdron-Lopez M, Colombo MI, Coombs GH, et al. Autophagy in protists. Autophagy 2011; 7:127-58.

7. Brennand A, Gualdron-Lopez M, Coppens I, Rigden DJ, Ginger ML, Michels PA. Autophagy in parasitic protists: unique features and drug targets. Mol Biochem Parasitol 2011; 177:83-99.

8. Ghosh D, Walton JL, Roepe PD, Sinai AP. Autophagy is a cell death mechanism in Toxoplasma gondii. Cell Microbiol 2012; 14:589-607.

9. Besteiro S, Brooks CF, Striepen B, Dubremetz JF. Autophagy protein Atg3 is essential for maintaining mitochondrial integrity and for normal intracellular development of Toxoplasma gondii tachyzoites. PLoS Pathog 2011; 7:e1002416.

10. Jayabalasingham B, Bano N, Coppens I. Metamorphosis of the malaria parasite in the liver is associated with organelle clearance. Cell Res 2010; 20:1043-59.

11. Coppens I. Metamorphoses of malaria: the role of autophagy in parasite differentiation. Essays Biochem 2011; 51:127-36.

12. Kitamura K, Kishi-Itakura C, Tsuboi T, Sato S, Kita K, Ohta N, et al. Autophagy-related Atg8 localizes to the apicoplast of the human malaria parasite Plasmodium falciparum. PLoS One 2012; 7:e42977.

13. Eickel N, Kaiser G, Prado M, Burda PC, Roelli M, Stanway RR, et al. Features of autophagic cell death in Plasmodium liver-stage parasites. Autophagy 2013; 9:568-80.

14. van Dooren GG, Marti M, Tonkin CJ, Stimmler LM, Cowman AF, McFadden GI. Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of Plasmodium falciparum. Mol Microbiol 2005; 57:405-19.

15. van Dooren GG, Waller RF, Joiner KA, Roos DS, McFadden GI. Traffic jams: protein transport in Plasmodium falciparum. Parasitol Today 2000; 16:421-7.

16. Ponts N, Yang J, Chung DW, Prudhomme J, Girke T, Horrocks P, et al. Deciphering the ubiquitin-mediated pathway in apicomplexan parasites: a potential strategy to interfere with parasite virulence. PLoS One 2008; 3:e2386.

17. Lang T, Schaeffeler E, Bernreuther D, Bredschneider M, Wolf DH, Thumm M. Aut2p and Aut7p, two novel microtubule-associated proteins are essential for delivery of autophagic vesicles to the vacuole. EMBO J 1998; 17:3597-607.

18. Teter SA, Eggerton KP, Scott SV, Kim J, Fischer AM, Klionsky DJ. Degradation of lipid vesicles in the yeast vacuole requires function of Cvt17, a putative lipase. J Biol Chem 2001; 276:2083-7.

19. Nakatogawa H, Ichimura Y, Ohsumi Y. Atg8, a ubiquitin-like protein required for autophagosome formation, mediates membrane tethering and hemifusion. Cell 2007; 130:165-78.

20. Waller RF, Reed MB, Cowman AF, McFadden GI. Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. EMBO J 2000; 19:1794-802.

21. Yeh E, DeRisi JL. Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage Plasmodium falciparum. PLoS Biol 2011; 9:e1001138.

22. Baker DA. Malaria gametocytogenesis. Mol Biochem Parasitol 2010; 172:57-65.

23. Okamoto N, Spurck TP, Goodman CD, McFadden GI. Apicoplast and mitochondrion in gametocytogenesis of Plasmodium falciparum. Eukaryot Cell 2009; 8:128-32.

24. Babbitt SE, Altenhofen L, Cobbold SA, Istvan ES, Fennell C, Doerig C, et al. Plasmodium falciparum responds to amino acid starvation by entering into a hibernatory state. Proc Natl Acad Sci U S A 2012; 109:E3278-87.

25. Yamamoto A, Tagawa Y, Yoshimori T, Moriyama Y, Masaki R, Tashiro Y. Bafilomycin A1 prevents maturation of autophagic vacuoles by inhibiting fusion between autophagosomes and lysosomes in rat hepatoma cell line, H-4-II-E cells. Cell Struct Funct 1998; 23:33-42.

26. Totino PR, Daniel-Ribeiro CT, Corte-Real S, de Fatima Ferreira-da-Cruz M. Plasmodium falciparum: erythrocytic stages die by autophagic-like cell death under drug pressure. Exp Parasitol 2008; 118:478-86.

27. Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, Tarning J, et al. Artemisinin resistance in Plasmodium falciparum malaria. N Engl J Med 2009; 361:455-67.

28. Washburn MP, Wolters D, Yates JR, 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 2001; 19:242-7.

29. Talman AM, Blagborough AM, Sinden RE. A Plasmodium falciparum strain expressing GFP throughout the parasite's life-cycle. PloS one 2010; 5:e9156.

30. Kabeya Y, Kamada Y, Baba M, Takikawa H, Sasaki M, Ohsumi Y. Atg17 functions in cooperation with Atg1 and Atg13 in yeast autophagy. Mol Biol Cell 2005; 16:2544-53.

31. Dawaliby R, Mayer A. Microautophagy of the nucleus coincides with a vacuolar diffusion barrier at nuclear-vacuolar junctions. Mol Biol Cell 2010; 21:4173-83.

32. Mijaljica D, Prescott M, Devenish RJ. The intricacy of nuclear membrane dynamics during nucleophagy. Nucleus 2010; 1:213-23.

33. Kraft C, Deplazes A, Sohrmann M, Peter M. Mature ribosomes are selectively degraded upon starvation by an autophagy pathway requiring the Ubp3p/Bre5p ubiquitin protease. Nat Cell Biol 2008; 10:602-10.

34. Ossareh-Nazari B, Bonizec M, Cohen M, Dokudovskaya S, Delalande F, Schaeffer C, et al. Cdc48 and Ufd3, new partners of the ubiquitin protease Ubp3, are required for ribophagy. EMBO Rep 2010; 11:548-54.

35. Kamada Y, Funakoshi T, Shintani T, Nagano K, Ohsumi M, Ohsumi Y. Tor-mediated induction of autophagy via an Apg1 protein kinase complex. J Cell Biol 2000; 150:1507-13.

36. Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, et al. Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. Science 2003; 299:705-8.

37. Jackson KE, Pham JS, Kwek M, De Silva NS, Allen SM, Goodman CD, et al. Dual targeting of aminoacyl-tRNA synthetases to the apicoplast and cytosol in Plasmodium falciparum. Int J Parasitol 2012; 42:177-86.

38. Deponte M, Hoppe HC, Lee MC, Maier AG, Richard D, Rug M, et al. Wherever I may roam: Protein and membrane trafficking in P. falciparum-infected red blood cells. Mol Biochem Parasitol 2012; 186:95-116.

39. Axe EL, Walker SA, Manifava M, Chandra P, Roderick HL, Habermann A, et al. Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. J Cell Biol 2008; 182:685-701.

40. Hayashi-Nishino M, Fujita N, Noda T, Yamaguchi A, Yoshimori T, Yamamoto A. A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation. Nat Cell Biol 2009; 11:1433-7.

41. Yla-Anttila P, Vihinen H, Jokitalo E, Eskelinen EL. 3D tomography reveals connections between the phagophore and endoplasmic reticulum. Autophagy 2009; 5:1180-5.

42. Klionsky DJ, Elazar Z, Seglen PO, Rubinsztein DC. Does bafilomycin A1 block the fusion of autophagosomes with lysosomes? Autophagy 2008; 4:849-950.

43. Sullivan DJ, Jr., Gluzman IY, Russell DG, Goldberg DE. On the molecular mechanism of chloroquine's antimalarial action. Proceedings of the National Academy of Sciences of the United States of America 1996; 93:11865-70.

44. Gonzalez-Noriega A, Grubb JH, Talkad V, Sly WS. Chloroquine inhibits lysosomal enzyme pinocytosis and enhances lysosomal enzyme secretion by impairing receptor recycling. The Journal of cell biology 1980; 85:839-52.

45. Alano P. Plasmodium falciparum gametocytes: still many secrets of a hidden life. Mol Microbiol 2007; 66:291-302.

46. Trager W, Jensen JB. Human malaria parasites in continuous culture. 1976. J Parasitol 2005; 91:484-6.

47. de Koning-Ward TF, Fidock DA, Thathy V, Menard R, van Spaendonk RM, Waters AP, et al. The selectable marker human dihydrofolate reductase enables sequential genetic manipulation of the Plasmodium berghei genome. Mol Biochem Parasitol 2000; 106:199-212.

48. Wu Y, Sifri CD, Lei HH, Su XZ, Wellems TE. Transfection of Plasmodium falciparum within human red blood cells. Proc Natl Acad Sci U S A 1995; 92:973-7.

49. Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, et al. Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. Genome Res 2004; 14:2308-18.

50. Topolska AE, Lidgett A, Truman D, Fujioka H, Coppel RL. Characterization of a membrane-associated rhoptry protein of Plasmodium falciparum. The Journal of biological chemistry 2004; 279:4648-56.

51. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976; 72:248-54.

52. Cervantes S, Stout PE, Prudhomme J, Engel S, Bruton M, Cervantes M, et al. High content live cell imaging for the discovery of new antimalarial marine natural products. BMC Infect Dis 2012; 12:1.

53. McDonald WH, Ohi R, Miyamoto DT, Mitchison TJ, Yates JR, 3rd. Comparison of three directly coupled HPLC MS/MS strategies for identification of

proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. Int J Mass spectrom 2002; 219:245-51.

54. Zhang Y, Wen Z, Washburn MP, Florens L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. Anal Chem 2010; 82:2272-81.

55. Ponts N, Saraf A, Chung DW, Harris A, Prudhomme J, Washburn MP, et al. Unraveling the ubiquitome of the human malaria parasite. J Biol Chem 2011; 286:40320-30.

56. Alexa A, Rahnenfuhrer J. topGO: topGO: Enrichment analysis for Gene Ontology. 2010.

57. Gentry J, Long L, Gentleman R, Falcon S, Hahne F, Sarkar D, et al. Rgraphviz: Provides plotting capabilities for R graph objects.

58. Wilkinson L. venneuler: Venn and Euler Diagrams. . 2011.

Chapter 3

Degradation of histone proteins in *Plasmodium falciparum* by the autophagy machinery

[1,2]Serena Cervantes, [2]Evelien M. Bunnik, [2]Jacques Prudhomme, [2]Karine G. Le Roch

[1]Graduate Program in Cell, Molecular, and Developmental Biology, University of California, Riverside, CA, USA 92521, [2]Department of Cell Biology and Neuroscience, University of California, Riverside, CA, USA 92521

Summary

During the trophozoite stage of the human malaria parasite histone proteins are depleted. Chromatin structure has been shown to play an important role in gene regulation. Removal of histone proteins allows DNA transcription and replication machinery to access the DNA. The molecular process that regulates histone degradation in *Plasmodium falciparum* remains unknown. We hypothesize that an autophagic-like process envelops histone proteins into vesicles, which then transport these histones to the digestive food vacuole for degradation. Using inhibitors of the autophagy pathway, we show histone proteins accumulate in vesicles outside of the nucleus. Future studies will be necessary to validate these findings. This study highlights the unique transport system the malaria parasite employs to degrade histone proteins.

Abstract

During the asexual blood stage of the human malaria, *Plasmodium falciparum,* there is a significant decrease of histone proteins at the trophozoite stage. Since removal of histone proteins allows effector proteins to access DNA, it seems that *P. falciparum* has developed chromatin structure changes as an alternative mechanism to regulate transcriptional activity and initiation of DNA replication. Thus, histone protein removal is likely to be an important prerequisite for cell cycle progression. Here, we studied the molecular processes responsible for the time-dependent removal and degradation of histone proteins from the nucleus. We determined that the mechanism involved in degrading bulk amounts of histone protein is part of the autophagy machinery. Parasite cultures treated with bafilomycin A1, an inhibitor of autophagosome-lysosome fusion, displayed cell cycle arrest and an accumulation of histone proteins outside the nucleus. This is the first report of a rudimentary set of autophagy proteins participating in histone degradation. This previously unidentified, parasite-specific, pathway may lead to the development of agents to reduce cell cycle progression and transmission from human to mosquito.

Introduction

The human malaria parasite remains a major public health burden in poverty stricken countries. Recent estimates by the World Health Organization show that over 200 million cases of malaria and over 600,000 malarial deaths occurred in 2012 [1]. Currently, no vaccine is available and the most recent malaria vaccine candidate, RTS,S/AS01E, had a dismal efficacy of 16.8% over a 4-year period [2]. Treatment of malaria infections is largely dependent on a single antimalarial drug, artemisinin, which is the recommended and most effective drug therapy today. Resistance to artemisinin was first reported in 2008 on the Thai-Cambodia border [3] and has now spread to four countries (Burma, Cambodia, Thailand and Viet Nam). If these resistant strains appear in Africa, the consequences could be dire [1]. Understanding the complex life cycle of the human malaria parasite *Plasmodium falciparum* to uncover novel drug targets therefore remains a high priority.

The life cycle stage responsible for disease in humans is the erythrocytic cell cycle. During this 48-hour cell cycle, the parasite invades a red blood cell and develops through ring, trophozoite and schizont stages to release 16-32 daughter cells into the bloodstream. One of the processes that are critical for parasite development is gene expression. Disruption of mechanisms that control gene expression could thus serve as a potent strategy to fight malaria infections. Gene regulation in *P. falciparum* is ambiguous, but is likely to rely greatly on

posttranscriptional regulation and epigenetic mechanisms. The *P. falciparum* genome contains about a third of the number of transcription-associated proteins typically found in the genome of a free-living eukaryote, while it encodes a relatively high number of genes involved in regulating chromatin structure and mRNA metabolism [4,5]. Indeed, dramatic changes in chromatin structure have been observed during the transcriptionally active trophozoite stage of parasite's erythrocytic cell cycle [6,7], suggesting an important role of histone turnover in *P. falciparum* gene regulation.

Chromatin is a condensed DNA:protein complex that allows DNA to be stored in the confined space of the nucleus while still being accessible to the transcription machinery. The basic packaging unit, the nucleosome, consists of approximately 147 base pair of DNA wrapped around an octamer of histone proteins: two each of H2A, H2B, H3, and H4. In other eukaryotic organisms, histone displacement is associated with transcription and DNA replication. Higher rates of histone turnover are observed over coding regions with high transcriptional activity and other active DNA elements [8] [9]. Changes in histone deposition are mainly local, while global histone levels within the nucleus remain relatively stable during the cell cycle. In *P. falciparum* however, changes in histone abundance occur at a global level. Histone protein abundance in the ring stage is high but decreases by two to three folds during DNA replication in the early trophozoite stage, before reaching high levels again at the schizont stage

[7]. This histone depletion at the trophozoite stage has been validated at both the proteomic and chromatin structure level [7] [6] [10] [11]. Furthermore, it has been shown that post-translational modifications (PTM) of histone tails can influence the turnover rates of nucleosomes. Interestingly, global histone analysis revealed that the malaria parasite has a high content of acetylated lysine residues [12] [13] that could be linked to a significantly faster turnover rate. In this study, we investigated how the unusual depletion of histones during the erythrocytic cycle of the malaria parasite is regulated.

To maintain homeostasis of histone protein levels in yeast, non-chromatin bound histones are rapidly degraded in a Rad53 kinase-dependent manner by the ubiquitin-proteosome (UPS) pathway [14]. In addition to the UPS pathway, the autophagy pathway has been shown to remove damaged DNA that contains histone H1 and the variant histone γH2AX, a marker of DNA double-stranded breaks [15]. The autophagy pathway is a membrane-mediated process that envelopes material and transports vesicles to the lysosome where cargo proteins are recycled back to amino acids. Autophagy proteins were first identified in yeast under nutrient starvation, but the autophagy machinery has since been shown to be capable of selective and organelle specific degradation. The molecular mechanisms that regulate micronucleophagy require a number of autophagy-related (ATG) genes, but remains poorly understood [16] [17] [18].

In *P. falciparum,* mass spectrometry analysis of histone PTMs failed to detect the modifications associated to histone turnover using the classical proteasome system [19] [13] [12]. We recently performed a functional analysis of the autophagy pathway in *P. falciparum* and discovered that autophagy is likely to be involved in protein turnover [Cervantes, manuscript submitted]. To study the molecular processes involved in histone degradation, we treated parasites with various inhibitors of both the UPS and autophagy pathway and analyzed the impact on histone degradation. Treatment of late ring staged parasite cultures with bafilomycin, a lysosomotropic agents that inhibit autophagosome and lysosome fusion [20] [21] [22], resulted in an accumulation of histone proteins. Histone degradation by the autophagy pathway was validated by localization experiments. Our results suggest that the autophagy pathway may be responsible for histone degradation. This overlooked pathway may thus regulate a key developmental event during the parasite asexual cell cycle.

Results

Inhibitors of the autophagy pathway affect histone protein degradation

Global histone proteins levels in *P. falciparum* are depleted at 18 hours post invasion (hpi) during the early trophozoite stage [6] [10] [11]. To determine which pathway plays a crucial role in histone degradation, we incubated tightly synchronized parasites for 6 hours with various inhibitors of the ubiquitin-proteasome system and the autophagy pathway (listed in Table 3.1). Treatment of parasites with the proteasome inhibitor MG-132 to reduce the amount of degradation of ubiquitinated substrates did not have an effect on histone protein levels (data not shown). Moreover, MG-132 treatment did not result in the detection of an enrichment of ubiquitinated residues on histone tails [Saraf 2013 unpublished]. Similarly, addition of the ubiquitin isopeptidase inhibitor NSC 632839 did not change histone protein levels (Figure 3.1). A combination of leupeptin and pepstatin A slightly increased histone protein levels (Figure 3.1).

Table 3.1. *Inhibitors used to treat parasite cultures and their effect on histone protein levels.*

Pathway	Molecule	Concentration	Effect	Histone protein
Ubiquitin-proteasome pathway	MG-132	400 nM	proteasome inhibitor	Unchanged
	NSC 632839	30 μM	ubiquitin isopeptidase inhibitor	Unchanged
Proteases	E-64	25 μM	cysteine protease inhibitor	Small increase
	Leupeptin/Pepstatin A	150 μM/20 μM	serine and cysteine protease inhibitors	Small increase
Vesicle transport	Nocodazole	75 μM	micotubule polymerization inhibitor	Increase
Autophagy pathway	Bafilomycin A1	75 nM	inhibits autophagosome-lysosome fusion	Increase
	Wortmannin	100 nM	phosphoinositide 3-kinases	Increase

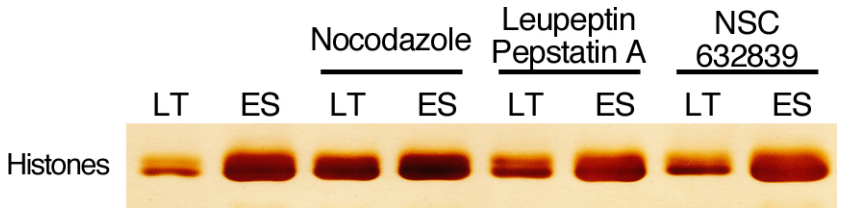


Figure 3.1. *Effect of degradation pathway inhibitor on global histone levels.* Histone proteins were extracted from equal amounts of synchronized parasites at the late trophozoite (LT) and early schizont (ES) stages, resolved on SDS-PAGE and detected by silver staining. Parasites were treated for 6 hours with 75 μM nocodazole, 150 μM leupeptin and 50 μM pepstatin A, and 30 μM NSC 632839. An increase of histone protein levels is observed when cultures were treated with nocodazole.

The protease inhibitor E-64 had a similar effect on histone levels in *P. falciparum* as leupeptin and pepstatin A (data not shown). However, when parasites were treated with nocadozole, which interferes with microtubule polymerization, an accumulation of histone protein levels was observed (Figure 3.1). These data led us to hypothesize that degradation of histone proteins may be dependent on vesicle-mediated transport via microtubules. To determine if the autophagy pathway may play a role in histone degradation, we incubated cultures with rapamycin and of wortmannin, both targeting phosphatidylinositol 3-kinases (PI3Ks). In the canonical autophagy pathway, PI3Ks are responsible for detecting nutrient levels and inducing autophagy. Rapamycin had no effect on histone protein levels, probably because *P. falciparum* does not have a homolog to the serine threonine kinase named “mammalian target of rapamycin” (mTor). Interestingly, the inhibitory concentration at 50% (IC_{50}) of rapamycin was 3.16 μM, a relatively high concentration compared to the 10-500 nM that is commonly

used to induce autophagy. On the other hand, incubation of parasites with wortmannin resulted in an increase of histone proteins. Although the kinase that regulates the autophagy machinery in *P. falciparum* remains unknown, these results suggest that PI3K is involved in histone degradation.

To validate the role of the autophagy machinery in histone degradation, histone protein accumulation was studied after incubation of parasites with an inhibitor of the fusion between autophagosome vesicles and the lysosome.. An equal number of synchronized parasites were incubated with bafilomycin A1 for 3 hours and collected at the late ring, early trophozoite, and late trophozoite stage. Histone precipitations were immunoblotted for histone H3. An additional coomassie stained gel was used as a loading control. Parasite cultures incubated with bafilomycin A1 showed an increase in histone H3 levels (Figure 3.2). This data suggests that inhibiting autophagosome-lysosome fusion causes an increase of histone protein in the malaria parasite.

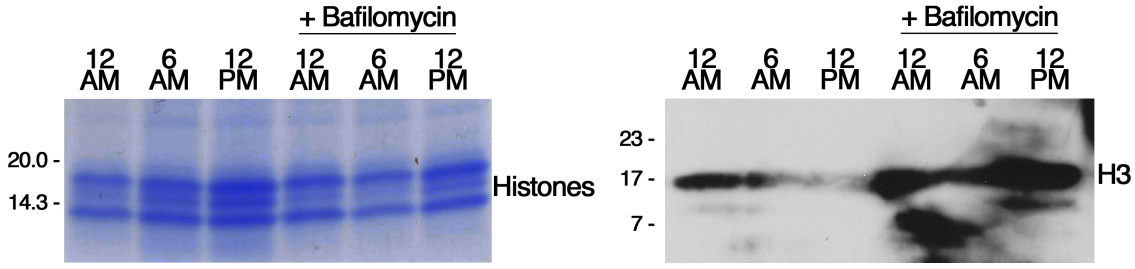
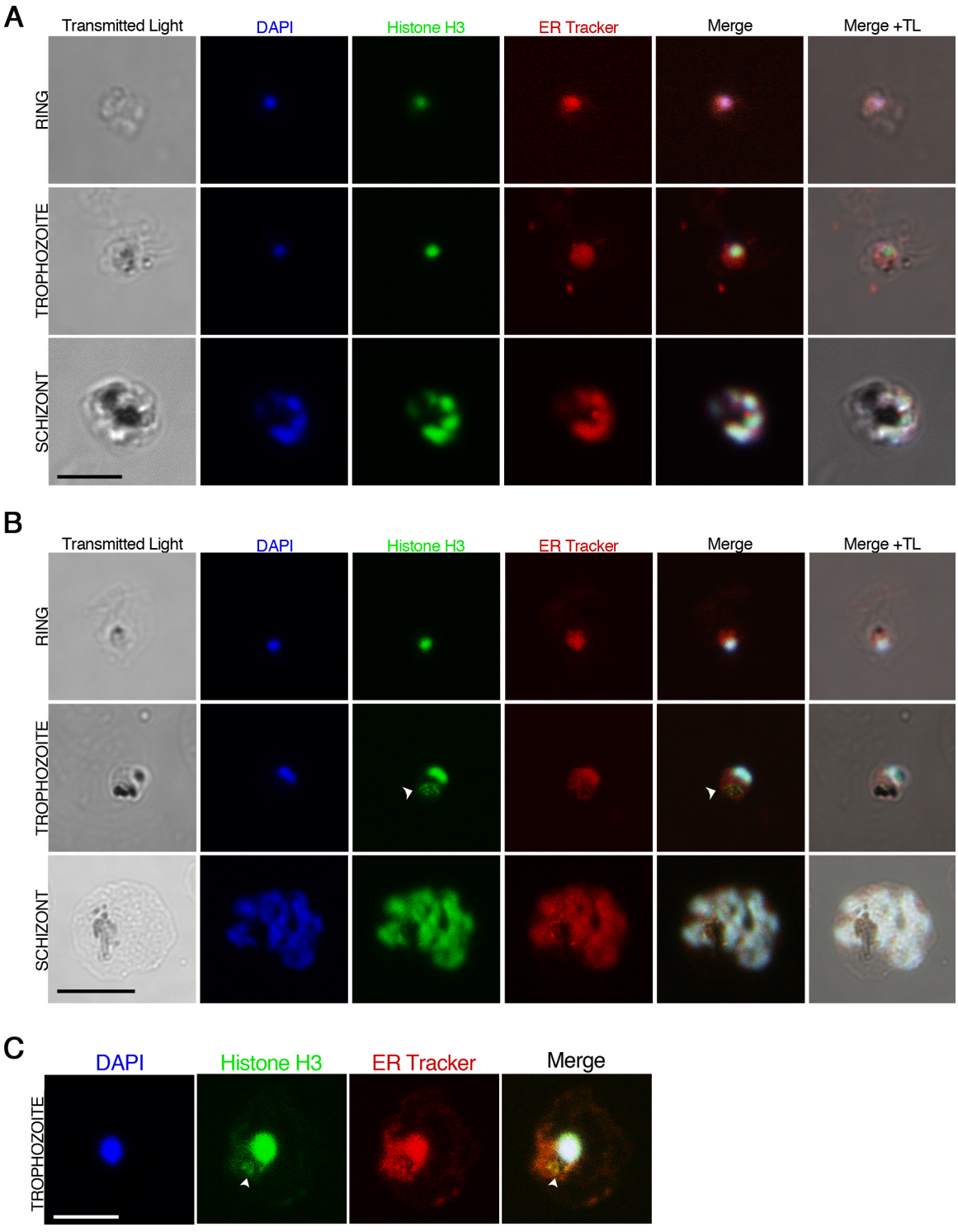


Figure 3.2. *Western blot analysis of histone H3 on histone precipitations.* Parasite cultures were tightly synchronized and treated with or without bafilomycin A1. Cultures were collected at the late ring (12 AM), early trophozoite (6 AM), and late trophozoite (12 PM) stages. Histones were precipitated and equal amounts of protein were loaded onto an SDS-PAGE. The coomassie stained gel was used as a loading control and an additional gel was immunoblotted for histone H3.

Localization of histone H3 in vesicles outside of the nucleus

To visualize the localization of histone proteins during the degradation process, we incubated synchronized parasites with bafilomycin A1 and nocodazole for 3 hours and subsequently performed immunofluorescence imaging by immunolabeling histone H3 on ring, trophozoite, and schizont stage parasites. Untreated parasites immunolabeled with histone H3 displayed overlap of histone proteins with the DNA marker DAPI (Figure 3.3A). A similar picture was observed for ring-stage and schizont-stage parasites treated with bafilomycin A1 (Figure 3.3B) or nocodazole (Figure 3.3C). However, treatment of parasites at the trophozoite stage, when histone proteins are degraded, resulted in localization of histone H3 both in the nucleus and in small vesicles outside of the nucleus. These results suggest that histones are exported out of the nucleus during the trophozoite stage, and that this process is dependent on vesicle transport.

Figure 3.3. *Extranuclear localization of histone H3 in trophozoite-stage parasites upon bafilomycin A1 or nocodazole treatment.* Parasite cultures were tightly synchronized and incubated with ER tracker (red) before immunolabeling with histone H3 (green). A) Untreated parasites show histone H3 localized with DNA marker DAPI (blue) at all stages of the asexual cycle. B) Parasites treated with 75 nM bafilomycin A1 for 3 hours before harvest. Histone H3 and DAPI show overlap at the ring and schizont stages. At the trophozoite stage histone H3 can be observed outside of the nucleus in small vesicles, indicated by white arrowheads. C) Trophozoite stage parasites incubated with 75 μM nocodazole display histone H3 staining outside of the nucleus in small vesicles. Scare bar indicates 5 μm.


A
Transmitted Light
DAPI
Histone H3
ER Tracker
Merge
Merge +TL
RING
TROPHOZOITE
SCHIZONT
B
Transmitted Light
DAPI
Histone H3
ER Tracker
Merge
Merge +TL
RING
TROPHOZOITE
SCHIZONT
C
DAPI
Histone H3
ER Tracker
Merge
TROPHOZOITE

Discussion

In *P. falciparum*, a large depletion of histone proteins has been observed at the trophozoite stage of the erythrocytic cell cycle [7] [11,26]. In this study, we show that this degradation of histones is most likely regulated through an autophagy-like process. At the trophozoite stage, treatment of parasites with an inhibitor of the downstream autophagy pathway resulted in accumulation of histone proteins outside the nucleus. Our data therefore suggest that proteins involved with the canonical autophagy pathway have adapted in *P. falciparum* to remove histone proteins in a time-dependent manner. Given the timing of histone degradation at the transcriptionally most active trophozoite stage, the removal of histone proteins may be a prerequisite for the global increase in transcription rates and initiation of DNA replication at this stage.

Autophagy is known to be involved in degradation of nuclear components. In yeast, piecemeal microautophagy of the nucleus (PMN) occurs at the nucleus-vacuole junction, where nonessential portions of the nucleus are pinched off and released into the vacuole [17]. Downstream ATG proteins are required for the PMN events, including proteins involved in vacuole enclosure and vesicle fusion. In addition, cells that undergo nuclear damage can have nuclear waste byproducts, such as histone variant γH2AX, removed by autophagy [15].

The organization of the nuclear envelope and nuclear pores in *P. falciparum* has recently been under investigation. Electron microscopy has been used to

construct a 3D model of the *P. falciparum* nucleus, showing that the number of nuclear pores increases at the trophozoite stage [27]. These nuclear pores appear to be located nearby loose euchromatin regions and not adjacent to heterochromatin. Comparative genomic studies have shown that *Plasmodium* parasites lack clear orthologues of most nuclear pore components [28]. Identification of a domain containing phenylalanine-glycine (FG) pairs of amino acids have suggested that nuclear pore components contain FG-repeats, and using this model a novel nuclear pore component (PF3D7_1446500) was identified [11]. More recently, an unusual nucleoporin of *P. falciparum,* PfSec13, was found to be a fusion between Sec13 and Nup145c of yeast and to associate with euchromatin [29]. In yeast, Sec13 forms a heterodimer with Sec31 to form a cage that will cause ER membrane to bud into COPII-coated vesicles, which are required for ER to golgi transport [30]. Moreover, a recent study showed that parasites communicate via exosome-like vesicles that are capable of transferring genes between individual parasites [31]. It thus seems likely that nuclear pore complexes associate with euchromatin, possibly assisting in budding off portions of the nuclear membrane into vesicles. Although future studies will be required to investigate the formation, cargo, and transport of these vesicles, we speculate that histone proteins may be encapsulated at the nuclear membrane and transported to the lysosome to be degraded.

Additional experiments will be required to validate our hypothesis of an autophagic-like process regulating histone depletion in the malaria parasite. ATG8 is the standard autophagosome marker and is essential to autophagosome formation. In yeast, ATG8 becomes lipidated and incorporated into autophagosome membranes, and is essential for autophagy in eukaryotic systems. Using a parasite strain that is transiently transfected with a GFP-PfATG8 fusion construct, we are planning to immunolabel histone H3 to confirm the localization of histone proteins within PfATG8 vesicles by immunofluorescence microscopy. To deduce whether PfATG8-coated vesicles transport histone proteins outside the nucleus and ultimately fuse with the lysosome to degrade the histones, we will perform immuno-transmission electron microscopy, using gold-labeled antibodies directed against PfATG8 and histone H3.Preliminary experiments to immunoprecipitate GFP-PfATG8 and assess the association of histone with PfATG8 by immunoblot were unsuccessful. We will therefore extract PfATG8 vesicles using ultra-centrifugation, isolate their content and subsequently identify cargo proteins by mass spectrometry. . In order for proteins to be enveloped by a membrane, a chaperone or scaffold protein may be involved to bind the highly basic histone. To identify proteins associated with histone H3, we will treat trophozoite stage parasites with bafilomycin A1 to induce accumulation of histone outside the nucleus, isolate the cytoplasmic protein fraction, perform immunoprecipitation of histone H3 and subsequently analyze

H3-associated proteins by mass spectrometry. Overall, these future experiments will strengthen our current data and clarify the role of autophagy in histone degradation.

Conclusion

In this study we demonstrate that histone proteins are depleted at the trophozoite stage by an autophagic-like process. Using nocodazole, an inhibitor of microtubule polymerization, and bafilomycin A1, an inhibitor of autophagosome-lysosome fusion, we show that histone proteins accumulate outside of the nucleus. This data suggests that histone proteins may be enveloped within a vesicle and transported to the digestive food vacuole. We speculate that removal of histone proteins is a prerequisite for gene transcription and DNA replication. Future studies will be required to identify a histone chaperone protein and validate that histone proteins are transported to the lysosome for degradation.

Materials and Methods

P. falciparum parasite culture

The following reagents were obtained through the MR4 as part of the BEI Resources Repository, NIAID, NIH: *Plasmodium falciparum* 3D7, MRA-102, deposited by DJ Carucci; Dd2, MRA-156, deposited by TE Wellems; D10 ACP(leader)-GFP, MRA-568, deposited by AF Cowman; W2, MRA-157, deposited by DE Kyle; 3D7HT-GFP, MRA-1029, deposited by A Talman and R Sinden; and NF54, MRA-1000, deposited by M Dowler, Walter Reed Army Institute of Research. *P. falciparum* strains were cultured in human O^+ erythrocytes as previously described [23]. The D10 ACP-GFP transfectant strain was supplemented with 100 nM pyrimethamine [24]. The GFP-PfAtg8 transfectant strain was supplemented with 5 nM WR99210. Cultures were treated with 75 nM bafilomycin, 400 nM MG-132, 75 μM nocodazole, 30 μM NSC 632829, 25 μM E-64, or 50 μM of pepstatin A and 150 μM leupeptin for 6 hours.

Immunofluorescence staining

Parasites cultures were rinsed in 0.015% saponin with incomplete media and air dried onto slides. ER-Tracker™ Red (Molecular Probes) was added for 30 minutes at a final concentration of 1 μM. Parasites were fixed with 4% paraformaldehyde in PBS for 15 minutes and permeablized with 0.1% Triton x-

100 for 10 minutes. Anti-histone H3 (abcam 1791) antibody was incubated in IFA buffer (2% BSA, 0.05% Tween-20, 100 mM glycine, 3 mM EDTA, 150 mM NaCl, 1x PBS) at 1/200 for 1.5 hours at room temperature. Slides were washed with 0.1% Tween-20 before incubation with secondary antibody (goat anti-rabbit IgG, FITC conjugate (AP132F)) in IFA buffer for 1 hour. Images were acquired with the fluorescent confocal Leica SP2 microscope.

Histone Precipitation

Parasites were extracted from RBCs by 0.15% saponin treatment. Nuclei were released in lysis buffer (25 mM Tris HCl pH 7.8, 1 mM EDTA, 0.25% [v/v] Igepal® CA-630, 2 mM PMSF, 20 mM NEM, and Roche complete mini EDTA-free protease inhibitor cocktail) [6], and washed with 0.8 M NaCl before histone acid extraction with 0.25 M HCl [25]. Histones were precipitated by trichloroacetic acid and resuspended in SDS-PAGE loading buffer.

Immunoprecipitation

Parasites were extracted from red blood cells (RBCs) with 0.15% saponin treatment. Immunoprecipitation buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 1 mM PMSF, 1% Triton x-100, and Roche protease inhibitor tablet) was added to parasite pellet, followed by mechanical lysis of parasites by syringe and needle. After centrifugation for 15 minutes at 13,000 g, parasite lysates were

collected and PureProteome magnetic protein A beads (Millipore) were added to preclear lysate for 2 hours. Precleared lysate was collected, the anti-GFP antibody (Abcam, ab290) was added to a final concentration of 5 μg/ml and the antibody-lysate suspension was incubated overnight at 4ºC. Immunocomplexes were collected by 30-minute incubation with protein A beads. Magnetic beads and bound protein complexes were washed with four wash buffers (A: 1x PBS, 1% NP40, 1 mM EDTA, and 1% BSA; B: buffer A with 0.5 M NaCl; C: 1x PBS, 1% NP40, 1 mM EDTA; D: 1x PBS, 1 mM EDTA). Bound proteins were eluted with 0.1 M glycine pH 2.8 for 5 minutes at room temperature and neutralized to 100 mM Tris-HCl.

Western blot

Immunoprecipitation samples were resolved on a 16.5% tris-tricine gel (Bio-rad) and transferred to PDVF membrane. Proteins were subsequenty detected by immunoblotting, using the following antibodies at the indicated concentrations: rabbit anti-Histone H3 (Abcam 1791) 1:3000; rabbit anti-GFP (Abcam, ab290) 1:3500; rabbit anti-PfAtg8 (ThermoFisher) 1:1500; goat anti-rabbit IgG HRP conjugate (Bio-rad) 1:25,000.

References

1. WHO U (2012) World Malaria Report 2012. Geneva, Switzerland and New York, USA.

2. Olotu A, Fegan G, Wambua J, Nyangweso G, Awuondo KO, et al. (2013) Four-year efficacy of RTS,S/AS01E and its interaction with malaria exposure. The New England journal of medicine 368: 1111-1120.

3. Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, et al. (2009) Artemisinin resistance in Plasmodium falciparum malaria. The New England journal of medicine 361: 455-467.

4. Coulson RM, Hall N, Ouzounis CA (2004) Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. Genome research 14: 1548-1554.

5. Gardner MJ, Hall N, Fung E, White O, Berriman M, et al. (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419: 498-511.

6. Ponts N, Harris EY, Prudhomme J, Wick I, Eckhardt-Ludka C, et al. (2010) Nucleosome landscape and control of transcription in the human malaria parasite. Genome research 20: 228-238.

7. Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, et al. (2004) Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. Genome research 14: 2308-2318.

8. Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, et al. (2007) Dynamics of replication-independent histone turnover in budding yeast. Science 315: 1405-1408.

9. Deal RB, Henikoff JG, Henikoff S (2010) Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. Science 328: 1161-1164.

10. Ponts N, Harris EY, Lonardi S, Le Roch KG (2011) Nucleosome occupancy at transcription start sites in the human malaria parasite: a hard-wired evolution of virulence? Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 11: 716-724.

11. Oehring SC, Woodcroft BJ, Moes S, Wetzel J, Dietz O, et al. (2012) Organellar proteomics reveals hundreds of novel nuclear proteins in the malaria parasite Plasmodium falciparum. Genome biology 13: R108.

12. Saraf A, Cervantes S, Ponts N, Sardiu ME, Chung DW, et al. Global Patterns of Histone Modifications throughout the Malaria Parasite Cell Cycle. Submitted.

13. Trelle MB, Salcedo-Amaya AM, Cohen AM, Stunnenberg HG, Jensen ON (2009) Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum. Journal of proteome research 8: 3439-3450.

14. Singh RK, Kabbaj MH, Paik J, Gunjan A (2009) Histone levels are regulated by phosphorylation and ubiquitylation-dependent proteolysis. Nature cell biology 11: 925-933.

