

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


Diversity of Antigenic Secretion in Apicomplexa Parasites and Its Role in Plasmodium Falciparum Malaria

Citation

Pelle, Karell Guemmegne. 2013. Diversity of Antigenic Secretion in Apicomplexa Parasites and Its Role in Plasmodium Falciparum Malaria. Doctoral dissertation, Harvard University.

Link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:11004915

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

DIVERSITY OF ANTIGENIC SECRETION IN APICOMPLEXA PARASITES AND ITS ROLE IN *PLASMODIUM FALCIPARUM* MALARIA

A dissertation presented

by

Karell Guemmegne Pellé

to

The Department of Immunology and Infectious Diseases

in partial fulfillment of the requirements for

the degree of Doctor in Philosophy

in the subject of

Biological Sciences and Public Health

Harvard University

Cambridge, Massachusetts

May 2013

© 2013 - Karell Guemmegne Pellé

All rights reserved

Dissertation Advisor:

Professor Matthias Marti **Karell Guemmegne Pellé**

ABSTRACT

DIVERSITY OF ANTIGENIC SECRETION IN APICOMPLEXA PARASITES AND ITS ROLE IN PLASMODIUM FALCIPARUM MALARIA

Apicomplexan parasites are responsible for some of the most devastating human and veterinarian diseases and are parasites of great economic importance. Apicomplexa include *Plasmodium, Toxoplasma* and *Babesia* species. The pathogenic mechanisms developed by Apicomplexa parasites, in particular those that reside in a parasitophorous vacuole, involve considerable changes to the host cell, including the expression of variable surface proteins required for immune evasion. In *Plasmodium falciparum* infections, host cell remodeling is responsible for disease symptomology and severity in the human host. This work represents a multi-faceted study of antigenic secretion and the role of secreted antigens in pathogenesis. We study in detail the mechanisms of antigen secretion in Apicomplexa parasites. By use of comparative genomics, we find Plasmodium export element (PEXEL)-like motifs in a subset of *Cryptosporidium* and *Babesia* secreted proteins. However, in *Babesia* the motif functions as a spherical body targeting sequence, suggesting that secretory mechanisms in Apicomplexa are adapted to the parasite's intracellular lifestyle. To elucidate the relationship and function of exported antigens, we first focused on *P. falciparum* to determine gene co-expression modules. We found that *in vivo*, export modules are composed of constitutively or variably expressed genes, the latter group associated with patient clinical phenotypes. We then focused on a novel gene family called "*phist*" and show, using transcriptional expression profiling, its role in *P. falciparum* cytoadherence. In total, we demonstrate that antigen secretion is an evolutionary mechanism in Apicomplexa parasites and that variant expression of the genes encoding these antigens may allow parasites to adapt to environmental stresses.

-TABLE OF CONTENTS-

CHAPTER I: INTRODUCTION

I. APICOMPLEXAN PARASITES 3

II. INSIGHT FROM SEQUENCING APICOMPLEXAN GENOMES 4

III. APICOMPLEXAN PATHOGENESIS 5

A. APICOMPLEXAN PATHOGENESIS: REMODELING THE HOST CELL 5

B. PLASMODIUM FALCIPARUM PATHOGENESIS DURING CEREBRAL MALARIA 9

IV. PROTEIN TRAFFICKING IN APICOMPLEXA PARASITES 11

A. LESSONS FROM THE PAST: TARGETING TO THE APICOPLAST 11

B. TARGETING TO THE HOST CELL: PLASMODIUM FALCIPARUM PROTEIN EXPORT 14

C. CONCLUSION 16

V. MECHANISMS OF ANTIGENIC VARIATION IN PLASMODIUM FALCIPARUM MALARIA 16

A. VARIANT SURFACE ANTIGENS 17

B. GENETICS OF VARIANT EXPRESSION 18

1. Partitioning the genome: importance of the telomeres 18

2. Transcriptional regulation 19

C. EPIGENETIC MECHANISMS OF VARIANT EXPRESSION 20

VI. SUMMARY OF AIMS 22

VII. REFERENCES 23

CHAPTER II: SHARED ELEMENTS OF HOST-TARGETING PATHWAYS AMONG APICOMPLEXAN PARASITES OF DIFFERING LIFE STYLES

I. INTRODUCTION 36

II. RESULTS 37

A. DEFINITION OF APICOMPLEXAN SECRETOMES 37

A. IDENTIFICATION OF A PEXEL-LIKE MOTIF IN BABESIA AND CRYPTOSPORIDIUM 40

B. CHARACTERIZATION OF A BABESIA HOST-TARGETING PATHWAY 46

C. THE BABESIA PEXEL-LIKE MOTIF EXPORTS A REPORTER IN TRANSGENIC PLASMODIUM FALCIPARUM 51

III. DISCUSSION 53

IV. MATERIALS AND METHODS 61

V. REFERENCES 66

CHAPTER III: TRANSCRIPTIONAL PROFILING OF GENES AT THE HOST-PARASITE INTERFACE IDENTIFIES EXPRESSION MODULES INVOLVED IN MALARIA PATHOGENESIS

I. INTRODUCTION 73

II. RESULTS 76

A. TRANSCRIPTIONAL PROFILING OF GENES AT THE HOST-PARASITE INTERFACE 76

B. PROPERTIES OF CONSTITUTIVE AND VARIABLE EXPORT MODULES 77

C. MODULE VALIDATION BY QRT-PCR IN CEREBRAL MALARIA PATIENTS 80

D. MODULE VALIDATION BY QRT-PCR IN A IN VITRO MODEL OF CEREBRAL MALARIA 83

E. A PROTEIN IN VARIABLE CLUSTER I IS EXPORTED TO THE SURFACE 86

F. EPIGENETIC MARKS OF REGULATION ARE ENRICHED IN VARIABLE MODULES 89

G. FUNCTIONAL ENRICHMENT OF EXPORTOME GENE MODULES INCLUDE "PROTEIN BINDING" AND "PROTEIN FOLDING" 92

H. VARIABLE EXPORT MODULES ARE ASSOCIATED WITH PATHOLOGY IN CM PATIENTS 93

III. DISCUSSION 95

IV. SUBJECTS AND METHODS 99

V. REFERENCES 105

CHAPTER IV: CHARACTERIZATION OF A NOVEL GENE FAMILY IN PLASMODIUM FALCIPARUM, PHISTA

I. INTRODUCTION 115

II. RESULTS 120

A. ABSENCE OF PHISTA GENE SELECTION IN CULTURE-ADAPTED PARASITES 120

B. PHISTA GENE EXPRESSION IS MODULATED IN CYTOADHERENCE-SWITCHED CLONES AND PARASITES FROM CHILDREN WITH CEREBRAL MALARIA 122