15. Park YE, Hayashi YK, Bonne G, Arimura T, Noguchi S, et al. (2009) Autophagic degradation of nuclear components in mammalian cells. Autophagy 5: 795-804.

16. Roberts P, Moshitch-Moshkovitz S, Kvam E, O'Toole E, Winey M, et al. (2003) Piecemeal microautophagy of nucleus in Saccharomyces cerevisiae. Mol Biol Cell 14: 129-141.

17. Krick R, Muhe Y, Prick T, Bredschneider M, Bremer S, et al. (2009) Piecemeal microautophagy of the nucleus: genetic and morphological traits. Autophagy 5: 270-272.

18. Millen JI, Krick R, Prick T, Thumm M, Goldfarb DS (2009) Measuring piecemeal microautophagy of the nucleus in Saccharomyces cerevisiae. Autophagy 5: 75-81.

19. Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AM, et al. (2009) Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America 106: 9655-9660.

20. Mizushima N, Yamamoto A, Hatano M, Kobayashi Y, Kabeya Y, et al. (2001) Dissection of autophagosome formation using Apg5-deficient mouse embryonic stem cells. The Journal of cell biology 152: 657-668.

21. Shintani T, Klionsky DJ (2004) Autophagy in health and disease: a double-edged sword. Science 306: 990-995.

22. Levine B, Kroemer G (2008) Autophagy in the pathogenesis of disease. Cell 132: 27-42.

23. Trager W, Jensen JB (2005) Human malaria parasites in continuous culture. 1976. J Parasitol 91: 484-486.

24. de Koning-Ward TF, Fidock DA, Thathy V, Menard R, van Spaendonk RM, et al. (2000) The selectable marker human dihydrofolate reductase enables sequential genetic manipulation of the Plasmodium berghei genome. Mol Biochem Parasitol 106: 199-212.

25. Longhurst HJ, Holder AA (1997) The histones of Plasmodium falciparum: identification, purification and a possible role in the pathology of malaria. Parasitology 114 (Pt 5): 413-419.

26. Florens L, Liu X, Wang Y, Yang S, Schwartz O, et al. (2004) Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. Molecular and biochemical parasitology 135: 1-11.

27. Weiner A, Dahan-Pasternak N, Shimoni E, Shinder V, von Huth P, et al. (2011) 3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite Plasmodium falciparum. Cellular microbiology 13: 967-977.

28. Neumann N, Lundin D, Poole AM (2010) Comparative genomic evidence for a complete nuclear pore complex in the last eukaryotic common ancestor. PLoS One 5: e13241.

29. Dahan-Pasternak N, Nasereddin A, Kolevzon N, Pe'er M, Wong W, et al. (2013) PfSec13 is an unusual chromatin-associated nucleoporin of Plasmodium falciparum that is essential for parasite proliferation in human erythrocytes. Journal of cell science 126: 3055-3069.

30. Stagg SM, Gurkan C, Fowler DM, LaPointe P, Foss TR, et al. (2006) Structure of the Sec13/31 COPII coat cage. Nature 439: 234-238.

31. Regev-Rudzki N, Wilson DW, Carvalho TG, Sisquella X, Coleman BM, et al. (2013) Cell-Cell Communication between Malaria-Infected Red Blood Cells via Exosome-like Vesicles. Cell 153: 1120-1133.

Conclusion

In recent years, the malaria field has made substantial progress in understanding the mechanisms that regulate gene expression in the parasite. The data presented in this dissertation contribute to the understanding of parasite gene regulation, and support a model in which gene transcription during the asexual stage is regulated by opening and compacting of chromatin structure. When chromatin is “closed” during the ring and schizont stage, transcription rates are at low levels. Conversely, at the trophozoite stage DNA is “open”, the nucleus is larger in size, and transcription rates are high. With increasing support that the human malaria parasite is constitutively ready for active transcription, it becomes clear that the space required to access DNA must be regulated. Therefore, one of the interesting future directions for malaria research will be to focus on the movement of lipids from the host erythrocyte to the nuclear envelope at the early trophozoite stage. Understanding how the nuclear envelope (NE) enlarges to allow increases in transcription and initiation of DNA replication will be an important piece of the gene regulation puzzle.

Previous studies focused heavily on deciphering the ‘histone code’ in *P. falciparum*, with hopes that parasite specific post-translational modifications (PTMs) would regulate gene expression. Surprisingly, many genes in the parasite genome appear to be marked as transcriptionally active with high levels of

acetylation PTMs during the erythrocytic stage [1, 2]. Conversely, only a small number of genes with transcriptionally silencing PTMs, such as methylation, were detected. Methylation of histones appears to be restricted to virulence factors involved in antigenic variation that are clustered in subtelomeric regions [3]. Only a single *var* gene is expressed at a time, and switches in *var* variant expression contribute to evasion of detection by host immunity. In particular, methylation of H3K36 seems essential for transcriptional repression of *var* genes [4]. Our comprehensive map of PTMs data correlates with these previous findings. Interestingly, we detected additional novel PTMs sites on the globular domain of histone proteins, although their functions remain unknown. Characterization of these PTMs will be necessary to determine if they play a role in gene regulation.

From examining mass spectrometry data obtained to study histone PTMs, it appeared that histone proteins show a decrease in abundance at the trophozoite stage. Various groups with multiple experimental approaches have observed this phenomenon [5, 6]. Preliminary data indicated that the autophagy pathway may play a role in histone degradation, but at the time only a few publications examined autophagy in *Plasmodium*. The first research article describing the function of Atg8, an autophagosome membrane marker, in *P. falciparum* was published in 2012 and suggested PfAtg8 was involved with apicoplast biogenesis at the merozoite stage [7]. We made a similar observation, and in addition, we continued to localize and characterize PfAtg8 throughout the

entire erythrocytic stage and determined that the autophagy pathway is multifunctional. In addition to apicoplast biogenesis, the autophagy pathway is involved with, but not limited to, vesicle traffic, protein catabolism, and export and import into the host RBC. Moreover, we showed that the autophagy machinery is responsible for the transport of histone proteins before degradation by the food vacuole. This novel parasite pathway answers our hypothesis of the autophagy machinery regulating histone depletion, but it has created numerous questions regarding transport in *P. falciparum.*

Transport mechanisms remain ambiguous in *P. falciparum*, because the parasite has atypical features compared to model organisms. Within a single infected erythrocyte, there are additional organelles within the parasite (rhoptries, micronemes, the digestive food vacuole, and apicoplast) and sites within the host RBC (the parasitophorous vacuole and membrane, Maurer's clefts, and the cytoplasm and plasma membrane) that require proper transport of proteins. Transport from the nucleus is currently becoming a hot topic in cell biology. In 2012, the canonical view of nucleocytoplasmic transport changed from nuclear pore complexes (NPCs) being the sole gatekeepers to include nuclear envelope budding [8]. Large ribonucleoprotein particles that are too large to pass through the NPC, egress from the nucleus using a mechanism that was first discovered in the egress of herpes viruses. Instead of the virus disrupting the nuclear lamina during egress, it was shown that this pathway is an endogenous nuclear export

pathway. One year later, *P. falciparum* was found to communicate to each other and transfer DNA via extracellular vesicles that are exosome-like [9]. These exciting breakthroughs highlight gaps in vesicle traffic from the nucleus and will lead to subsequent characterization of nuclear transport systems. Deciphering the mechanisms that regulate vesicular traffic to and from the nucleus will remain a high priority within the field.

The atypical transport that occurs within the parasite may play a significant role in cell cycle progression and gene regulation. One area that has been largely overlooked is movement of lipids within the parasite and its contribution to NE expansion. The volume of the trophozoite nucleus is 14 times larger than the nucleus at the ring stage [10]. This dynamic change of the NE can be attributed to an increase in NPCs and lipids. The 3D structure of the nucleus throughout the asexual cycle was examined by an emerging electron microscope technique, known as 'slice and view', that allows acquisition of serial surface images using a scanning electron microscope [11]. At the ring stage a few NPC are clustered closely together, but NPC biogenesis at the trophozoite stage increases numbers up to 58 pores per nucleus, which are distributed evenly around the NE. NPC biogenesis stops before schizogony and subsequent mitotic divisions leave daughter merozoites with a few clustered NPCs. Although in *silico* analysis of the *Plasmodium* genome has been unsuccessful in identifying orthologues of most NPCs [12], studies are underway to identify novel nuclear pore components [5]. It

would be interesting to identify the composition in ring and trophozoite nuclei to determine the main protein contributor of membrane expansion. Speculation has been made that nuclear size does not determine the packing of the genetic matter, but whether nuclear size promotes gene transcription has not been addressed.

Currently, little is known about the mechanisms that regulate nuclear membrane remodeling in model organisms. NE lipid composition of rat liver is mostly composed of phospholipids at 65%, 15% are free fatty acids, 10% cholesterol, and the remaining lipids are cholesterol ester, diacylglycerol, and triacyglycerol [13]. To date the lipid composition of *P. falciparum's* NE has not been studied and remains unknown. However, alterations in RBC membrane phospholipid organization [14] and *Plasmodium* uptake of phospholipids from RBC and human serum have been observed [15]. Blocking lipid transport may cause the NE to remain small and refrain chromatin from obtaining an "open" state. By restricting the size of the nucleus DNA transcription and replication effectors cannot access transcriptionally permissive DNA due to chromatin compaction. A recent report showed starved parasites enter a hibenatory state at the mid-trophozoite stage and genes involved in glycolysis, translations, transcription, and apicoplast and mitochondrial function appear to remain normal, albeit at a slowed rate [16].

In addition, apicoplast-minus parasites, generated using antibiotics and supplementing cultures with isopentenyl pyrophosphate (IPP), can be grown indefinitely and suggest the only essential function of the apicoplast is isoprenoid precursor synthesis [17]. Apicomplexan protists rely on the prokaryotic MEP/DOXP/non-mevalonate pathway for synthesizing isoprenoid precursors instead of the canonical mevalonate pathway most eukaryotes and all mammals use [18]. Interestingly, the downstream isoprenoid products and their function during the asexual cycle are unknown in *P. falciparum*. However, IPP in plants are building blocks to sterols, which stabilize plasma membranes [19]. If synthesizing isoprenoid precursors is the essential function of the apicoplast, could IPP be converted to sterols in order to stabilize the host RBC plasma membrane? Is there an exchange of sterol and phospholipids that allows the parasite to utilize the phospholipids from the host RBC plasma membrane to enlarge the nuclear envelope? To address these questions, future work should focus on determining the lipid transport mechanisms required to enlarge the NE. Although *P. falciparum* is far from being considered a model organism, understanding the atypical transport mechanisms that the human malaria parasite employs could impact the malaria field as well as the vesicle transport field.

References

1. Salcedo-Amaya, A.M., et al., *Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(24): p. 9655-60.

2. Saraf, A., et al., *Global Patterns of Histone Modifications throughout the Malaria Parasite Cell Cycle.* Submitted.

3. Flueck, C., et al., *Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors.* PLoS pathogens, 2009. **5**(9): p. e1000569.

4. Jiang, L., et al., *PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum.* Nature, 2013.

5. Oehring, S.C., et al., *Organellar proteomics reveals hundreds of novel nuclear proteins in the malaria parasite Plasmodium falciparum.* Genome biology, 2012. **13**(11): p. R108.

6. Le Roch, K.G., et al., *Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle.* Genome research, 2004. **14**(11): p. 2308-18.

7. Kitamura, K., et al., *Autophagy-related Atg8 localizes to the apicoplast of the human malaria parasite Plasmodium falciparum.* PLoS One, 2012. **7**(8): p. e42977.

8. Speese, S.D., et al., *Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling.* Cell, 2012. **149**(4): p. 832-46.

9. Regev-Rudzki, N., et al., *Cell-Cell Communication between Malaria-Infected Red Blood Cells via Exosome-like Vesicles.* Cell, 2013. **153**(5): p. 1120-33.

10. Bannister, L.H., G. Margos, and J.M. Hopkins, *Molecular Approaches to Malaria*. Making a home for Plasmodium post-genomics: ultrastructural organization of the blood stages, ed. I.W. Sherman. Vol. bath:28626. 2005, Washington, D.D.: ASM Press.

11. Weiner, A., et al., *3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite Plasmodium falciparum.* Cellular microbiology, 2011. **13**(7): p. 967-77.

12. Neumann, N., D. Lundin, and A.M. Poole, *Comparative genomic evidence for a complete nuclear pore complex in the last eukaryotic common ancestor.* PLoS One, 2010. **5**(10): p. e13241.

13. Khandwala, A.S. and C.B. Kasper, *The fatty acid composition of individual phospholipids from rat liver nuclear membrane and nuclei.* The Journal of biological chemistry, 1971. **246**(20): p. 6242-6.

14. Maguire, P.A., J. Prudhomme, and I.W. Sherman, *Alterations in erythrocyte membrane phospholipid organization due to the intracellular growth of the human malaria parasite, Plasmodium falciparum.* Parasitology, 1991. **102 Pt 2**: p. 179-86.

15. Moll, G.N., et al., *Phospholipid uptake by Plasmodium knowlesi infected erythrocytes.* FEBS letters, 1988. **232**(2): p. 341-6.

16. Babbitt, S.E., et al., *Plasmodium falciparum responds to amino acid starvation by entering into a hibernatory state.* Proc Natl Acad Sci U S A, 2012. **109**(47): p. E3278-87.

17. Yeh, E. and J.L. DeRisi, *Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage Plasmodium falciparum.* PLoS Biol, 2011. **9**(8): p. e1001138.

18. Hoeffler, J.F., et al., *Isoprenoid biosynthesis in higher plants and in Escherichia coli: on the branching in the methylerythritol phosphate pathway and the independent biosynthesis of isopentenyl diphosphate and dimethylallyl diphosphate.* The Biochemical journal, 2002. **366**(Pt 2): p. 573-83.

19. Rohmer, M., *The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants.* Natural product reports, 1999. **16**(5): p. 565-74.

Appendices

Chapter 1a

High-content live cell imaging with RNA probes: advancements in high-throughput antimalarial drug discovery

Serena Cervantes[1], Jacques Prudhomme[2], David Carter[3], G. Gopi Krishna[4], Qian Li[6], Young-Tae Chang[4, 5] and Karine G. Le Roch[2*]

[1] Cell, Molecular, and Developmental Biology Graduate Program, University of California - Riverside, CA, 92521, USA. [2] Department of Cell Biology and Neurosciences, University of California - Riverside, CA, 92521, USA. [3] Institute for Integrative Genome Biology, University of California - Riverside, CA, 92521, USA. [4] Department of Chemistry, & NUS MedChem Program of the Office of Life Sciences, National University of Singapore, Singapore 11754. [5] Laboratory of Bioimaging Probe Development, Singapore Bioimaging Consortium, Agency for Science, Technology and Research (A*STAR), Biopolis, Singapore 138667. [6] Department of Chemistry, New York University, New York, NY 10003, USA.

Summary

Malaria remains a major public health issue in developing countries. With increasing drug resistant parasite strains and no available vaccine, continuing the search for novel antimalarial compounds remains a high priority for malaria research. Antimalarial drug assays first utilized radioactive compounds and current screens employ DNA fluorescent dyes to measure parasite propagation. Additional tests are needed to examine parasite morphology and host red blood cell integrity. In this study, we screened a library of RNA probes and found 3 to fluoresce live *P. falciparum*. Using a high-content imaging platform, we can efficiently examine the effects of unknown compounds on individual parasites. Overall, we established a robust antimalarial screening method using high-content live cell imaging with RNA probes.

Abstract

Background: Malaria, a major public health issue in developing nations, is responsible for more than one million deaths a year. The most lethal species, *Plasmodium falciparum*, causes up to 90% of fatalities. Drug resistant strains to common therapies have emerged worldwide and recent artemisinin-based combination therapy failures hasten the need for new antimalarial drugs. Discovering novel compounds to be used as antimalarials is expedited by the use of a high-throughput screen (HTS) to detect parasite growth and proliferation. Fluorescent dyes that bind to DNA have replaced expensive traditional radioisotope incorporation for HTS growth assays, but do not give additional information regarding the parasite stage affected by the drug and a better indication of the drug's mode of action. Live cell imaging with RNA dyes, which correlates with cell growth and proliferation, has been limited by the availability of successful commercial dyes.

Results: After screening a library of newly synthesized stryrl dyes, we discovered three RNA binding dyes that provide morphological details of live parasites. Utilizing an inverted confocal imaging platform, live cell imaging of parasites increases parasite detection, improves the spatial and temporal resolution of the parasite under drug treatments, and can resolve morphological changes in individual cells.

Conclusion: This simple one-step technique is suitable for automation in a microplate format for novel antimalarial compound HTS. We have developed a new *P. falciparum* RNA high-content imaging growth inhibition assay that is robust with time and energy efficiency.

Background

Malaria continues to be a major public health issue in many parts of the developing world [1]. Each year 300 to 500 million new clinical cases are officially reported. In the mid-fifties, the World Health Organization (WHO) launched a worldwide malaria eradication campaign using effective and inexpensive therapeutics and insecticides in designated malaria-infected areas. The program resulted in the elimination of endemic malaria in developed countries and a significant reduction of cases in developing parts of the world. The emergence of chloroquine-resistant parasites and DDT-resistant mosquito vectors has led to a reappearance and spread of malaria in most of the developing world. With the absence of an efficient vaccine, worldwide resistance to all commonly used antimalarial drugs (quinine, aminoquinolines and antifolate derivates), and the concern of emerging resistance to our last defense against this disease (artemisinin-based combination therapies) [2], there is a dire need for new antimalarial strategies.

One approach to the discovery of new therapeutic agents involves the identification of inhibitory small molecules through whole parasite-based screens. In such assays, large collections of small molecule libraries can be tested against parasite growth in culture. For decades, accurate and reliable quantitative assessment of the drug effects on parasite growth has been achieved by blood smear microscopic examinations and *in vitro* measurement of parasitic uptake of

radioactive substrates [3-5]. However, these methods are relatively expensive, require multi-step procedures and are time-consuming. They become impractical as technology advances and the volume of small molecule libraries increases. Over the past few years, a number of new techniques have been developed to improve the cost and compatibility for today's automated high-throughput screening (HTS) facilities. These techniques include colorimetric assay [6]; such as, fluorescence-based assays that measure parasite nucleic acids using fluorescent dyes (e.g. Hoechst [7, 8], PicoGreen® [9], SYBR Green I [10, 11], YOYO-1 [12] and DAPI [13]), or stable expression of chimeric fluorescent protein [14]. All these techniques have been used and adapted in automated or semi-automated HTS analyses using microplate readers or flow cytometry, and have proven to be reliable and cost-effective [12, 13, 15, 16]. However, while these techniques allow the quantification of parasite growth in human erythrocyte cultures, they detect an average response from the whole population and are inapt to efficiently detect the drug effect at the morphological level. Morphological analysis can give additional information regarding the parasite stage affected by the drug and eventually an indication of the drug's mode of action. For such analyses, investigators are still compelled to examine Giemsa-stained infected blood smears under brightfield microscopy [15-17]. While today this methodology can be completed using a semi-automatic image analysis

system [18], it requires a laborious multi-step preparation of fixed, stained blood smears and an experienced technician.

Recently, the development of high-throughput cellular imaging has emerged as a crucial tool to allow the integration of biologically complex effects into drug discovery. Such techniques can overcome the limitations of cellular-based high-throughput screening that measure a “simple” survival count by detecting morphological changes of individual cells in a microplate well. Cellular imaging technologies have been used in all stages of drug discovery including target discovery or mode-of-action studies [19]. When confocal microscopes and image-analysis tools are combined, cellular imaging platforms can facilitate the detection of cytotoxicity or cellular phenotypic changes in a cell population, and lead to the discovery of new drug targets [20]. Membrane permeable fluorescent molecules have been one of the most viable tools for live cell imaging technology [13]. Several permeable dyes, such as nucleic acid dyes (e.g. Hoechst), are commercially available and have been used in malaria high-throughput screening. While DNA measurement has been shown to be efficient in detecting parasite growth in the high-throughput screening format, the use of such dyes for parasite structural analysis is limited. Specific RNA quantification, which also correlates with cell growth and proliferation, could significantly increase parasite detection and improve the spatial and temporal resolution of the malaria parasite under drug treatment. However, cell-permeable RNA specific dyes are not

readily available and are limited in their successful use, even though the exploitation of both permeate DNA and RNA dyes would be particularly valuable for high-throughput screening and high-content imaging in *Plasmodium*.

To identify malaria parasite specific RNA probes that could improve live cell microscopy imaging, we have screened a combinatorial library of 125 fluorescence styryl molecules [21, 22] with infected red blood cells in a microplate format. Microplates were analyzed using the BD Pathway HT, an automated confocal imaging workstation. This inverted confocal microscope with temperature and carbon dioxide regulation chamber allows suspension cultures to be viewed without a fixation step. Fourteen dyes were found to have fluorescence intensity comparable or superior to DAPI staining. Out of the fourteen dyes, three dyes displayed a higher affinity to RNA. They exhibit a specific RNA staining pattern representative of the different morphological stages of the malaria parasite erythrocytic cycle. Using the identified RNA dyes and a high-throughput confocal imaging system, we have developed a sensitive and simple one-step fluorescence-based assay for use in *Plasmodium* high-content imaging. Inhibitory concentrations of known antimalarials using our image-based assay were similar to current nucleic acid dye and spectrophotometer assays. This new technology is expected to enhance the malaria drug discovery program and could eventually be adapted as an automated solution to screen parasitemia in infected patients.

Results

Screening of styryl dye library and identification of three RNA dyes.

To identify new parasite RNA dyes, we screened a library of styryl compounds. Styryl dyes were synthesized and purified using a condensation reaction between methylated-pyridine derivatives and aldehydes, as previously described [21, 22]. Cultures were diluted in microplates and incubated with the styryl compounds for 30 minutes at 37°C in the dark. Live cell imaging was performed using the Pathway HT microscope with temperature regulated at 37°C. This automated spinning disk confocal microscope with a motionless stage allows a monolayer of erythrocytes to be observed without fixation and minimal sample disruption. The Pathway HT was programmed to auto-focus to acquire transmitted light and fluorescent images in a special optics 96-well microplate format. Our Pathway HT is customized with high efficiency Semrock bandpass filters, and samples were imaged with all fluorescence combinations of Semrock DAPI, GFP, FRET, CFP, YFP, and Texas Red BrightLine filter sets to measure fluorescence intensity of the initial 125 RNA probes screened (Figure a1.1). From the initial 125 stryrl dyes, fourteen RNA-probes were found to have fluorescent intensities similar or superior to 4'-6-diamidino-2-phenylindole (DAPI) and SYBR Green DNA staining.

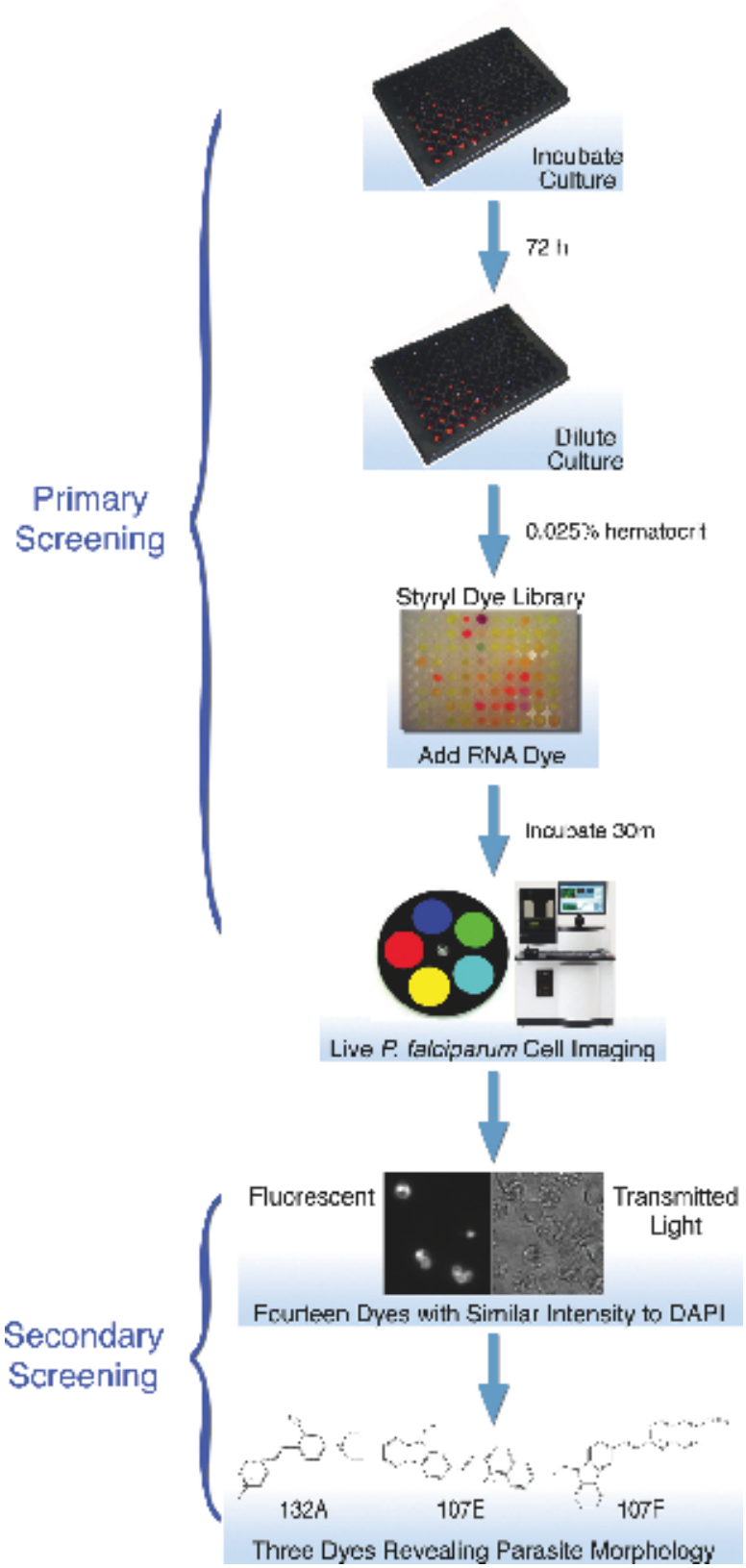


Figure a1.1. *Design of the RNA imaging P. falciparum growth assay.* Cultures were incubated in a 96-well microplate with antimalarial compounds in complete medium for 72 hrs under a standard gas environment at 37°C. After the drug incubation period, cultures were diluted to a 0.025% hematrocrit to form a monolayer of erythrocytes in a 96-well microplate. RNA dyes were added and incubated at 37°C for 30 min to stain parasites in infected erythrocytes. An automated fluorescent microscope, the Pathway HT, was used to detect and obtain images of fluorescent stained parasites, as described in Materials and Methods. Transmitted light images were also obtained to verify erythrocyte infection. Fourteen dyes were found to have similar intensity to DAPI and SYBR Green I in the primary screen. 132A, 107E, and 107F dyes were found to have a higher affinity to RNA, displaying a diffused staining throughout the parasite, in the secondary screening

Out of the 14 dyes selected, three exhibited a higher affinity to RNA and displayed characteristic nucleolus targeting with a diffused staining pattern throughout the cytoplasm in all intraerythrocytic stages observed. Fluorescence properties of RNA-probes 132A and 107E were analyzed at the reported excitation 500 and emission 610 [21]. We found the YFP 500/24 excitation filter and Texas Red 624/40 emission filter gave the best fluorescent signal with dye 132A; while the long Stokes shift CFP 438/24 excitation and YFP 542/27 emission filter combination worked best for 107E and 107F (Figure a1.1). To determine the optimum working dyes concentration 10μM, 7.5μM, 5μM, and 2.5μM were tested in infected erythrocytes. Fluorescent images taken with a 5μM concentration provided an optimum signal to noise ratio. The three dyes demonstrated a signal to noise ratio of 14:1 for 107E, 11:1 for 107F, and 16:1 for 132A. The signal to noise ratio of 132A supports previous findings of the coupound's fluorescent emission fold changes to RNA [21]. Little photobleaching and no visible change in cell morphology due to toxicity were detected during the analyses (Figure a1.2A). The selected RNA dyes were sensitive enough to identify small-sized parasites, merozoites, and provided detailed images to differentiate between asexual, trophozoite, and early sexual stages of the malaria parasite (Figure a1.2B).

To determine if the RNA probes specifically target RNA, an RNA versus DNA specificity solution assay was performed. 132A, 107E, and 107F at a 10μM

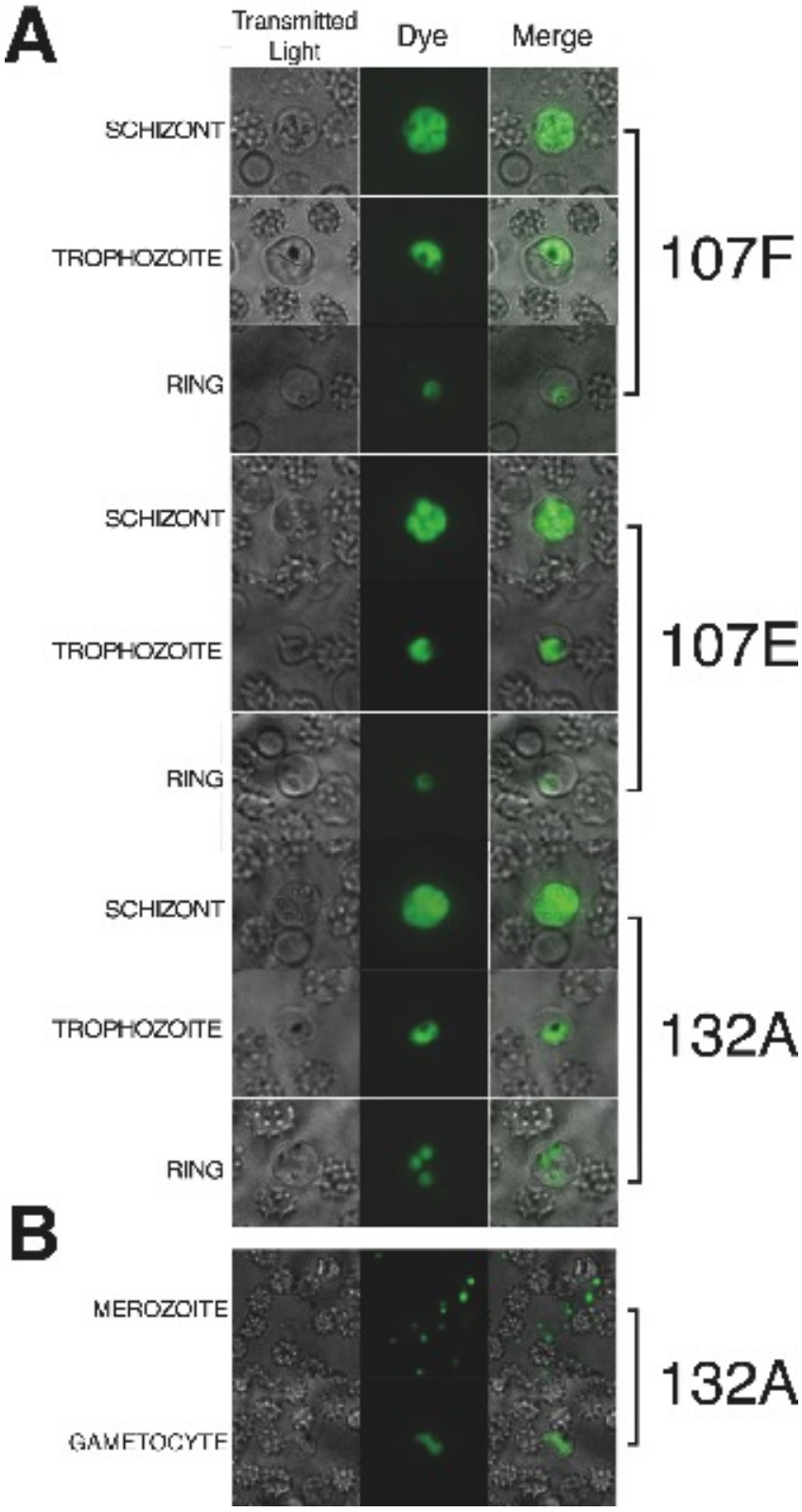


Figure a1.2. *Morphological analyses of malaria parasite with RNA probes.* A) RNA dye staining of *P. falciparum* during schizont, trophozoite, and ring stages of erythrocyte life cycle were colored in green. RNA dyes were obtained using excitation 438/24 and emission 542/27 for 107F and 107E, and a red filter excitation 500/24 and emission 624/40 for 132A. B) RNA dyes are sensitive enough to detect merozoites and gametocytes in culture.

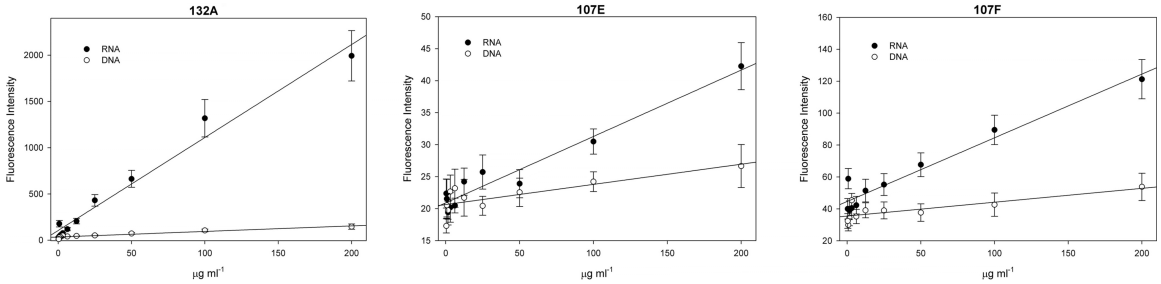


Figure a1.3. *RNA versus DNA specificity solution assay.* Fluorescence intensity of two-fold serial dilutions of *P. falciparum* RNA and DNA, with a 10μM concentration of RNA probes. Dyes display a higher fluorescence intensity with RNA than DNA.

concentration were added to two-fold serial dilutions of *P. falciparum* RNA and DNA. Fluorescence intensity as a function of RNA and DNA concentrations, displayed a higher response when bound to RNA when compared to DNA (Figure a1.3). These results correspond to better nucleoli targeting in the live cell [21]. Furthermore, 132A exhibited a higher response in fluorescence intensity when bound to RNA in a sample containing DNA (Figure a1.4).

To compare selectivity and sensitivity of our selected dyes with classical DNA staining, cell cultures were counterstained with the nucleic acid stain DAPI and Hoechst. Dyes 107E, 107F, and 132A efficiently permeated infected erythrocytes without the aid of lysis buffers or detergents. 132A displayed greater permeable properties and sensitivity in detecting parasites than DAPI (Figure a1.5A), which is used in current HT antimalarial screens. High-content imaging can benefit by using both DNA and RNA dyes to assist in parasite stage analyses.

Synchronized cell cultures stained with 107E, 107F, and 132A RNA dyes, and

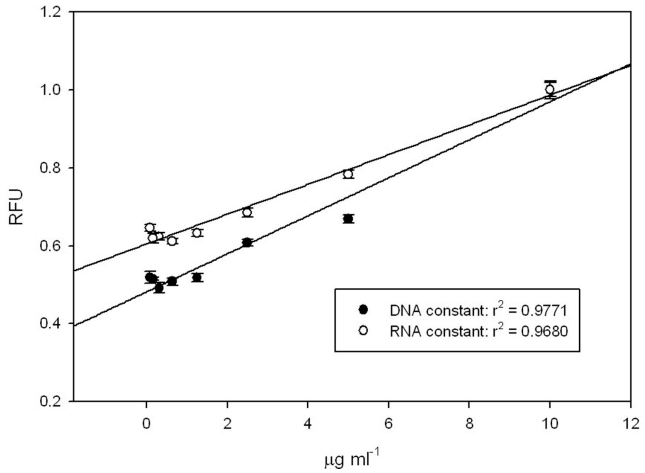
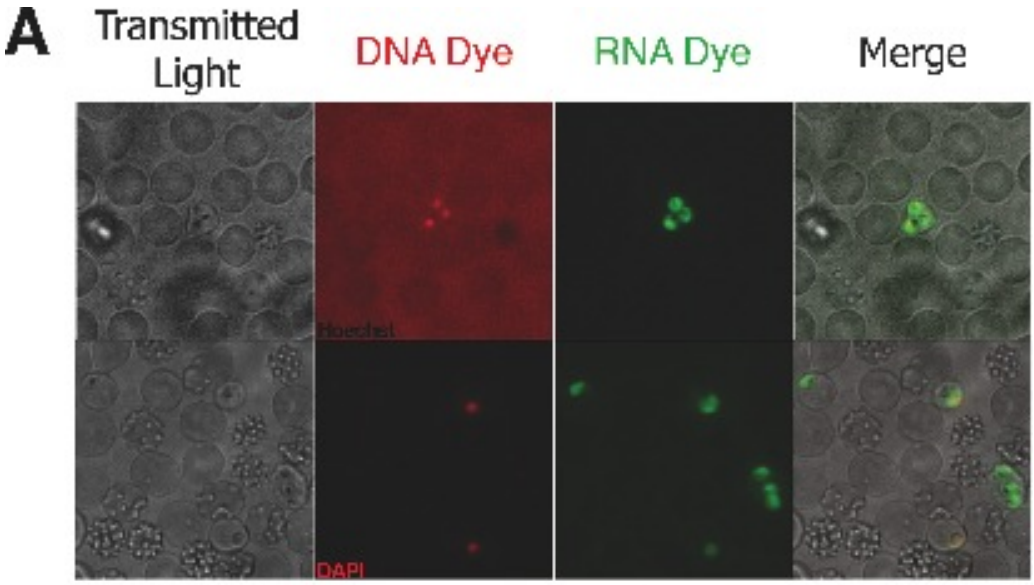
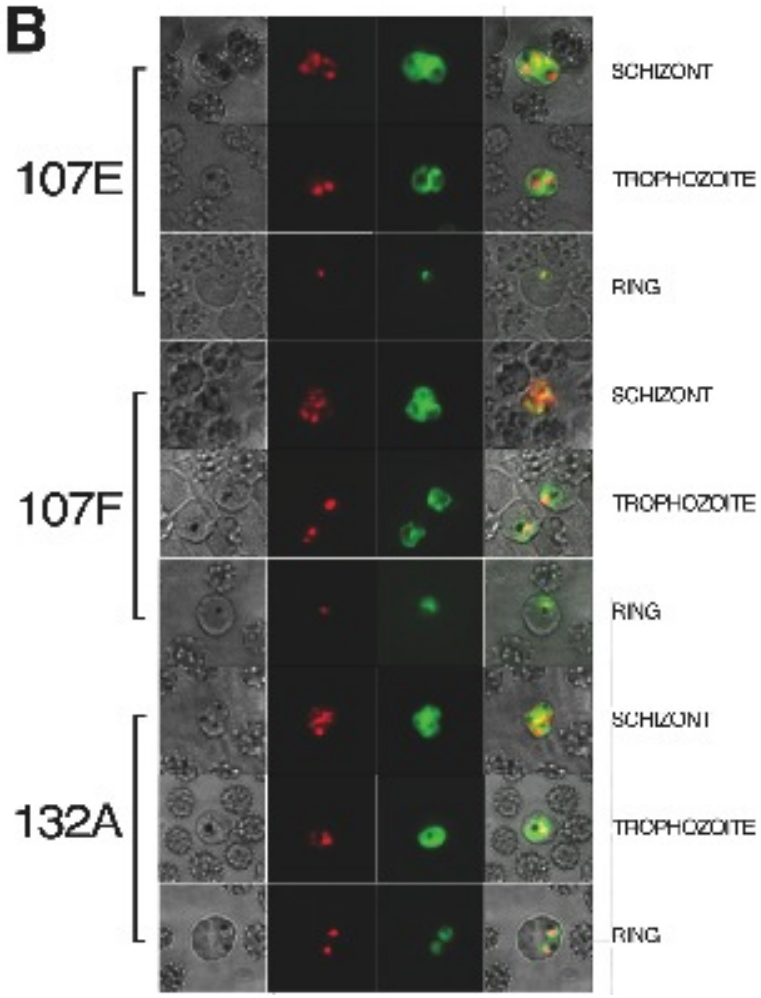


Figure a1.4. *DNA: RNA binding assay.* Graph represents 132A staining of RNA and DNA in a mixed population, with one nucleic acid kept at a steady state and the other serially diluted, where the maximum fluorescence intensity was set to one. Steady state RNA was observed to have a higher fluorescence intensity than serially diluted RNA, indicating 132A has a higher fluorescent signal when bound to RNA.

the DNA-specific dye DAPI, allowed the detection of all parasite asexual stages (Figure a1.5B). Co-localization was observed in the nucleus, but RNA dyes targeted the nucleolus and diffused throughout the parasite allowing the visualization of parasite morphology at every erythrocytic stage of the malaria parasite. These data demonstrate the capability of RNA dyes to permeate parasite membranes more efficiently than DAPI staining and significantly increases the quality of the parasite phenotypic analysis.

Figure a1.5. *Counter-staining analyses using fluorometric DNA and RNA dyes.*
A) Counter-staining with DNA dyes demonstrates the sensitivity, permeability and the adequate signal to noise ratio of RNA dyes. Hoechst staining has a significant amount of background compared to 132A. 132A detected parasites that were overlooked by DAPI staining. B) Counter staining with nucleic acid dye DAPI, in red, and RNA dyes 107F, 107E, and 132A, in green, at schizont, trophozoite, and ring stages of erythrocytic cycle. Images obtained by the Pathway HT were grayscale; red and green colors were artificially added to emphasize localization.


A
Transmitted Light
DNA Dye
RNA Dye
Merge
Hoechst
DAPI


B
107E
107F
132A
SCHIZONT
TROPHOZOITE
RING
SCHIZONT
TROPHOZOITE
RING
SCHIZONT
TROPHOZOITE
RING

Relationship between parasitemia and fluorescence intensity and validation of RNA dyes for their use in a high-throughput malaria growth assay.

To determine the relationship between infected erythrocytes and relative fluorescence intensity when stained with RNA dye, we compared the fluorescence intensity to the percentage of parasite-infected red blood cells. Serial dilutions of infected erythrocytes with 5μM of RNA dye were prepared in triplicate and analyzed using the Pathway HT. As shown in Figure a1.5, an almost linear relationship between the percentages of infected red blood cells and the fluorescence signal was observed with an R^2 value of 0.868, 0.959 and 0.916 for 132A, 107E and 107F, respectively. To further validate our assay, we calculated

the Z-factor values for each of the selected dyes. The Z-factor is a statistical characteristic for HTS assays used to compare and evaluate the quality of high-throughput assays [25]. A Z-factor value equal to one is an ideal assay, while values between 1 and 0.5 are considered excellent. The Z-factor was

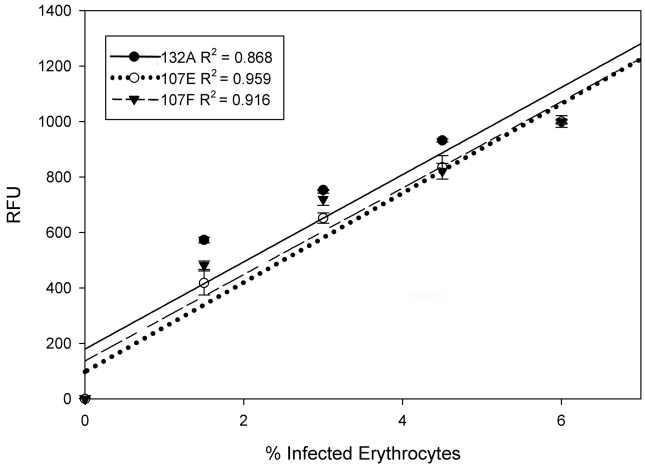


Figure a1.6. *Evaluation of the percentage of P. falciparum-infected erythrocytes using fluorescence intensity obtained by the Pathway HT.*
Serial dilution of synchronized infected cultures (6% parasitemia) of *P. falciparum* 3D7 strain were conducted as

determined for each of our three selected dyes in a 96-well microplate and exhibited 0.689, 0.706, and 0.728 Z'- values for 107E, 107F, and 132A, respectively. These data demonstrate the ability to determine parasitemia based on relative fluorescence units (RFU), and validates the use of RNA probes for high-throughput live cell imaging.

To ensure that the activities of antimalarials measured by RNA dye and image analysis were consistent with established parasite growth inhibition

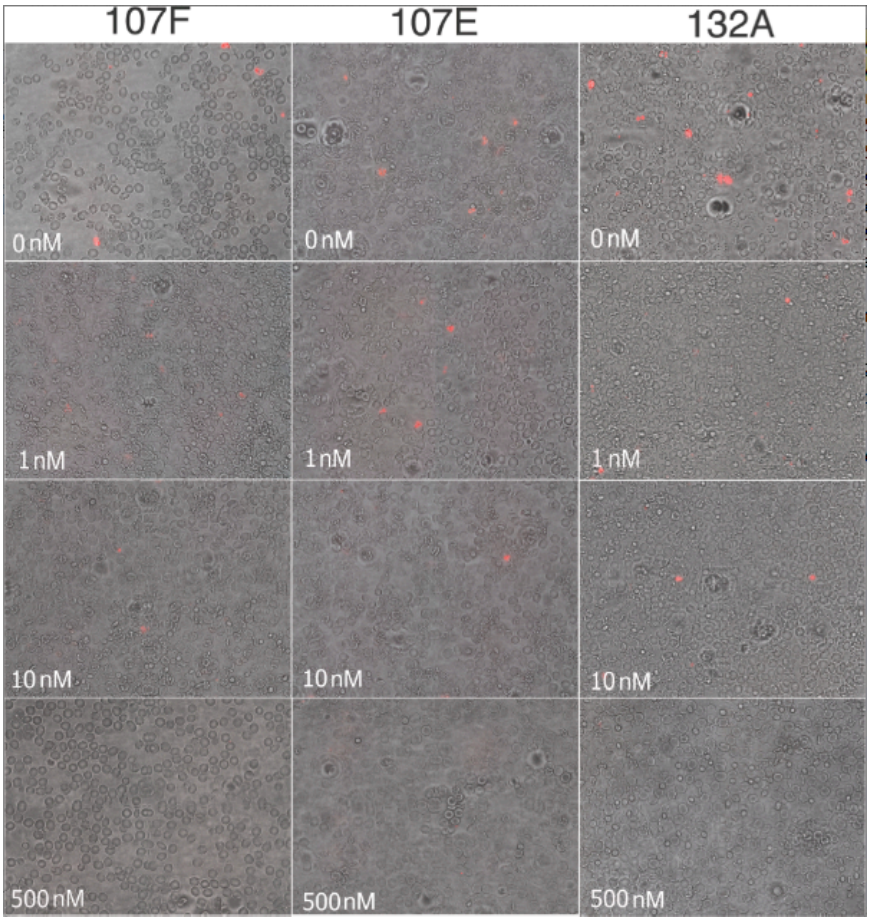


Figure a1.7. *Merged montage images.* A tiling of two by two images of transmitted light and fluorescent parasites, were taken with the Pathway HT. Decreased parasitemia with increasing concentration of chloroquine was observed with the 3D7 strain.

assays, we measured the IC_{50} values of known antimalarials (chloroquine and artemisinin), and of chloroquine sensitive and resistant parasite strains using the Pathway HT. The Pathway allows the detection of fluorescent parasites based on set parameters and numbers them according to size and intensity. Furthermore, the Pathway HT montage function compiles multiple frames into one tiled image that provides a large area of observation (Figure a1.7). These data can then be used to determine percentage of parasite growth compared to drug dose. IC_{50} values obtained using RNA dye high-throughput confocal microscopy assay were comparable to IC_{50} values measured by the widely used SYBR Green assay (Figure a1.8). Data were further validated by microscopic analyses of Giemsa-stained blood smears (data not shown). While the IC_{50} values obtained with the three selected dyes are closely related to each other, they did not produce identical results. This can be explained by the differences observed in signal to noise ratios and the small disparate Z'-values obtained for each dye. To further validate the sensitivity of our assay, we calculated the minimum parasitemia detected by the Pathway HT above background. The Pathway HT detected 1 single parasite out of 12,000 RBCs, corresponding to a 0.0083% parasitemia, confirming the sensitivity of the high-content imaging assay.

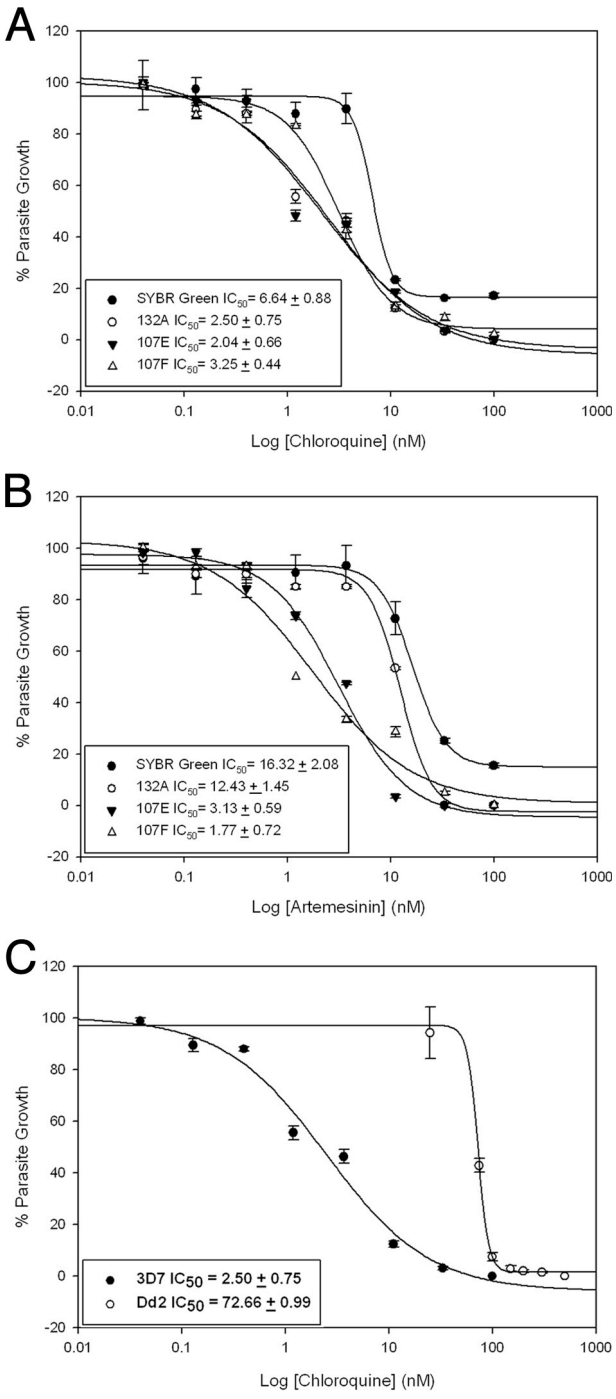
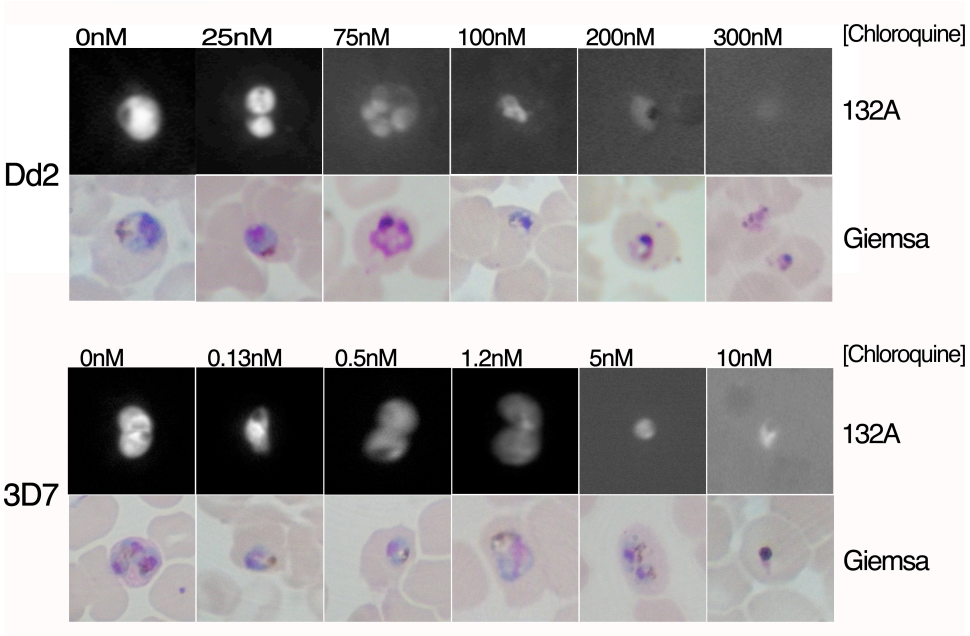


Figure a1.8. *Validation of parasite growth inhibition measurement by known antimalarials.* IC_{50} values for 3D7 parasite cultures treated with chloroquine (A) and artemesinin (B) were calculated using images obtained by the Pathway HT and compared to the classical SYBR Green assay. C) IC_{50} values determined with our high-content assay using fluorometric RNA dye 132A for parasite chloroquine sensitive 3D7 and chloroquine resistant Dd2 strains.

To confirm that the developed high-content live cell imaging assay can combine IC_{50} values with morphological examination of the parasite under drug treatment, images obtained with parasite Dd2 and 3D7 strains were compared to Giemsa-stained blood smears (Figure a1.9). Morphological phenotypes of parasites observed between the two techniques were highly similar. Using the fluorometric RNA staining dye 132A, parasites display a loss of nucleolus staining, and a decrease in signal intensity with an increase in drug concentration. All together, these data demonstrate that RNA dyes are sufficiently sensitive and reproducible to identify the inhibition effects of antimalarial compounds in a HTS screening and high-content imaging format.

Figure a1.9. *Morphological analysis of 3D7 and Dd2 parasites using the Pathway HT.* Parasite synchronized cultures were incubated for 72 hours with various concentrations of chloroquine, live-stained with the fluorometric RNA dye 132A and fixed-stained with Giemsa. Detailed images of infected erythrocytes show a decrease in fluorescence intensity, loss of nucleolus staining, and change of parasite morphology with increasing drug concentration.



Discussion

In the past few years, cellular imaging has emerged as a critical tool to efficiently integrate the cellular complexity into drug discovery. Continuous technical improvements in fluorescence, confocal microscopy platforms, and image analysis have enhanced the detection and improved the resolution by which an individual molecule can act in a cell. High-content cellular imaging screenings recently lead to the discovery of small molecule inhibitors of novel cellular targets in human cancer cell lines [20]. Technological progress observed in cellular imaging toward human diseases needs to be adapted to the malaria parasite to enhance its drug discovery program. It is essential to integrate a spatial and temporal phenotypic analysis of the parasite under drug pressure in order to identify small molecule inhibitors that block major specific parasitic events (e.g. parasite invasion, division or egress). These types of phenotypic studies will improve the biological relevance of malaria HTS.

To address this need, we screened a combinatorial library of fluorescent styryl molecules, which contain RNA probing compounds that have been successfully identified in human cell lines [21, 22]. Using a confocal microscope platform, the Pathway HT, we identified three RNA dyes with strong imaging properties relative to commercially available DNA dyes in live parasites. The selected dyes, 107E, 107F, and 132A are membrane permeable, diffuse throughout the cytoplasm to display parasite morphology, target nucleoli, and can

be used in conjunction with DAPI staining. Out of the three dyes 132A exhibited the best signal to noise ratio and Z'-value, therefore this dye appears to have the best properties and will be used in future assays. Such characteristics allow the observation of RNA quantity and distribution in relation to the organization of the DNA within the parasite. Simultaneous DNA and RNA staining significantly increases the cell structural organization and facilitates the morphological analysis of *Plasmodium* in culture. While GFP fluorescence proteins have been exploited as tools to facilitate the dynamic behaviors of proteins and parasite phenotype examination in a real-time manner, these techniques require the use of modified cell lines, which are not the optimal choice for a large, small molecule screen against different *Plasmodium* drug resistant and sensitive isolates. In addition, optimum DNA staining, which has been successful in malaria HTS, requires the addition of lysis buffer and/or a fixation step that prevents the observation of live parasites and increases the screen's complexity. Although it is possible using DAPI staining to see numerous nuclei in close enough proximity to identify schizont stages, the differentiation between other stages of the erythrocyte life cycle is more challenging. The use of RNA dye together with DNA staining provides a dynamic morphology of all parasite erythrocytic cycle stages. It gives accurate information on parasite RNA content and its transcriptional activity, as well as information on cytotoxicity effects on erythrocytes and parasites.

Materials and Methods

P. falciparum Culture

3D7 and Dd2 *P. falciparum* malaria parasites (MRA-102, 156, MR4, ATCC® Manassas, Virginia*)* were cultured in human type O+ erythrocytes in complete medium (RPMI 1640, 10 mg/mL Gentamicin (Gibco), 1.36 g/L Hypoxantine (Acros), 1M HEPES (Sigma), 7.5% Sodium Bicarbonate (Gibco), 20% Glucose (MP Biomedical), 1M NaOH (Sigma), 20% Albumax (Gibco), 5% human serum) as previously described [23]. Cultures were maintained in 25-cm^2 flasks (Corning) at a volume of 10 ml and were gassed for 30 s with an environment of 3% CO_2, 1% O_2, and 96% N_2, then incubated at 37°C. Synchronization of culture was achieved through sorbitol lysis of mature stage using 5% sorbitol (Fisher) fine-tuned by another lysis 8 hours later [24].

Dye Preparation

Microwell plates already containing a library of RNA-probes were diluted in DMSO to 10mM. A screening of 125 different RNA-probes were performed by diluting infected erythrocytes to 0.025% hematrocrit with a 6% parasitemia into optical bottom 96-well assay plates (Costar #3614, Corning, NY) containing 240μl of complete medium. RNA-probes were added to a final concentration of 1μM, 5μM, 7.5μM, and 10μM with a total DMSO level of 0.5%. Plates were incubated in the dark for 30 minutes at 37°C. During microscopic analysis, the

Atto Pathway HT (BD Biosciences Bioimaging, Rockville, MD) temperature was regulated at 37°C. Each well was fluorescently imaged using fluorescence combinations of Semrock (Rochester, NY) DAPI, CFP, GFP, YFP and Texas Red BrightLine filter sets. We found that 5μM was the optimal working concentration.

RNA versus DNA specificity solution assay

DNA was extracted from *P. falciparum* using an adapted phenol/chloroform extraction. Infected erythrocytes were harvested and brought to a 50% hematocrit in PBS. Samples were incubated with cell lysis solution (Promega #A7933, Madison, WI) for 10 minutes at room temperature, followed by centrifugation. After removing the supernatant, pellet was incubated at 55ºC for an hour in lysis buffer containing 4M guanidine HCl (Promega), 10% SDS (Promega), and 20mg/ml Proteinase K (New England Biolabs), and left overnight at 4ºC. DNA was then extracted using Phenol/chloroform/isoamyl alcohol (Sigma) following the standard procedures.

RNA was extracted from infected erythrocytes with Trizol LS (Life Technology), as previously described in Le Roch *et. al.* 2003 [23].

To scan fluorescence emission, the SpectraMAX GeminiEM (MDS Analytical Technologies, Toronto, Canada) and the SoftMax Pro program were used. Using black, clear bottom, 96-well assay plates (Costar #3904), 10 two-fold serial dilutions were performed with extracted DNA and RNA starting at 200μg/ml. Steady state concentrations of RNA or DNA at 10μg/ml, were mixed

with two-fold serial dilutions of the opposing nucleic acid. 132A, 107E, and 107F RNA probes were added to wells at a 10μM concentration. Endpoint readings with an excitation of 500nm, cutoff at 515nm and emission of 605nm for 132A, and excitation of 435nm, cutoff at 605nm and emission of 610nm for 107E and 107F. DNA and RNA concentration as a function of fluorescence intensity was plotted with SigmaPlot.

Counter Stain

Synchronized infected erythrocytes were stained with nuclear dye DAPI (Molecular Probes #D21490, Eugene, OR) or Hoechst 34580 (Molecular Probes #H21486, Eugene, OR) diluted in H_2O and added to reach a final concentration of 20ng/μL or 5μg/ml respectively. RNA specific dyes were then added to reach a concentration of 5μM and plates were incubated with both dyes in the dark at 37°C for 30 minutes. Images were taken using the Atto Pathway HT with the temperature regulated at 37°C. A macro, an instruction set for the Pathway, was designed to take images of the same frame using the DAPI filter wheel, then the corresponding filter for RNA dye, then transmitted light, auto focusing for each image, moving to another frame in the same well, and to repeat in subsequent wells.

Relative Fluorescence

Synchronized infected erythrocytes were diluted with uninfected erythrocytes in complete media (i.e. 0%, 1.5%, 3%, 4.5%, 6%). Diluted probes were added to reach a final concentration of 5 μM, incubated in the dark for 30 min at 37°C and the BD Pathway HT measured fluorescence intensity. Data analyses were performed using Microsoft Excel 2004 for Mac. Data points were graphed by 1) calculation of the mean of triplicates per sample condition, 2) subtraction of the fluorescence background, and 3) conversion to relative fluorescence units.

Quantitative assay evaluation

Three 96-well microplates with parasite-infected erythrocytes were used as a positive control, and three microplates with uninfected erythrocytes were used as a negative control. All plates were diluted to a 0.025% hematocrit, and 6% parasitemia. 107E, 107F, and 132A dyes were added to microplates and the Pathway HT determined fluorescence intensity. The formula $z' = 1 - (3SD_+ + 3SD_-)/|Ave_+ - Ave_-|$ was used to caculate the Z' value. Where SD_+ represents the positive control standard deviation, SD_- the negative control standard deviation, Ave_+ the mean value of the positive control, and Ave_- the mean value of the negative control.

IC_{50}

3D7 parasite cultures were incubated with 0.1nM, 0.4nM, 1.2nM, 3.7nM, 11.1nM, 33.3nM, 100nM, and 500nM concentrations of chloroquine or artemisinin (Sigma #C6628-25G, St. Louis, MO) for 72 hours. Dd2 parasite cultures were incubated with 5nM, 25nM, 75nM, 100nM, 150nM, 200nM, 300nM, and 500nM concentrations of chloroquine. RNA dyes were added to reach 5μM final concentration and incubated in the dark for 30 minutes at 37°C. Microscopic analysis was performed at a regulated temperature of 37°C. Montage images of 4 by 4 frames were acquired with transmitted light and the corresponding filter for RNA dyes. Images were merged to verify infected erythrocytes by having the fluorescent target overlap with visible hemazoin in transmitted light images. The Pathway HT calculated the parasitemia by using the Region of Interest (ROI) function to count fluorescent parasites. Parasitemia was determined by manual counts of fluorescent images and Giemsa-stained blood smears, and compared to Pathway HT counts. Samples of identical cultures were taken to perform a SYBR Green assay as previously described [16]. Data were analyzed using Microsoft Excel for Mac, and graphs were plotted using SigmaPlot 10 (Systat).

List of Abbreviations

HTS: high-throughput screen; **RFU:** relative fluorescence units

Authors' Contributions

YTC and KLR conceived the study. SC carried out screening and microscopic analysis. JP and SC prepared samples and collected spectrometer data. DC participated in the assay design. QL synthesized original library of probes, and GGK synthesized selected compounds. SC, JP, YTC and KLR contributed analyzing experimental studies. KLR drafted the manuscript. All authors read, edited and approved the final manuscript.

Acknowledgements

We thank Duk-Won Doug Chung and Nadia Ponts for critically revising the manuscript and MR4 for providing us with malaria parasites (contributed Thomas Wellems, Daniel Carucci, and Alister Craig). This work was supported by funding from the University of California Riverside Bridge to Doctorate Fellowship (SC) and the National University of Singapore (NUS) (Young Investigator Award: R-143-000-353-123) (YTC).

Conclusion

Our IC_{50} results clearly demonstrate the feasibility of RNA directed fluorescence-based assays for high-throughput screening of antimalarials. Furthermore, these type of HTS and high-content imaging assay analyses can be defined and programmed automatically on the Pathway HT, avoiding the need of an experienced or biased microscopist to analyze the images. This new *Plasmodium* RNA dye growth inhibition and high-content imaging assay is a robust, cost effective, fast, and simple one-step technique suitable for automation in 96- to 384-well plates. It will be an ideal technique for a secondary screen validation, help to determine the stage specificity of drugs and potentially indicate a drug's mode of action. When combined with a similar mammalian screen, we should be able to identify the parasite specific compounds. It is also important to highlight that the sensitivity of such technique should allow the method to be adapted for the identification of infected blood samples of patients.

References

1. Guerra CA, Gikandi PW, Tatem AJ, Noor AM, Smith DL, Hay SI, Snow RW: **The limits and intensity of Plasmodium falciparum transmission: implications for malaria control and elimination worldwide**. *PLoS Med* 2008, **5**(2):e38.

2. Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, Fukuda MM: **Evidence of artemisinin-resistant malaria in western Cambodia**. *N Engl J Med* 2008, **359**(24):2619-2620.

3. Desjardins RE, Canfield CJ, Haynes JD, Chulay JD: **Quantitative assessment of antimalarial activity in vitro by a semiautomated microdilution technique**. *Antimicrob Agents Chemother* 1979, **16**(6):710-718.

4. Elabbadi N, Ancelin ML, Vial HJ: **Use of radioactive ethanolamine incorporation into phospholipids to assess in vitro antimalarial activity by the semiautomated microdilution technique**. *Antimicrob Agents Chemother* 1992, **36**(1):50-55.

5. Iber PK, Pavanand K, Wilks NE, Colwell EJ: **Evaluation of in vitro drug sensitivity of human Plasmodium falciparum by incorporation of radioactive isoleucine**. *J Med Assoc Thai* 1975, **58**(11):559-566.

6. Druilhe P, Moreno A, Blanc C, Brasseur PH, Jacquier P: **A colorimetric in vitro drug sensitivity assay for Plasmodium falciparum based on a highly sensitive double-site lactate dehydrogenase antigen-capture enzyme-linked immunosorbent assay**. *Am J Trop Med Hyg* 2001, **64**(5-6):233-241.

7. Davis WC, Wyatt CR, Hamilton MJ, Goff WL: **A rapid, reliable method of evaluating growth and viability of intraerythrocytic protozoan hemoparasites using fluorescence flow cytometry**. *Mem Inst Oswaldo Cruz* 1992, **87 Suppl 3**:235-239.

8. Smeijsters LJ, Zijlstra NM, Franssen FF, Overdulve JP: **Simple, fast, and accurate fluorometric method to determine drug susceptibility of Plasmodium falciparum in 24-well suspension cultures**. *Antimicrob Agents Chemother* 1996, **40**(4):835-838.

9. Corbett Y, Herrera L, Gonzalez J, Cubilla L, Capson TL, Coley PD, Kursar TA, Romero LI, Ortega-Barria E: **A novel DNA-based microfluorimetric method to evaluate antimalarial drug activity**. *Am J Trop Med Hyg* 2004, **70**(2):119-124.

10. Bennett TN, Paguio M, Gligorijevic B, Seudieu C, Kosar AD, Davidson E, Roepe PD: **Novel, rapid, and inexpensive cell-based quantification of antimalarial drug efficacy**. *Antimicrob Agents Chemother* 2004, **48**(5):1807-1810.

11. Smilkstein M, Sriwilaijaroen N, Kelly JX, Wilairat P, Riscoe M: **Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening**. *Antimicrob Agents Chemother* 2004, **48**(5):1803-1806.

12. Weisman JL, Liou AP, Shelat AA, Cohen FE, Guy RK, DeRisi JL: **Searching for new antimalarial therapeutics amongst known drugs**. *Chem Biol Drug Des* 2006, **67**(6):409-416.

13. Baniecki ML, Wirth DF, Clardy J: **High-throughput Plasmodium falciparum growth assay for malaria drug discovery**. *Antimicrob Agents Chemother* 2007, **51**(2):716-723.

14. Kadekoppala M, Kline K, Akompong T, Haldar K: **Stable expression of a new chimeric fluorescent reporter in the human malaria parasite Plasmodium falciparum**. *Infect Immun* 2000, **68**(4):2328-2332.

15. Plouffe D, Brinker A, McNamara C, Henson K, Kato N, Kuhen K, Nagle A, Adrian F, Matzen JT, Anderson P *et al*: **In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen**. *Proc Natl Acad Sci U S A* 2008, **105**(26):9059-9064.

16. Prudhomme J, McDaniel E, Ponts N, Fenical W, Jensen P, Le Roch K: **Marine actinomycetes: A new source of compounds against the human malaria parasite, Plasmodium falciparum.** . *PLoS ONE* 2008, **June 2008**.

17. Le Roch KG, Johnson JR, Ahiboh H, Chung DW, Prudhomme J, Plouffe D, Henson K, Zhou Y, Witola W, Yates JR *et al*: **A systematic approach to understand the mechanism of action of the bisthiazolium**

compound T4 on the human malaria parasite, Plasmodium falciparum. *BMC Genomics* 2008, **9**:513.

18. Le MT, Bretschneider TR, Kuss C, Preiser PR: **A novel semi-automatic image processing approach to determine Plasmodium falciparum parasitemia in Giemsa-stained thin blood smears**. *BMC Cell Biol* 2008, **9**:15.

19. Lang P, Yeow K, Nichols A, Scheer A: **Cellular imaging in drug discovery**. *Nat Rev Drug Discov* 2006, **5**(4):343-356.

20. Tanaka M, Bateman R, Rauh D, Vaisberg E, Ramachandani S, Zhang C, Hansen KC, Burlingame AL, Trautman JK, Shokat KM *et al*: **An unbiased cell morphology-based screen for new, biologically active small molecules**. *PLoS Biol* 2005, **3**(5):e128.

21. Li Q, Chang YT: **A protocol for preparing, characterizing and using three RNA-specific, live cell imaging probes: E36, E144 and F22**. *Nat Protoc* 2006, **1**(6):2922-2932.

22. Li Q, Kim Y, Namm J, Kulkarni A, Rosania GR, Ahn YH, Chang YT: **RNA-selective, live cell imaging probes for studying nuclear structure and function**. *Chem Biol* 2006, **13**(6):615-623.

23. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ *et al*: **Discovery of Gene Function by Expression Profiling of the Malaria Parasite Life Cycle**. *Science* 2003.

24. Lambros C, Vanderberg JP: **Synchronization of Plasmodium falciparum erythrocytic stages in culture**. *J Parasitol* 1979, **65**(3):418-420.

25. Zhang JH, Chung TD, Oldenburg KR: **A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays**. *J Biomol Screen* 1999, **4**(2):67-73.

Chapter 2a

High content live cell imaging for the discovery of new antimalarial marine natural products

Serena Cervantes[1], E Paige Stout[2,3], Jacques Prudhomme[1], Sebastian Engel[2,] Matthew Bruton[1], Michael Cervantes[1], David Carter[4], Young Tae-Chang[5,6], Mark E Hay[3], William Aalbersberg[7], Julia Kubanek [2,3], Karine Le Roch[1]

[1]Department of Cell Biology and Neuroscience, University of California Riverside, Riverside, CA 92521, USA. [2]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA. [3]School of Biology, Georgia Institute of Technology, Atlanta, GA 30332, USA. [4]Institute for Integrative Genome Biology, University of California, Riverside, CA 92521, USA. [5]Laboratory of Bioimaging Probe Development, Singapore Bioimaging Consortium, Biopolis 138667, Singapore. [6]Agency for Science, Technology and Research (A*STAR), Biopolis 138667, Singapore, [7]Institute of Applied Sciences, University of the South Pacific, Suva, Fiji,

Summary

There are drug resistant strains of malaria found worldwide to all available drug therapies. Screening chemical and natural products libraries is crucial to identify novel antimalarials. Traditionally, antimalarial screens are based on growth proliferation assays after 72 hours of incubation. We employed the RNA dyes, previously identified by our group, as part of a secondary screen. This high-content live cell-imaging (HCLCI) can be used to screen chemical libraries and identify parasite specific inhibitors with limited host cytotoxic effects. In addition, we can detect morphological changes to parasites, which provide a lead to the potential mode of action.

Abstract

Background: The human malaria parasite remains a burden in developing nations. It is responsible for up to one million deaths a year, a number that could rise due to increasing multi-drug resistance to all antimalarial drugs currently available. Therefore, there is an urgent need for the discovery of new drug therapies. Recently, our laboratory developed a simple one-step fluorescence-based live cell-imaging assay to integrate the complex biology of the human malaria parasite into drug discovery. Here we used our newly developed live cell-imaging platform to discover novel marine natural products and their cellular phenotypic effects against the most lethal malaria parasite, *Plasmodium falciparum.*

Methods: A high content live cell imaging platform was used to screen marine extracts effects on malaria. Parasites were grown *in vitro* in the presence of extracts, stained with RNA sensitive dye, and imaged at timed intervals with the BD Pathway HT automated confocal microscope.

Results: Image analysis validated our new methodology at a larger scale level and revealed potential antimalarial activity of selected extracts with a minimal cytotoxic effect on host red blood cells. To further validate our assay, we investigated parasite's phenotypes when incubated with the purified bioactive natural product bromophycolide A. We show that bromophycolide A has a strong

and specific morphological effect on parasites, similar to the ones observed from the initial extracts.

Conclusion: Collectively, our results show that high-content live cell-imaging (HCLCI) can be used to screen chemical libraries and identify parasite specific inhibitors with limited host cytotoxic effects. All together we provide new leads for the discovery of novel antimalarials.

Background

Malaria remains a major public health issue in developing nations. In 2006, the World Health Organization reported approximately 250 million cases of malaria, which caused 1 million deaths a year [1]. Despite such a high number of deaths per year, malaria is a curable disease and traditional medicinal plants have been used for treatment since antiquity. Native Peruvians used the bark of the *Cinchona succirubra* (Rubiaceae) tree for centuries before quinine was isolated from it in 1820 [2]. Its semi-synthetic derived compound, chloroquine (CQ), became the prophylactic treatment for malaria in 1947 and was the most effective treatment until CQ-resistant strains appeared in 1957. In 1972, a new natural product, artemisinin, was isolated from *Artemisia annua,* a plant used in traditional Chinese medicine for over 2000 years [3]. Artemisinin-based combination therapies (ACTs) are currently our last resort in combating malaria infection. Unfortunately, the first ACT-resistant strains appeared in Cambodia in 2009 and hasten the need for new antimalarials [4]. In the long history of drug discovery against the human malaria parasite it is clear that natural products have outlived many synthetic drugs and remain a valuable resource in identifying efficient and long lasting novel antimalarials.

One successful approach in discovering new chemical and natural therapeutic agents against the malaria parasite is based on high-throughput screening (HTS) on the whole organism. Large collections of small molecule libraries can be

tested directly against parasite growth in culture [5] [6-8]. Traditionally, the [3^{H}]hypoxanthine incorporation assay was the gold standard to determine, *in vitro,* the drug susceptibility of the malaria parasite [9]. This method has now been largely replaced by less hazardous, cost and labor efficient DNA dye intercalation assays (SYBR Green I [10,11], Pico green [12], 4′, 6-diaminino-2-phenylindole (DAPI) [13]), assays using quantum dots labeling late stage infected erythrocytes [14], and assays using parasites that stably express cytoplasmic firefly luciferase [15,16]. While the various DNA dye assays are capable of quantifying parasite growth, they are limited to testing a simple survival count and do not efficiently detect the effect of drug treatment at the morphological level or provide information of a potential drug's cytotoxicity. Current screenings with parasite strains expressing green fluorescence protein (GFP) have facilitated the observation of the dynamic behaviors of parasite phenotype in a real-time manner. However, these techniques require the use of a modified cell line for all screening purposes.

Recently, we developed a semi-automated RNA fluorescence-based high-content live cell-imaging (HCLCI) assay that has multiple advantages [17]. It is a fast, simple and a one-step fluorescence-based assay that can be used with any type of *Plasmodium* laboratory and field isolate strains. It can detect a very low number of live parasites, their morphological stages and their transcriptional activities. When high-quality bioimaging microscopes and image- analysis tools

are combined, these screening platforms can facilitate the detection of cytotoxicity or cellular phenotypic changes in the parasite population and its host cell. Therefore, this assay can potentially lead to the discovery of novel drugs with novel modes of action and a hint toward the identification of their cellular targets.

In this study, we have used our new live cell confocal imaging platform to identify new parasite inhibitors from crude extracts and their specific morphological effects on *Plasmodium falciparum.* Our natural product library was first extracted from a rich source of various sessile marine organisms (e.g., red macroalgae, benthic cyanobacteria, cnidarians, and sponges). Sessile marine organisms are exposed to numerous pathogens and predators in seawater and have developed effective natural chemical defenses [18,19]. Their natural products have been shown to have biological activity against a wide range of human cancer cell lines and several infectious agents [20,21].

To screen our marine natural product library for antimalarial activity, we first employed a standard SYBR Green I based assay. Extracts that inhibited parasite growth at low concentrations were then subjected to a secondary HCLCI screen to examine parasite morphologies and the toxicity of natural product treatment on host red blood cells. Collectively, this live cell-imaging assay improves the spatial and temporal resolution of the malaria parasite screening assay, and provides

valuable information about the health of host cells under drug treatment. Potent novel antimalarials were identified and they produced phenotypic irregularities in the malaria parasite, but had minimal cytotoxic effects on host red blood cells. Further validation of our methodology was achieved using the newly purified natural compound bromophycolide A. All together our data demonstrate that this streamlined assay can be used to discover novel and efficient antimalarials.

Results and Discussion

Crude extract library preparation

To investigate novel antimalarial compounds, we prepared a library of crude extracts derived from various marine organisms including red, green, and brown macroalgae, cyanobacteria, cnidarians, and sponges collected from the Fiji Islands. Figure a2.1, a schematic flow chart, illustrates the key steps in our screen. Extracts were first separated by reversed-phased column chromatography and four fractions ranging in compound polarity were collected for each individual organism. All fractions were tested for antimalarial activity.