C. A SUBSET OF PHISTA GENES ARE VARIANTLY EXPRESSED 123

D. THE EXPRESSION OF SOME PHISTA MAY BE UNDER EPIGENETIC REGULATION 127

E. SUBCELLULAR LOCALIZATION OF PHISTA DOMAINS IN DEVELOPING PARASITES 129

III. DISCUSSION 131

IV. MATERIALS AND METHODS 134

V. REFERENCES 137

CHAPTER V: CONCLUDING REMARKS

I. SUMMARY AND GENERAL DISCUSSION 145

II. FUTURE PERSPECTIVES 148

III. REFERENCES 150

ANNEX

-FIGURE INDEX-

CHAPTER II

Figure 2- 1: Protein secretion and trafficking in Apicomplexa. 39

Figure 2- 2: Identification and analysis of putative export motifs in Babesia and Cryptosporidium.. 45

Figure 2- 3: Analysis of smORF-YFP and SBP2-YFP expression in transgenic Babesia parasites. 48

Figure 2- 4: Analysis of transgenic Babesia parasites expressing mutant N-termini fused to YFP. 50

Figure 2- 5: Experimental validation of the Babesia PEXEL-like export motif function by heterologous expression in P. falciparum. 52

Figure 2- 6: Revisiting the Plasmodium PEXEL and the evolution of protein export in Apicomplexa. 59

CHAPTER III

Figure 3- 1: Gene co-expression modules and their characteristics. 79

Figure 3- 2: Transcriptional co-expression of variable export modules in P. falciparum parasites from CM patients. 81

Figure 3- 3: var transcript abundance by Ups group in parasites from CM patients. 82

Figure 3- 4: Transcriptional co-expression of variable export modules 1, 2 and 10 in P. falciparum parasites with cytoadhesive properties. 84

Figure 3- 5: Var transcript abundance by Ups group in cytoadherence-selected parasites. 85

Figure 3- 6: Subcellular localization of PF70 in P. falciparum infected erythrocytes. 88

Figure 3- 7: Representation of GO functions of export module gene that do not belong to the exportome. 92

Table 3- 1: Variable export modules are associated with epigenetic marks of expression regulation. 91

Table 3- 2: Variable export modules are associated with disease in Malawian cohort of CM. 94

CHAPTER IV

FIGURE 4- 1: LACK OF SELECTION OF PHISTA GENES CULTURE-ADAPTED PARASITE CLONES.. 121
FIGURE 4- 2: EXPRESSION OF PHISTA GENES IN UNSELECTED AND CULTURE-ADAPTED PARASITES..... 124
FIGURE 4- 3: EXPRESSION OF PHISTA GENES IN CYTOADHERENCE-SELECTED AND CEREBRAL MALARIA PARASITES..... 125
FIGURE 4- 4: PHISTA GENE EXPRESSION IN CM, CYTOADHERENCE-SELECTED AND CULTURE-ADAPTED PARASITES..... 126
FIGURE 4- 5: DISTRIBUTION OF VARIANT EXPRESSION IN THE PHISTA GENE FAMILY.. 128
FIGURE 4- 6: SUBCELLULAR LOCALIZATION OF THREE PHISTA DOMAINS DURING BLOOD STAGE DEVELOPMENT 130

-ACKNOWLEDGEMENTS-

My journey up till defending my dissertation has undoubtedly involved many people along the way. Some are no longer with us today. Others have left their mark in sometimes the subtlest ways. I am indebted to all of them.

I first would like to thank my dissertation advisor Dr. Matthias Marti. I joined the lab at a time when it was just forming and I am ever grateful to him for the opportunity to be one of the foundation rocks of his group. This in combination with his mentoring has allowed me to develop into a thorough, team-driven and productive scientist.

I would like to express my deepest appreciation to Dr. Barbara Burleigh who convinced and pushed me to apply to Ph.D. programs after doing a summer internship in her laboratory during my junior year in college. Without her support and her belief in my potential, I doubt I would have built the confidence to apply for Ph.D. programs, let alone at Harvard University.

My dissertation work has had the scientific input of many bright individuals. First, from the Marti group, I would like to thank all the members, past and present: Dr. Pierre-Yves Mantel, Dr. Kathrin Buchholz, Dr. Mythili Aingaran, Regina Joice, Dr. Sue Law, Dr. Sandra Nilsson, Dr. Johan Ankarklev, Evan Meyer, Ilana Goldowitz, Dr. Nicanor Obaldia, Kassey Dantzler, Deepali Ravel, Thomas Burke and Kelly DeLorenzo. They have been invaluable in the execution and progress of my work. In particular, I would like to acknowledge Dr. Pierre-Yves Mantel, my “francophone alliance”, who I worked closely with during my rotation and has been a great motivator and friend; and Dr. Mythili Aingaran, my “bay-mate”, a true friend, an older sister, who has always showed support and encouragement and whose real-life anecdotes I will dearly miss. To Dr. Johan Ankarklev, I have to thank for helping me prepare for my thesis defense and for reviewing my dissertation.