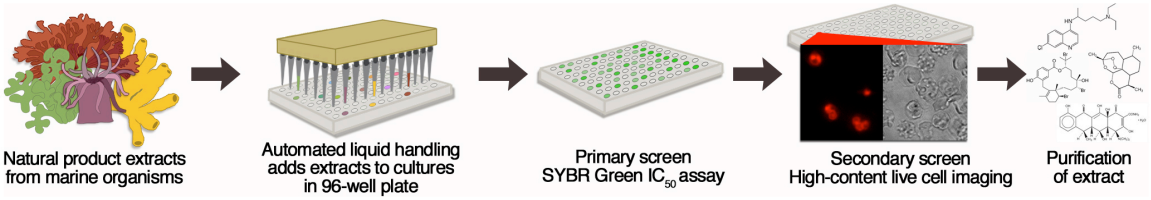


Figure a2.1. *A schematic diagram of natural products antimalarial screen.* Marine organisms were collected from the reef off the Fijian Islands. Extracts were aliquoted into 96-well microtiter plates and cultures were subsequently added. A SYBR Green I based assay was used as a primary screen of 2685 extracts. Extracts that inhibited *P. falciparum* growth were subjected to a high-content live cell imaging secondary screen

Primary bioactivity screen of natural products

Our extract library was screened for antimalarial activity using a standard high-throughput (HT) SYBR Green I based assay. We employed the Biomek automated liquid handling system to aliquot extracts into three different concentrations (100, 33, and 11 µg/ml) in 96-well plates. *P. falciparum* cultures

were added and incubated for 72 h before being subjected to a HT assay to evaluate parasite growth by relative fluorescence intensity of the nucleic acid dye SYBR Green I. To validate the quality of our HT SYBR Green I assay we calculated its Z-factor, which is a high-throughput assay quality assessment. An ideal Z-factor is 1.0, and below a 0.5 is a marginal assay. Our primary screen Z-factor value was 0.669, indicating an excellent primary assay to determine parasite growth inhibition based on relative fluorescence units. From the initial library of 2685 marine extracts examined, 27 extracts inhibited parasite growth below 11 µg/ml. With approximately 1% of extracts found to be potent inhibitors and based on our Z-factor score, our primary assay is efficient and effective for screening our library and marine organisms are a novel source for new antimalarial therapies. While this assay gives us an indication of the parasite survival, it does not provide any indication of the drug effects on the parasite morphology and potential toxicity on host red blood cells (RBCs).

Secondary high-content live cell imaging screen

To further validate antimalarial activity, extracts that inhibited parasite growth below 11µg/ml were selected and subjected to a secondary HCLCI screen. By using this technique, we could easily visualize our live parasites in host red blood cells. For this secondary screen, we used the Bioscience Pathway HT, an inverted confocal microscope with a motionless stage in a climate-controlled environment together with a fluorescent RNA selective probe (132A). This RNA

probe has previously been shown to permeate infected erythrocytes without additional lysis buffers or detergents [17]. When compared to DNA binding dyes (DAPI, Hoechst) used in current HT antimalarial screens, our selected RNA probe diffuses throughout the cytoplasm and provides detailed live cell staining of *P. falciparum.* Using this methodology we were able to evaluate the parasitemia, morphology and transcriptional activity under drug treatment. In addition to that information, the transmitted light images provide valuable information on potential toxic effects of the extracts on host RBC morphology. For example, it has been previously observed that erythrocytes crenate and transform into echinocytes when cell membranes are perturbed [24]. A characteristic of an altered RBC membrane is the formation of small knoblike spicules on the membrane surface that causes the cell to transform from a biconcave disk, discocyte, into a sea urchin-like shape, echinocyte. There are 3 stages of echinocyte transformation. In stage 1 RBCs start to have an irregular contour disk shape, stage 2 cells are still flat with spicules, and stage 3 cells have an even distribution of 10–30 spicules on the surface [25]. Transmitted light images provide insight whether extracts compromise host RBC integrity and therefore parasite survival, or if the extracts specifically target parasites. Cytotoxic effects of marine extracts on RBC were evaluated by the degree and amount of echinocyte transformation relative to untreated cultures. Extracts that caused cultures to have a 50% increase of

stage 3 echinocytes, when compared to untreated cultures, were considered toxic.

From the secondary screen, 10 of these extracts were observed to have minimal echinocyte transformation when compared to untreated parasite culture (data not shown). This second set of selected extracts most likely has a specific antimalarial effect with limited toxicity on RBC integrity.

Determination of antimalarial specificity of our selected extracts

To further validate our screening procedure, we determined the IC_{50} values of our selected extracts using the classical SYBR Green I Assay. We found that their exact IC_{50} ranged from 6.65 ng/ml to 0.85 µg/ml (Figure a2.2). Next we compared the antimalarial IC_{50} values to the IC_{50} values of these extracts against various infectious agents (e.g. methicillin-resistant *Staphylococcus aureus* (MRSA); vancomycin-resistant *Enterococcus faecium (VREF);* and both wild type and amphotericin B-resistant *Candida albicans* (WTCA and ARCA, respectively). MRSA is a pathogenic bacterium that causes skin and soft-tissue infections and is associated with sepsis and necrotizing pneumonia. VREF is another infectious bacterium that commonly causes urinary tract infections and can cause neonatal sepsis and meningitis. *C. albicans* is a fungus that lives in the gut of 80% of the human population, but can be lethal in immunocompromised patients. Most extracts were inactive against WTCA and ARCA. Results show that our selected

extracts inhibit growth of the *P. falciparum* 3D7 strain at 10–200-fold lower concentrations compared to other infectious agents (Table a2.1). Overall, our data obtained from the microscopic analysis of parasite and host RBC and other infectious agents indicated that our selected extracts are most active against *P. falciparum.*

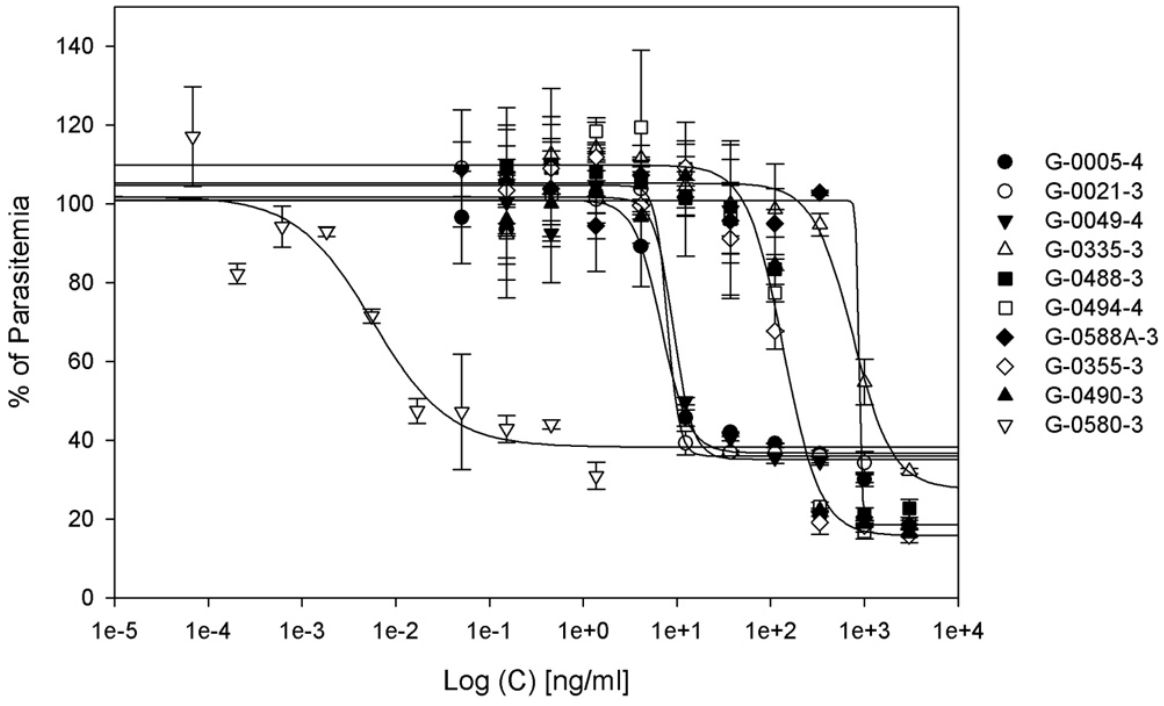


Figure a2.2. *IC_{50} graph of potent antimalarial extracts.* An inhibitory concentration at 50% growth (IC_{50}) graph of extracts inhibiting parasites at concentrations 6.65 ng/ml to 0.85 µg/ml

Table a2.1. *IC_{50} values of seven extracts in P. falciparum and other microbes.* Extracts are considered a 'hit' if IC_{50} <100 µg/mL, and extracts with >250 µg/mL are considered not active. IC_{50} values for 3D7 *P. falciparum* strain are significantly lower than other microbes (MRSA: methicillin-resistant *Staphylococcus aureus;* VREF: vancomycin- resistant *Enterococcus faecium;* WTCA: wild type *Candida albicans;* ARCA: amphotericin-B-resistant *Candida albicans).* NT indicates not tested, values for *P. falciparum* 3D7 samples are from duplicates, and the values for MRSA, VREF, WTCA, or ARCA screens are from a single experiment. Unit values are in µg/mL.

	3D7	MRSA	VREF	WTCA	ARCA
G-0005-4	0.680 ± 0.13	NT	16.4	227	156
G-0021-3	0.790 ± 0.17	11.9	15.2	>250	250
G-0049-4	0.916 ± 0.19	8.5	7.1	221	225
G-0335-3	0.775 ± 0.25	62.5	>250	>250	>250
G-0355-3	0.131 ± 0.02	1.9	69.2	>250	>250
G-0488-3	0.119 ± 0.03	0.5	>250	1.08	1.11
G-0490-3	0.163 ± 0.02	1.26	67.61	>250	>250
G-0494-4	0.137 ± 0.02	0.5	0.5	>250	>250
G-0580-3	0.0055 ± 0.01	92.26	153.11	>250	>250
G-0588A-3	0.847 ± 0.09	39.9	107.5	150	150

Phenotypic analysis of parasites incubated with extracts

To further investigate the potential mode of action of potent extracts, we examined *Plasmodium* morphology within 24 h of treatment. Synchronized parasite cultures were incubated with IC_{80} concentrations of extracts at the early and mature stages of the erythrocytic cycle. As shown in Figure a2.3A, untreated parasites displayed a robust fluorescent signal with an average of 1000 relative fluorescence units (RFUs) per frame, indicating a high level of RNA and consequently an active transcriptional state. In the transmitted light images, hemozoin, the byproduct of hemoglobin catabolism, can be observed in all

stages as a dark pigmented region within the parasites. We also detected a tight and continuous cell cycle progression during our first 24 h of observation. In contrast to this pattern, when treated with sponge extract G-0490-3, parasite growth was inhibited at the late ring and early trophozoite stage (Figure a2.3B). Parasites treated at the mature stage displayed a decrease of RNA fluorescence signal intensity to an average of 400 RFUs per frame in ring stage parasites. Following our phenotypic observations, it is highly possible that extract G-0490-3 targets a protein or a pathway essential to the early stage of the parasite intraerythrocytic cycle. This phenotype has been observed with *Cinchona* alkaloid treatment, which arrest development of *P. falciparum* at the mature trophozoite stage [3]. While extract G-0490-3 appeared to be ring stage specific, the cyanobacteria extract G-0580-3 produced a different effect on parasite phenotype (Figure a2.3C). When this extract was added to ring stage parasites, we observed enlarged vacuoles adjacent to hemozoin. When the same extract was added at the mature stage, multiple abnormal vacuoles appeared (16 h) and parasites were unable to egress properly (24 h). Giemsa stained blood smears confirmed the appearance of large vacuoles. This phenotype has been previously observed with inhibitors targeting protease in the parasite food vacuole. Enlarged food vacuoles are observed due to an accumulation of undegraded globin [26]. All together these data suggest that G-0580-3 may target a similar pathway. Similar phenotypes were observed with treatment with samples G-0005-4 (Figure

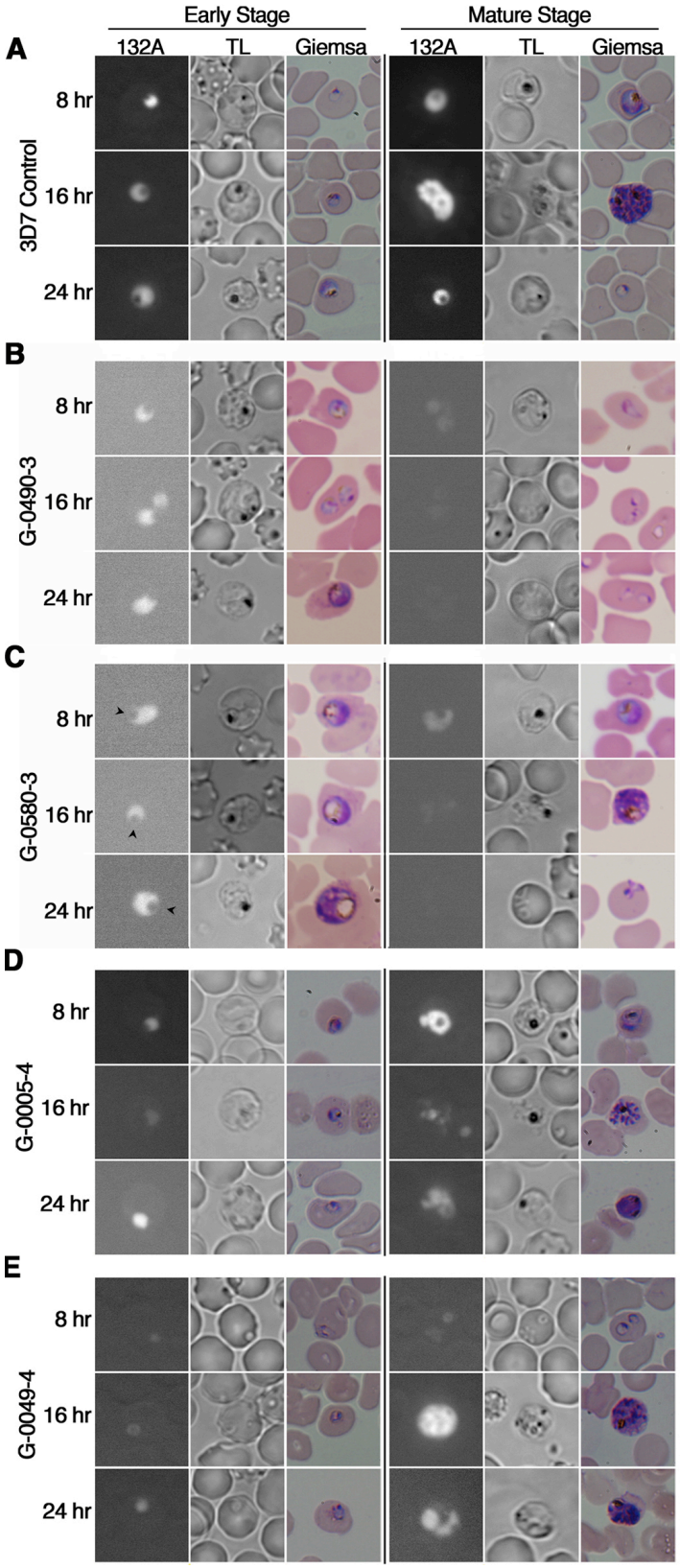


Figure a2.3. *Morphological analysis of P. falciparum with IC_{80} concentrations of extracts.* Infected erythrocytes were synchronized and live parasites were stained with RNA probe 132A. **A**) are images of the *P. falciparum* 3D7 strain without any drug treatment at the ring and mature stage (control). **B**) displays the phenotype of extract G-0490-3, stunted growth of parasites at the trophozoite stage. **C**) displays the phenotype of extract G-0580-3, arrowheads indicate the enlarged food vacuoles at the early stages of the cell cycle. **D** and **E**), extract G-0005-4 and G-0049-4 respectively, inhibit parasite life cycle progression, decrease hemozoin, and hinder egression

3d), G-0049-4 (Figure a2.3E), and G-0021-3 (not shown) and all were extracted from the red macroalga *Callophycus serratus.* To investigate the morphological effects of our selected extracts, we further examined their effects on chloroquine-

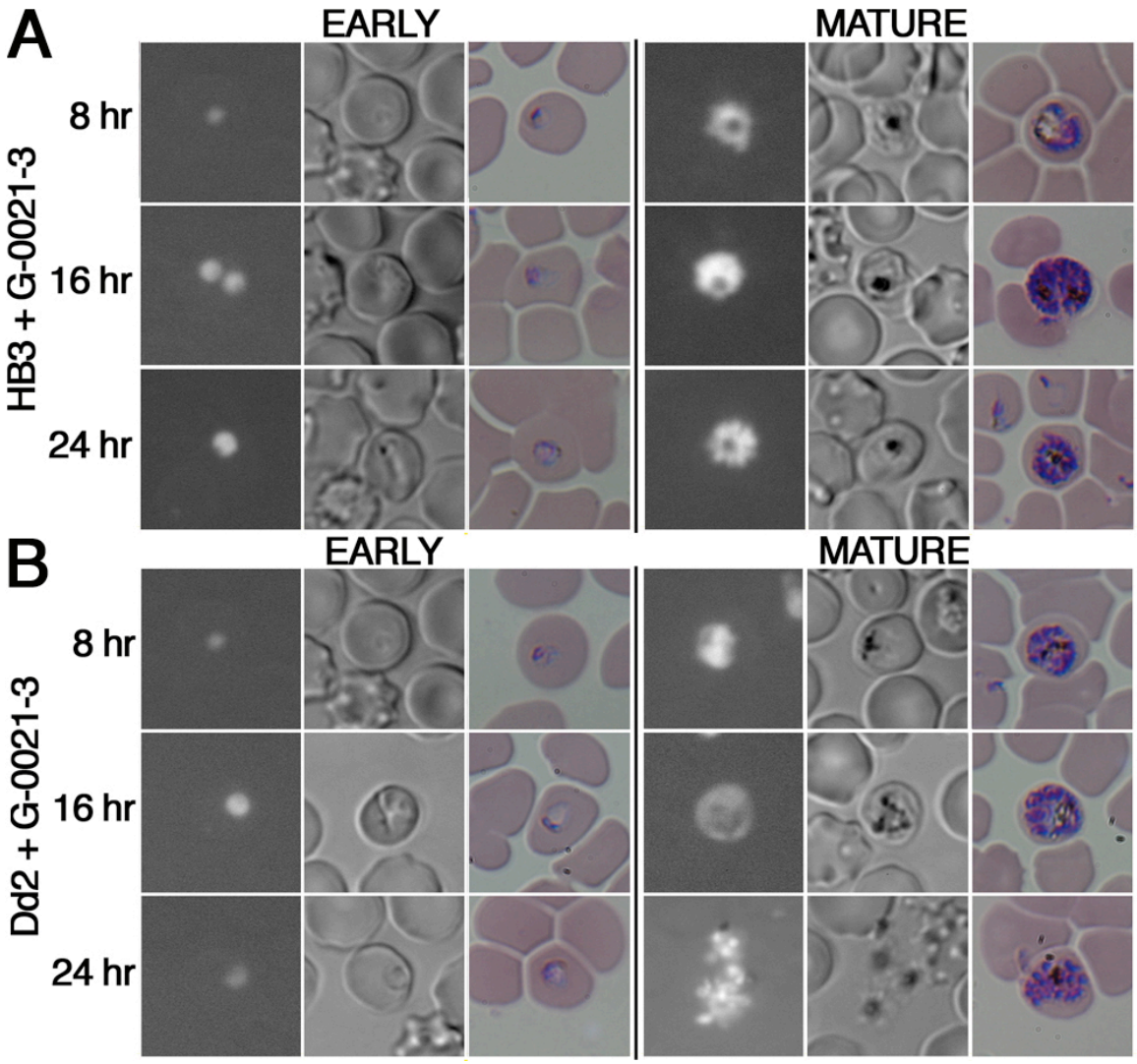


Figure a2.4. *Morphological analysis of various strains of P. falciparum.* Treatments with extract G-0021-3 on synchronized parasites at the ring and mature stage. A) chloroquine sensitive strain, HB3; and B) chloroquine resistant strain, Dd2

resistant, Dd2 (Figure a2.4a), and chloroquine-sensitive, HB3 (Figure a2.4b) strains. In Dd2, HB3 and the reference 3D7 strain, we observed the almost identical morphological phenotype of stunted hemozoin formation and cell cycle arrest. Collectively, by using the HCLCI assay we can observe phenotypic changes that can be used as a guide to identifying antimalarial properties on drug resistant and sensitive parasite strains.

Verification of method using initial extracts from purified bioactive compound bromophycolide A

The purified bioactive compound, (bromophycolides (A-U)), present in G-0005-4, G- 0049–4, and G-0021-3 extracts isolated from *C. serratus* [20, 27-27] was shown to inhibit parasite growth at a sub-micromolar range (IC_{50} value of 0.7 μM [28]; IC_{80} value of 2.8μM). To understand the morphological effects of pure bromophycolide A on *P. falciparum* 3D7 cultures, we synchronized parasites and treated cultures at its IC_{80} value at the early and late stage of the parasite erythrocytic cycle (Figure a2.5). Using our high-content live cell-imaging assay, we detected an immediate arrest of cell cycle progression at every stage tested, a decrease in RNA intensity and an enlargement of the food vacuole in the parasite. Furthermore we observed a hindrance of hemozoin formation from the ring to trophozoite stage. All together, our newly developed screen is able to

identify new antimalarials and suggest starting hypotheses for understanding the molecular processes targeted by extracts and their purified bioactive compounds.

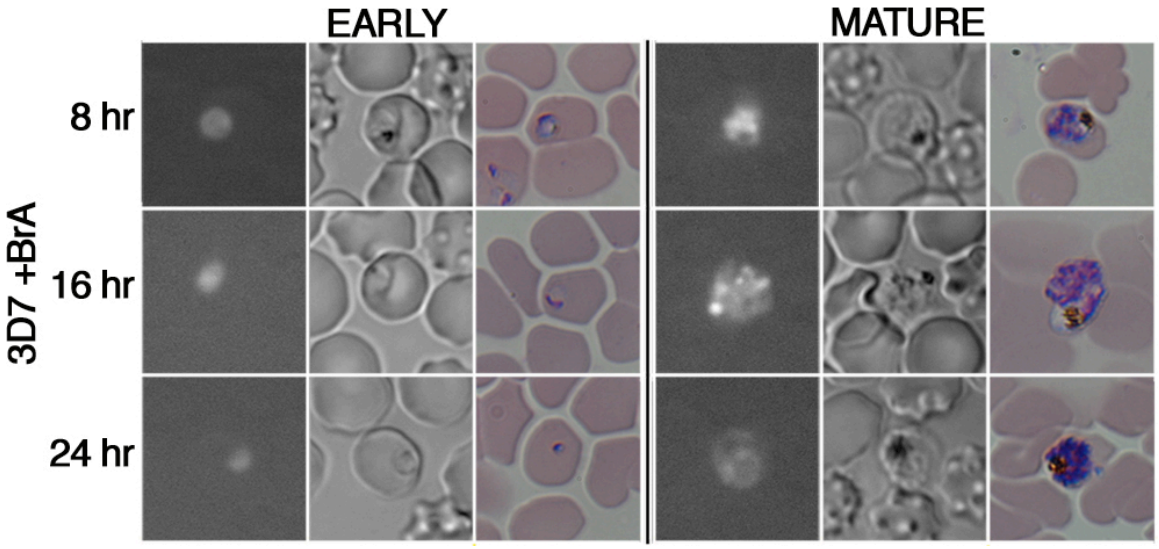


Figure a2.5. *Morphological analysis of P. falciparum with bromophycolide A.* Images of synchronized parasites at the ring and mature stage with bromophycolide A treatment at IC_{80} concentrations. Untreated control sample is shown in Figure a2.3A

Within the past decade significant advancements in high-throughput antimalarial drug screening have efficiently reduced time, labor and costs. However, it remains a challenge to determine the mode of action of active compounds [30]. We propose in this work to use our HCLCI assay to screen large libraries for antimalarial properties. The HCLCI assay is straightforward and allows for semi-automated image analysis of parasite drug resistant and sensitive strains without laborious steps (ex. blood smears, fixation, centrifugation, or aspiration steps). An additional advantage of our assay is the ability to observe cytotoxic effects on host red blood cell. Our method can detect parasite morphological changes to

select for specific phenotypes, and further select the extracts for the purification of a single bioactive compound with activity against malaria parasite growth. We observed a distinct phenotype of cell cycle arrest and lack of hemozoin formation when cultures were incubated with extracts from which bromophycolide A was purified and thus pure bromophycolide A as well. These observations indicated that bromophycolide A's potential mode of action may prevent hemozoin formation. Further investigation using a fluorescent coumarin tagged bromophycolide A for sub-cellular localization and molecular target identification studies validated that heme crystallization was disrupted by the natural product bromophycolide A [31].

Conclusions

In this study we investigated 2685 extracts from sessile marine organisms collected in Fijian waters and successfully identified nine selective extracts that exhibited parasite growth inhibition without compromising RBC membrane integrity. Furthermore we were able to validate that the purified bioactive compound bromophycolide A extracted from the red macroalga *C. serratus,* had similar morphological effects compared to the initial tested crude extracts. Collectively, we demonstrate that by using our newly developed screening we can quickly select a desired phenotype from crude extracts to act as a guide to the purification of a single antimalarial compound. Specific parasite phenotypes can provide insight on how extracts act on the whole organism, thus presenting a starting point to test hypotheses regarding the potential drug's mode of action. In the antimalarial drug field, targeting the parasite at different stages of the life cycle by combinatorial drug treatment is highly recommended to prevent drug resistance. By implementing our screen to classical drug discovery methods, we can expedite the discovery of molecules that can target distinct parasitic stages and lead to new combination drug therapies.

Methods

Parasite culture

3D7, HB3 and Dd2 *P. falciparum* malaria parasites (MRA-102, 155, 156, MR4, ATCC® Manassas, Virginia) were cultured in human type O + erythrocytes in complete medium (RPMI 1640 w/L-Glutamine, w/o Phenol Red (Gibco), 0.043 mg/mL Gentamicin (Gibco), 0.014 mg/mL Hypoxantine (Acros), 38.5 mM HEPES (Promega), 0.18% Sodium Bicarbonate (Gibco), 0.2% Glucose (MP Biomedical), 2.6 mM NaOH (Sigma), 0.20% Albumax (Gibco), 5% human serum) as previously described [22]. Cultures were maintained in 25-cm2 flasks (Corning) at a volume of 10 ml and were gassed for 30 s with an environment of 3% CO2, 1% O2, and 96% N2, then incubated at 37°C. Synchronization of culture was achieved through sorbitol lysis of mature stage using 5% sorbitol (Fisher) fine-tuned by another lysis 8 h later (27).

Marine organism extracts

Each marine organism was exhaustively extracted with methanol, and the extracts were combined, filtered, and reduced *in vacuo.* The crude extract was separated with HP20ss resin into four fractions, eluting with (1) 1:1 methanol: water, (2) 4:1 methanol: water, (3) 100% methanol, and (4) 100% acetone.

Plasmodium falciparum SYBR Green I-based fluorescence growth inhibition assay

Extracts were diluted into three different concentrations, 100, 33, and 11 µg/ml, by the Biomek FXP Laboratory Automation Workstation (Beckman Coulter, Inc., Brea, CA) into clear bottom 96-well plates (Costar #3904). *P. falciparum* cultures were added at a 2.5% hematocrit with a 1% parasitemia to 96-well plates and incubated for 72 h at 37°C. Plates were frozen at –80°C overnight. Lysis buffer (20 mM Tris–HCl, 5 mM EDTA, 0.008% saponin, 0.08% Triton X-100, and 0.2 µl/ml SYBR Green I dye (Invitrogen #S7585)) was added to thawed plates and incubated at 37°C. Microtiter plate was read using the Spectra Max Gemini EM reader (Molecular Devices). Data were analyzed with SoftMax Pro v5 (Molecular Devices Software, Inc., Sunnyvale, CA). The formula z' = 1 - (3SD+ + 3SD-)/lAve + – Ave-l was used to calculate the Z' value. Where SD + represents the positive control standard deviation, SD- the negative control standard deviation, Ave + the mean value of the positive control, and Ave- the mean value of the negative control.

Bacterial growth inhibition assay

Methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococcus faecium* (VREF) as test pathogens were used to perform antibacterial assays. As previously described [20], these bacteria were grown in

GYT media (1 g glucose, 2.5 g yeast extract, 5.0 g tryptone, 1 L H2O) overnight at 37°C. Cultures were diluted to an optical density of 0.04-0.06, diluted 10-fold, and then 250µg/ml of extracts were added to microtiter plates. Extracts were serially diluted for a total of 8 dilutions, and the plates were incubated for 16–18 h at 37°C. The optical density was then measured at 600 nm using a Molecular Devices Emax microplate reader and the minimum inhibitory concentration (MIC) was calculated for each compound using the analysis program SOFTmax PRO.

Fungal growth inhibition assay

As previously described [20] amphotericin B-resistant *Candida albicans* was grown overnight at 30°C in RPMI media (Invitrogen). The indicator Alamar Blue 100× (TREK Diagnostic Systems) was added to the cell suspension at 1 × 104 cells/mL, and then cells were added to microtiter plates. Extracts were added to 96-well plates at 250 µg/ml, serially diluted for a total of 8 dilutions, and incubated for 12–15 h at 37°C. The visual assessment of colorimetric changes was done using a BioTek model ELx800 microplate reader to detect the altered oxidation state of Alamar Blue. Plotting the absorbance versus compound concentration indicated cell proliferation.

High-content live cell imaging assay, Montage analysis

Samples were prepared in optical bottom 96-well assay plates (Costar #3614, Corning, NY). Parasite cultures were diluted to 0.025% hematocrit with a 6% parasitemia in complete medium and 240 µl was added to each well. RNA-probe 132A was incubated with samples for 30 min at 37°C, in the dark at a final concentration of 5 µM. The BD Pathway HT (BD Biosciences Bioimaging, Rockville, MD) incubation chamber was set at 37°C and 5% CO2. Individual wells were fluoresced by Semrock (Rochester, NY) Texas Red BrightLine filter sets. A 3 × 3 tiling composed montage images to examine approximately 1500 red blood cells.

IC_{80} phenotypic assay

Parasite cultures were synchronized using 5% sorbitol as previously described [23]. Cultures were then diluted to 1.67% hematocrit with a 6% parasitemia for ring stage and 3% mature stage in complete medium to reach a final volume of 240 µl. Treatment with extracts or bromophycolide A at their respective IC_{80} values were added at 0, 8, and 16 h, and image analysis occurred at 24 h.

List of Abbreviations

HCLCI: high-content live cell-imaging; **RBC**: red blood cell.

Authors' contributions

KLR conceived the study and helped to draft the manuscript. JP and MB prepared samples for the primary screen. SC and DC participated in the design of the HCLCI assay. SC and MC prepared samples and carried out microscopic analysis. JK, MEH and WA provided extracts. SE and EPS established the library and purified extracts. SC drafted the manuscript, and all authors read and approved the final manuscript.

Acknowledgements

This work was supported by the U.S. National Institutes of Health's International Cooperative Biodiversity Groups program (grant No. U01 TW007401). We would like to thank D. Montgomery for critically editing our manuscript.

References

1. Organization WH: **World Malaria Report**. *Geneva, Switzerland, 2009.* 2009.

2. JE H: **Drug-resistant malaria**. *Trends Parasitol* 2005, **21**(11):494–498.

3. White NJ: **Qinghaosu (Artemisinin): The Price of Success**. *Science* 2008, **320**:330–334.

4. Noedl H, Youry Se, Schaecher K, Smith BL, Socheat D, Fukuda MM, the Artemisinin Resistance in Cambodia 1 (ARC1) Study Consortium: **Evidence of Artemisinin-Resistant Malaria in Western Cambodia**. *N Engl J Med* 2008, **359**:2619–2620.

5. Plouffe D, Brinker A, McNamara C, Henson K, Kato N, Kuhen K, Nagle A, Adrián F, Matzen JT, Anderson P, Nam TG, Gray NS, Chatterjee A, Janes J, Yan SF, Trager R, Caldwell JS, Schultz PG, Zhou Y, Winzeler EA: **In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen**. *Proc Natl Acad Sci U S A* 2008, **105**(26):9059–9064.

6. Weisman JL, Liou AP, Shelat AA, Cohen FE, Guy RK, DeRisi JL: **Searching for new antimalarial therapeutics amongst known drugs**. *Chem Biol Drug Des* 2006, **67**(6):409–416.

7. Guiguemde WA, Shelat AA, Bouck D, Duffy S, Crowther GJ, Davis PH, Smithson DC, Connelly M, Clark J, Zhu F, Jiménez-Díaz MB, Martinez MS, Wilson EB, Tripathi AK, Gut J, Sharlow ER, Bathurst I, El Mazouni F, Fowble JW, Forquer I, McGinley PL, Castro S, Angulo-Barturen I, Ferrer S, Rosenthal PJ, Derisi JL, Sullivan DJ, Lazo JS, Roos DS, Riscoe MK, Phillips MA, Rathod PK, Van Voorhis WC, Avery VM, Guy RK: **Chemical genetics of Plasmodium falciparum**. *Nature* 2010, **465**(7296):311–315.

8. Gamo FJ, Sanz LM, Vidal J, de Cozar C, Alvarez E, Lavandera JL, Vanderwall DE, Green DV, Kumar V, Hasan S, Brown JR, Peishoff CE, Cardon LR, Garcia-Bustos JF: **Thousands of chemical starting points for antimalarial lead identification**. *Nature* 2010, **465**(7296):305–310.

9. Desjardins RE, Canfield CJ, Haynes D, Chulay JD: **Quantitative assessment of antimalarial activity in vitro by a Semiautomated Microdilution Technique**. *Antimicrob Agents Chemother* 1979, **16**(6):710–718.

10. Smilkstein M, Sriwilaijaroen N, Kelly JX, Wilairat P, Riscoe M: **Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening.** *Antimicrob Agents Chemother* 2004, **48**(5):1803–1806.

11. Johnson JD, Denull RA, Gerena L, Lopez-Sanchez M, Roncal NE, Waters NC: **Assessment and continued validation of the malaria SYBR Green I-based fluorescence assay for use in malaria drug screening**. *Antimicrob Agents Chemothery* 2007, **51**(6):1926–1933.

12. Corbett Y, Herrera L, Gonzalez J, Cubilla L, Capson TL, Coley PD, Kursar TA, Romero LI, Ortega-Barria E: **A novel DNA-based microfluorimetric method to evaluate antimalarial drug activity**. *Am J Trop Med Hyg* 2004, **70**(2):119–124.

13. Baniecki ML, Wirth DF, Clardy J: **High-throughput Plasmodium falciparum growth assay for malaria drug discovery**. *Antimicrob Agents Chemother* 2007, **51**(2):716–723.

14. Ku MJ, Dossin FM, Choi Y, Moraes CB, Ryu J, Song R, Freitas-Junior LH: **Quantum dots: a new tool for anti-malarial drug assays**. *Malar J* 2011, **10**:118.

15. Lucumi E, Darling C, Jo H, Napper AD, Chandramohanadas R, Fisher N, Shone AE, Jing H, Ward SA, Biagini GA, Degrado WF, Diamond SL, Greenbaum DC: **Discovery of potent small molecule inhibitors of multi-drug resistant P. falciparum using a novel miniaturized high-throughput luciferase-based assay.** *Antimicrob Agents Chemother* 2010.

16. Buchholz K, Burke TA, Williamson KC, Wiegand RC, Wirth DF, Marti M: **A high-throughput screen targeting malaria transmission stages opens new avenues for drug development**. *J Infect Dis* 2011, **203**(10):1445–1452.

17. Cervantes S, Prudhomme J, Carter D, Gopi KG, Li Q, Chang YT, Le Roch KG: **High-content live cell imaging with RNA probes: advancements in high-throughput antimalarial drug discovery.** *BMC Cell Biology* 2009, **10**:45.

18. Lane AL, Kubanek J: **Secondary metabolite defenses against pathogens and biofoulers**. *Algal Chemical Ecology, Amsler CD (ed), Springer-Verlag, Berlin* 2008:229–243.

19. Engel S, Jensen PR, Fenical W: **Chemical ecology of marine microbial defense**. *J Chem Ecol* 2002, **28**(10):1971–1985.

20. Kubanek J PAC, Snell TW, Giese RA, Hardcastle K, Fairchild C, Aalbersberg W, Raventos-Suarez C, Hay ME: **Antineoplastic diterpene-benzoate macrolides from the Fijian red alga Callophycus serratus**. *Org Lett* 2005, **7:**5261–5264.

21. Stout EP, Hasemeyer AP, Lane AL, Davenport T. Engel S, Hay ME, Fairchild CR, Prudhomme J, Le Roch K, Aalbersberg W, Kubanek J: **Antibacterial neurymenolides from the Fijian red alga Neurymenia fraxinifolia.** *Org Lett* 2009, **11**:225–228.

22. Trager W, Jensen JB: **Human malaria parasites in continuous culture**. *Science* 1976, **193**(4254):673–675.

23. Lambros C, Vanderberg JP: **Synchronization of Plasmodium falciparum erthrocytic stages in culture**. *J Parasitol* 1979, **65**(3):418–420.

24. Suwalsky M, Kuhajda FP, Villena F, Sotomayor CP: **Effects of lithium of the human erythrocyte membrane and molecular models**. *Biophys Chem* 2007, **129**:36–42.

25. Brecher G, Bessis M: **Present status of spiculed red cells and their relationship to the discocyte-echinocyte transformation: a critical review**. *Blood* 1972, **40**(3):333–344.

26. Rosenthal PJ: ***Plasmodium falciparum*: effects of proteinase inhibitors on globin hydrolysis by cultured malaria parasites**. *Exp Parasitol* 1995, **80**(2):272–281.

27. Lin AS, Stout EP, Prudhomme J, Le Roch K, Fairchild CR, Franzblau SG, Aalbersberg W, Hay ME, Kubanek J: **Bioactive bromophycolides R-U from the Fijian red alga Callophycus serratus.** *J Nat Prod* 2010, **73**(2):275–278.

28. Lane AL, Stout EP, Lin AS, Prudhomme J, Le Roch K, Fairchild CR, Franzblau SG, Hay ME, Aalbersberg W, Kubanek J: **Antimalarial bromophycolides J-Q from the Fijian red alga Callophycus serratus.** *J Org Chem* 2009, **3**(74(7)):2736–2742.

29. Kubanek J, Prusak AC, Snell TW, Giese RA, Fairchild CR, Aalbersberg W, Hay ME: **Bromophycolides C-I from the Fijian red alga Callophycus serratus.** *J Nat Prod* 2006, **69**(5):731–735.

30. Wells TN, Duffy PE: **When is enough enough? The need for a robust pipeline of high-quality antimalarials**. *Discovery Med* 2010, **9**(48):389–398.

31. Stout EP, Cervantes S, Prudhomme J, France S, La Clair JJ, Le Roch K, Kubanek J: **Bromophycolide a targets heme crystallization in the human malaria parasite *Plasmodium falciparum***. *ChemMEDChem 2011*, **6**(9):1572–1577.

Chapter 3a

Bromophycolide A Targets Heme Crystallization in the Human Malaria Parasite *Plasmodium falciparum*

E. Paige Stout[†], Serena Cervantes[§], Jacques Prudhomme[§], Stefan France[†], James J. La Clair[||], Karine Le Roch[§] and Julia Kubanek[*,†,‡]

[†]School of Chemistry and Biochemistry and [‡]School of Biology, Georgia Institute of Technology, Atlanta, GA, USA 30332. [§]Department of Cell Biology and Neuroscience, University of California Riverside, Riverside, CA, USA 92521. [||]Xenobe Research Institute, P. O. Box 4073, San Diego, California, 92164-4073, USA

Summary

The ability for *P. falciparum* to mutate quickly allows drug resistance to develop faster than most other eukaryotic pathogens. Identifying novel antimalarial compounds, instead of derivatives from current antimalarial drugs, is imperative. Macrocyclic meroditerpenes were previously isolated from the tropical red macroalga *Callophycus serraturs* and found to have antimalarial properties. The natural product bromophycolide A was found to be the single bioactive compound. By adding a fluorescent probe onto bromophycolide A, we determined that inhibition of heme crystallization is the mode of action. Hemoglobin is the main nutrient source for parasites during the erythrocytic stage and when catabolized it creates toxic heme that needs to crystalize to become nontoxic. Disruption of heme crystallization is a physicochemical process and is not part of any parasite enzymatic reaction, therefore eliminating the possibility of rapid resistance development.

Abstract

Malaria is a major human health concern in many tropical and sub-tropical regions where current antimalarial drugs are becoming increasingly ineffective due to emerging and spreading resistance. This shortcoming has created a need for new leads with defined activities. Using a combination of cellular and molecular approaches, this study identifies a mode of action for a unique class of antimalarial macrocyclic meroditerpenes isolated from the tropical red macroalga *Callophycus serratus*. Derivatization of the natural product bromophycolide A into fluorescent probes facilitated the identification of a non-enzymatic mode of action against the human malaria parasite *Plasmodium falciparum*. Through a combination of structure-activity relationship studies, confocal fluorescent microscopy of live parasites, UV-visible spectroscopic analyses, and affinity biochemical techniques, bromophycolide A was found to suppress proliferation of the human malaria parasite by inhibition of heme crystallization. This inhibition prevents detoxification of heme to hemozoin in parasite food vacuoles, and ultimately resulting in parasite death. New classes of natural products that target heme crystallization are excellent candidates for further drug development due to the fact that heme crystallization is a physicochemical rather than an enzymatic process, removing some avenues for rapid evolution of drug resistance within the parasite. These studies identify a new class of natural products that target heme

detoxification in both drug-sensitive and drug-resistant *P. falciparum* and suggest new leads to alleviate drug resistant infections.

Introduction

Plasmodium falciparum, the most deadly human malaria parasite, poses a major threat to human health worldwide with over 500 million clinical cases and 1-2 million deaths annually (*1*). Natural products and their synthetic derivatives have provided the greatest number of successful antimalarial treatments to date, representing approximately 65% of prescribed drugs (*2*). Quinine, discovered from cinchona tree bark, has been used to treat malaria since the 17th century and was the primary antimalarial drug until it was replaced by chloroquine, a synthetic derivative, in the 1940s (*3*). Chloroquine became the mainstay antimalarial agent until resistant strains began to appear over a decade after its introduction. Artemisinin, isolated from the plant *Artemisia annua* used in traditional Chinese medicine, ushered in a new wave of antimalarials and became the most potent and rapid-acting drug available (*4-5*). Several artemisinin synthetic derivatives have since been developed, and artemisinin-based combination therapies are currently being used throughout the world to treat this parasitic disease. However, artemisinin-resistant strains have recently been reported (*6-7*), and new antiparasitic drugs are urgently needed to combat these strains.

Therapeutic development for malaria is complex due to the intricate host-guest interactions during the plasmodial life cycle. During its maturation, the malaria parasite is transferred to human hosts via mosquitoes, and after a

transitory stage in the liver, the parasite resides in the hosts' red blood cells. While residing in red blood cells, the parasite catabolizes hemoglobin within its food vacuole for its nutritional benefit. This hemoglobin catabolism can be lethal as it releases free heme molecules which are toxic to the parasite (*8*). However, the parasite overcomes heme toxicity through the crystallization of free heme into non–toxic hemozoin, a non–enzymatic process which is crucial to parasite survival and thus an excellent antimalarial drug target (*9*). Chloroquine has been shown to inhibit heme crystallization (*10*), and structurally related antimalarial drugs such as quinine, amodiaquine and mefloquine have also been proposed to disrupt hemozoin formation. Resistance to chloroquine has been reported to occur through multiple mutations in a *P. falciparum* transmembrane protein that, when mutated, pumps chloroquine out of the food vacuole, significantly reducing the concentration of the drug (*11*). The primary mode of action for artemisinin remains in debate, although several hypotheses have been reported, including alkylation of heme by carbon-centered free radicals (*12-13*), interference with proteins such as the sarcoplasmic/endoplasmic calcium ATPase (*14*), as well as damaging of normal mitochondrial functions (*15*). The identification of molecular targets for bioactive natural products is essential for understanding their modes of action, developing new drug leads, and anticipating evolution of resistance.

Results and Discussion

We previously discovered a unique group of brominated macrocyclic meroditerpenes from the Fijian red macroalga *Callophycus serratus* (*16-17*) and showed that this class of natural products is cytotoxic to *P. falciparum* (*18-19*). The most abundant natural product of this class with sub–micromolar inhibition of *P. falciparum* proliferation, bromophycolide A (**1**), induces a strong arrest of the parasite erythrocytic cycle (*20*), and was therefore used to develop a probe for both molecular imaging and molecular target identification studies to gain insights on potential modes of action of **1** against *P. falciparum*.

In order to maintain antimalarial activity while not compromising the integrity of potential receptor binding, attempts initially focused on connecting a linker through the tertiary alcohol at C11 since changes at this position resulted in only minor differences in activity, based on structure–activity relationship (SAR) trends gleaned from the known *C. serratus* natural products (see Figures a3.1 and a3.2). Unfortunately, all synthetic efforts failed at this position, likely due to a combination of poor reactivity of tertiary alcohols and steric hindrance. From a synthetic perspective, the most chemically reactive and accessible position in **1** is the phenolic hydroxyl at C18. Because the C18 phenol was conserved in all *C. serratus* metabolites and thus SAR data was lacking, we acetylated the phenol in **1** to test for changes in activity. Fortuitously, the acetylated derivative **7** resulted in an increase in activity (IC_{50} value of 241 nM, Tables 1 and 2). Therefore, two

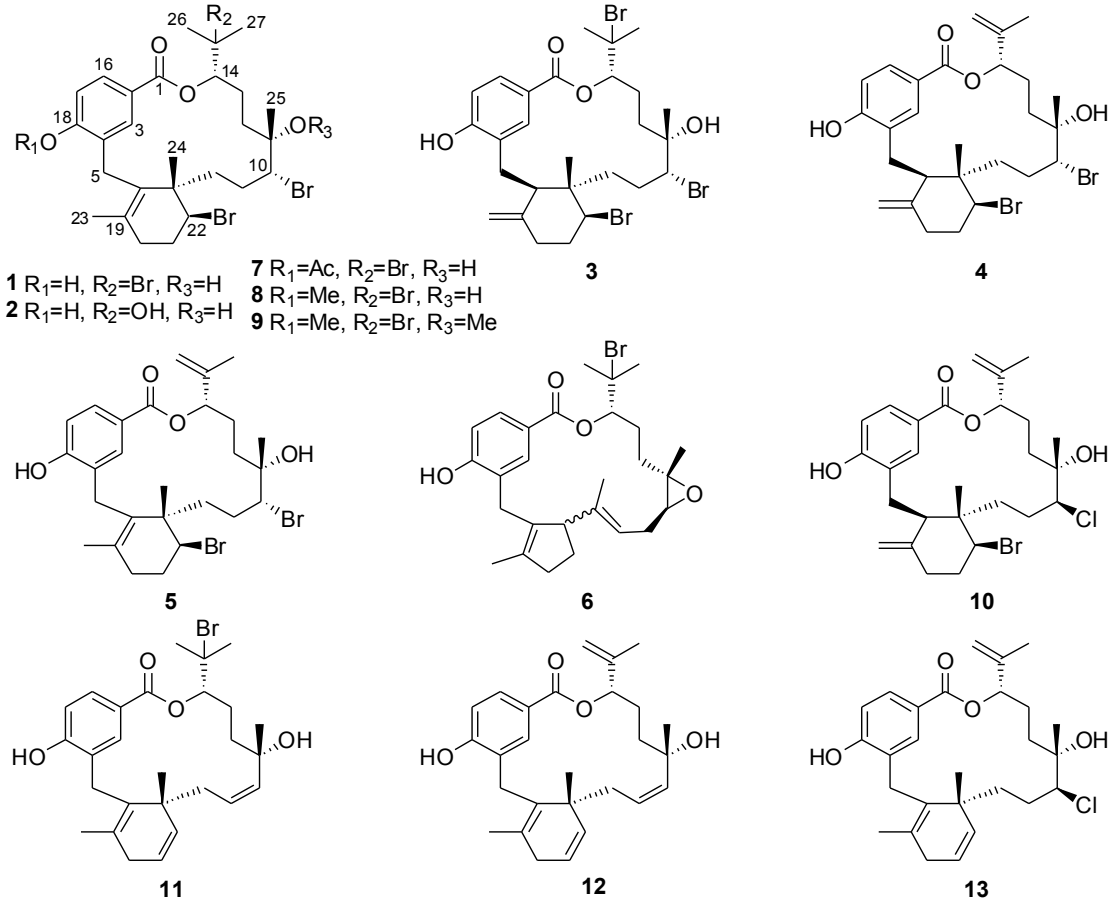


Figure a3.1. Structures of bromophycolides A (**1**), C (**2**), D (**3**), E (**4**), M (**5**), and debromophycolide A (**6**) isolated from *Callophycus serratus* (*16-18*) and semi-synthetic derivatives of bromophycolide A (**7-13**).

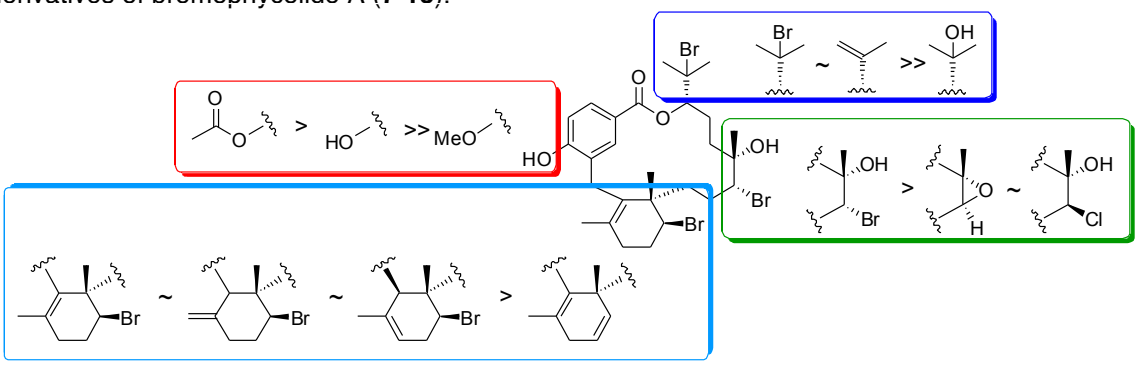


Figure a3.2. Structure-activity relationship summary of *C. serratus* natural products and bromophycolide A (**1**) semi-synthetic derivatives. Greater potency is indicated by ">", lesser potency by "<", similar potency by "~".

Table a3.1. Antimalarial activities of bromophycolides A (**1**), C (**2**), D (**3**), E (**4**), M (**5**), debromophycolide A (**6**), semi-synthetic derivatives of bromophycolide A (**7-13**), and probes **15** and **17**.

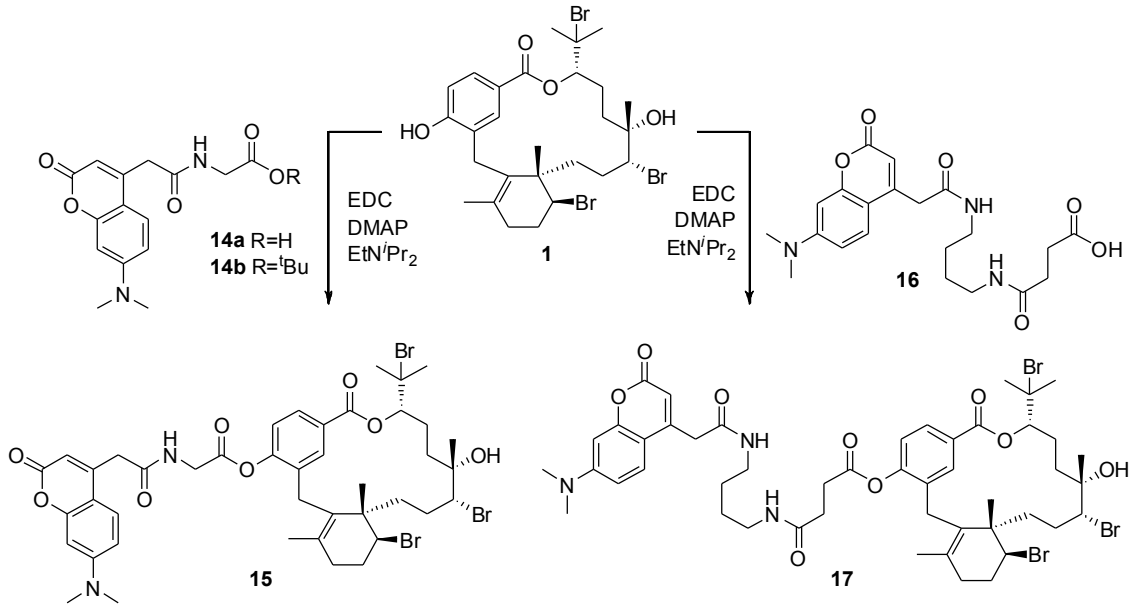
cmpd	*P. falciparum* (3D7)
1	0.66
2	56
3	0.35
4	0.82
5	0.55
6	>100
7	0.24
8	23
9	21
10	3.1
11	4.6
12	3.5
13	2.7
15	0.27
17	0.53

Table a3.2. Efficacies of bromophycolide A (**1**), 18-OAc-bromophycolide A (**7**) and amodiaquine (AMQ) against three *Plasmodium falciparum* strains, in inhibition of heme crystallization, and cytotoxicity against healthy human cells.

cmpd	3D7 IC_{50} (nM)[a]	Dd2 IC_{50} (nM)[b]	HB3 IC_{50}	IC_{50} for heme crystallization	Cytotoxicity (IC_{50} mM)		
					Vero	J774	HepG2
1	658	497	565	2.5 ± 0.13	34.6	19.0	21.3
7	241	304	431	2.0 ± 0.15	74.0	24.7	38.5
AMQ	7.8[d]	–	8.5[d]	1.2 ± 0.19	–	–	–

[a]3D7 and HB3 = chloroquine-sensitive parasites
[b]Dd2 = chloroquine-resistant parasite
[c]IC_{50} values in Ncokazi & Egan (*24*) were reported as molar equivalents of drug to heme
[d]From Hawley *et al.* (*39*)

fluorescent immunoaffinity fluorescent (IAF) tags (*21*) with varying chain lengths were each synthetically appended to **1** through an acetyl linkage (Scheme 1). Probe **15** exhibited an IC_{50} of 271 nM, similar to that of **7**. Furthermore, both tags **14a** and **14b** were completely inactive against *P. falciparum*, exhibiting IC_{50} values >100 μM and confirming that the activity observed in probe **15** was a result of the natural product and not the IAF tag.



Scheme 1. Probes **15** and **17** were prepared by coupling acids **14** and **16**, respectively, with bromophycolide A (**1**). EDC = *N'*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide; DMAP = 4-dimethylaminopyridine; EtN^iPr_2 = *N,N*-diisopropylethylamine; DMF = dimethylformamide.

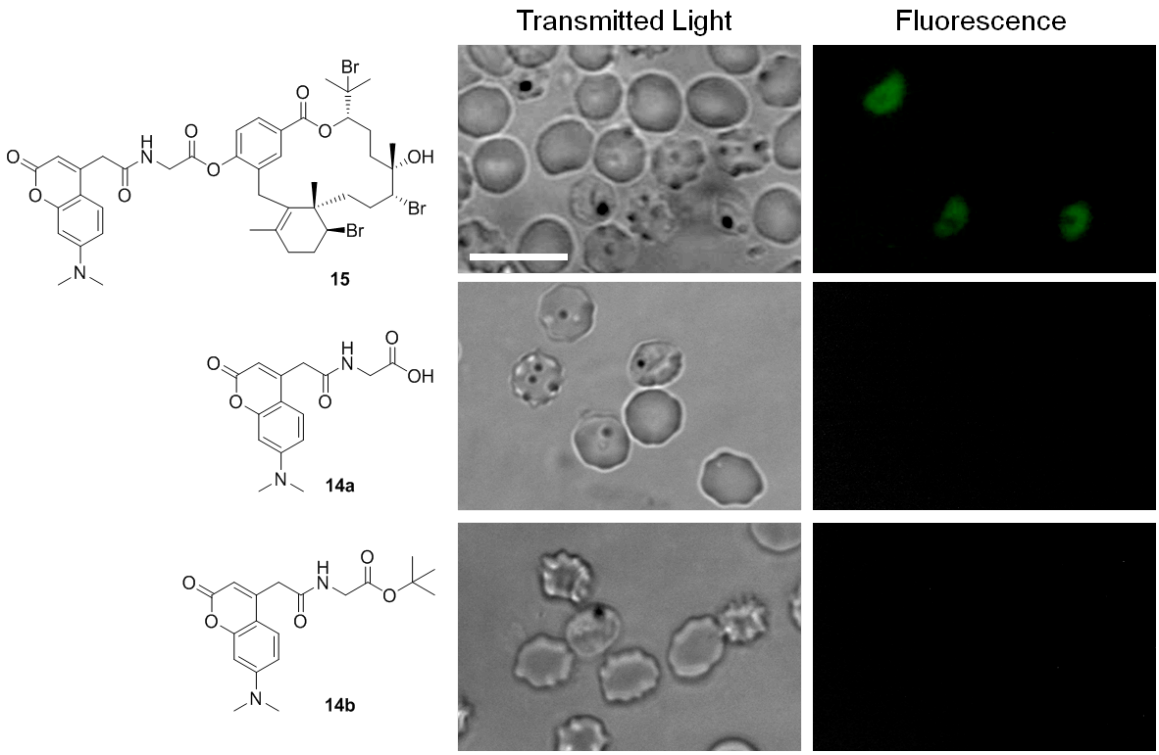


Figure a3.3. Confocal microscopy images of *live* mixed-stage *P. falciparum*–infected erythrocytes incubated with either 10 μM bromophycolide A probe **15**, 10 μM control tags **14a** or **14b** in complete medium. Hemozoin crystals were identified in the transmitted light image by dark crystals. Probe **15** localized only in *P. falciparum*–infected erythrocytes, as seen in the fluorescent image. No fluorescence was observed when *P. falciparum*–infected erythrocytes were incubated with either **14a** or **14b**. Scale bar denotes 10 μm.

With active probes in hand, we turned to confocal microscopy to study the intracellular distribution of probe **15** in *P. falciparum*–infected erythrocytes. Fluorescence was observed only in *P. falciparum*–infected erythrocytes (Figure a3.3) which were distinguished by the presence of hemozoin (dark crystals, Hz). No fluorescence was observed when *P. falciparum*–infected erythrocytes were incubated with either tag **14a** or **14b**, offering further evidence that **15** provided an effective mimic of natural product **1** (Figure a3.3). In order to determine its

specific intracellular target, the histological localization of probe **15** was compared with that of fluorescent organelle and lipid probes by co–incubation and imaging using multichannel fluorescence confocal microscopy. Incubation of infected erythrocytes with both **15** and Nile Red, a neutral lipid stain, repeatedly showed co–localization of the two stains, suggesting that **15** was associating with intracellular neutral lipids (Figure a3.4). Pisciotta established that heme crystallization occurs inside neutral lipid nanospheres, which encompass hemozoin in *P. falciparum* food vacuoles (*22*). Therefore, molecular imaging studies with **15** and infected erythrocytes provided the first suggestion that **1** could be targeting heme crystallization.

To further explore possible heme–bromophycolide *chemical* interactions, the absorbance spectra of ferriprotoporphyrin IX (Fe(III)PPIX), commercially available heme, was recorded in 20 mM HEPES, pH 7.4 containing 40% DMSO, conditions that were previously demonstrated to evaluate monomeric Fe(III)PPIX(*23*). Fe(III)PPIX gives a characteristic absorption spectrum with an intense Soret band at 402 nm, whose intensity was decreased by up to 45% by addition of **1**, suggesting that **1** forms a saturable complex with Fe(III)PPIX (Figure a3.5a).

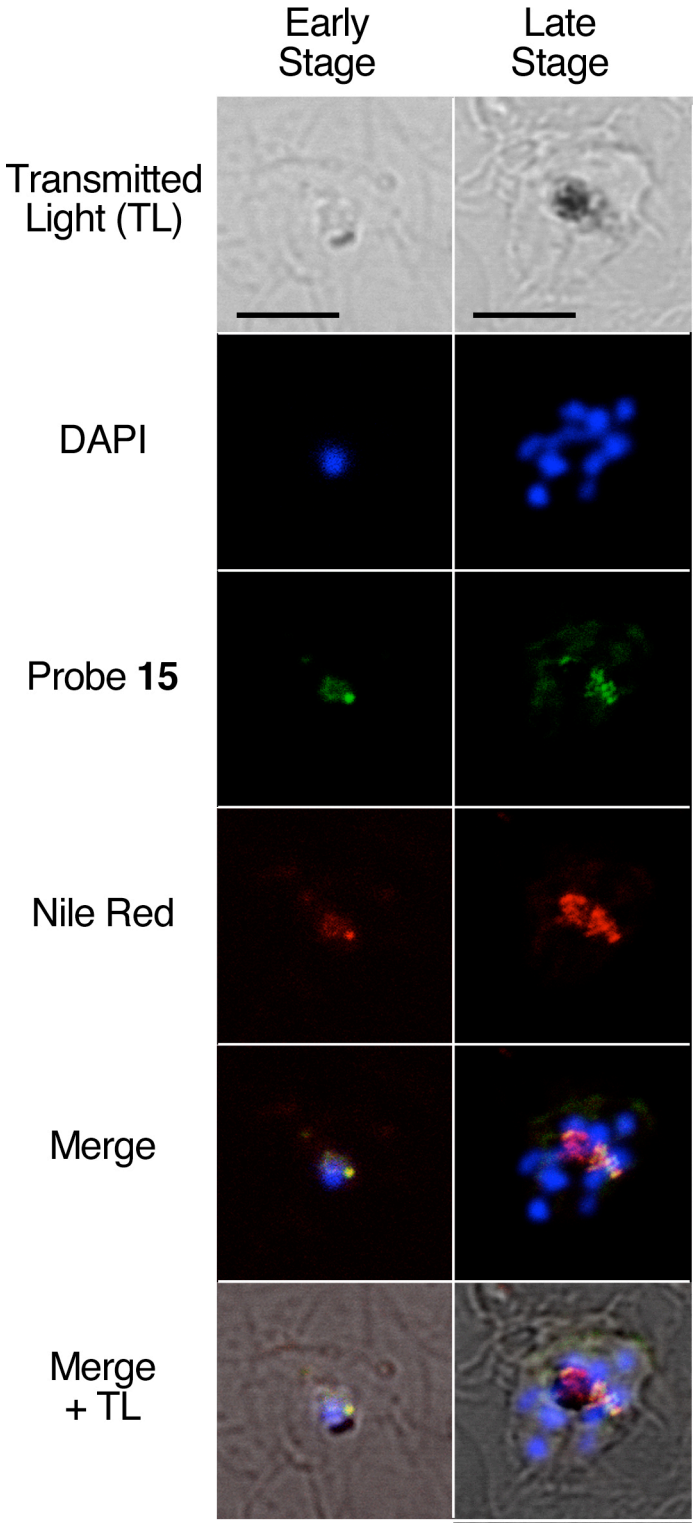


Figure a3.4. Confocal images of fixed parasites incubated with 2 μM probe **15** (green), 30 nM Nile Red (red), and 70 nM DAPI (blue). Co–localization of **15** and Nile Red (neutral lipids) is indicated by the yellow merged image. Interestingly, while probe **15** emission spectra showed a maximum at 450±10 nm (blue channel) in both complete medium and methanol, **15** fluoresced in the green channel after parasite fixation with paraformaldehyde. The latter effect could result from solvatochromic effects while localized within the lipid environment. Bars denote 5 μm.

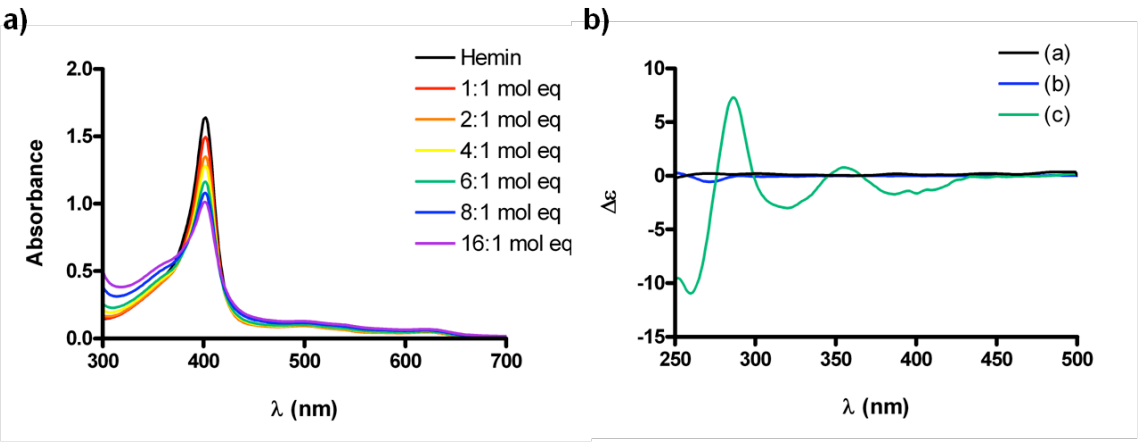


Figure a3.5. Spectroscopic changes observed when Fe(III)PPIX (hemin) is incubated with various concentrations of bromophycolide A (**1**). a) Changes in the Soret band (402 nm) of hemin. The absorbance (Abs) at 402 nm decreases as concentration of **1** increases. Molar equivalent (mol eq) ratios are shown as bromophycolide A (**1**) to hemin. b) CD spectra of hemin and **1**. Hemin is an achiral molecule and does not have a CD signal (black line (a), 12 μM). The chiral **1** gives a weak Cotton effect at 270 nm (blue line (b), 72 μM), and addition of **1** to a solution of hemin results in induced Cotton effects around 400 and 350 nm (green line (c), 12 μM hemin + 72 μM **1**). All spectra have been corrected for dilution by subtracting a reference cell. Conditions: 23 ºC, 40% DMSO, 20 mM HEPES buffer, pH 7.4.

Using a 1:1 complexation model between Fe(III)PPIX and **1** (*23*), the association constant for the Fe(III)PPIX–(**1**) complex was calculated to have a log *K* value of 4.52 ± 0.18, suggesting favorable interactions between these molecules (Supporting Information). This Fe(III)PPIX–(**1**) association constant is comparable to other known heme crystallization inhibitors, such as chloroquine and quinine, whose log *K* values are 5.52 ± 0.03 and 4.10 ± 0.02, respectively (*23*). In addition, circular dichroism (CD) analyses indicated induced Cotton effects at 400 and 350 nm under similar conditions used for UV-vis titration experiments, further supporting a measurable association between heme and **1**, as the achiral Fe(III)PPIX alone does not produce a CD signal (Figure a3.5b).

Potential interactions of **1** with heme include hydroxyl group coordination to the iron center, π-π complexation, and hydrophobic interactions between heme and **1**. The increased potency of **1** (C18-OH) and **7** (C18-OAc) over **8** (C18-OMe) supports a hypothesis of coordination of the phenolic oxygen or adjacent carbonyl with the iron center in heme since modification at this position resulted in significant changes in antimalarial activity, although π-π complexation and hydrophobic interactions between heme and **1** are also possible. Fe–OH coordination of bromophycolides with heme may require that the C18–OAc derivative **7** undergoes deacylation by esterases within the parasite, making **7** a prodrug of the hydrophobic natural product **1**. This is seemingly unsupported by the cellular studies with probe **15** as shown in Figures a3.3 and a3.4, as release of the tags would fail to deliver the observed localization. However, proliferation assays need 72 h of compound–parasite incubation, whereas microscopy studies require 1-2 h of probe exposure.

Because heme crystallization is a non-enzymatic process, parasite-free colorimetric high-throughput screens for heme crystallization inhibitors have been developed (*24*) and refined (*25*). Both **1** and **7** were tested for heme crystallization inhibition compared to the potent antimalarial drug amodiaquine as positive control. Bromophycolide A (**1**) inhibited heme crystallization with an IC_{50} value of 2.5 molar equivalents of drug to heme, and **7** was slightly more effective than **1**, with an IC_{50} of 2.0 molar equivalents, consistent with growth inhibition

assays (Table a3.2); the IC_{50} for amodiaquine (1.2 molar equivalents) was consistent with previously reported values (*24*). Collectively, these data provided compelling evidence of heme crystallization inhibition as a molecular target of **1**.

With one possible mode of action established for **1**, potential protein targets within the malaria parasite were also explored. Taking advantage of the dual utility of probes **15** and **17** as cellular visualization agents and for affinity chromatography, *P. falciparum* lysates were screened for bromophycolide–binding proteins. A selective monoclonal antibody (mAb) against the IAF tag (*26*) was available for immunoprecipitation (IP) with *P. falciparum* lysates incubated with probes **15** and **17**. Repeated IP experiments failed to result in isolation of any bromophycolide–binding proteins. This anti–IAF–mAb affinity method has been successful with IP of protein targets in multiple systems in our labs, from cancerous cells (*26-27*) to micro-invertebrates (*28*). While a negative IP experiment cannot prove the lack of a protein target, these data, along with the observation that probe **15** localized in a low protein environment (i.e., neutral lipid nanospheres), provide a strong suggestion that a mode of action of **1** involves heme crystallization inhibition. This hypothesis is further supported by recent studies which indicate that samples of heme isolated from *P. falciparum* trophozoites, a feeding stage, had minimal protein content but were rich in neutral lipids (*22*).

To date, we have isolated 33 unique *C. serratus* natural products whose IC_{50} values ranged from 0.3 to >100 μM, providing a small library to analyze SAR trends (representative natural products **1–6**, Figure a3.1) (*16-19, 29-30*). Several semi–synthetic derivatives of **1** were also prepared (**7–13**) to further enhance SAR studies and to provide insights for potential future designs of bromophycolide–inspired synthetic compounds. When examining the natural product SAR trends, the most striking observation related to the diterpene head. While replacing the bromine at C15 with an isoprene functionality caused little change in antimalarial activity, substitution of a hydroxyl group at C15 resulted in dramatic loss of activity (Table a3.1), suggesting that a hydrogen bond donating group is poorly suited in this position (Figure a3.2).

Moving to functional groups in the aliphatic macrocycle, synthetic modifications at C10 and C11 also did not significantly affect activity, and replacing the bromine at C10 with chlorine resulted in a slight decrease in activity, although it should be noted that the configuration at C10 was inverted (**4** vs. **10**). Regioisomerization of the double bond in the cyclohexene ring in both natural products and synthetic derivatives also demonstrated only minor alterations in activity (Figure a3.2). Interestingly, elimination of all bromines from **1** only moderately decreased the IC_{50} value from 0.65 to 3.5 μM (**1** vs. **12**) whereas the natural product **6** was completely inactive. The final portion of **1** to investigate was the *p*–hydroxybenzoate group. As previously discussed,

acylation of the phenol in **1** to **7** resulted in a slight increase in activity. Surprisingly, methylation of the phenol in **1** to **8** led to a 30–fold decrease in activity. Collectively, these data suggest that positions C15 and C18 have the most striking effects on antimalarial activity.

The discovery and development of heme crystallization inhibitors remains an excellent malaria drug target because heme crystallization is a physicochemical process and therefore requires the development of independent export mechanisms by the parasite (*31*). Drug resistance most commonly involves mutations in a drug's target protein or proteins which regulate efflux in order to remove the drug (*32*), as the latter was shown for chloroquine resistance. Given that both **1** and **7** effectively inhibited chloroquine-resistant parasites (Dd2) with IC_{50} values of 497 and 304 nM, respectively (Table a3.2), it is clear that the efflux systems in chloroquine-resistant parasites do not recognize bromophycolides. Furthermore, **1** and **7** exhibited a respective 50- and 300-fold increase in activity for *P. falciparum* parasites over human Vero cells (Table a3.2), providing support for these molecules as antimalarial leads.

The molecular structures of the bromophycolides are notably different from quinine, chloroquine, and other heme crystallization inhibitors, thus suggesting a new class of therapeutic leads. Moreover, red algae may represent an excellent source for the discovery of novel antimalarials, as a recent study revealed a common evolutionary lineage for red algae and the malaria parasite (*33*). As

historically evident, natural products continue to provide a rich resource for the discovery of unique antimalarial drugs, and these studies now suggest a new class of marine meroditerpenes as discreet inhibitors of *P. falciparum*.

Methods

General

Commercially available reagents were purchased from Sigma-Aldrich (Milwaukee, WI) or VWR Scientific (Brisbane, CA), and used as received. All reactions were performed under a nitrogen atmosphere unless otherwise noted. Reactions were monitored using thin-layer chromatography (TLC) using Silica Gel 60 F254 plates from EMD Chemicals (San Diego, CA), and visualized with a UV lamp at 254 nm. Solvents used for HPLC and LC-MS were HPLC or Optima grade (Fisher Scientific). NMR solvents were purchased from Cambridge Isotope Laboratories. 2,5-Dimethylfuran was used as an internal standard for quantitative analyses by ^{1}H NMR spectroscopy (*34*).

Structural analyses

NMR spectra were recorded on a Bruker DRX-500 instrument in $CDCl_3$ using a 5 mm inverse detection probe for ^{1}H, COSY, HSQC, and HMBC experiments, and a 5 mm broadband probe for ^{1}H and ^{13}C experiments, and referenced to residual $CHCl_3$ (7.24 and 77.0 ppm, for ^{1}H and ^{13}C, respectively). ^{13}C NMR spectra were not collected for compounds < 1.0 mg due to inadequate sensitivity. In these cases, structures were determined by analysis of ^{1}H NMR and high-resolution mass spectra and comparison to the 33 previously identified natural products for which extensive 1D and 2D NMR spectral data were available, since synthetic

modifications occurred at positions throughout the macrocycle for which well-resolved ^{1}H NMR chemical signals were apparent. High-resolution mass spectra were generated using electrospray ionization with an Applied Biosystems QSTAR-XL hybrid Quadrupole-Time-of-Flight tandem mass spectrometer and Analyst QS software in the mass spectral facilities at Georgia Institute of Technology. LC-MS analyses were conducted using a Waters 2695 HPLC with Waters 2996 photodiode array UV detection and Micromass ZQ 2000 mass spectrometer with electrospray ionization. LC-MS chromatography was achieved with an Alltech Alltima C_{18} (3 µm, 2.1 × 150 mm) column, using gradient mobile phases of aqueous acetonitrile with 0.1% acetic acid.

18-O-acetyl-bromophycolide A (7)

Acetic anhydride (10 µl, 108 µmol) was added dropwise to a solution of bromophycolide A (**1**) (1.0 mg, 1.5 µmol) in anhydrous pyridine (50 µl) and allowed to stir at rt for 4 h. The reaction was concentrated *in vacuo* and the product was purified by reversed-phase HPLC using a gradient of aqueous MeCN (Alltech Alltima C_{18}, 10 × 250 mm, 5 µm) to yield **7** as a colorless oil (0.8 mg, 75%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ: 8.08 (s, 1H); 7.93 (d, 1H, *J* = 10 Hz); 7.16 (d, 1H, *J* = 10 Hz); 4.67 (dd, 1H, *J* = 5, 15 Hz); 4.46 (t, 1H, *J* = 7.5 Hz); 3.48 (dd, 1H, *J* = 5, 15 Hz); 3.41 (d, 1H, *J* = 20 Hz); 3.17 (d, 1H, *J* = 20 Hz); 2.25-2.29 (m, 2H); 2.22 (s, 3H); 1.85-2.18 (m, 7H); 1.81 (s, 3H); 1.8 (s, 3H); 1.68-1.78 (m,

3H); 1.38 (s, 3H); 1.29 (s, 3H); 1.26 (s, 3H); 1.08 (m, 1H). HRESIMS: $[M + H]^+$ *m/z* 705.0467 (calcd for $C_{29}H_{40}Br_3O_5$, 705.0425); [M + Na] *m/z* 727.0170 (calcd for $C_{29}H_{39}Br_3O_5Na$, 727.0245).

18-methoxy-bromophycolide A (8)

A solution of **1** (3.0 mg, 4.5 µmol) in THF (50 µl) was added dropwise to a suspension of sodium hydride (60% dispersion in mineral oil pre-washed with Et_2O, 0.54 mg, 14 µmol) in THF (150 µl) at 0 ºC. The mixture was stirred for 30 min at the same temperature. Methyl iodide (27 µl of a 0.5 M solution in THF) was added dropwise to the solution at 0 ºC and the resulting mixture was stirred for 7 h at rt. Saturated NH_4Cl (aq) was added to the solution at 0 ºC, followed by extraction with EtOAc (3 × 300 µl). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo*. Purification of the product by reversed-phase HPLC using a gradient of aqueous MeOH (Zorbax C_{18}, 9.4 × 250 mm, 5 µm) yielded **8** as a colorless oil (2.5 mg, 81%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.95 (s, 1H); 7.92 (d, 1H, *J* = 10 Hz); 6.89 (d, 1H, *J* = 10 Hz); 4.65 (dd, 1H, *J* = 5, 10 Hz); 4.48 (dd, 1H, *J* = 5, 12 Hz); 3.93 (s, 3H); 3.44 (m, 1H); 3.42 (d, 1H, *J* = 18 Hz); 3.21 (d, 1H, *J* = 18 Hz); 2.22-2.33 (m, 3H); 1.84-2.17 (m, 7H); 1.82 (s, 3H); 1.81 (s, 3H); 1.68-1.80 (m, 3H); 1.38 (s, 3H); 1.30 (s, 3H); 1.27 (s, 3H); 1.07 (m, 1H). ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 165.9, 161.4, 133.1, 130.6, 130.4, 130.0, 127.5, 121.0, 109.7, 80.4, 73.9, 72.4, 67.3, 61.1,

55.6, 43.5, 37.6, 34.3, 33.7, 33.6, 31.6, 31.0, 30.5, 29.5, 28.6, 28.3, 26.1, 21.0. HRESIMS: $[M + H]^+$ *m/z* 677.0404 (calcd for $C_{28}H_{40}Br_3O_4$, 677.0477).