Second, I would like to show my appreciation to the members of neighbouring labs. First, Dr. Dyann Wirth who has been supportive and provided essential criticism of my work. Dr. Manoj Duraisingh, who has given me frank advice about science but most importantly about career options, often at times when I needed it the most. Also, Dr. Jeffrey Dvorin, Dr. Rays Jiang, Dr. Brad Coleman, Dr. Catherine Merrick, Dr. Amy Bei, Caeul Lim, Dr. Christof Gruring, Dr. Markus Ganter, Dr. Kacey Caradonna, Mwai Ndegwa, Dr. Saurabh Patel, Dr. Adi Paul, Selasi Dankwa, Dr. Amanda Lukens, Dr. Elizabeth Eagan, Dr. Usheer Kanjee: they have all been encouraging and contributed immensely to my research. In particular, I'd like to specially thank Dr. Ulf Ribacke, who has been a great friend and scientific advisor.

My dissertation work also involved many collaborators. Working with them has been a memorable learning experience and I am blessed to have been coached by them. These are Dr. Curtis Huttenhower and Dr. David Allred.

Recognition also has to be given to the administrative team, which has helped smooth my experience as a doctoral student: Andrea Sabaroff, Holly Southern, Ruth Kenworthy, Steve Heim, Joann Garrido, Soledad Marc, and Christopher Helmuth. I would also like to thank funding agencies, including a doctoral fellowship from the New England Biolabs, Dr. Adeoye Olukotun through the Wanda Lane Buck Fellowship of HSPH, and the Harvard Alumni for Global Action.

I also would like to thank the BPH program director Dr. Marianne Wessling-Resnick. She has been an indelible driving force in the managing of the program and has supervised the successful completion of BPH students. Also, many thanks are directed to my Dissertation Advisory Committee (DAC): Dr. Karl Munger, Dr. M-J Gubbels and Dr. Tiffany Horng. They were instrumental in seeing my progress through their constructive feedback and their guidance. Also, I would like to thank the examiners on my Dissertation Defense Committee: Dr. M-J Gubbels, Dr. Barbara Burleigh, Dr. Sarah Fortune and Dr. Johanna Daily.

I would like to acknowledge my fellow BPH classmates with whom I started my doctoral journey in the summer of 2007. Prerna Bhargava, Ediz Calay, Celene Chang, Khanichi Charles, Chris Ford,

Franziska Graf, Regina Joice, Sihao Liu and Daria Van Tyne: I have enjoyed getting to know all of them and see them grow as scientist and in their personal lives. Ediz, I will surely miss our random morning encounters on the “66 Bus”! I wish them all the best in your future endeavors!

Boston Massachusetts is not my home, but a few people have helped me adopt this city as my own. Lana Dinic, I am so thankful of our friendship. She has become almost like a sister to me and I cannot wait to see what life holds for her! Cinthya Torres, an amazing woman with a bright future. I am thankful for our growing friendship and I hope it will continue to grow for many years to come. Catherine Hammons, I cherish our fun times back in the beginning days at HSPH. I had the best time in Boston with her! Kim Brolin, I have never met such a superwoman! Balancing Ph.D., motherhood, and family life! I can only strive to be like her. I enjoyed and will miss our long coffee breaks. Colby Wells, she introduced me to Boston and welcomed me into her home. I am grateful to have her as a friend. Nono Ayivi and Hirity Shimellis: they are my confidants! I cannot express how I cherish them both and how they have helped me through some really tough times during this journey.

Most importantly, none of this would have happened without the continuous and unconditional love, patience and support from my family, to whom this thesis is dedicated. They have advised me during the toughest times and helped me stay strong in my faith. I am indebted to them. First, I would like to express my heart-felt gratitude to my parents Dr. Roger and Esther Pellé. They have been my rock, my home base, the only two people I always turn to in this world for strength and peace; second, my brothers, Stéphane and Brice for leading by example and for pushing me to always seek better; third, their wives, Joelle and Serwaa, for being the sisters I never had; and finally, my niece and nephew, Nailah and Ayden, for bringing pure love and fun in my life.

-DEDICATION-

This thesis is dedicated to my family, the light in my life.

CHAPTER I

Introduction

CONTRIBUTIONS

This introduction was composed by Karell Pellé.

I. Apicomplexa parasites

Apicomplexa is a phylum of unicellular parasites that cause some of the most devastating human and livestock diseases. The name Apicomplexa stems from the fact that they harbor a unique set of secretory organelles, the apical complex involved in invasion. They are ubiquitously spread around the globe and make up over 4500 different species of parasite (Levine et al., 1988). This notorious group can be split into different classes based on their phylogenetic relationship and host-tropisms. Parasites transmitted by blood-sucking insects that infect erythrocytes, the Haemosporidians, include species like *Babesia* and *Theileria*, which cause severe disease in cattle (Bishop et al., 2004; Homer et al., 2000; Mosqueda et al., 2012). More notable are *Plasmodium spp.,* which cause malaria with the most severe form *Plasmodium falciparum* claiming the life of 1-2 million people per year (Murray et al., 2012). Another group, the Coccidians are parasites that infect the intestinal tract and are transmitted by the fecal-oral route (Adl et al., 2005). The Coccidians include *Toxoplasma* responsible for toxoplasmosis and *Cryptosporidium* causing the diarrheal disease cryptosporidiosis, both diseases are often severe in immunocompromised individuals (Current and Garcia, 1991; Esch and Petersen, 2013).

Apicomplexans are intracellular parasites. Unlike other organisms, they undergo drastic changes in shape during the transition between hosts, in general vertebrates and invertebrates. Their life cycle is divided between sexual and asexual reproduction. Although they reside in a variety of host cells, from dendritic cells to erythrocytes, the mechanism of invasion is conserved (Huynh et al., 2004). Unlike other pathogens that are “taken up” by or induce phagocytosis (i.e. *Leishmania* and *Trypanosoma cruzi*), Apicomplexa actively invade their host cells (Burleigh and Woolsey, 2002; Ueno and Wilson, 2012). Invasion follows a series of events: first, parasites propel themselves across the surface of the host cells through substrate-dependent motility termed “gliding motility” they then initiate attachment through receptor-ligand interactions and finally penetrate the cell by forming a moving tight junction with the host plasma membrane (Keeley and Soldati, 2004; Sibley, 2010; Suss-Toby et al., 1996; Wetzel et al., 2004). The formation of a viable intracellular niche relies on the secretion of parasite factors from the apical

complex, a set of electron dense specialized secretory organelles localized as their name suggests at the apical end of the cell, during invasion (Carruthers and Sibley, 1997; Chen et al., 2004). These are the micronemes, rhoptries, and dense granules. After invasion, parasites remain free in the host cytoplasm (i.e. *T. parva)*, while others immediately establish an intracellular niche in a non-lytic parasitophorous vacuole (PV) sometimes for a very short time, a matter of minutes in the case of *Babesia bovis*, or until egress more than 40 hours later as in *P. falciparum* infections (Asada et al., 2012; Wickham et al., 2003). Parasites then go through multiple rounds of replication to form schizonts. They then induce the release of merozoites in the extracellular space, allowing further invasion of susceptible cells and thus closing the asexual cycle. Parasites can also differentiate into sexual stages, female and male gametes, which will then fertilize to form oocysts or, less frequently invasive stage zygotes (ookinetes). Ookinetes then undergo one round of meiosis to release haploid invasive sporozoites, while oocyst forms are spread in the environment.

II. Insight from sequencing apicomplexan genomes

The recent explosion of genomics-based research has revealed just how uniquely adapted apicomplexan parasites are to their host, but also how closely-related parasites may have developed similar mechanisms to survive host and even vector challenges. Comparison of genomes from different organisms, those closely and even distantly related has allowed the scientific community to identify homologs and therefore extrapolate gene function and pathways from known organisms to those recently sequenced. Classical pathways such as DNA replication, mitochondrial energy efflux and protein secretion have been easily discovered via comparative genomics. In particular, the assembly of metabolic pathways that occur in the apicoplast revealed for the first time the essential nature of this unique organelle for apicomplexan parasite survival in the host (Seeber and Soldati-Favre, 2010). Genomic data has also provided clues into virulence processes. For apicomplexan parasites, members of large gene families are often found in

tandem at the chromosome ends and these genes are often undergoing diversification and/or involved in antigenic variation (Brayton et al., 2007; Gardner et al., 2002). In *Plasmodium,* this has been well characterized with *var* and *rif* genes involved in parasite immune evasion and host cell binding and in *Babesia* with the variably expressed surface antigen family, VESA (Aley et al., 1984; Brayton et al., 2007).

Sequencing data has also revealed high degree of conservation within the Apicomplexa phylum. *Toxoplasma, Cryptosporidium* and *Eimeria* species all express orthologs of a *Plasmodium* membrane protein, thrombospondin-related anonymous protein or TRAP, involved in gliding motility, an important process preceding invasion characteristic of parasites of this phylum (Kappe et al., 1999). However, many genes in apicomplexan parasites have no known homologs in non-apicomplexan species and remain functionally uncharacterized. New discoveries such as the vestigial apicoplast has launched the development of compounds for potential anti-malarial chemotherapies and more recently discoveries of *Plasmodium* export pathways has opened up doors to new ways of identifying and targeting parasite gene products that may play a role in pathogenesis and ultimately disease outcome (Fichera and Roos, 1997).

III. Apicomplexan pathogenesis

A. Apicomplexan pathogenesis: remodeling the host cell

Apicomplexa parasites are successful at establishing infection and surviving in the host. Within the host, parasites must overcome multiple barriers, including mucus layers, blood-brain barriers, and the epithelia, often eliciting at the same time host immunological responses and physical stress (Plattner and Soldati-Favre, 2008). The damages caused by these parasites in their respective hosts have also been felt in both human and livestock deaths and economic losses. There are no licensed vaccines against human apicomplexa diseases to date. Thus, it is important to understand the mechanisms employed by these parasites to survive and assure

propagation of their progeny. These pathogenic cellular processes are sometimes parasite-specific but in general involve dramatic changes to the host cell.

Invasion by Apicomplexa parasites and the establishment of residence in the infected cell is accompanied by extensive remodeling of the host cell. In *P. falciparum* infections, these events include the dramatic loss of the red blood cell normal discoid shape associated with changes in membrane rigidity, the creation of new permeation pathways to adapt to nutrient deprivation and the presentation of parasite ligands at the surface for cytoadherence (Plattner and Soldati-Favre, 2008). Remodeling, in fact, begins right at invasion when the content of the apical complex of invading merozoites is secreted to form the PV. For parasites like *T. gondii*, parasitophorous vacuole membrane (PVM) remodeling is crucial for survival as tachyzoite stages infect almost any nucleated cell. PV remodeling allows for the escape from lysosome fusion with the PV and vacuole acidification. Assuredly, during invasion host transmembrane (TM) proteins are excluded from the invaginating membrane, thus resulting in a non-fusogenic vacuole (Mordue et al., 1999). Sorting of host TM proteins also occurs in *P. falciparum* invasion of erythrocytes (Lauer et al., 2000).

Whereas micronemes initiate apical attachment and entry in the host cell, rhoptries and dense granules are released and mainly modify the host cell during invasion. Rhoptry proteins and membranous material are released directly into the host cytosol as vesicles, which then coalesce and fuse with the invaginating host plasma membrane (PM) that surrounds the invading parasite (Hakansson et al., 2001). Some of these proteins are immunosuppressive and allow parasites to evade immunity. In *T. gondii* mouse infection, the rhoptry bulb serine-threonine kinase TgROP18 is secreted in complex with membrane whorls in the vacuolar space and quickly associates with the PVM (El Hajj et al., 2007). This kinase phosphorylates p47 immunity related GTPases-IRGa6 and IRGb6, which in ROP18-deficient parasites form oligomers that bind the PVM, hence destabilizing the vacuole and killing the parasite (Taylor et al., 2006). On top of its immunosuppressive role, TgROP18 also enhances parasite growth and replication thus aiding parasite virulence (El Hajj et al., 2007; Taylor et al., 2006).