11, 18-dimethoxy-bromophycolide A (9)

A solution of **8** (2.0 mg, 3.0 µmol) in THF (25 µl) was added dropwise to a suspension of sodium hydroxide (60% dispersion in mineral oil pre-washed with Et_2O, 0.48 mg, 12 µmol) in THF (50 µl) at 0 ºC. The mixture was stirred for 30 min and allowed to reach rt. Methyl iodide (25 µl of a 0.5 M solution in THF) was added dropwise to the solution at rt and the resulting mixture was stirred overnight. Saturated NH_4Cl (aq) was added to the solution, followed by extraction with EtOAc (3 × 200 µl). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo*. Purification of the product by reversed-phase HPLC using a gradient of aqueous MeOH (Zorbax C_{18}, 9.4 × 250 mm, 5 µm) yielded **9** as a colorless oil (1.0 mg, 49%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.94 (d, 1H, *J* = 10 Hz); 7.88 (s, 1H); 6.89 (d, 1H, *J* = 10 Hz); 4.96 (dd, 1H, *J* = 7, 12 Hz); 4.50 (m, 1H); 3.93 (s, 3H); 3.62 (s, 3H); 3.53 (d, 1H, *J* = 18 Hz); 3.34 (d, 1H, *J* = 18 Hz); 3.07 (dd, 1H, *J* = 5, 10 Hz); 2.47-2.55 (m, 1H); 2.29 (m, 2H); 2.12-2.19 (m, 1H); 1.83-1.94 (m, 5H); 1.74 (s, 3H); 1.72 (s, 3H); 1.51 (s, 3H); 1.32-1.42 (m, 3H); 1.16 (s, 3H); 0.89 (s, 3H). ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 166.3, 161.6, 133.3, 131.8, 131.0, 130.1, 128.1, 120.6, 109.5, 78.6, 68.7, 65.2, 63.8, 59.9, 55.5, 43.1, 39.9, 31.8, 30.9, 30.4, 29.3, 29.0, 28.1, 24.6, 24.5, 23.9,

20.7, 19.5, 19.4. HRESIMS: $[M + H]^+$ *m/z* 691.0584 (calcd for $C_{29}H_{42}Br_3O_4$, 691.0633).

10-(S)-Cl-bromophycolide E (10)

Lithium chloride (8 µl, 8.2 µmol, of a 1M solution in warmed DMF) was added dropwise to a solution of bromophycolide E (**4**) (1.6 mg, 2.7 µmol) in DMF (100 µl) at rt. The reaction was allowed to stir at 70 ºC for 6 h. The reaction was allowed to cool and deionized H_2O was added (0.5 ml) to the solution, followed by extraction with EtOAc (3 × 0.5 ml). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo* and the product was purified by reversed-phase HPLC using a gradient of aqueous MeOH (Zorbax C_{18}, 9.4 × 250 mm, 5 µm) to yield **10** as a white powder (0.13 mg, 8.8%). ^1H NMR ($CDCl_3$, 500 MHz) δ: 7.83 (s, 1H); 7.68 (d, 1H, *J* = 10 Hz); 6.74 (d, 1H, *J* = 10 Hz); 5.38 (s, 1H); 5.26 (br s –OH); 5.14 (m, 1H); 4.95 (s, 1H); 4.92 (s, 1H); 4.85 (s, 1H); 4.34 (dd, 1H, *J* = 5, 12 Hz); 4.03 (d, 1H, *J* = 12 Hz); 3.12 (dd, 1H, *J* = 15, 20 Hz); 2.57 (d, 1H, *J* = 20 Hz); 2.47 (m, 1H); 2.29-2.38 (m, 3H); 1.99-2.10 (m, 5H); 1.91 (m, 2H); 1.79 (s, 3H); 1.66-1.75 (m, 3H); 1.37 (s, 3H); 0.96 (s, 3H). HRESIMS $[M – H]^-$ *m/z* 537.1283 (calcd for $C_{27}H_{35}BrClO_4$, 537.1407).

$\Delta^{9,10}$- $\Delta^{21,22}$-Bromophycolide A (11).

Lithium chloride (3.8 mg, 90 µmol) was added to a solution of **1** (10 mg, 15 µmol) in DMF (1 ml) at rt. The reaction was allowed to stir at 70 ºC for 6 h. The reaction

was allowed to cool and deionized H_2O was added (5 ml) to the solution, followed by extraction with EtOAc (3 × 5 ml). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo* and the product was purified by reversed-phase HPLC using a gradient of aqueous MeOH (Zorbax C_{18}, 9.4 × 250 mm, 5 µm) to yield **11** as a white powder (0.1 mg, 1.3%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ: 8.37 (s, 1H); 7.68 (d, 1H, *J* = 10 Hz); 6.74 (d, 1H, *J* = 10 Hz); 5.42 (m, 1H); 5.41 (m, 1H); 5.11 (br s, –OH); 4.63 (m, 2H); 3.81 (dd, 1H, *J* = 8, 17 Hz); 3.65 (d, 1H, *J* = 20 Hz); 3.21 (d, 1H, *J* = 20 Hz); 2.51 (m, 2H); 1.92-2.10 (m, 5H); 1.74-1.82 (m, 1H); 1.45 (s, 3H); 1.39 (s, 3H); 1.37 (s, 3H); 1.32 (s, 3H); 1.28 (s, 3H); 1.20 (m, 1H). HRESIMS $[M - H]^-$ *m/z* 501.1442 (calcd for $C_{27}H_{34}BrO_4$, 501.1640).

$\Delta^{9,10}$- $\Delta^{21,22}$-Bromophycolide M (12)

Lithium chloride (3.8 mg, 90 µmol) was added to a solution of **1** (10 mg, 15 µmol) in DMF (1 ml) at rt. The reaction was allowed to stir at 70 ºC for 6 h. The reaction was allowed to cool and deionized H_2O (5 ml) was added to the solution, followed by extraction with EtOAc (3 × 5 ml). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo* and the product was purified by reversed-phase HPLC using a gradient of aqueous MeOH (Zorbax C_{18}, 9.4 × 250 mm, 5 µm) to yield **12** as a white power (0.15 mg, 2.4%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ:7.94 (s, 1H); 7.75 (d, 1H, *J* = 10 Hz); 6.75 (d, 1H, *J* = 10 Hz);

5.70 (br d, 1H, *J* = 10 Hz); 5.33 (br d, 1H, *J* = 12 Hz); 5.14 (br s,–OH); 5.10 (br d, 1H, *J* = 10 Hz); 5.00 (s, 1H); 4.89 (s, 1H); 4.61 (m, 1H); 3.62 (d, 1H, *J* = 18 Hz); 3.20 (d, 1H, *J* = 18 Hz); 2.64-2.78 (m, 2H); 2.59 (m, 1H); 1.78 (s, 3H); 1.63-1.72 (m, 4H); 1.54 (s, 3H); 1.45 (s, 3H); 1.30-1.37 (m, 2H); 1.24 (m, 1H); 1.13 (s, 3H). HRESIMS: $[M - H]^-$ *m/z* 421.2430 (calcd for $C_{27}H_{33}O_4$, 421.2378).

10-(S)-Cl-$\Delta^{21,22}$-bromophycolide M (13)

Lithium chloride (3.8 mg, 90 µmol) was added to a solution of **1** (10 mg, 15 µmol) in DMF (1 ml) at rt. The reaction was allowed to stir at 70 ºC for 6 h. The reaction was allowed to cool and deionized H_2O was added (5 ml) to the solution, followed by extraction with EtOAc (3 × 5 ml). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo* and the product was purified by reversed-phase HPLC using a gradient of aqueous MeOH (Zorbax C_{18}, 9.4 × 250 mm, 5 µm) to yield **13** as a white powder (2.0 mg, 29%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ: 8.01 (s, 1H); 7.79 (d, 1H, *J* = 8 Hz); 6.79 (d, 1H, *J* = 8 Hz); 5.77 (m, 1H); 5.46 (d 1H, *J* = 10 Hz); 5.39 (br s, –OH); 5.30 (m, 1H); 5.00 (s, 1H); 4.88 (s, 1H); 3.85 (d, 1H, *J* = 10 Hz); 3.49 (d, 1H, *J* = 17 Hz); 3.35 (d, 1H, *J* = 17 Hz); 2.75 (d, 1H, *J* = 22 Hz); 2.64 (d, 1H, *J* = 22 Hz); 1.90-2.02 (m, 3H); 1.77 (s, 3H); 1.63-1.75 (m, 5H); 1.51 (s, 3H); 1.23 (s, 3H); 1.16 (s, 3H); 0.85 (m, 1H). ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 165.5, 157.5, 143.1, 135.3, 131.9, 130.8, 129.8, 129.3, 125.4, 122.6, 122.4, 115.0, 111.1, 76.2, 74.6, 72.2, 40.7, 37.4, 33.7, 31.9,

30.2, 28.7, 28.3, 27.6, 26.6, 20.5, 18.9. HRESIMS $[M - H]^-$ *m/z* 457.1126 (calcd for $C_{27}H_{34}ClO_4$, 457.2145).

Bromophycolide A 'short-linker' probe (15)

N'-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide (EDC) (11 µl, 60 µmol) was added to a solution of 2,5-dioxopyrrolidin-1-yl-2-(2-(7-(dimethylamino)-2-oxo-2H-chromen-4-yl)acetamido)acetate (**14**) (9.1 mg, 30 µmol), 4-dimethylaminopyridine (DMAP) (6 µl, 6 µmol, of a 1M solution in DMF), and **1** (5 mg, 7.5 µmol) in 3:2 DMF/CH_2Cl_2 (0.3 ml), followed by slow addition of *N,N*-diisopropylethylamine (EtN^iPr_2, 20 µl). The reaction was allowed to stir at rt for 24 h and then was quenched with saturated $NaHCO_3$ solution (0.5 ml). The solution was extracted with EtOAc (3 × 0.75 ml). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo* and the product was purified by reversed-phase HPLC using a gradient of aqueous MeCN and 0.2% TFA (Alltech Alltima C_{18}, 10 × 250 mm, 5 µm) to yield **15** (1.6 mg, 22%). ^1H and ^{13}C NMR data table has been provided in Supporting Information. HRESIMS: $[M + H]^+$ *m/z* 949.1286 (calcd for $C_{42}H_{52}Br_3N_2O_8$, 949.1274).

IAF tag 16

A three-step process was used to prepare tag **16**. HATU (570 mg, 1.5 mmol) was added to a solution of 7-dimethylamino-4-coumarin acetic acid (247.2 mg, 1

mmol) and EtN^iPr_2 (0.42 ml, 3 mmol) in anhydrous DMF (3 mL). After 1 h at rt, *N*-Boc-1,4-diaminobutane (225.3 mg, 1.2 mmol) in anhydrous DMF (1 ml) was added and the mixture was stirred for 4 h at rt. The contents were dried via rotary evaporation. The crude product was suspended in DCM (10 ml) and applied to a Dry Column Vacuum Chromatography (DCVC) column containing 80 g of Silica Gel (15–40 μm Silica Gel 60 from EM Sciences) and eluted with 5% MeOH in EtOAc. Drying via rotary evaporation provided the Boc-protected amide (395.2 mg) along with approximately 10% of 7-dimethylamino-4-methylcoumarin (a product of the decarboxylation of 7-dimethylamino-4-coumarin acetic acid). This mixture was then dissolved in dry CH_2Cl_2 (10 ml) using ultrasonification and treated with TFA (2 ml) at rt. TLC analysis showed clean removal of the Boc protecting group within 4 h. The resulting mixture was dried by the flow of Argon and the residue was rotary evaporated twice from EtOAc (20 ml) and toluene (10 ml) and then dissolved in anhydrous DMF (5 ml) containing EtN^iPr_2 (0.84 mL, 6 mmol). Solid succinic anhydride (250.1 mg, 2.5 mmol) was added and the mixture was stirred at rt. After 12 h, the mixture was dried via rotary evaporation and **16** was obtained by silica flash chromatography using a gradient consisting of 2:1 hexanes:EtOAc, 1:1 hexanes:EtOAc, EtOAc, 1:10 MeOH:EtOAc and 1:5 MeOH:EtOAc. Tag **16** (317.1 mg, 76%) was obtained as a light yellow powder after evaporation of selected fractions. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 12.05 (br s, –OH); 8.18 (t, –NH, J = 5.3); 7.81 (t, –NH, J = 5.5); 7.52 (d, 1H, J = 9.0);

6.72 (dd, 1H, J = 2.3, 9.0); 6.54 (d, 1H, J = 2.3); 5.98 (s, 1H); 3.57 (s, 2H); 3.05 (m, 4H); 3.00 (s, 6H); 2.40 (t, 2H, J = 7.1); 2.27 (t, 2H, J = 7.0); 1.37 (m, 4H). ^{13}C NMR (DMSO-d_6, 100 MHz) δ: 173.9, 170.7, 167.6, 160.7, 155.4, 152.8, 151.4, 126.0, 109.4, 109.0, 108.2, 97.5, 38.5, 38.2, 30.0, 29.2, 26.6, 26.4. ESIMS $[M + H - H_2O]^+$ *m/z* 400.43; $[M + H]^+$ *m/z* 418.43; [M + Na] *m/z* 440.41; $[M - H]^-$ *m/z* 416.50.

Bromophycolide A 'long-linker' probe (17)

EDC (15 µl, 81 µmol) was added to a solution of 4-(4-(2-(7-(dimethylamino)-2-oxo-2H-chromen-4-yl)acetamido)butylamino)-4-oxobutanoic acid (**16**) (11 mg, 27 µmol), DMAP (10 µl, 10 µmol, of a 1M solution in DMF), and **1** (4.5 mg, 6.8 µmol) in 3:2 DMF/CH_2Cl_2 (0.3 ml), followed by slow addition of EtN^iPr_2 (20 µl). The reaction was allowed to stir at rt for 24 h and then was quenched with saturated $NaHCO_3$ solution (0.5 ml). The solution was extracted with EtOAc (3 × 0.75 ml). The combined organic layers were washed with brine, dried over $MgSO_4$, and evaporated *in vacuo* and the product was purified by reversed-phase HPLC using a gradient of aqueous MeCN and 0.2% TFA (Alltech Alltima C_{18}, 10 × 250 mm, 5 µm) to yield **17** (2.1 mg, 29%). ^{1}H and ^{13}C NMR data table has been provided in Supporting Information. HRESIMS: $[M + H]^+$ *m/z* 1062.2021 (calcd for $C_{48}H_{63}Br_3N_3O_9$, 1062.2114).

Parasite culture

3D7, HB3 and Dd2 *P. falciparum* malaria parasites (MRA-102, 155, 156, MR4, ATCC Manassas, Virginia) were cultured in human type O+ erythrocytes in complete medium (RPMI 1640, 0.043 mg/ml gentamicin (Gibco), 0.014 mg/ml hypoxanthine (Acros), 38.5 mM HEPES (Sigma), 0.18% sodium bicarbonate (Gibco), 0.20% glucose (MP Biomedical), 2.57 mM NaOH (Sigma), 0.21% albumax (Gibco), and 4.3% human serum as previously described (*35*). Cultures were maintained in 25 cm^2 cell culture treated flasks (Corning) at a volume of 10 ml and were gassed for 30 s with an environment of 3% CO_2, 1% O_2, and 96% N_2, then incubated at 37 °C. Synchronization of culture was achieved through sorbitol lysis of mature stage using 5% sorbitol (Fisher) and fine tuned by another lysis 8 h later.

Growth inhibition assay with Plasmodium falciparum

Antimalarial activity was determined with a SYBR Green based parasite proliferation assay, adapted from Smilkstein (*36*) and Bennett (*37*). *P. falciparum* parasites were cultured in human O+ erythrocytes as described in the previous section. Compounds were diluted in complete medium and 40 µl transferred to 96-well assay plates. To this, 80 µl of complete media with malaria infected erythrocytes were dispensed in each well, with final assay percentages at 2.5% hematocrit and 0.5% parasitemia. Uninfected erythrocytes were added to the background wells at the same final hematocrit percentage. Plates were incubated

for 72 h in a low oxygen environment (96% N_2, 3% CO_2, 1% O_2) in a modular incubation chamber. The plates were sealed and placed in a -80 ºC freezer overnight then thawed, and 120 μl of lysis buffer (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.08% Triton X-100, 0.008% saponin with 0.2 μl/ml Sybr Green I (Invitrogen)) was dispensed into each well and incubated at 37 ºC in the dark for 6 h. Fluorescence of each well was read with a Molecular Devices SpectraMAX Gemini EM at excitation: 495 nm, emission: 525 nm with 515 nm cut-off.

Fluorescence microscopy – live cell confocal imaging.

Samples were prepared in optical bottom 96-well assay plates (Costar #3614, Corning, NY). Parasite cultures were diluted to 0.025% hematocrit and 6% parasitemia in complete medium and 240 μl was added to each well. Probe **15** was incubated with samples for 1 h at 37°C, in the dark at a final concentration of 2 μM. The BD Pathway HT (BD Biosciences Bioimaging, Rockville, MD) incubation chamber was set at 37 °C and 5% CO_2.

Confocal microscopy with fixed cells

Mixed-stage or synchronized *P. falciparum* cultures at 6-8% parasitemia were incubated either with 2 μM probe **15**, 2 μM control probe **14a**, or 2 μM control probe **14b** for 1 h at 37 ºC. The following conditions were used for the sub-cellular labeling of parasites: 50 nM Lysosensor Green 189 for 1 min; 70 nM

DAPI for 10 min; 1 μM ER-Tracker Green for 30 min, 500 nM MitoTracker Red 580 for 1 h, 500 nM BODIPY 505/515 for 30 min or 30 nM Nile Red for 10 min. All stains were diluted fresh into complete medium and incubated at 37 ºC. Following incubation, infected red blood cells were washed in complete medium and 7 μl of the resulting thick pellet was added onto a glass slide and allowed to dry in the dark. After drying, samples were fixed using 4% paraformaldehyde for 20 minutes at room temperature before mounting. Images were obtained using a Leica TCS SP2 confocal microscope (Leica Microsystems Inc., Bannockburn, IL).

Heme-bromophycolide A UV-vis and CD spectroscopic analyses.

UV-vis spectroscopic analyses were performed similar to those reported by Egan and Ncokazi (*38*). A stock solution of 0.0012 M porcine hemin (Sigma) was prepared using molecular grade DMSO and stored in the dark. Working Fe(III)PPIX solutions (12 μM) were prepared fresh by mixing 20 μl of stock hemin solution with 0.8 ml DMSO and 1.2 ml of HEPES buffer (20 mM, pH 7.4) in a 10 mm quartz cuvette. A stock solution (5 mM) of bromophycolide A (**1**) was prepared in DMSO and was used to titrate the natural product into the cuvette. The titrations involved addition of four 5 μl aliquots, followed by four 10 μl aliquots and two 20 μl aliquots. The final two additions did not cause any more spectral changes. For circular dichroism experiments, both Fe(III)PPIX (12 μM) and **1** (72 μM) were recorded individually in 40% DMSO, 20 mM HEPES, pH 7.4

in 10 mm quartz cuvettes. Bromophycolide A (**1**) was titrated into a 10 mm quartz cuvette containing Fe(III)PPIX (12 μM) so that the molar equivalents of **1** to Fe(III)PPIX was 6:1. All spectra were corrected for dilution by subtracting a reference cell. UV spectra were recorded on a Jasco V-630 double-beam spectrophotometer. Circular dichroism measurements were acquired on a Jasco J-810 grating spectropolarimeter.

Heme crystallization inhibition assay.

Heme crystallization inhibition assays were performed as previously described (*24-25*). Briefly, 30 μl of a 0.3 mM solution of hemin in 0.1 M NaOH was added to rows B-H in a round-bottom 96-well microtiter plate, and 50 μl of the 0.3 mM hemin solution was added to row A. Amodiaquine and DMSO were used as positive and negative controls, respectively. Controls, **1**, or **7** were added to wells in row A in 10 μl DMSO (starting concentration of either 6, 5, or 4 mol eq relative to hemin) and serial diluted by half down each column, discarding 30 μl from the final row. Next, 10 μl of a 9.8 M sodium acetate solution (pH 4.8) was added to each well, and the plate was incubated for 2 h at 60 ºC. After cooling for 15 min, 75 μl of a 14% (v/v) pyridine solution in 20 mM HEPES (pH 7.5) was added to each well and the plate was centrifuged at 2000 rpm for 1 min. Next, 50 μl of each supernatant was transferred into a flat-bottom 96-well plate and the absorbance was measured at 405 nm using a microplate reader. The IC_{50} of

each compound was calculated using the dose concentration at 50% inhibition on a sigmoidal dose response curve generated using GraphPad Prism version 4.00.

Immunoprecipitation with P. falciparum cell lysates

A 0.15% saponin solution was used to lyse erythrocytes and isolate parasites, and samples were washed twice with PBS. Parasite pellet was incubated on ice for 5 minutes in cytoplasmic lysis buffer (20 mM HEPES pH 7.9, 10 mM KCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 0.5 mM PMSF, 0.65% Igepal, 1x Roche cocktail inhibitor tablet). Samples were centrifuged at 1500 g for 10 min at 4 ºC, supernatant was collected, and the resulting pellet was re-suspended with nucleus lysis buffer (20 mM HEPES pH7.9, 0.1 M NaCl, 0.1 mM EDTA, 0.1 mM EGTA, 1.5 mM $MgCl_2$, 1 mM DTT, 25% glycerol, 1 mM PMSF, 1x Roche cocktail inhibitor tablet) and incubated for 20 min at 4 ºC. After centrifugation at 6000 g for 10 min at 4 ºC, the supernatant was collected and combined with cytoplasmic proteins. The remaining pellet was sonicated in cytoplasmic lysis buffer, centrifuged, and then all supernatants were combined. The total protein concentration was determined to be 5 mg/ml total protein by Bradford analysis. Probe **15** or **17** (2 µM or 20 µM) was added to lysates for 1 or 4 h and incubated at 4 ºC with constant rotation. Next, 50 µl of Affigel 10 resin containing 3.5 mg/ml of the XRI-TF35 mAb was added to the probe/lysate mixture for an additional 2 h or 14 h with agitation at 4 ºC. As negative controls, tag **14a** (2 µM), **14b** (2 µM),

or 0.7% DMSO was added to lysates. Samples were then washed with wash buffer (0.5% Tritonx-100, 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM EDTA, 2 mM PMSF, 1x Roche cocktail inhibitor tablet) three times. Proteins were either eluted from the resin by incubation with 35 μl of 1 mg/ml of **14a** in RIPA buffer for 1 h at rt, or laemmli buffer was added directly to samples and boiled for 5 min at 100 ºC to elute proteins. Samples were resolved by 16% SDS-PAGE and detected with Coomassie stain or Silver stain. SDS-PAGE analyses of varying immunoprecipitation conditions with either probe **15**, **17**, or negative controls were identical as noted by a complete lack of protein.

Acknowledgements

This work was supported by the U.S. National Institutes of Health's International Cooperative Biodiversity Groups program (Grant U01 TW007401). The authors thank MR4 for providing malaria parasites contributed by D. Carucci, A. Craig and T. Wellems. We especially thank T.F. Molinski for thorough discussions and insights and for use of his circular dichorism spectropolarimeter. We thank D. Hostetler, F. Fernandez, M.C. Sullards, and D. Bostwick for mass spectroscopic analyses; L. Gelbaum for NMR assistance; and E. Zhao and B. Hammer for use of their microplate reader.

Conclusion

The need for novel antimalarial agents remains a high priority in the malaria field. Bromophycolide A was originally extracted from the red macroalga *Callophycus serratus,* and provides a unique class of meroditerpenes that inhibits parasite growth. Bromophycolide A was tagged with the fluorescent probe, coumarin, and we were able to localize this compound with hemezoin. Immunoprecipitations of the tagged bromophycolide A were unsuccessful at identifying an interacting parasite protein, but this compound did inhibit heme crystallization *in vitro*. Bromophycolide A is structurally different from quinine, and its derivates, that inhibit heme crystallization, and chloroquine resistant and sensitive parasites exhibit the same inhibitory concentration at 50 percent. These results lead to testing bromophycolide A in a mouse model, which had some promising results albeit with solubility and metabolic issues. This study presents an interesting outlook for antimalarial drug discovery because most natural products tested for antimalarial properties are from terrestrial organisms. We present a new avenue for investigating extracts from the ocean for novel compounds with antimalarial properties.

References

1. Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y., and Hay, S. I. (2005) The global distribution of clinical episodes of Plasmodium falciparum malaria, *Nature 434*, 214-217.

2. Newman, D. J., and Cragg, G. M. (2007) Natural products as sources of new drugs over the last 25 years, *J. Nat. Prod. 70*, 461-477.

3. Hyde, J. E. (2005) Drug-resistant malaria, *Trends Parasitol. 21*, 494-498.

4. Klayman, D. L., Lin, A. J., Acton, N., Scovill, J. P., Hoch, J. M., Milhous, W. K., and Theoharides, A. D. (1984) Isolation of artemisinin (qinghaosu) from Artemisia annua grouping in the United States, *J. Nat. Prod. 47*, 715-717.

5. Klayman, D. L. (1985) Qinghaosu (artemisinin) - an antimalarial drug from China, *Science 228*, 1049-1055.

6. Noedl, H., Se, Y., Schaecher, K., Smith, B. L., Socheat, D., Fukuda, M. M., and Consortium, A. R. C. S. (2008) Evidence of Artemisinin-Resistant Malaria in Western Cambodia, *N. Engl. J. Med. 359*, 2619-2620.

7. Dondorp, A. M., Nosten, F., Yi, P., Das, D., Phyo, A. P., Tarning, J., Lwin, K. M., Ariey, F., Hanpithakpong, W., Lee, S. J., Ringwald, P., Silamut, K., Imwong, M., Chotivanich, K., Lim, P., Herdman, T., An, S. S., Yeung, S., Singhasivanon, P., Day, N. P. J., Lindegardh, N., Socheat, D., and White, N. J. (2009) Artemisinin Resistance in Plasmodium falciparum Malaria, *N. Engl. J. Med. 361*, 455-467.

8. Goldberg, D. E., Slater, A. F. G., Cerami, A., and Henderson, G. B. (1990) Hemoglobin degradation in the malaria parasite Plasmodium falciparum - an ordered process in a unique organelle, *Proc. Natl. Acad. Sci. U. S. A. 87*, 2931-2935.

9. Rosenthal, P. J., and Meshnick, S. R. (1996) Hemoglobin catabolism and iron utilization by malaria parasites, *Mol. Biochem. Parasitol. 83*, 131-139.

10. Chou, A. C., Chevli, R., and Fitch, C. D. (1980) Ferriprotoporphyrin-IX fulfulls the criteria for identification as the chloroquine receptor of malaria parasites, *Biochemistry 19*, 1543-1549.

11. Fidock, D. A., Nomura, T., Talley, A. K., Cooper, R. A., Dzekunov, S. M., Ferdig, M. T., Ursos, L. M. B., Sidhu, A. B. S., Naude, B., Deitsch, K. W., Su, X. Z., Wootton, J. C., Roepe, P. D., and Wellems, T. E. (2000) Mutations in the P-falciparum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance, *Mol. Cell 6*, 861-871.

12. Robert, A., and Meunier, B. (1998) Is alkylation the main mechanism of action of the antimalarial drug artemisinin?, *Chem. Soc. Rev. 27*, 273-279.

13. Hong, Y. L., Yang, Y. Z., and Meshnick, S. R. (1994) The interaction of artemisinin with malarial hemozoin *Mol. Biochem. Parasitol. 63*, 121-128.

14. Eckstein-Ludwig, U., Webb, R. J., van Goethem, I. D. A., East, J. M., Lee, A. G., Kimura, M., O'Neill, P. M., Bray, P. G., Ward, S. A., and Krishna, S. (2003) Artemisinins target the SERCA of Plasmodium falciparum, *Nature 424*, 957-961.

15. Li, W., Mo, W. K., Shen, D., Sun, L. B., Wang, J., Lu, S., Gitschier, J. M., and Zhou, B. (2005) Yeast model uncovers dual roles of mitochondria in the action of arternisinin, *PLoS Genet. 1*, 329-334.

16. Kubanek, J., Prusak, A. C., Snell, T. W., Giese, R. A., Hardcastle, K. I., Fairchild, C. R., Aalbersberg, W., Raventos-Suarez, C., and Hay, M. E. (2005) Antineoplastic diterpene-benzoate macrolides from the Fijian red alga Callophycus serratus, *Org. Lett. 7*, 5261-5264.

17. Kubanek, J., Prusak, A. C., Snell, T. W., Giese, R. A., Fairchild, C. R., Aalbersberg, W., and Hay, M. E. (2006) Bromophycolides C-I from the Fijian red alga Callophycus serratus, *J. Nat. Prod. 69*, 731-735.

18. Lane, A. L., Stout, E. P., Lin, A. S., Prudhomme, J., Le Roch, K., Fairchild, C. R., Franzblau, S. G., Hay, M. E., Aalbersberg, W., and Kubanek, J. (2009) Antimalarial Bromophycolides J-Q from the Fijian Red Alga Callophycus serratus, *J. Org. Chem. 74*, 2736-2742.

19. Lin, A. S., Stout, E. P., Prudhomme, J., Le Roch, K., Fairchild, C. R., Franzblau, S. G., Aalbersberg, W., Hay, M. E., and Kubanek, J. (2010) Bioactive Bromophycolides R-U from the Fijian Red Alga Callophycus serratus, *J. Nat. Prod. 73*, 275-278.

20. Cervantes, S., Stout, E. P., Prudhomme, J., Engel, S., Bruton, M., Cervantes, M., Carter, D., Tae-Chang, Y., Hay, M. E., Aalbersberg, W., Kubanek, J., and

Le Roch, K. (2010) High content live cell imaging evaluation for the discovery of new antimalarial natural products, *Submitted to Antimicrobial Agents and Chemotherapy*

21. Alexander, M. D., Burkart, M. D., Leonard, M. S., Portonovo, P., Liang, B., Ding, X. B., Joullie, M. M., Gulledge, B. M., Aggen, J. B., Chamberlin, A. R., Sandler, J., Fenical, W., Cui, J., Gharpure, S. J., Polosukhin, A., Zhang, H. R., Evans, P. A., Richardson, A. D., Harper, M. K., Ireland, C. M., Vong, B. G., Brady, T. P., Theodorakis, E. A., and La Clair, J. J. (2006) A central strategy for converting natural products into fluorescent probes, *ChemBioChem 7*, 409-416.

22. Pisciotta, J. M., Coppens, I., Tripathi, A. K., Scholl, P. F., Shuman, J., Bajad, S., Shulaev, V., and Sullivan, D. J. (2007) The role of neutral lipid nanospheres in Plasmodium falciparum haem crystallization, *Biochem. J. 402*, 197-204.

23. Egan, T. J., Mavuso, W. W., Ross, D. C., and Marques, H. M. (1997) Thermodynamic factors controlling the interaction of quinoline antimalarial drugs with ferriprotoporphyrin IX, *J. Inorg. Biochem. 68*, 137-145.

24. Ncokazi, K. K., and Egan, T. J. (2005) A colorimetric high-throughput beta-hematin inhibition screening assay for use in the search for antimalarial compounds, *Anal. Biochem. 338*, 306-319.

25. Rush, M. A., Baniecki, M. L., Mazitschek, R., Cortese, J. F., Wiegand, R., Clardy, J., and Wirth, D. F. (2009) Colorimetric High-Throughput Screen for Detection of Heme Crystallization Inhibitors, *Antimicrob. Agents Chemother. 53*, 2564-2568.

26. Hughes, C. C., MacMillan, J. B., Gaudencio, S. P., Fenical, W., and La Clair, J. J. (2009) Ammosamides A and B Target Myosin, *Angewandte Chemie-International Edition 48*, 728-732.

27. Hughes, C. C., Yang, Y. L., Liu, W. T., Dorrestein, P. C., La Clair, J. J., and Fenical, W. (2009) Marinopyrrole A Target Elucidation by Acyl Dye Transfer, *J. Am. Chem. Soc. 131*, 12094-12096.

28. Stout, E. P., La Clair, J. J., Snell, T. W., Shearer, T. L., and Kubanek, J. (2010) Conservation of progesterone hormone function in invertebrate reproduction, *Proc. Natl. Acad. Sci. U. S. A. 107*, 11859-11864.

29. Lane, A. L., Stout, E. P., Hay, M. E., Prusak, A. C., Hardcastle, K., Fairchild, C. R., Franzblau, S. G., Le Roch, K., Prudhomme, J., Aalbersberg, W., and Kubanek, J. (2007) Callophycoic acids and callophycols from the Fijian red alga Callophycus serratus, *J. Org. Chem. 72*, 7343-7351.

30. Stout, E. P., Prudhomme, J., Le Roch, K., Fairchild, C. R., Franzblau, S. G., Aalbersberg, W., Hay, M. E., and Kubanek, J. (2010) Unusual antimalarial meroditerpenes from tropical red macroalgae, *Bioorg. Med. Chem. Lett. 20*, 5662-5665.

31. Dorn, A., Stoffel, R., Matile, H., Bubendorf, A., and Ridley, R. G. (1995) Malarial haemozoin beta-hematin supports heme polymerization in the absence of protein, *Nature 374*, 269-271.

32. Srivastava, I. K., Morrisey, J. M., Darrouzet, E., Daldal, F., and Vaidya, A. B. (1999) Resistance mutations reveal the atovaquone-binding domain of cytochrome b in malaria parasites, *Mol. Microbiol. 33*, 704-711.

33. Janouskovec, J., Horak, A., Obornik, M., Lukes, J., and Keeling, P. J. (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids, *Proc. Natl. Acad. Sci. U. S. A. 107*, 10949-10954.

34. Gerritz, S. W., and Sefler, A. M. (2000) 2,5-dimethylfuran (DMFu): An internal standard for the "traceless" quantitation of unknown samples via H-1 NMR, *J. Comb. Chem. 2*, 39-41.

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

36. Smilkstein, M., Sriwilaijaroen, N., Kelly, J. X., Wilairat, P., and Riscoe, M. (2004) Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening, *Antimicrob. Agents Chemother. 48*, 1803-1806.

37. Bennett, T. N., Paguio, M., Gligorijevic, B., Seudieu, C., Kosar, A. D., Davidson, E., and Roepe, P. D. (2004) Novel, rapid, and inexpensive cell-based quantification of antimalarial drug efficacy, *Antimicrob. Agents Chemother. 48*, 1807-1810.

38. Egan, T. J., and Ncokazi, K. K. (2004) Effects of solvent composition and ionic strength on the interaction of quinoline antimalarials with ferriprotoporphyrin IX, *J. Inorg. Biochem. 98*, 144-152.

39. Hawley, S. R., Bray, P. G., Mungthin, M., Atkinson, J. D., O'Neill, P. M., and Ward, S. A. (1998) Relationship between antimalarial drug activity, accumulation, and inhibition of heme polymerization in Plasmodium falciparum in vitro, *Antimicrob. Agents Chemother. 42*, 682-686.

Chapter 4a

Post-translational modifications in *Plasmodium*: More than you think!

Duk-Won Doug Chung, Nadia Ponts, Serena Cervantes, Karine G. Le Roch

Department of Cell Biology and Neuroscience, University of California, Riverside,

900 University Avenue, Riverside, CA 92521, USA

Abstract

Recent evidences indicate that transcription in Plasmodium may be hard-wired and rigid, deviating from the classical model of transcriptional gene regulation. Thus, it is important that other regulatory pathways be investigated as a comprehensive effort to curb the deadly malarial parasite. Research in post-translational modifications in Plasmodium is an emerging field that may provide new venues for drug discovery and potential new insights into how parasitic protozoans regulate their life cycle. Here, we discuss the recent findings of post-translational modifications in Plasmodium.

1. Introduction

Malaria is one of the deadliest infectious diseases of the world. Each year, malaria infects over 300 million people world-wide and causes an estimated one to two million deaths [1–3]. With the increase of drug resistance to most of the widely used anti-malarial drugs, it is imperative to better understand key regulatory elements driving the Plasmodium life cycle, eventually leading to the discovery of new drug targets. The sequencing of the Plasmodia genomes, along with subsequent comparative bioinformatics approaches, transcriptome and proteome analyses, have created a vast amount of information regarding protein prediction and their hypothetical functions in both the human host and mosquito vector [4–7].While microarray analyses have demonstrated a remarkable change in steady-state mRNA levels during parasite development, only relatively few regulatory motifs and transcription regulators have been uncovered so far [8]. In addition, unlike other organisms, there seems to be few transcriptional changes in Plasmodium following exposure to external stimuli [3,9,10]. These findings imply that parasite transcription could be hard-wired [10]. This rigidity in transcription suggests that post-transcriptional and post-translational mechanisms are likely to play major roles in regulating the parasite life cycle.

By definition, post-translational modifications (PTMs) must add or subtract a specific mass difference and not be particular to any one protein [11]. In addition to changing protein mass, PTMs can also alter the protein charge and

conformation. Such changes modify the protein's enzyme activity, binding affinity and hydrophobicity [12]. The spectrum of the PTMs within a cell is immense and varies with respect to specificity and abundance. For example, some types of PTMs such as phosphorylation and ubiquitination are universally employed to regulate a broad host of functions, have relatively high abundance, and have awide range of target substrates. On the other hand, some PTMs such as acetylation are highly specific in their roles, have a relatively low abundance, and may target only a few proteins or even one target substrate at a time. Furthermore, PTMs can either modify a protein at one specific amino acid residue or be associated with numerous different residues. The potential diversity of protein modifications seem limitless due to the fact that an individual protein can undergo a single modification or multiple types of modifications at several sites, possibly producing multiple protein isoforms, each with a unique biological activity. One can easily see why PTMs are responsible for a major increase in complexity from genome to proteome. For example, the human genome contains approximately 30,000 open reading frames but is predicted to give rise to roughly 1.8 million different protein variants [13].

Furthermore PTMs are reversible and thus provide the flexibility and adaptability that are essential for mediating rapid cellular responses to the cell's constantly changing conditions. Thus, posttranslational modifications are vital for the survival of all kinds of cells, with Plasmodium not being an exception. Due to

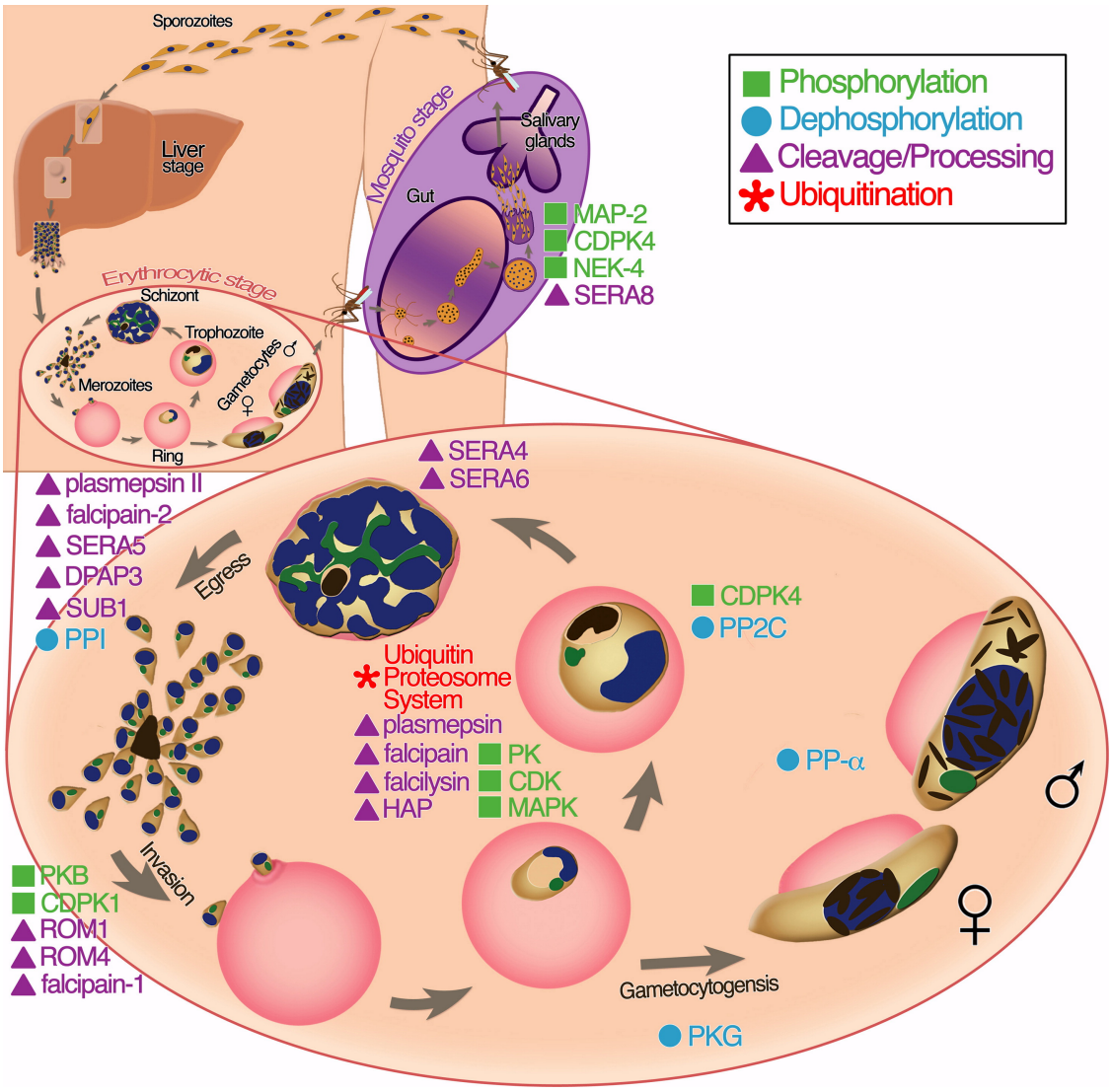


Figure a4.1. *General depiction of the roles of a few selected proteins involved in post-translational modifications during the life cycle of the malaria parasite.*

their diversity, essentiality, and wide-spread roles, post-translational modifications could present new major targets for effective and specific therapeutic intervention against the malarial parasite. Here we present an overview of the major post-translational modifications found so far in Plasmodium and discuss the recent developments of this rapidly expanding field. A summary

Modification	Type of protein	Protein name	Gene ID
Phosphorylation	Casein kinase	CK1	PF11_0377
	Cyclin-dependent kinase	PK5	MAL13P1.279
	Cyclin-dependent kinase	mrk	PF10_0141
	Serine/threonine protein kinase	PK7	PFB0605w
	Mitogen-activated protein kinase	map-1	PF14_0294
	Mitogen-activated protein kinase	map-2	PF11_0147
	Mitogen-activated protein kinase	map-2	PB000659.00.0
	Calcium-dependent kinase	CDPK4	PF07_0072
	cGMP-dependant protein kinase	PKG	PF14_0346
	NIMA-related kinase	nek-4	MAL7P1.100
	Calmodulin-like protein kinase	CDPK1	PFB0815w
	Protein kinase B	PKB	PFL2250c
	Protein kinase A	PKA	PFI1685w
	FIKK orphan kinases	Multiple proteins	Multiple genes
Dephosphorylation	Mg2+-dependent serine/ threonine protein phosphatase 2C	PP2C	Multiple genes
	VH1 family phosphatases	YVH1	Multiple genes
	Protein phosphatase 1	PP1	Multiple genes
	PPP-related protein serine/ threonine phosphatase	PP-Î±	Multiple genes
Lipidation	Prenyl modifier	PFT Î±-subunit	PFL2050w
	Prenyl modifier	PFT Î²-subunit	PF11_0483
Ubiquitination	Ubiquitin	pUB	PFL0585w
	Ubiquitin	UbS27a	PF13_0346
	Ubiquitin	UbL40	PF14_0027
	SUMO	PfSUMO	PFE0285c
	E2 conjugating enzyme	UBC13	PFE1350c
	Deubiquitylating/DeNeddylating enzyme	UCH54	PF14_0576
Cleavage/Processing	Serine protease	Subtilisin-1 (SUB1)	PFE0370c
		Subtilisin-2 (SUB2)	PF11_0381
		Subtilisin-3 (SUB3)	PFE0355c
		ROM-1	PF11_0150
		ROM-1	PB000352.00.0
		ROM-4	PFE0340c
		SP1	PF13_0118
	Cysteine protease	Falcipain-1	PF14_0553
		Falcipain-2	PF11_0165
		Falcipain-3	PF11_0162
		SERA-4	PFB0345c
		SERA-5	PFB0340c
		SERA-6	PFB0335c
		SERA-8	PFB0325c

Putative role
Unknown [20]
Cell-cycle regulation [25], [26] and [27]
Cell-cycle regulation [25], [26] and [27]
Proliferation and development [31]
Unknown [33]
Asexual cycle regulation [33]
Exflagellation in mosquito midgut [34]
DNA replication, ookinete gliding mobility, mosquito midgut invasion [35], [36] and [37]
Gametocytogenesis initiation [18]
Ookinete maturation [38]
Secretion of microneme contents, formation of tight moving junctions [40]
Parasite gliding [41]
Mediates erythrocyte anion channels and permeability of host plasma membrane [44] and [45]
Targets parasite proteins to host membranes [46]
Transcription elongation [55]
Nuclear protein activity [56]
Release of infection merozoites [57]
Cell-cycle control and signal transduction in sexual stages [50]
Unknown [88]
Unknown [88]
Polyubiquitin gene that provides the ubiquitin monomers that are covalently attached to proteins.
Cell-cycle regulation [106]
Ubiquitin moiety that provide the ubiquitin monomers for attachment to proteins. Cell-cycle regulation [108]
Ubiquitin moiety that provide the ubiquitin monomers for attachment to proteins. Cell-cycle regulation [108]
Ubiquitin-like peptide that is covalently attached to proteins for DNA repair, cell-cycle regulation,
nuclear localization and protein stability [110]
Was found to be phosphorylated by PfPK9. Role in cell-cycle has to be validated in Plasmodium[111]
Has dual deubiquitylating and deNeddylating activity.
May have roles in maintaining stable apical membrane epithelial Na+ channels [118]
Activates SERA proteins. Primary regulator of egress [143]
Shedding of MSP-1 and PfAMA-1 after invasion [151]
Ubiquitous role [144]
Shedding of adhesins [147]
May have roles in invasion [149]
Shedding of adhesins [146]
Cleavage of transit peptide [123]
Putative role in invasion [153]
Hemoglobinase & role in egress [132]
Hemoglobinase [132]
Expressed in the PV at the late schizont stage [140]
Role in merozoite egress [139]
Expressed in the PV at the late schizont stage [140]
Sporozoite release from oocystes [141]

	DPAP3	PFD0230c
	Calpain	MAL13P1.310
Metallo protease	Falcilysin	PF13_0322
Aspartic protease	Plasmepsin I	PF14_0076
	Plasmepsin II	PF14_0077
	Plasmepsin IV	PF14_0075
	Histo-aspartic protease (HAP)	PF14_0078

Table a4.1. *Genes that are putatively involved in post-translational modifications in Plasmodium.*

Primary regulator of egress [142]
Unknown [132]
Hemoglobinase [125]
Initiate degradation of hemoglobin [133]
Initiate degradation of hemoglobin & role in egress [133]
Cleavage of denatured globin [133]
Cleavage of denatured globin [133]

of the major proteins involved in regulating PTMs in Plasmodium is also presented in Figure a4.1 and Table a4.1.

2. Phosphorylation/dephosphorylation

Phosphorylation involves the reversible esterification of a phosphate group to an amino acid residue by protein kinases (PKs) that transfer a phosphoryl group from an ATP to hydroxyamino acid residues, mostly serine, threonine and tyrosine [14]. As the most highly studied covalent modification of proteins in eukaryotic cells, phosphorylation events can be linked to practically most functions within a cell: cell growth, cell differentiation, receptor activations, metabolic pathways, enzyme activities, cytoskeletal organization, chromatin remodeling, protein activations/inhibitions and protein–protein interactions. Approximately 1.5–2% of genes in an eukaryotic genome are PK family genes, which reflects the importance of phosphorylation events. While protein kinases catalyze the phosphorylation of protein residues, phosphatases hydrolyze the phosphoester bond of the modified amino acid, restoring the hydroxyamino acid

to its unphosphorylated state [15]. In eukaryotes, the number of protein phosphatases is relatively small when compared to that of protein kinases. However, additional phosphatase regulatory proteins mediate specific regulation of these enzymes and it is speculated that the number of phosphatase complexes involved in regulatory pathways may exceed the protein kinases repertoires [16].

2.1. Kinases

Several kinase inhibitors have been shown to inhibit Plasmodium development at different stages of the life cycle [9,17–20] validating the importance of phosphorylation in maintaining the parasite. Depending on the stringency applied, computational analyses retrieved 86–99 PK-related enzymes from the Plasmodium falciparum (Pf) genome [21,22]. Phylogenetic studies have demonstrated that most of the eukaryotic families of PKs are present in the Plasmodium genome with the exception of two groups, ste- 20 (STE) and the tyrosine protein kinases (TyrK) families. The STE family includes PKs involved in Mitogen-activated protein kinase (MAPK) cascades and plays a central role in transduction signals (see reference Ward et al., 2004 for further details). This group is evolutionarily conserved and their absence in the parasite genome may indicate that the mode of activation of the MAPK pathway in the parasite differs from other eukaryotes. The TyrKs are known to function in hormone-response

receptor-linked pathways essential for intercellular communication in multicellular organisms and are therefore not expected in the malaria parasite.

Several “orphan” PKs have also been identified in the Plasmodium kinome. They display only limited similarities with yeast or mammalian kinases. Examples of these atypical kinases include the NIMA-related kinase Nek1 (PFL1370W) [17], twenty FIKK PKrelated proteins, found only in apicomplexan parasites [22,23], and a family of calcium-dependant kinases (CDPKs) that are usually found in plants and alveolate but not in metazoans [24]. In the following section we briefly describe the key biological functions identified so far for a few selected Plasmodium kinases.

2.1.1. Cell-cycle regulation

In eukaryotes, kinases are known to play a major role in cell cycle progression. This is particularly the case for cyclin-dependent kinases (CDKs), several of which have been identified in the Pf genome [22]. Two Pf CDKs, PfPK5 (MAL13P1.279) and Pfmrk (PF10 0141), have been shown to be positively regulated by the binding of Plasmodium andmammaliancyclins as well as negatively regulated by CDK inhibitors [25–27]. These results demonstrate that regulatory activities of CDKs are well-conserved across species. While the targets of these Plasmodium kinases have not yet been identified, it is speculated that such proteins might play a major role in the regulation of the parasite

schizogony and its nuclear division cycles [28]. In vitro biochemical characterizations and crystal structure analyses have uncovered possible unique regulatory mechanisms of Plasmodium CDKs [27,29]. Such particularities indicate an atypical mode of DNA replication/mitosis in the parasite [30]. Furthermore, an atypical Pf orphan protein kinase, PfPK7 (PFB0605w), that exhibits maximal homology to a MEKKK and a fungal PKA has been shown to regulate parasite proliferation and development [31]. Discrepancies observed between parasite and host kinases may be exploited in the search of parasite-specific kinase inhibitors as potential anti-malarial drugs but would need to be further validated in vivo.

2.1.2. Cell proliferation and differentiation

MAPKs are known to be key players in signal transductions, cell development and differentiation in response to a variety of stimuli [32]. Though no classical MAPK kinase has been identified in the Plasmodium genome, two atypicalMAPKhomologs have been characterized, Pfmap-1 (PF14 0294) and Pfmap-2 (PF11 0147). While the exact role of Pfmap-1 is still speculated, Pfmap-2 appears to be essential for the completion of the parasite asexual erythrocytic cycle [33]. However in Plasmodium berghei (Pb), the Pfmap-2 orthologue seems to be essential to parasite exflagellation in the mosquitos midgut [34]. Discrepancies observed between these two Plasmodium species will need to be

further clarified. Nonetheless, it is possible to envision a complementation effect induced by species-specific Plasmodium kinases.

2.1.3. Sexual differentiation

Several identified Plasmodium kinases have been implicated in sexual differentiation. In Pb, a calcium-dependant kinase (CDPK4, PF07 0072 in Pf) has been shown to regulate gamete formation with the initiation of DNA replication, ookinate gliding mobilities, and mosquitos midgut invasion [35–37]. cGMP-dependant protein kinase (PFPKG, PF14 0346 in Pf) seems to be essential for mediating initiation of gametocytogenesis [18]. Finally, the Pb Pbnek-4 (MAL7P1.100 in Pf) has been shown to be critical for ookinate maturation [38]. The detection of an increased number of protein kinases involved in sexual differentiation demonstrates that PKs have a crucial role in regulating gametocytogenesis. These preliminary results provide a framework for identifying substrates of these Plasmodium enzymes to further comprehend the signalling cascade involved in sexual differentiation.

2.1.4. Parasite egress and invasion

In contrast to other pathogens, apicomplexans such as Plasmodium can quickly exit and enter the cell using their own machinery. Several steps involving a large collection of proteins are required in erythrocytic egress and invasion [39].

From this collection, a few protein kinases have been shown to have a role in controlling these parasitic processes. The calmodulin-like PK (CDPK1, PFB0815w in Pf) is required for the secretion of the parasite microneme contents and the formation of tight moving junctions [40]. Two components of the acto-myosin motor complex, the myosin A tail domain interacting protein (MTIP) and the glideosome-associated protein 45 (GAP45) have been identified to be CDPK1 substrates. In addition, PfPKB (PFL2250c) [41], an important member of the phosphatidylinositol 3-kinase-dependent signaling pathway, has been implicated to regulate parasite gliding mechanisms [42,43]. However, validating the role of PfCDPK1 and PfPKB in vivo remains to be established.

2.1.5. Host–parasite interaction

Recent works have begun to reveal the importance of Plasmodium PKs in modulating the erythrocyte membranes of infected hosts. PfPKA (PFI1685w) has been shown to alter both the activity of an erythrocyte anion channel and the permeability of the host plasma membrane [44,45]. Interestingly, the members of the Plasmodium FIKKs orphan kinases possess a Plasmodium export element (PEXEL) motif [46] that target parasite proteins to the host membrane [47,48]. Immuno-microscopy reveals that many FIKK proteins can be found in the erythrocyte cytoplasm and colocalized with Maurer's clefts protein. This finding suggests the importance of these FIKKs in the remodeling of the infected

erythrocyte membrane [46].

While it is increasingly apparent that a significant number of protein kinases regulate major parasite cell cycle and differentiation events, our understanding of their targeted substrates is extremely limited. In order to comprehend the function and essentiality of each kinase, a systematic biological approach involving reverse genetics, chemical genetics and phosphoproteome studies will need to be developed.

2.2. Phosphatases

A total of twenty-seven malaria protein phosphatases (PP) have been identified in the Pf genome (reviewed by [49]). The identified Plasmodium PP clustered with the four major established eukaryotic PP families: the Metallophosphatases (PPP), the Serine/Threonine Phosphatases (PPM), the Protein Tyrosine Phosphatases (PTP) and the NLI Interacting Factor-like phosphatases (NIF). Protein sequence information as well as biochemical characterizations studies of several of these PPs in Plasmodium have validated their phosphatase activities and revealed significant differences with other eukaryotic organisms [49–52].

The use of phosphatase inhibitors have shown that Plasmodium PPs are essential and may be involved in invasion [53] and cell growth [54]. Protein–protein interaction and localization studies have highlighted a role of Plasmodium

PPs in transcriptional elongation [55], nuclear protein activity [56] and the release of infectious merozoites [57]. Interestingly, highly conserved eukaryotic phosphatases seem to be either missing or too divergent to be detected in the Plasmodium genome. Among the missing phosphatases are the CDC25 homologue (known to play a major role in cell cycle control) [58], the cdc14 phosphatase (regulates mitotic events) [59] and tyrosine phosphatases [60] (though there is a report of a possible PRL tyrosine phosphatase) [61]. These findings further validate the phylogenetic distance observed between Plasmodium spp. and its vertebrate host. The identification of PP regulatory subunits, as well as their specific substrates, will need to be further investigated. However, it has become evident that these PPs are essential to the parasite and should be considered as potential targets for new anti-malarial strategies.

3. Acetylation

Acetylation is the addition of an acetyl functional group onto a protein substrate. Acetyltransferases transfer an acetyl group from acetyl coenzyme A onto conserved N-terminal lysine residues. The result of acetylation usually modifies DNA binding properties, protein stability, and protein–protein interactions.

Acetylation of actin at the N-terminus is a highly conserved PTM that has also been reported in Pf [62]. In addition to actin, Plasmodium histones are found

to be acetylated at their N-terminal lysine residues [63]. Histone PTMs are a vital part of the 'histone code' hypothesis that proposes specific combinations of PTMs that modify chromatin structure and act as platforms for the binding of transcriptional regulators of gene expression. Consistent with other eukaryotes, Plasmodium histone acetylation is believed to play a major role in transcriptional regulation. In a genome-wide investigation of histone modifications and their relationship with transcriptional activation/silencing, Cui et al. (2007) [64] employed a combination of immunoprecipitation and DNA microarray detection (ChIP-chip) to show that acetylation of histone H3 (H3K9ac) was associated with active genes across the Pf genome. PfGCN5 was the first characterized histone acetyltransferase (HAT) subunit identified in Plasmodium. Recombinant PfGCN5 displayed histone H3 acetylase activity in vitro and exists as a catalytic subunit of P. falciparum HAT complex. PfADA2, the yeast transcriptional coactivator homolog, has been characterized as another component of the trimeric catalytic core [65]. In vitro pull-down and yeast two-hybrid experiments suggest that PfADA2 and PfGCN5 are present in complex(es) and may have conserved chromatin remodeling functions. Natural compounds that inhibit HAT activity are being explored as potential antimalarials [66]. However, because HATs are conserved across species, toxicity studies will have to be further investigated.

Pf HDAC1 and PfSir2, two of the five putative Plasmodium Histone Deacetylases (HDACs), have been partially characterized in Plasmodium [67].

PfSir2, a yeast homolog of the silent information regulator 2 (SIR2) was found to bind to parasite telomeres and cause histone de-acetylation and silencing of the var multi-gene family [68]. Inhibitors of HDAC activity have also been investigated for potential antimalarial drugs [69,70].

4. Methylation

Methylation is catalyzed by methyltransferases that attach a methyl group onto its substrate. Such modification increases lipophilicity and reduces the substrate solubility in water. Methylation is associated with the regulation of gene expression and protein activity. In Plasmodium, the most commonly methylated subtrates are the parasite histones. Genomewide analysis of histone modification showed that H3K9me3 is associated with gene silencing [64]. Recently, chromatin immunoprecipitation of transcriptionally active var gene loci assessed the enrichment of histone H3 di- and trimethylation marks (H3K4me2 and H3K4me3) in the 5' flanking region [71].

At least nine Plasmodium histone methyltransferases (HMT) (containing a SET-domain, characteristic of histone lysine methyltransferase), and two Plasmodium demethylases (containing a JumonjiC-domain, characteristic of Histone Lysine Demethylases (HDMs)) have been identified in silico. Phylogenetic analysis divided putative histone lysine methyltransferases (HKMTs) into five subfamilies with different putative substrate specificities.

HKMTs and HDMs not only have substrate specificity but also specificity for different methyl states (mono-, di-, tri-) [72].

5. Lipidation

Lipidation is the covalent binding of a lipid group to a peptide chain and can change the activity and/or cellular localization of the modified protein. Often times, the attachment of a hydrophobic chain can help to anchor soluble proteins, or proteins with weak membrane-affinity, to the inner face of a membrane. Some examples of lipidation include N-myristoylation, prenylation, GPIanchor addition, and palmitoylation. Discussed below are a couple of post-translational modifications via lipidation that have been studied so far in Plasmodium.

5.1. GPI-anchoring

The glycosylphosphatidylinositol (GPI) anchor is a glycolipid moiety that is added to the C-terminal of proteins after translation. Such modified proteins are attached to the outer leaflet of the cell membrane. GPI-anchored proteins are ubiquitous among eukaryotic organisms and represent a very functionally diverse group that is being extensively studied in a wide-array of organisms. Anchoring with GPI is the major type of glycosylation that are found in Plasmodium [73]. It is commonly postulated that Plasmodium proteins are glycosylated by the erythrocytic machinery (combinatorial metabolism) during the exchanges that

occur between the parasitophorous vacuole and the erythrocyte cytoplasm [74]. In the human host, the parasite's GPI anchors have been linked to the pathobiology of the disease [75] and are associated with increased levels of TNF-α leading to a systemic inflammation reaction [76,77]. It has been found that resistance to malaria in endemic areas is associated with the production of circulant antibodies directed against the parasite's GPI anchors [75] and are common targets for vaccine research. Also, most P. falciparum merozoite surface proteins are GPI-anchored and are involved in erythrocyte recognition and attachment to erythrocytes before invasion (reviewed by [78]). GPIs may also be involved in postinvasion processes. The P. falciparum GPI-anchored rhoptry protein Pf34 (PFD0955w) is suspected to play a role in the formation of the parasitophorous vacuole [79,80].

In the mosquito host (Anopheles spp.), P. falciparum GPIs may play a direct role in immune response [81,82]. The immune response in Anopheles gambiae is elicited by the parasite GPIs resulting in a significant reduction in fecundity (diminished egg production) [83]. The role of such effect on the host–parasite interactions remains open to discussion. The authors propose that the parasite may use its GPIs to trigger the mosquito immune response and "mis-direct" the specificity of the response against other pathogen-associated molecular patterns rather than ookinetes or microneme proteins essential for parasitic processes. Furthermore, the mosquito stage GPI anchored surface

proteins P25 (PF10 0303) and P28 (PF10 0302) are shed during ookinete maturation and play an important role in midgut invasion [84–86]. Finally, the GPI anchor of the circumsporozoite protein (PFC0210c) plays a crucial role in sporogenesis [87]. Due to their general importance, GPIs and their biosynthesis pathway are being investigated as potential targets to antimalarial strategies.

5.2. Prenylation

Prenylation, also known as isoprenylation, is the posttranslational modification of proteins by covalent attachment near the carboxyl terminal of isoprenyl lipids, a 15 carbon farnesyl or a 20 carbon geranylgeranyl group [88]. The attachment of isoprenyl lipids creates a hydrophobic tail that promotes membrane association and plays an important role in cell signal transduction, vesicle trafficking, and cell-cycle progression [89]. Unlike animals, fungi and archeabacteria, which use the classical mevalonate pathway for isoprenoid synthesis, the Plasmodium synthesizes its isoprenoid precursors via the 1-deoxy-D-xylulose 5-phosphate (DOXP) pathway within its apicoplast, a plastid-like organelle [90].

Prenylation is mediated by three enzymes: protein farnesyltransferase (PFT), protein geranylgeranyltransferase type I and type II. PFT in Plasmodium has been characterized by partial purification of protein farnesyltransferase, PfPFT, and radiolabeling of prenylated proteins [88]. Prenylation precursors are

incorporated into asexual synchronized parasites in a stage-specific manner with the highest amount occurring from trophozoite to schizont, and schizont to ring transitions. Dolichylation, the addition of 11 isoprene units, has also been reported in P. falciparum during trophozoite and schizont stages [91].

Pharmaceutical companies have invested into the development of PFT inhibitors for the treatment of cancer with a few drugs in clinical trials. Interestingly, PFT inhibitors have been found to be potent antimalarials [92] and validate farnesylation as essential in the malaria parasite. Drug development for antimalarials is taking a 'piggy-back' approach since the concentrations needed to inhibit parasites are significantly lower than that of mammalian cells [89].

5.3. Palmitoylation

Palmitoylation is the covalent attachment of fatty acids to cysteine residues of proteins, often giving soluble proteins (or proteins with weak membrane avidity) a hydrophobic membrane anchor [93]. In P. falciparum, a 45 kDa gliding-associated protein (GAP45) was found to be both palmitoylated and N-myristoylated [94]. It is hypothesized that GAP45 may play a role in binding acto-myosin motors to the outer face of the inner membrane complex, which is implicated as the underlying force driving both gliding motility and host cell invasion in Plasmodium merozoites.

6. Ubiquitination

Ubiquitin is a highly conserved 76 amino acid peptide found in eukaryotic organisms. Beyond the more familiar association with protein degradation, the modification of proteins by ubiquitin conjugation is known to serve as a regulatory signal for cell proliferation, cell-stress response, transcription, cell death, DNA repair, intracellular trafficking, endocytosis and signal transduction [95]. The reversible conjugation of ubiquitin to the lysine residues in target proteins is controlled by a series of enzymes: ubiquitin activating enzymes (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin ligases (E3) [96]. Ubiquitin is activated by E1 and transferred to E2. Then, ubiquitin is either transferred to a monomeric E3 that catalyzes ubiquitination of the target substrate or ubiquitinated E2 forms a complex with the E3 to catalyze ubiquitination of the substrate. Deubiquitinating enzymes (DUBs) serve to reverse ubiquitin-conjugation by removing ubiquitin from substrate proteins and also help to replenish the free ubiquitin pool.

Diverse forms of ubiquitin modifications have been reported, each potentially mediating a specific function. K48-linked polyubiquitin chains to substrates often serve as signals for targeted protein degradation via the ubiquitin/proteasome system [97]. In Plasmodium, several studies have validated the essentiality of the proteasome and protein turnover in regulating the cell cycle progression. A collection of proteasome inhibitors has shown promising results in impeding the parasites [98–103].

Apart from protein degradation, poly-ubiquitination via K63- linkages seems to play important roles in DNA damage tolerances, endocytosis, ribosomal protein synthesis, and inflammatory response [104]. In addition to poly-ubiquitination, proteins can either be mono-ubiquitinated, which is the attachment of a single ubiquitin to a protein, or multi-ubiquitinated, which is the attachment of individual ubiquitin to a substrate at multiple sites. Both mono- and multi-ubiquitination are reported to have nonproteolytic roles such as endocytosis [105] and DNA repair [106]. In the following sections, we briefly describe what has been reported so far concerning ubiquitination in Plasmodium.

6.1. Ubiquitin and ubiquitin-like proteins

The P. falciparum poly-ubiquitin gene, PfpUB (PFL0585w) is present as a single-copy on chromosome 12 with five tandem repeats of the ubiquitin open reading frame [107]. Translation of PfpUB comprises of five ubiquitin monomers that have sequence identities of no less than 94% to that of other eukaryotic species. RT-PCR and northern analysis revealed that steady-state transcript levels of PfpUB are expressed at all stages of the intraerythrocytic cycles with significant increases at the late trophozoite and schizont stages. Under heat shock, polypeptide levels of PfpUB and ubiquitinated adducts showed dramatic increases without a significant increase in steady-state transcript levels. This observation suggests that the heat shock response appears to be maintained at

the level of translation [107].

Two additional ubiquitin moieties, UbS27a and UbL40, fused to the ribosomal proteins L40 (PF13 0346) and S27a (PF14 0027) have also been identified in the P. falciparum genome [108,109]. Expression data indicate that these ubiquitin genes are expressed throughout the P. falciparum life cycle [5,7]. Interestingly, though the Pf UbL40 shows high protein sequence fidelity with other eukaryotic species, Pf UbS27a harbors significant divergences within the ubiquitin domain [109].

In addition to ubiquitin, ubiquitin-like proteins (UBLps) have been identified as modifiers of cellular-processes. Though a number of UBLps (ISG15, FAT10, UFM1, FUB1) that are typical in higher eukaryotes were not found in Plasmodium, gene expression data suggests that SUMO, NEDD8, HUB1, URM1 and ATG8 are expressed at all life-cycle stages [108,110].

Recently, Issar et al. (2008) [111] were the first to investigate and characterize SUMO within P. falciparum. Amino acid sequence comparisons revealed that P. falciparum gene PFE0285c (PfSUMO) has significant sequence homology to that of known SUMO orthologs from other eukaryotic organisms such as yeast, human, mouse and Saccharomyces pombe, while immunoblot analysis confirmed the presence of SUMO in P. falciparum. In addition, enzyme homology searches have identified P. falciparum orthologs of all the necessary members of the SUMO pathway. Using LC-MS/MS analysis, more than 20

putative SUMO substrates of P. falciparum proteins were identified. These protein substrates varied from histones to transcription factors to RNA helicases, which implicates a wideranging regulatory scope of SUMO. Immunofluorescence assays indicate that PfSUMO localizes to distinctive subcellular compartments within the P. falciparum and also in the host cell cytoplasm within the parasite-derived structures called Maurer's clefts.

6.2. Ubiquitin enzymes

Computational studies have identified over a hundred proteins predicted to be involved in the reversible conjugation of ubiquitin or ubiquitin-like proteins in P. falciparum alone [108,110].

6.2.1. Ubiquitin activating enzymes (E1)

Eight putative E1 have been identified in the parasite genome [108]. While primary sequence identity has been observed in the core ubiquitin activating enzyme domain, sequences outside of this core diverge rapidly as the functional requirements for these E1 enzymes change to specifically interact with their respective E2 conjugating enzymes. Sequence analysis indicates the existence of Plasmodium E1 paralogs for UBA1 and UBA1-like proteins, and UBA2, UBA3, UBA4, ATG7 proteins, which mediate the activation of ubiquitin-like proteins SUMO, NEDD8, URM1, and ATG8, respectively.

6.2.2. Ubiquitin conjugating enzymes (E2)

Fourteen putative E2 paralogs were found in P. falciparum [108]. They exhibited extensive conservation with other eukaryotic E2 proteins. Gene expression data for nine of the fourteen Plasmodium E2s reveal a diverse pattern of steady-state mRNA at different stages of the intraerythrocytic cycle, suggesting the existence of a temporal profile of delivering ubiquitin or UbLps to different E3s, indicating a potential additional level of temporal control in ubiquitination during the parasite's life cycle.

Recently, a *Pf* homolog (PfUBC13) of the E2 ubiquitinconjugating enzyme 13 (UBC13) was characterized and found to be a substrate of the Pf protein kinase PfPK9 [112]. Reverse-phase HPLC and in vitro ubiquitination assay show that PfPK9 phosphorylates PfUBC13 at S106 and suppresses ubiquitin conjugating activity. Though the physiological role of PfUBC13 is unknown, the highlyconserved UBC13 (coupled with an ubiquitin E2-variant protein) assembles K63-linked ubiquitin chains [113], which mediate nonproteolytic pathways [114]. UBC13's conjugating activity regulates various cellular processes such asDNArepair [115], tumor suppressor p53 activity [116], and mitotic progression [117].

6.2.3. Ubiquitin ligating enzymes (E3)

Though Plasmodium E1 and E2 enzymes exhibited strong conserved homology with other E1 and E2 enzymes from other eukaryotes, Plasmodium E3 ligases were found to be highly divergent and the most abundant. Within the 54 putative E3 ligases identified, all superfamilies (HECT, RING, U-box, and cullin) of E3 ligases are represented within P. falciparum with E3 RING finger proteins making up the majority of the Plasmodium E3 ligases [108].

Functional annotation analysis reveals that these Plasmodium ligases have a wide array of potential roles including cell cycle regulation, trafficking, DNA repair, chromatin structure, and mRNA transport. However several of these proteins seem to be specific to the Apicomplexa phylum.While the functional analysis of these parasite-specific putative E3s will need to be further validated in vitro and in vivo, two selected E3 ligases seem to be essential to the parasite erythrocytic cycle and have already shown in vitro ubiquitination activity validating further the importance of this pathway in parasite development (Chung and Le Roch, unpublished data).

6.2.4. Deubiquitinating enzymes

While ubiquitin ligases catalyze the attachment of ubiquitin and UbLs, DUBs hydrolyze ubiquitin and UbLps [118]. Depending on the computational tools employed, 18 or 29 Plasmodium DUBs were found [108,110]. Five distinct gene families were identified: the ubiquitin C-terminal hydrolases (UCHs); the

ubiquitin-specific peptidases (USPs/UBPs); the ovarian tumor (OTU) domain proteins; the Josephin or Machado-Joseph disease (MJD) proteins and the JAMM(Jab1/MPN domain-associated metalloisopeptidase) domain proteins. In addition, DUBs for UbLps including SUMO, ATG8, and NEDD8 have also been identified.

PfUCH54 was the first DUB characterized in P. falciparum [119]. Using electrophilic probes that detect enzymes capable of removing ubiquitin and other ubiquitin-like proteins, PfUCH54 was found to possess both deubiquitination and deneddylation activity. Using known active site residues and crystal structure of homologous DUBs, PfUCH54 was found to have high homology to UCHL3 [119], an enzyme that has been reported to also have dual deubiquitinating and deneddylating activity [120].

Though the function of PfUCH54 is not known, mouse UCHL3 is required to maintain a stable apical membrane epithelial sodium channel, facilitating the dynamic recycling of sodium channels at the apical surface [121]. Though functional analysis studies will need to further validate the roles of ubiquitination pathways in Plasmodium, E3 ligases and DUBs may represent a good target for new therapeutic interventions due to their diversity and potential involvement in parasite-specific pathways.

7. Protein cleavage and processing

Proteases are major virulence factors in parasitic diseases as largely reviewed by [122] and [123]. Five main classes of proteases have been identified: (1) cysteine proteases, (2) serine proteases, (3) threonine proteases, (4) aspartic proteases, and (5) metalloproteases. A computational analysis identified 92 proteases in the genome of P. falciparum, 83 of them being transcribed during the erythrocytic cell cycle and 67 being translated during the parasite's life cycle [124]. Proteases are commonly involved in a wide array of biological processes such as lysosomal proteolysis, precursor protein processing or trafficking.

In Plasmodium, trafficking is a vital biological process, especially with regards to targeting proteins to the apicoplast. Trafficking pathways leading to the apicoplast have not yet been elucidated but are known to involve the presence of both a signal peptide and a transit peptide that are processed along the pathway [125].Wuet al. identified a serine protease Signal Peptidase 1 (SP1, PF13 0118) that could be responsible for the cleavage of signal peptides [124]. In addition, metalloprotease falcilysin (PF13 0322 in P. falciparum) [126] has been recently implicated in transit peptide cleavage [127].

Another important parasite-specific pathway is the PEXELmediated pathway that targets parasite proteins to the surface of infected erythrocytes. This pathway is thought to promote parasite evasion from the host immune system [128]. Chang et al. (2008) demonstrated that exported proteins are

processed at their N-terminal end [129]. This processing involves acetylation and cleavage of the PEXEL motif in the endoplasmic reticulum. The authors suggest that this N-terminal processing may be more generally utilized for many exported soluble proteins. However, the protease responsible for such cleavage remains to be identified.

In Plasmodium, various proteases play key roles in hemoglobin degradation, egress and invasion during the parasite erythrocytic cycle. The roles of such proteases were usually determined from inhibitor studies, which emphasize the potential to target such enzymes by anti-malarial drugs.

7.1. Hemoglobin degradation

The degradation of hemoglobin involves various proteases called hemoglobinases [130] and provides the essential amino acids that are taken up by the parasite [122,131,132].

The cysteine proteases falcipain-2 (PF11 0165) and falcipain-3 (PF11 0162) are known to play a role in the early steps of hemoglobin degradation within the parasite food vacuole (see [133] for a review). In addition to these cysteine proteases, degradation of hemoglobin seems to involve metallo (falcilysin) and aspartic proteases such as plasmepsins [131]. Four plasmepsins are present and active in the food vacuole of P. falciparum, plasmepsin I (PF14 0076), II (PF14 0077), IV (PF14 0075) and a histo-aspartic

protease (HAP, PF14 0078) [134–136]. It was postulated that plasmepsins I and II are matured and released in the food vacuole by protein cleavage [134,137]. A recent study demonstrated that falcipain-2 and falcipain-3 are responsible for plasmepsins preprocessing and that auto-processing can occurwhenfalcipain activities are inhibited, providing an alternative pathway to activate plasmepsins [138].

7.2. Parasite egress

Parasite proteases are also known to trigger the degradation of parasite and host membranes, leading to the egress of infectious parasites. The role of these proteases in cell egress has been recently reviewed [139]. Briefly, the proteases implicated in parasite egress are falcipain II, plasmepsin II and putative papain-like SERA proteases. SERA proteins are a family of nine members, which are activated by a subtilisin-like serine protease SUB1 (PFE0370c). Among them are SERA-4 (PFB0345c), SERA-5 (PFB0340c, [140]) and SERA-6 (PFB0335c), which are essential proteins that are expressed in the parasitophorous vacuole of the late trophozoite and schizont stage (erythrocytic cycle) [141]. Additionally, SERA-8 (PFB0325c), another member of the SERA protein family, is essential for sporozoite release from oocytes (mosquito stage) [142].

The involvement of the subtilisin-family serine protease PfSUB1 and the

cysteine protease dipeptidyl peptidase 3 (DPAP3, PFD0230c) as primary regulators of parasite egress have been recently evidenced using serine and cysteine protease inhibitors [143]. Just prior to egress, the essential serine protease PfSUB1 is discharged from the exonemes into the parasitophorous vacuole space [144]. Inhibition of both DPAP3 and PfSUB1 blocks the processing of the serine repeat antigen protein SERA-5, which correlates with the inhibition of membrane rupture. In Plasmodium, two other subtilisin serine proteases of unknown function can be found: PfSUB2 (PF11 0381) and PfSUB3 (PFE0355c) (reviewed in [145]). PfSUB2 is believed to play an important role in the erythrocytic cycle while PfSUB3 appears to be more ubiquitous as it is expressed during the asexual blood stage, in gametocytes and in sporozoites [6].

7.3. Parasite invasion

The critical role of proteases in merozoite invasion of erythrocytes has been largely studied and reviewed [39,146]. Briefly, invasion involves contact, interaction and junction between the merozoite and the red blood cell surface via adhesins and various GPI-anchored proteins, such as PfAMA-1 and MSPs followed by the active entry of the parasite within the host. In order for invasion to be completed, these interactions are interrupted by the shedding of the protein coat covering the merozoite surface made of various MSPs and adhesins. The proteases involved in such removal are called sheddases. For example, the

adhesins Duffy binding ligand erythrocyte-binding antigen (DBL-EBP) 175 (EBA-175, MAL7P1.176) is shed from the merozoite at around the point of invasion by the rhomboid protease PfROM4 (PFE0340c) [147]. Rhomboid proteases are ubiquitous intramembrane serine proteases (see [148–150] for an extensive description of the different types of rhomboid proteases). Also, sporozoite invasion of hepatocytes is reported to be mediated by the shedding of PfAMA-1 and the thrombospondin-related adhesive protein (TRAP) by a serine protease [151].

Double cleavage of MSPs is required for invasion and probably involves serine proteases 9 See Harris, Yeoh et al. (2005) [152] for further details 0. Very recently, it was suggested that MSP-1, MSP-6 and MSP-7 undergo proteolytic maturation catalyzed by PfSUB1 before egress, maturation that is essential for further processing of MSPs involved in invasion [153].