Secretion from dense granules on the other hand, is thought the mediate the maturation of the PV and parasite replication and occurs within the first hours of PV formation (Carruthers and Sibley, 1997; Dubremetz et al., 1993; Torii et al., 1989). The *P. falciparum* vaccine candidate Pf155/RESA (ring-expressed surface antigen) a resident dense granule protein, binds and alters host cell PM physical properties (Aikawa et al., 1990). In fact, during invasion RESA is incorporated in the cytoskeleton of infected cells by forming stable bonds with spectrin tetramers. These interactions increase resistance to heat shock and increase host cell membrane rigidity, which ultimately suppresses any additional invasion events (Pei et al., 2007a). Other dense granule components include the putative *P. falciparum* translocon for exported protein or PTEX (de Koning-Ward et al., 2009). This macromolecular complex consists of at least five proteins, two hypothetical proteins PTEX 150 (PTEX-150) and PTEX 88 (PTEX88), a heat shock protein 101 (HSP101), thioredoxin 2 (TRX2) and exported protein 2 (EXP2), a bacterial like pore forming protein (de Koning-Ward et al., 2009). This adenosine triphosphate (ATP)-powered translocon is released as early as five minutes after invasion and anchors in the membrane of the PV (Bullen et al., 2012; Riglar et al., 2013). Proteins that are synthesized when the parasite is completely encapsulated and destined to localize beyond the PVM are thought to translocate through this pore, which begins to be functional in ring stages.

Considerable changes in cell permeability also occur during *Plasmodium* erythrocyte infections. The permeation of the RBC membrane and the PVM allows import of otherwise non-available amino acids that *Plasmodium* requires for growth (*Plasmodium* is auxotrophic for purines) and metabolized hemoglobin as an energy source (Decherf et al., 2004; Desai et al., 2000; Gero and O'Sullivan, 1990; Martin and Kirk, 2007). Import of organic and inorganic solutes was recently shown to occur through a parasite-derived channel, the Plasmodial Surface Anion Channel (PSAC) composed in part by the parasite cytoadherence-linked asexual protein 3 (CLAG 3) formally thought to play a role in cytoadherence (Desai, 2012; Nguitragool et al., 2011). Host membrane permeability changes also occur in *T. gondii* infections and this parasite in common with *Plasmodium,* cannot synthesize purines *de novo.* Evidently, parasite pores and transporters are embedded in the PVM and allow the efflux of small molecules (Schwab et al., 1994).

As *Plasmodium* develops in a truly bona fide parasitophorous vacuole, other structural changes can be observed. These include the tubulovesicular networks (TVN), extensions of the PVM into the host cell periphery and the Maurer's Clefts (MCs) membrane-derived secretory compartments where parasite exported proteins transit before transferring to the host cell surface (Lauer et al., 1997; Tilley et al., 2008). In growing trophozoite stages, the parasite also exports factors to the host cytoplasm and surface including immunovariant adhesins *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). Surface expression of PfEMP1 is a crucial process for *Plasmodium* as it allows trophozoites and schizonts to avoid splenic clearance by adhering to the endothelium of microcapillaries in various tissues through receptor-ligand interactions. This process activates the endothelium igniting local pro-inflammatory immune responses (Turner et al., 1998; Turner et al., 1994). In the brain, endothelial activation is thought to lead to a signaling cascade where host adhesion molecules such as intracellular adhesion molecule 1 (ICAM-1) are upregulated, amplifying the number of parasitized RBCs binding to brain endothelial cells and pro-inflammatory cytokines and radicals are released in the local microenvironment causing the disruption of the blood-brain barrier (Adams et al., 2002).

Host cell remodeling is a very important process in Apicomplexa infections, in particular *Plasmodium* malaria. Changes to the RBC require secretion and sorting of host and parasite proteins at invasion and during development. It is important to note that differences in genetic backgrounds of host cell types can affect the extent of membrane modifications undertaken by the parasite in the host and therefore affect the extent of pathogenesis. In fact, the hemoglobinopathy found in sickle cell anemia patients abolishes trafficking of parasite factors to the RBC surface (Cyrklaff et al., 2011). Therefore overall, Apicomplexa pathogenesis and ultimately disease prognosis, is a complex system that is composed in part by parasite and host factors.

B. *Plasmodium falciparum* pathogenesis during cerebral malaria

P. falciparum malaria caused over 600,000 deaths and over 219 million infections in 2010 alone. The burden of this disease is mostly carried by children under the age of 5 in sub-Saharan Africa (over 85% of all deaths) (WHO, 2012). Pedriatic malaria often manifests itself as CM and this is partly due to immune naiveté (Marsh et al., 1995). The World Health Organization defines CM as the presence of *P. falciparum* parasitemia and coma with no other apparent causes of altered consciousness (WHO, 2000). In field studies, there is a high false positive rate of children meeting CM clinical definition due to low specificity of diagnosis and due to the heterogeneity of CM cases (Taylor et al., 2004). In fact, children whose coma is attributable to CM alone (and to no other common infections such as bacterial meningitis) and who died of clinically defined CM, can be divided in three distinct groups: CM1, presence of iRBC sequestration only, CM2 presence of iRBC sequestration with intravascular and perivascular pathology (microthrombi and hemorrhages), and CM3 absence of sequestration (Dorovini-Zis et al., 2011). Currently, the combination of peripheral blood and ophthalmoscopic detection (or retinopathy) of *P. falciparum* iRBCs in retinal vessels as a discriminant of CM and severe malaria without any histological evidence of cerebral damage has increased the reliability of CM diagnosis significantly (Lewallen et al., 2000; White et al., 2001). The definitive diagnosis so far (CM with a positive retinopathy result) can only be reached through a serological test for plasma concentrations of a secreted parasite protein, histidine-rich protein 2 (HRP2) (98% sensitive; 94% specific) or by post-mortem examination of the brain for the presence of sequestered parasites (Seydel et al., 2012; Milner et al., 2012).

The cause of CM is still not well defined due to the heterogeneity of CM cases diagnosed (as described above) and conflicting data. However, it is well accepted that blood-brain barrier (BBB) dysfunction is involved in CM progression and has been described in pediatric CM (CM1-3, although CM3 are thought to have died of causes not necessarily linked to a central nervous system etiology) and experimental models (Dorovini-Zis et al., 2011). The BBB is a selectively permeable structure that lines cerebral blood vessels and is composed of specialized endothelial cells (ECs) and astrocytes. It regulates the passage of ion and nutrients as well as other

molecules from the blood stream into to the brain parenchyma (Abbott et al., 2006). During CM, the BBB is damaged and brain homeostasis is no longer maintained. The mechanism leading to BBB dysfunction is a complex system but includes an inflammatory response, mechanical occlusion and finally BBB breakdown. Parasitized RBCs cytoadhere to ECs through PfEMP1-ICAM-1 interactions (Berendt et al., 1992; Newbold et al., 1999). These interactions stimulate the release of pro-inflammatory cytokines such as TNFα, which activates local endothelial cells resulting in endothelial ICAM-1 upregulation (Turner et al., 1994). The process is potentiated through the recruitment of platelets to the endothelium as these bridge iRBC binding to the endothelium (Wassmer et al., 2004). The accumulation of adherent parasites and platelets ultimately occludes blood vessels resulting in focal ischemia and hypoxia (Newton et al., 1994). As damaged cells undergo apoptosis, the integrity of tight junctions between ECs is disrupted and the BBB finally breaks down (Shikani et al., 2012; Susomboon et al., 2006). This allows foreign and toxic plasma material to permeate the parenchyma, activating macrophages and leukocytes of the central nervous system (CNS) and leading to edema, intracranial pressure build up and coma (Brown et al., 1999). In children with CM, BBB dysfunction can cause life-long neurological sequelae such as motor disorders (hemiparesis and quadriparesis) as well as blindness, epilepsy and aphasia (Brewster et al., 1990).

The sequestration of parasites in the brain is thought to occur through iRBC binding to ICAM-1 in cerebral vessels specifically through PfEMP1 encoded by a unique set of UpsA type *var* genes. This was first shown in field studies where iRBCs from CM patients co-localized with ICAM-1 in cerebral vessels post-mortem and showed selective preference for ICAM-1 binding as opposed to CD36 which can bind almost every parasite and has limited expression in the brain (Newbold et al., 1999; Turner et al., 1994; Hasler et al., 1990). Recently, the UpsA encoded PfEMP1s were identified by surveying *var* gene expression in parasites after panning human cerebral tissue and parasites from Tanzanian children with CM (Avril et al., 2012; Claessens et al., 2012; Lavstsen et al., 2012).

A running hypothesis in malaria, and this is also true for other diseases, stipulates that disease severity is due to the presence of multiple parasite genotypes in the human host in

particular during seasons of high transmission. However, in two retrospective studies on children who meet the clinical definition of CM, only single genotypes can be detected and these are the same circulating in peripheral blood, in the brain and in multiple other organs. Mixed infections however, associated with malaria patients who died of non-malarial causes. Notably, in these studies, not one single parasite genotype was associated with malaria (Milner et al., 2012; Montgomery et al., 2006). This suggests that CM is caused by other factors then mixed *P. falciparum* infections and favors the role of a parasite pathogenic process such as antigenic variation through which switching of surface expressed PfEMP1 determines the route to severe disease (CM and PAM) versus uncomplicated malaria.

IV. Protein trafficking in apicomplexan parasites

Taking-over the host cell is not an easy task. The many changes that occur in the host cell, scavenging of nutrient sources and even hijacking of members of host cell pathways requires unrequested access to the host cell. For this, protein secretion of parasites factors beyond the confines of the PVM is an essential process not only for growth but also immune evasion. Crossing multiple membranes in particular in transcriptionally and protein biosynthesis inactive cells like mature erythrocytes for the Haemosporidians *Babesia* and *Plasmodium* can be challenging. In *Plasmodium falciparum* it is estimated that approximately 8-10% of proteins (about 400) are exported and a large number have not been functionally characterized (Boddey et al., 2013; Sargeant et al., 2006). To understand protein targeting to the host, we look back at protein trafficking to the apicoplast for clues as to understand how parasite factors make it across multi-membranous structures.

A. Lessons from the past: targeting to the apicoplast

Apicomplexan parasites contain a unique organelle called the apicoplast. The apicoplast is a relic plastid that originated from endosymbiont green algae (Kohler et al., 1997). Four membranes

envelop the apicoplast, with the outermost originating from the endomembrane system. Interestingly, this plastid is found in all Apicomplexa except *Cryptosporidium.* This organelle contains its own circular genome, one that has evolved to retain only some essential function from the ancestral bacteria. In fact, the genome has shrunk from 150 to 35kb losing in the process photosynthetic and nucleomorph functions (Marechal and Cesbron-Delauw, 2001). The apicoplast is a vital organelle for api-parasites. It functions in metabolic processes such as fatty acid, heme and isoprenoid synthesis as well as general housekeeping function to maintain plastid gene expression. This plastid's genome however does not play a large role in the organelle's function. In fact, the apicoplast proteome is largely encoded by the nuclear genome (Waller et al., 1998). Gene products entering the secretory pathway are targeted post-translationally to the apicoplast through the recognition of bi-partite N-terminal leader sequences consisting of a classical secretory signal sequence followed by a plastid transit peptide. Genetic studies in *T. gondii* have shown that the transit peptide is necessary and required for apicoplast targeting, as the expression of the signal sequence of an apicoplast protein alone targets a fluorescent marker to the PV (DeRocher et al., 2000; Harb et al., 2004). Similarly, the expression of the signal sequence from a secreted protein, erythrocyte binding antigen-EBA 150, in lieu of the signal sequence from an apicoplast-targeted protein properly targets to the plastid (Tonkin et al., 2006). Consistently, in *Plasmodium* similar conclusions were made by site-directed mutagenesis of apicoplast putative targeting domains (signal sequence followed by plastid transit peptide) of a nuclear-encoded acyl-carrier protein (Foth et al., 2003).