A role of cysteine proteases (falcipain) in erythrocyte invasion has also been proposed but remains elusive. Inhibitors of falcipain- 1 (PF14 0553) block invasion of host red blood cells [154] whereas parasites with a disrupted falcipain-1 gene were not affected in terms of asexual growth [155]. More recently, the presence of an endogenous cysteine protease inhibitor in Pf, falstatin, has been identified [156]. The presence of such an endogenous inhibitor may be another level of proteolytic control of certain parasite/host proteases in order to facilitate erythrocyte invasion.

Proteolytic processing is finally involved in the maturation of proteins secreted by the merozoite-specific organelles, rhoptries. Rhoptry-associated proteins (RAP) are believed to play a role in invasion and are activated upon proteolytic cleavage. For example, the rhoptry-associated protein RAP-1 is maturated in vivo by multiple processing steps [157–159].

8. Concluding remarks

In addition to the PTMs presented above, there are several other PTMs in Plasmodium that are being investigated but, due to spaceconstraints, are unable to be fully expanded upon in this review. However, we would like to briefly mention a few. For example, a Plasmodium cathepsin-C-like protein is reported to be modified via O-sulfonation [160]. Also, polyglutamylation of tubulin was found in the microtubule organizing centers and post-mitotic microtubular structures of Plasmodium [161]. Lastly, several chaperone proteins, involved in the proper folding of proteins, have been described in Plasmodium and are reported to be involved in roles such as trafficking [162–165].

Today, transcriptional studies make up a significant amount of publications that are dedicated to understanding mechanisms regulating the Plasmodium developmental cycle. However, recent evidences suggest that Plasmodium may not follow the classical transcriptional model and may have rigid transcription machinery, which may indicate the need to reallocate investigative efforts into

other fields in order to better understand the parasite's life-cycle regulation. With the exception of phosphorylation/dephosphorylation, post-translational modifications in Plasmodium have been understudied and are largely overshadowed by the classical view that proteins are mainly regulated at the transcriptional level. With further investigation, post-translational regulation may reveal to be a bigger factor in parasite development than previously thought.

For example, though it is apparent that proteases are necessary to destabilize host cell membranes during invasion and egress, the mechanisms of how these proteases are activated and tightly controlled temporally are yet unclear. However, there are increasing reports that PTMs may play a major role in both malarial invasion and egress. As already described above, it is suggested that additional proteases, such as SUB proteases, may act as regulators of egress by processing, and thereby activating, distinct effector proteins or signal transduction pathways [139,144]. In addition, it is reported that Pf protein kinases, such as CDPK, may also be involved in regulating invasion and egress [19].

Besides having amajor regulatory role within the parasite, PTMs may also play key roles outside the parasite by manipulating the host's signaling pathways thereby usurping normal cellular processes for survival and escape from immune responses. Currently, there are an increasing number of reports that show PTMs to play an integral role in host–pathogen interactions, or cross-talk, in bacteria, viruses and protozoa [166–168]. For example, it has been shown that host

organisms employ a wide range of posttranslational modifications to initiate their immune responses in order to avoid invasion by pathogens. For instance, host cells use ubiquitination in defense strategies as a way to degrade parasitic proteins and also to activate inflammatory and anti-apoptotic genes in a non-degradative fashion [169]. As a countermeasure, it has been shown that pathogens are able to avoid and exploit these ubiquitination defense strategies by secreting proteins of their own that either deubiquitinate or inhibit the host ubiquitinating enzymes [169,170]. It is also reported that bacterial pathogens are able to provoke histone modifications (via acetylation and phosphorylation) and chromatin remodeling in host cells, thereby manipulating the host's transcriptional programming and diminish the host innate immune response [167].

The utilization of PTMs within host–pathogen interactions for the purposes of pathogenic growth and immune evasion is an emerging field, which may prove to be both widespread and diverse in the mechanisms at work. The utilization of PTMs to manipulate host cell signaling and immune response may also be true in Plasmodium. In P. falciparum, it is reported that kinases and phosphatases were found among the 320+ proteins predicted to be secreted from the parasite [48]. More recently, over 30 proteins have been validated to be secreted from the Pf, of which 27 proteins are novel extracellular proteins [171]. Several post-translational modifiers, such as kinases, phosphatases, and proteases, were confirmed to be among the secreted proteins. Furthermore, preliminary functional

analysis suggests that these secreted proteins are possibly involved in immune evasion and signaling [171].

Our current understanding of the roles and extent of PTMs within Plasmodium is still greatly limited due to the relative novelty and inherent challenges of this particular field. More genomewide approaches, along with their proper biological validation, will provide significant advances in filling the gaps of our present understanding of the regulatory mechanisms driving this deadly parasite. Though there is still much more to uncover, current reports collectively show that Plasmodium PTMs have vital roles in all aspects of the parasites' life cycle, including host–pathogen interactions. Because of the numerous types and virtually incalculable combinations of PTMs, the possibilities of regulation at the post-translational level are vast. In addition to being vitally essential, many Plasmodium PTM proteins are both highly divergent and specific to the parasite, making Plasmodium PTM proteins excellent candidates for drug targeting. Thus, Plasmodium PTMs may open new venues for drug discovery and may prove to be more significant in the regulation of this deadly parasite than previously acknowledged.

References

[1] Miller LH, Greenwood B. Malaria—a shadow over Africa. Science 2002;298(5591):121–2.

[2] Breman JG, Egan A, Keusch GT. The intolerable burden of malaria: a new look at the numbers. Am J Trop Med Hyg 2001;64(1–2 Suppl.):iv–vii.

[3] Guerra CA, Gikandi PW, Tatem AJ, Noor AM, Smith DL, Hay SI, et al. The limits and intensity of Plasmodium falciparum transmission: implications for malaria control and elimination worldwide. PLoS Med 2008;5(2):e38.

[4] Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, et al. Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 2002;419(6906):498–511.

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

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

[7] Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, et al. Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. Genome Res 2004;14(11):2308–18.

[8] Coulson RM, Hall N, Ouzounis CA. Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. Genome Res 2004;14(8):1548–54.

[9] Le Roch KG, Johnson JR, Ahiboh H, Chung DW, Prudhomme J, Plouffe D, et al. A systematic approach to understand the mechanism of action of the bisthiazolium compound T4 on the human malaria parasite, Plasmodium falciparum. BMC Genomics 2008;9:513.

[10] Ganesan K, Ponmee N, Jiang L, Fowble JW, White J, Kamchonwongpaisan S, et al. A Genetically Hard-Wired Metabolic Transcriptome in Plasmodium falciparum Fails to Mount Protective Responses to Lethal Antifolates. PLoS Pathogens 2008;4(11):e1000214.

[11] Hoffman MD, Sniatynski MJ, Kast J. Current approaches for global posttranslational modification discovery and mass spectrometric analysis. Anal Chim Acta 2008;627(1):50–61.

[12] Clark RS, Bayir H, Jenkins LW. Posttranslational protein modifications. Crit Care Med 2005;33(12 Suppl.):S407–9.

[13] Jensen ON. Modification-specific proteomics: characterization of posttranslational modifications by mass spectrometry. Curr Opin Chem Biol 2004;8(1):33–41.

[14] Reinders J, Sickmann A. State-of-the-art in phosphoproteomics. Proteomics 2005;5(16):4052–61.

[15] Sickmann A, Meyer HE. Phosphoamino acid analysis. Proteomics 2001;1(2):200–6.

[16] Bollen M. Combinatorial control of protein phosphatase-1. Trends Biochem Sci 2001;26(7):426–31.

[17] Dorin D, Le Roch K, Sallicandro P, Alano P, Parzy D, Poullet P, et al. Pfnek-1, a NIMA-related kinase from the human malaria parasite Plasmodium falciparum biochemical properties and possible involvement in MAPK regulation. Eur J Biochem 2001;268(9):2600–8.

[18] McRobert L, Taylor CJ, Deng W, Fivelman QL, Cummings RM, Polley SD, et al. Gametogenesis in malaria parasites is mediated by the cGMP-dependent protein kinase. PLoS Biol 2008;6(6):e139.

[19] Kato N, Sakata T, Breton G, Le Roch KG, Nagle A, Andersen C, et al. Gene expression signatures and small-molecule compounds link a protein kinase to Plasmodium falciparum motility. Nat Chem Biol 2008;4(6):347–56.

[20] Knockaert M, Gray N, Damiens E, Chang YT, Grellier P, Grant K, et al. Intracellular targets of cyclin-dependent kinase inhibitors: identification by affinity chromatography using immobilised inhibitors. Chem Biol 2000;7(6):411–22.

[21] Anamika, Srinivasan N, Krupa A. A genomic perspective of protein kinases in Plasmodium falciparum. Proteins 2005;58(1):180–9.

[22] Ward P, Equinet L, Packer J, Doerig C. Protein kinases of the human malaria parasite Plasmodium falciparum: the kinome of a divergent eukaryote. BMC Genomics 2004;5(1):79.

[23] Schneider AG, Mercereau-Puijalon O. A new Apicomplexa-specific protein kinase family: multiple members in Plasmodium falciparum, all with an export signature. BMC Genomics 2005;6(1):30.

[24] Zhang XS, Choi JH. Molecular evolution of calmodulin-like domain protein kinases (CDPKs) in plants and protists. J Mol Evol 2001;53(3):214–24.

[25] Merckx A, Le Roch K, Nivez MP, Dorin D, Alano P, Gutierrez GJ, et al. Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite Plasmodium falciparum. J Biol Chem 2003;278(41):39839–50.

[26] Li Z, Le Roch K, Geyer JA, Woodard CL, Prigge ST, Koh J, et al. Influence of human p16(INK4) and p21(CIP1) on the in vitro activity of recombinant Plasmodium falciparum cyclin-dependent protein kinases. Biochem Biophys Res Commun 2001;288(5):1207–11.

[27] Le Roch K, Sestier C, Dorin D, Waters N, Kappes B, Chakrabarti D, et al. Activation of a Plasmodium falciparum cdc2-related kinase by heterologous p25 and cyclin H. Functional characterization of a P. falciparum cyclin homologue. J Biol Chem 2000;275(12):8952–8.

[28] Graeser R, Wernli B, Franklin RM, Kappes B. Plasmodium falciparum protein kinase 5 and the malarial nuclear division cycles. Mol Biochem Parasitol 1996;82(1):37–49.

[29] Holton S, Merckx A, Burgess D, Doerig C, Noble M, Endicott J. Structures of P.falciparum PfPK5 test the CDK regulation paradigm and suggest mechanisms of small molecule inhibition. Structure 2003;11(11):1329–37.

[30] Schrevel J, Asfaux-Foucher G, Bafort JM. Ultrastructural study of multiple mitoses during sporogony of Plasmodium b. berghei. J Ultrastruct Res 1977;59(3):332–50.

[31] Dorin-Semblat D, Sicard A, Doerig C, Ranford-Cartwright L, Doerig C. Disruption of the PfPK7 gene impairs schizogony and sporogony in the human malaria parasite Plasmodium falciparum. Eukaryot Cell 2008;7(2):279–85.

[32] Raman M, Chen W, Cobb MH. Differential regulation and properties of MAPKs. Oncogene 2007;26(22):3100–12.

[33] Dorin-Semblat D, Quashie N, Halbert J, Sicard A, Doerig C, Peat E, et al. Functional characterization of both MAP kinases of the human malaria parasite Plasmodium falciparum by reverse genetics. Mol Microbiol 2007;65(5):1170–80.

[34] Tewari R, Dorin D, Moon R, Doerig C, Billker O. An atypical mitogen-activated protein kinase controls cytokinesis and flagellar motility during male gamete formation in a malaria parasite. Mol Microbiol 2005;58(5):1253–63.

[35] Billker O, Dechamps S, Tewari R, Wenig G, Franke-Fayard B, Brinkmann V. Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. Cell 2004;117(4):503–14.

[36] Siden-Kiamos I, Ecker A, Nyback S, Louis C, Sinden RE, Billker O. Plasmodium berghei calcium-dependent protein kinase 3 is required for ookinete gliding motility and mosquito midgut invasion. Mol Microbiol 2006;60(6):1355–63.

[37] Ishino T, Orito Y, Chinzei Y, Yuda M.Acalcium-dependent protein kinase regulates Plasmodium ookinete access to the midgut epithelial cell. MolMicrobiol 2006;59(4):1175–84.

[38] Reininger L, Billker O, Tewari R, Mukhopadhyay A, Fennell C, Dorin-Semblat D, et al. A NIMA-related protein kinase is essential for completion of the sexual cycle of malaria parasites. J Biol Chem 2005;280(36):31957–64.

[39] Cowman AF, Crabb BS. Invasion of red blood cells by malaria parasites. Cell 2006;124(4):755–66.

[40] Moskes C, Burghaus PA, Wernli B, Sauder U, Durrenberger M, Kappes B. Export of Plasmodium falciparum calcium-dependent protein kinase 1 to the parasitophorous vacuole is dependent on three N-terminal membrane anchor motifs. Mol Microbiol 2004;54(3):676–91.

[41] Kumar A, Vaid A, Syin C, Sharma P. PfPKB, a novel protein kinase B-like enzyme from Plasmodium falciparum: I. Identification, characterization, and possible role in parasite development. J Biol Chem 2004;279(23):24255–64.

[42] Vaid A, Sharma P. PfPKB, a protein kinase B-like enzyme from Plasmodium falciparum: II. Identification of calcium/calmodulin as its upstream activator and dissection of a novel signaling pathway. J Biol Chem 2006;281(37):27126–33.

[43] Vaid A, Thomas DC, Sharma P. Role of Ca2+/calmodulin-PfPKB signaling pathway in erythrocyte invasion by Plasmodium falciparum. J Biol Chem 2008;283(9):5589–97.

[44] Merckx A, Bouyer G, Thomas SL, Langsley G, Egee S. Anion channels in Plasmodium-falciparum-infected erythrocytes and protein kinase A. Trends Parasitol 2009;25(3):139–44.

[45] Merckx A, Nivez MP, Bouyer G, Alano P, Langsley G, Deitsch K, et al. Plasmodium falciparum regulatory subunit of cAMP-dependent PKA and anion channel conductance. PLoS Pathog 2008;4(2):e19.

[46] Nunes MC, Goldring JP, Doerig C, Scherf A. A novel protein kinase family in Plasmodium falciparum is differentially transcribed and secreted to various cellular compartments of the host cell. Mol Microbiol 2007;63(2):391–403.

[47] Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science 2004;306(5703):1930–3.

[48] Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estrano C, et al. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science 2004;306(5703):1934–7.

[49] Wilkes JM, Doerig C. The protein-phosphatome of the human malaria parasite Plasmodium falciparum. BMC Genomics 2008;9:412.

[50] Li JL, Baker DA. A putative protein serine/threonine phosphatase from Plasmodium falciparum contains a large N-terminal extension and five unique inserts in the catalytic domain. Mol Biochem Parasitol 1998;95(2):287–95.

[51] Mamoun CB, Sullivan Jr DJ, Banerjee R, Goldberg DE. Identification and characterization of an unusual double serine/threonine protein phosphatase 2C in the malaria parasite Plasmodium falciparum. J Biol Chem 1998;273(18):11241–7.

[52] Dobson S, May T, Berriman M, Del Vecchio C, Fairlamb AH, Chakrabarti D, et al. Characterization of protein Ser/Thr phosphatases of the malaria parasite. Plasmodium falciparum: inhibition of the parasitic calcineurin by cyclophilincyclosporin complex Mol Biochem Parasitol 1999;99(2):167–81.

[53] Ward GE, Fujioka H, Aikawa M, Miller LH. Staurosporine inhibits invasion of erythrocytes by malarial merozoites. Exp Parasitol 1994;79(3):480–7.

[54] Yokoyama D, Saito-Ito A, Asao N, Tanabe K, Yamamoto M, Matsumura T. Modulation of the growth of Plasmodium falciparum in vitro by protein serine/threonine phosphatase inhibitors. Biochem Biophys Res Commun 1998;247(1):18–23.

[55] Mamoun CB, Goldberg DE. Plasmodium protein phosphatase 2C dephosphorylates translation elongation factor 1beta and inhibits its PKC-mediated nucleotide exchange activity in vitro. Mol Microbiol 2001;39(4):973–81.

[56] Kumar R, Musiyenko A, Cioffi E, Oldenburg A, Adams B, Bitko V, et al. A zinc-binding dual-specificity YVH1 phosphatase in the malaria parasite, Plasmodium falciparum, and its interaction with the nuclear protein, pescadillo. Mol Biochem Parasitol 2004;133(2):297–310.

[57] Blisnick T, Vincensini L, Fall G, Braun-Breton C. Protein phosphatase 1, a Plasmodium falciparum essential enzyme, is exported to the host cell and implicated in the release of infectious merozoites. Cell Microbiol 2006;8(4):591–601.

[58] Rudolph J. Cdc25 phosphatases: structure, specificity, and mechanism. Biochemistry 2007;46(12):3595–604.

[59] Trinkle-Mulcahy L, Lamond AI. Mitotic phosphatases: no longer silent partners. Curr Opin Cell Biol 2006;18(6):623–31.

[60] Fauman EB, Saper MA. Structure and function of the protein tyrosine phosphatases. Trends Biochem Sci 1996;21(11):413–7.

[61] Pendyala PR, Ayong L, Eatrides J, Schreiber M, Pham C, Chakrabarti R, et al. Characterization of a PRL protein tyrosine phosphatase from Plasmodium falciparum. Mol Biochem Parasitol 2008;158(1):1–10.

[62] Schmitz S, Grainger M, Howell S, Calder LJ, Gaeb M, Pinder JC, et al. Malaria parasite actin filaments are very short. J Mol Biol 2005;349(1):113–25.

[63] Miao J, Fan Q, Cui L, Li J, Li J, Cui L. The malaria parasite Plasmodium falciparum histones: organization, expression, and acetylation. Gene 2006;369:53–65.

[64] Cui L, Miao J, Furuya T, Li X, Su XZ, Cui L. PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in Plasmodium falciparum. Eukaryote Cell 2007;6(7):1219–27.

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

[66] Cui L, Miao J, Furuya T, Fan Q, Li X, Rathod PK, et al. Histone acetyltransferase inhibitor anacardic acid causes changes in global gene expression during in vitro Plasmodium falciparum development. Eukaryot Cell 2008;7(7):1200–10.

[67] Joshi MB, Lin DT, Chiang PH, Goldman ND, Fujioka H, Aikawa M, et al. Molecular cloning and nuclear localization of a histone deacetylase homologue in Plasmodium falciparum. Mol Biochem Parasitol 1999;99(1):11–9.

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

[69] Andrews KT, Tran TN, Wheatley NC, Fairlie DP. Targeting histone deacetylase inhibitors for anti-malarial therapy. Curr Top Med Chem 2009;9(3):292–308.

[70] Dow GS, Chen Y, Andrews KT, Caridha D, Gerena L, Gettayacamin M, et al. Antimalarial activity of phenylthiazolyl-bearing hydroxamate-based histone deacetylase inhibitors. Antimicrob Agents Chemother 2008;52(10):3467–77.

[71] Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R, Scherf A. 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol Microbiol 2007;66(6):1296–305.

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

[73] Gowda DC, Gupta P, Davidson EA. Glycosylphosphatidylinositol anchors represent the major carbohydrate modification in proteins of intraerythrocytic stage Plasmodium falciparum. J Biol Chem 1997;272(10):6428–39.

[74] Haldar K, Samuel BU, Mohandas N, Harrison T, Hiller NL. Erythrocytic vacuolar rafts induced by malaria parasites. Curr Opin Hematol 2001;8(2):92–7.

[75] Naik RS, Branch OH, Woods AS, Vijaykumar M, Perkins DJ, Nahlen BL, et al. Glycosylphosphatidylinositol anchors of Plasmodium falciparum: molecular characterization and naturally elicited antibody response that may provide immunity to malaria pathogenesis. J Exp Med 2000;192(11):1563–76.

[76] Schofield L, Hackett F. Signal transduction in host cells by a glycosylphosphatidylinositol toxin of malaria parasites. J Exp Med 1993;177(1):145–53.

[77] Patel SN, Lu Z, Ayi K, Serghides L, Gowda DC, Kain KC. Disruption of CD36 impairs cytokine response to Plasmodium falciparum glycosylphosphatidylinositol and confers susceptibility to severe and fatal malaria in vivo. J Immunol 2007;178(6):3954–61.

[78] Nasir ud D, Ahmad I, Hoessli DC, Walker-Nasir E, Choudhary MI. Glycosylphosphatidylinositol (GPI) anchored proteins of Plasmodium falciparum: Antigenic determinants and role of sugar moieties in the GPI anchor. Current Organic Chemistry 2007;11(7):609–18.

[79] Proellocks NI, Kovacevic S, Ferguson DJ, Kats LM, Morahan BJ, Black CG, et al. Plasmodium falciparum Pf34, a novel GPI-anchored rhoptry protein found in detergent-resistant microdomains. Int J Parasitol 2007;37(11):1233–41.

[80] Gilson PR, Nebl T, Vukcevic D, Moritz RL, Sargeant T, Speed TP, et al. Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite Plasmodium falciparum. Mol Cell Proteomics 2006;5(7):1286–99.

[81] Lim J, Gowda DC, Krishnegowda G, Luckhart S. Induction of nitric oxide synthase in Anopheles stephensi by Plasmodium falciparum: mechanism of signaling and the role of parasite glycosylphosphatidylinositols. InfectImmun 2005;73(5):2778–89.

[82] Akman-Anderson L, Olivier M, Luckhart S. Induction of nitric oxide synthase and activation of signaling proteins in Anopheles mosquitoes by the malaria pigment, hemozoin. Infect Immun 2007;75(8):4012–9.

[83] Arrighi RB, Debierre-Grockiego F, Schwarz RT, Faye I. The immunogenic properties of protozoan glycosylphosphatidylinositols in the mosquito Anopheles gambiae. Dev Comp Immunol 2009;33(2):216–23.

[84] Blanco AR, Paez A, Gerold P, Dearsly AL, Margos G, Schwarz RT, et al. The biosynthesis and post-translational modification of Pbs21 an ookinete-surface protein of Plasmodium berghei. Mol Biochem Parasitol 1999;98(2):163–73.

[85] del Carmen Rodriguez M, Gerold P, Dessens J, Kurtenbach K, Schwartz RT, Sinden RE, et al. Characterisation and expression of Pbs25, a sexual and sporogonic stage specific protein of Plasmodium berghei. Molecular and Biochemical Parasitology 2000;110(1):147–59.

[86] Baton LA, Ranford-Cartwright LC. Do malaria ookinete surface proteins P25 and P28 mediate parasite entry into mosquito midgut epithelial cells? Malar J 2005;4(1):15.

[87] Wang Q, Fujioka H, Nussenzweig V. Mutational analysis of the GPI-anchor addition sequence from the circumsporozoite protein of Plasmodium. Cell Microbiol 2005;7(11):1616–26.

[88] Chakrabarti D,DaSilva T, Barger J, Paquette S, Patel H, Patterson S, et al. Protein farnesyltransferase and protein prenylation in Plasmodium falciparum. J Biol Chem 2002;277(44):42066–73.

[89] Eastman RT, Buckner FS, Yokoyama K, Gelb MH, Van Voorhis WC. Thematic review series: lipid posttranslational modifications. Fighting parasitic disease by blocking protein farnesylation. J Lipid Res 2006;47(2):233–40.

[90] Rohrich RC, Englert N, Troschke K, Reichenberg A, Hintz M, Seeber F, et al. Reconstitution of an apicoplast-localised electron transfer pathway involved in the isoprenoid biosynthesis of Plasmodium falciparum. FEBS Lett 2005;579(28):6433–8.

[91] D'Alexandri FL, Kimura EA, Peres VJ, Katzin AM. Protein dolichylation in Plasmodium falciparum. FEBS Lett 2006;580(27):6343–8.

[92] Fletcher S, Cummings CG, Rivas K, Katt WP, Horney C, Buckner FS, et al. Potent, Plasmodium-selective farnesyltransferase inhibitors that arrest the growth of malaria parasites: structure-activity relationships of ethylenediamine-analogue scaffolds and homology model validation. J Med Chem 2008;51(17):5176–97.

[93] Baekkeskov S, Kanaani J. Palmitoylation cycles and regulation of protein function (Review). Mol Membr Biol 2009;26(1):42–54.

[94] Rees-Channer RR, Martin SR, Green JL, Bowyer PW, Grainger M, Molloy JE, et al. Dual acylation of the 45 kDa gliding-associated protein (GAP45) in Plasmodium falciparum merozoites. Mol Biochem Parasitol 2006;149(1):113–6.

[95] Hershko A, Ciechanover A. The ubiquitin system. Annu Rev Biochem 1998;67:425–79.

[96] Laney JD, Hochstrasser M. Substrate targeting in the ubiquitin system. Cell 1999;97(4):427–30.

[97] Kerscher O, Felberbaum R, Hochstrasser M. Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu Rev Cell Dev Biol 2006;22:159–80.

[98] Gantt SM, Myung JM, Briones MR, Li WD, Corey EJ, Omura S, et al. Proteasome inhibitors block development of Plasmodium spp. Antimicrob Agents Chemother 1998;42(10):2731–8.

[99] Lindenthal C, Weich N, Chia YS, Heussler V, Klinkert MQ. The proteasome inhibitor MLN-273 blocks exoerythrocytic and erythrocytic development of Plasmodium parasites. Parasitology 2005;131(Pt 1):37–44.

[100] Reynolds JM, El Bissati K, Brandenburg J, Gunzl A, Mamoun CB. Antimalarial activity of the anticancer and proteasome inhibitor bortezomib and its analog ZL3B. BMC Clin Pharmacol 2007;7:13.

[101]Mordmuller B, Fendel R, Kreidenweiss A, Gille C, Hurwitz R, Metzger WG, et al. Plasmodia express two threonine-peptidase complexes during asexual development. Molecular and Biochemical Parasitology 2006;148(1):79–85.

[102] Prudhomme J, McDaniel E, Ponts N, Bertani Sp, Fenical W, Jensen P, et al. Marine Actinomycetes: A New Source of Compounds against the Human Malaria Parasite. PLoS ONE 2008;3(6):e2335.

[103] Kreidenweiss A, Kremsner PG, Mordmuller B. Comprehensive study of proteasome inhibitors against Plasmodium falciparum laboratory strains and field isolates from Gabon. Malar J 2008;7:187.

[104] Mukhopadhyay D, Riezman H. Proteasome-independent functions of ubiquitin in endocytosis and signaling. Science 2007;315(5809):201–5.

[105] Hicke L. Protein regulation by monoubiquitin. Nat Rev Mol Cell Biol 2001;2(3):195–201.

[106] Hofmann K. Ubiquitin-binding domains and their role in the DNA damage response. DNA Repair, 2009.

[107] Horrocks P, Newbold CI. Intraerythrocytic polyubiquitin expression in Plasmodium falciparum is subjected to developmental and heat-shock control. Mol Biochem Parasitol 2000;105(1):115–25.

[108] Ponts N, Yang J, Chung D-WD, Prudhomme J, Girke T, Horrocks P, et al. Deciphering the Ubiquitin-Mediated Pathway in Apicomplexan Parasites: A Potential Strategy to Interfere with Parasite Virulence. PLoS ONE 2008;3(6):e2386.

[109] Catic A, Ploegh HL. Ubiquitin—conserved protein or selfish gene? Trends Biochem Sci 2005;30(11):600–4.

[110] Ponder EL, Bogyo M. Ubiquitin-like modifiers and their deconjugating enzymes in medically important parasitic protozoa. Eukaryot Cell 2007;6(11):1943–52.

[111] Issar N, Roux E, Mattei D, Scherf A. Identification of a novel post-translational modification in Plasmodium falciparum: protein sumoylation in different cellular compartments. Cell Microbiol 2008;10(10):1999–2011.

[112] Philip N, Haystead TA. Characterization of a UBC13 kinase in Plasmodium falciparum. Proc Natl Acad Sci U S A 2007;104(19):7845–50.

[113] Hofmann RM, Pickart CM. Noncanonical MMS2-encoded ubiquitinconjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. Cell 1999;96(5):645–53.

[114] Pickart CM. Ubiquitin enters the new millennium. Mol Cell 2001;8(3):499–504.

[115] Brusky J, Zhu Y, Xiao W. UBC13, a DNA-damage-inducible gene, is a member of the error-free postreplication repair pathway in Saccharomyces cerevisiae. Curr Genet 2000;37(3):168–74.

[116] Laine A, Topisirovic I, Zhai D, Reed JC, Borden KL, Ronai Z. Regulation of p53 localization and activity by Ubc13. Mol Cell Biol 2006;26(23):8901–13.

[117] Bothos J, Summers MK, Venere M, Scolnick DM, Halazonetis TD. The Chfr mitotic checkpoint protein functions with Ubc13-Mms2 to form Lys63-linked polyubiquitin chains. Oncogene 2003;22(46):7101–7.

[118] Singhal S, Taylor MC, Baker RT. Deubiquitylating enzymes and disease. BMC Biochem 2008;9 Suppl. 1:S3.

[119] Artavanis-Tsakonas K, Misaghi S, Comeaux CA, Catic A, Spooner E, Duraisingh MT, et al. Identification by functional proteomics of a deubiquitinating/deNeddylating enzyme in Plasmodium falciparum. Mol Microbiol 2006;61(5):1187–95.

[120] Wada H, Kito K, Caskey LS, Yeh ET, Kamitani T. Cleavage of the C-terminus of NEDD8 by UCH-L3. Biochem Biophys Res Commun 1998;251(3):688–92.

[121] Butterworth MB, Edinger RS, Ovaa H, Burg D, Johnson JP, Frizzell RA. The deubiquitinating enzyme UCH-L3 regulates the apical membrane recycling of the epithelial sodium channel. J Biol Chem 2007;282(52):37885–93.

[122] McKerrow JH, Rosenthal PJ, Swenerton R, Doyle P. Development of protease inhibitors for protozoan infections. Curr Opin Infect Dis 2008;21(6):668–72.

[123] Armstrong PB. Proteases and protease inhibitors: a balance of activities in host–pathogen interaction. Immunobiology 2006;211(4):263–81.

[124] Wu Y, Wang X, Liu X, Wang Y. Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite. Genome Res 2003;13(4):601–16.

[125] Waller RF, Reed MB, Cowman AF, McFadden GI. Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. Embo J 2000;19(8):1794–802.

[126] Eggleson KK, Duffin KL, Goldberg DE. Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite Plasmodium falciparum. J Biol Chem 1999;274(45):32411–7.

[127] Ponpuak M, Klemba M, Park M, Gluzman IY, Lamppa GK, Goldberg DE. A role for falcilysin in transit peptide degradation in the Plasmodium falciparum apicoplast. Mol Microbiol 2007;63(2):314–34.

[128] Craig A, Scherf A. Molecules on the surface of the Plasmodium falciparum infected erythrocyte and their role in malaria pathogenesis and immune evasion. Mol Biochem Parasitol 2001;115(2):129–43.

[129] Chang HH, FalickAM,Carlton PM, Sedat JW, DeRisi JL, MarlettaMA.N-terminal processing of proteins exported by malaria parasites. Mol Biochem Parasitol 2008;160(2):107–15.

[130] Goldberg DE. Hemoglobin degradation. Curr Top Microbiol Immunol 2005;295:275–91.

[131] Francis SE, Banerjee R, Goldberg DE. Biosynthesis and maturation of the malaria aspartic hemoglobinases plasmepsins I and II. J Biol Chem 1997;272(23):14961–8.

[132] Lew VL, Tiffert T, Ginsburg H. Excess hemoglobin digestion and the osmotic stability of Plasmodium falciparum-infected red blood cells. Blood 2003;101(10):4189–94.

[133] Rosenthal PJ. Cysteine proteases of malaria parasites. Int J Parasitol 2004;34(13–14):1489–99.

[134] Banerjee R, Liu J, Beatty W, Pelosof L, Klemba M, Goldberg DE. Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. Proc Natl Acad Sci U S A 2002;99(2):990–5.

[135] Coombs GH, Goldberg DE, Klemba M, Berry C, Kay J, Mottram JC. Aspartic proteases of Plasmodium falciparum and other parasitic protozoa as drug targets. Trends Parasitol 2001;17(11):532–7.

[136] Wyatt DM, Berry C. Activity and inhibition of plasmepsin IV, a new aspartic proteinase from the malaria parasite, Plasmodium falciparum. FEBS Lett 2002;513(2–3):159–62.

[137] Francis SE, Sullivan Jr DJ, Goldberg DE. Hemoglobin metabolism in the malaria parasite Plasmodium falciparum. Annu Rev Microbiol 1997;51:97–123.

[138] Drew ME, Banerjee R, Uffman EW, Gilbertson S, Rosenthal PJ, Goldberg DE. Plasmodium food vacuole plasmepsins are activated by falcipains. J Biol Chem 2008;283(19):12870–6.

[139] Blackman MJ. Malarial proteases and host cell egress: an 'emerging' cascade. Cell Microbiol 2008;10(10):1925–34.

[140] Pang XL, Mitamura T, Horii T. Antibodies reactive with the N-terminal domain of Plasmodium falciparum serine repeat antigen inhibit cell proliferation by agglutinating merozoites and schizonts. Infect Immun 1999;67(4):1821–7.

[141] Miller SK, Good RT, Drew DR, Delorenzi M, Sanders PR, Hodder AN, et al.Asubset of Plasmodium falciparum SERA genes are expressed and appear to play an important role in the erythrocytic cycle. J Biol Chem 2002;277(49):47524–32.

[142] Aly AS, Matuschewski K. A malarial cysteine protease is necessary for Plasmodium sporozoite egress from oocysts. J Exp Med 2005;202(2):225–30.

[143] Arastu-Kapur S, Ponder EL, Fonovic UP, Yeoh S, Yuan F, Fonovic M, et al. Identification of proteases that regulate erythrocyte rupture by the malaria parasite Plasmodium falciparum. Nat Chem Biol 2008;4(3):203–13.

[144] Yeoh S, O'Donnell RA, Koussis K, Dluzewski AR, Ansell KH, Osborne SA, et al. Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. Cell 2007;131(6):1072–83.

[145] Withers-Martinez C, Jean L, Blackman MJ. Subtilisin-like proteases of the malaria parasite. Mol Microbiol 2004;53(1):55–63.

[146] O'Donnell RA, Blackman MJ. The role of malaria merozoite proteases in red blood cell invasion. Curr Opin Microbiol 2005;8(4):422–7.

[147] O'Donnell RA, Hackett F, Howell SA, Treeck M, Struck N, Krnajski Z, et al. Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. J Cell Biol 2006;174(7):1023–33.

[148] Baker RP, Wijetilaka R, Urban S. Two Plasmodium rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. PLoS Pathog 2006;2(10):pe113.

[149] Lemberg MK, Freeman M. Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases. Genome Res 2007;17(11):1634–46.

[150] Srinivasan P, Coppens I, Jacobs-Lorena M. Distinct roles of Plasmodium rhomboid 1 in parasite development and malaria pathogenesis. PLoS Pathog 2009;5(1):e1000262.

[151] Silvie O, Franetich JF, Renia L, Mazier D. Malaria sporozoite: migrating for a living. Trends Mol Med 2004;10(3):97–100, discussion 100-1.

[152] Harris PK, Yeoh S, Dluzewski AR, O'Donnell RA, Withers-Martinez C, Hackett F, et al. Molecular identification of amalaria merozoite surface sheddase. PLoS Pathog 2005;1(3):241–51.

[153] Koussis K, Withers-Martinez C, Yeoh S, Child M, Hackett F, Knuepfer E, et al. A multifunctional serine protease primes the malaria parasite for red blood cell invasion. Embo J 2009.

[154] Greenbaum DC, Baruch A, Grainger M, Bozdech Z, Medzihradszky KF, Engel J, et al. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. Science 2002;298(5600):2002–6.

[155] Eksi S, Czesny B, Greenbaum DC, Bogyo M, Williamson KC. Targeted disruption of Plasmodium falciparum cysteine protease, falcipain 1, reduces oocyst production, not erythrocytic stage growth.MolMicrobiol 2004;53(1):243–50.

[156] Pandey KC, Singh N, Arastu-Kapur S, Bogyo M, Rosenthal PJ. Falstatin, a cysteine protease inhibitor of Plasmodium falciparum, facilitates erythrocyte invasion. PLoS Pathog 2006;2(11):e117.

[157] Bushell GR, Ingram LT, Fardoulys CA, Cooper JA. An antigenic complex in the rhoptries of Plasmodium falciparum. Mol Biochem Parasitol 1988;28(2):105–12.

[158] Howard RF, Reese RT. Plasmodium falciparum: hetero-oligomeric complexes of rhoptry polypeptides. Exp Parasitol 1990;71(3):330–42.

[159] Howard RF, Schmidt CM. The secretary pathway of Plasmodium falciparum regulates transport of p82/RAP1 to the rhoptries. Mol Biochem Parasitol 1995;74(1):43–54.

[160] Medzihradszky KF, Darula Z, Perlson E, Fainzilber M, Chalkley RJ, Ball H, et al. O-sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes. Mol Cell Proteomics 2004;3(5):429–40.

[161] Fennell BJ, Al-shatr ZA, Bell A. Isotype expression, post-translational modification and stage-dependent production of tubulins in erythrocytic Plasmodium falciparum. Int J Parasitol 2008;38(5):527–39.

[162] Banumathy G, Singh V, Pavithra SR, Tatu U. Heat shock protein 90 function is essential for Plasmodium falciparum growth in human erythrocytes. J Biol Chem 2003;278(20):18336–45.

[163] Pavithra SR, Kumar R, Tatu U. Systems analysis of chaperone networks in the malarial parasite Plasmodium falciparum. PLoS Comput Biol 2007;3(9):1701–15.

[164] Mouray E, Moutiez M, Girault S, Sergheraert C, Florent I, Grellier P. Biochemical properties and cellular localization of Plasmodium falciparum protein disulfide isomerase. Biochimie 2007;89(3):337–46.

[165] Saridaki T, Sanchez CP, Pfahler J, Lanzer M. A conditional export system provides new insights into protein export in Plasmodium falciparum-infected erythrocytes. Cell Microbiol 2008;10(12):2483–95.

[166] Randow F, Lehner PJ. Viral avoidance and exploitation of the ubiquitin system. Nat Cell Biol 2009;11(5):527–34.

[167] Hamon MA, Cossart P. Histone modifications and chromatin remodeling during bacterial infections. Cell Host Microbe 2008;4(2):100–9.

[168] Stulemeijer IJ, Joosten MH. Post-translational modification of host proteins in pathogen-triggered defence signalling in plants. Mol Plant Pathol 2008;9(4):545–60.

[169] Munro P, Flatau G, Lemichez E. Bacteria and the ubiquitin pathway. Curr Opin Microbiol 2007;10(1):39–46.

[170] Edelmann MJ, Kessler BM. Ubiquitin and ubiquitin-like specific proteases targeted by infectious pathogens: emerging patterns and molecular principles. Biochim Biophys Acta 2008;1782(12):809–16.

[171] Singh M, Mukherjee P, Narayanasamy K, Arora R, Gupta S, Natarajan K, et al. Proteome analysis of plasmodium falciparum extracellular secretory antigens at asexual blood stages reveals a cohort of proteins with possible roles in immune modulation and signaling. Mol Cell Proteomics 2009.