The path of nuclear encoded proteins to the apicoplast is challenging. Proteins have to cross all four membranes of this relic organelle. Several models have been proposed to explain the translocation process of proteins in the secretory pathway to the apicoplast; some include the apicoplast residing in the ER, others through vesicular trafficking from the Golgi apparatus. Current models propose that upon entering the ER, the signal peptide is cleaved by the signal peptidase, exposing the transit peptide. ER-mediated vesicles harboring these proteins bud from the ER membrane to fuse with the apicoplast outer-membrane, bypassing Golgi entry. This was supported by data in *T. gondii* and *P. falciparum* using Brefeldin A as an inhibitor of trafficking

between ER and Golgi (DeRocher et al., 2005; Tonkin et al., 2006). Upon fusion, the content of the vesicles is released in the luminal space between the two outermost membranes. The next translocation steps across the membranes remain controversial in part because of lack of experimental data. However two models exist. The first “free-crossing” stipulates that proteins traverse the membranes through Tic and Toc translocons with the help of ATPase chaperone proteins; the second relies on vesicular migration across the membranes (Foth and McFadden, 2003; Kalanon et al., 2009). Once the proteins reach the lumen of the apicoplast, the transit peptide is cleaved (probably by an apicoplast homolog of the plant stromal processing-protease) to release mature proteins (Parsons et al., 2007; van Dooren et al., 2001).

The plastid transit peptides do not bear specific motifs as compared to other targeting sequences like the ER-retention signal KDEL (lys-asp-glu-leu). However, like other plant transit peptides a secondary structure is evident: they are 25-125aa in length and rich in basic and hydroxylated amino acids in particular serines and threonines and are found at the N-terminus downstream of the signal sequence (von Heijne and Nishikawa, 1991). The topology, sequence and position of apicoplast transit peptides have been used to predict not only the number of proteins targeted to the organelle but also to predict their function. Interestingly, comparative analysis of apicoplast pathways, in organisms where sequencing has been performed, shows that the function of this organelle has specified along lineages. As such, enzymes involved in heme and fatty acid metabolism in *Plasmodium spp.* are not found in *Theileria spp.* (Sato, 2011). In terms of isoprenoid metabolism, it is notable that species that have retained this pathway reside in a PVM, i.e. *Toxoplasma* and *Plasmodium,* whereas those that lost it remain free in the host cell cytosol, i.e. *Theileria parva* and *Theileria annulata* (Gardner et al., 2005). This is thought to be a result of host cell tropisms and hints that the apicomplexan phylum has had to adapt to host and possibly even vector challenges.

B. Targeting to the host cell: *Plasmodium falciparum* protein export

Intracellular parasites export multiple proteins in the host cell environment. Exported proteins play a major role in host cell remodeling, cytoadherence, immune evasion and even nutrient scavenging. Therefore, export is essential for parasite survival. In fact, a large genetic deletion screen for parasite genes involved in host cell remodeling identified genes that mediate cytoadherence, affect host cell membrane rigidity and even genes essential for *in vitro* culture growth (Maier et al., 2008). These genes could also likely be important for parasite growth and survival *in vivo* as changes in host cell rigidity and cytoadherence play a major role in pathogenesis. However, the process itself is challenging for Apicomplexa parasites that reside within the confines of a PV. Exported proteins need to cross the parasite plasma membrane as well as PVM. The best-studied system among apicomplexans with regard to protein secretion and export into the host cell is the human malarial parasite, *Plasmodium falciparum*.

Export of *Plasmodium* proteins is mediated by the recognition of a N-terminal bipartite leader sequence composed of the signal sequence followed by a pentameric export motif RxLxE/Q/D. This export motif has been called HTS (Host Targeting Signal) or PEXEL (*Plasmodium* Export Element) (Hiller et al., 2004; Marti et al., 2004). Several hundred proteins were predicted by bioinformatics to contain this positionaly conserved export motif (Sargeant et al., 2006; Boddey et al., 2013). The PEXEL motif was functionally validated to play a role in export by tracking the localization of GFP-tagged N-termini or mutant versions of few candidate proteins (some previously known to be out in the RBC and others unknown) (Boddey et al., 2013; Sargeant et al., 2006; van Ooij et al., 2008). The PEXEL motif is found on both soluble and membrane proteins and is located 20-30 amino acids downstream of the signal sequence. Unlike apicoplast-targeted proteins, most PEXEL proteins possess a recessed N-terminal signal sequence. The PEXEL motif is cleaved in the ER by an ER-residing protease called plasmepsin V (Boddey et al., 2010; Russo et al., 2010). Cleavage was shown to occur after the well-conserved Leucine at position three, releasing a mature protein with a starting N-terminal sequence of xE/Q/D (Chang et al., 2008). Unprocessed proteins in the ER were also shown to bypass plasmepsin V cleavage. Bhattacharjee and colleagues recently reported that the PEXEL

motif binds phosphatidylinositol-3-phosphate (PI(3)P) in the ER prior to and independent of protease processing (Bhattacharjee et al., 2012a; Bhattacharjee et al., 2012b). Binding occurs in PI(3)P-rich regions of the ER. The malarial PEXEL was previously shown to resemble both in sequence and function an export motif found in the eukaryote *Phytophtora infestans,* of sequence RxLR...DEER and also bound PI(3)P, thus this finding was not entirely surprising (Bhattacharjee et al., 2006; Whisson et al., 2007). Unprocessed and mature protein are then thought to traffick through vesicles to the parasite PM and secreted into the PV space. Here, proteins bind to chaperone proteins HSP101 and are escorted to an ATP-driven translocon complex (de Koning-Ward et al., 2009). Agreeably, proteins need to be unfolded in order to pass through the pore (Gehde et al., 2009). This process is true for soluble proteins but remained unresolved for proteins containing TMs until recent work on PNEPs.

In addition to signal-mediated targeting to the host, a number of proteins in *Plasmodium* are exported but do not contain a PEXEL motif. These proteins, called PEXEL-negative proteins, or PNEPs include REX1, REX2, MAHRP1, MAHRP2 and SBP1 which all localize to the MCs (Spielmann and Gilberger, 2010). PNEPs, unlike PEXEL proteins, do not have a signal peptide but contain a hydrophobic domain localized some 200 amino acids from the N-terminus (Haase et al., 2009; Spielmann et al., 2006). For a long time it wasn't known how PEXEL proteins with TM domains or PNEPs were exported across the PM and the PVM. Recently, it was shown that both PNEP and PEXEL-dependent trafficking routes converge. Precisely, Grüring and colleagues showed that a matured-PEXEL protein with a PNEP TM is exported in the host. The same was true for the addition of a PEXEL domain (signal sequence + PEXEL) to a PNEP protein. In both scenarios, the authors saw a protein solubility shift in fractionation experiments. This was evidence for a translocation event (Grüring et al., 2012). Thus, it is proposed that PNEPs and PEXEL proteins with TMs are trafficked through vesicles to the PM where they are inserted; they are then unfolded and bound by a chaperone (potentially HSP101) and handed to the translocon. This model also implies that both PNEP and exported TM proteins take the same route as soluble PEXEL proteins to enter the host cytoplasmic environment.

C. Conclusion

There are similarities between apicoplast-targeting and targeting to the host cell. In the case of *Plasmodium falciparum*, proteins entering the ER are rapidly sorted to enter different routes. The fork to the apicoplast requires recognition of basic N-terminal peptide early in the secretory process giving proteins eligibility to enter the four-membrane plastid through translocons. The fork to the red blood cell on the other hand is somewhat more diversified. A translocon mediates the translocation of proteins across the PVM but the recognition of a N-terminal peptide the PEXEL motif is sufficient but not required. In all cases, more work needs to be done to identify the players that function at the many steps of trafficking: chaperones, membrane receptors, enzymes, etc. n addition, the travels of proteins into the host environment in other encapsulated Apicomplexa parasites is not well-understood.

V. Mechanisms of Antigenic variation in *Plasmodium falciparum* malaria

The progresses in immunology these last decades have showed the amount of diversity that can be generated in antibody responses. A mammalian host can produce up to 10^10 distinct antibody idiotypes via class switch recombination and maturation. These two sophisticated processes rely on the shuffling of conserved and adaptable immunoglobulin domains (Alberts B, 2002). With the addition of immunological memory, not only at the individual level but at the population level, one is left to ask how are parasites able to escape from sterilizing host defenses when the host has likely been previously exposed to the same or similar antigens. One strategy is antigenic variation. Antigenic variation is the presentation of ever-changing population phenotypes to assure transmission and evasion from host immunity. For this reason, antigenic variation is a complex process. The parasite, with its limited genome, must continually switch from one antigenic type to the next while avoiding exhaustion of its own antigen repertoire. There are many mechanisms that are known to govern antigenic variation some are controlled at the transcriptional level by genetic processes and others by epigenetics.

A. Variant surface antigens

The first indication of antigenic variation in the malaria field was by Cox in 1959 where he observed relapses in *P. berghei* infections in mice (Cox, 1959). This was later confirmed in the context of rhesus monkey malaria, and falciparum malaria in squirrel monkeys (Brown and Brown, 1965; Hommel et al., 1983). The relapses observed in these experiments have been linked to the immune system's recognition of antigens from one *Plasmodium* family, which is expressed on the surface of infected erythrocytes, PfEMP1 (Leech et al., 1984). This clonally variant gene family is composed of about 60 members that alter the antigenic and cytoadherent properties of the host cell and ultimately pathology (Kraemer and Smith, 2006; Smith et al., 1995). However, PfEMP1 is not the only antigen expressed at the surface of infected erythrocytes. Other proteins have been found on the host cell surface and play an important role in adherence and changes to host cell deformability. The knob associated histidine rich protein, KAHRP, for instance is localized underneath the host cell membrane at "knobs", parasite-induced protrusions of the cell membrane where other parasitic proteins like PfEMP1 accumulate (Horrocks et al., 2005). In addition, transcriptional data has provided proof of variegated expression of other genes than *vars* in both culture-adapted strains and peripheral blood of malaria disease individuals. Such genes include the two-transmembrane domain (2TM) superfamily composed of *rifins, stevors,* and *Pfmc-2TM*. All three families encode proteins that are exported to the host cell membrane. They contain two-transmembrane topology and regions of hypervariability in the loops between the TM domains (Lavazec et al., 2006). These regions face the extracellular side of the host membrane and this diversity is thought to have arisen through selection by the immune system (Lavazec et al., 2006). *Rif* or repetitive intersperced family, encode more than 130 RIFINSs and are found on the host cell membrane. They are clonally variant with a switching rate close to that of *var* genes (Fernandez et al., 1999; Kyes et al., 1999). *Stevor* or subtelomeric variable open reading frame, is a smaller gene family of about 40 members and related in sequence and structure to *rifins*. They are also differentially expressed in various clonal lines (Lavazec et al., 2007). Interestingly, STEVORs were found to co-localize with MSP-1 in free merozoites and also at the "docking" site of merozoites as they invade erythrocytes. However

their role in invasion is still unclear (Blythe et al., 2008; Khattab et al., 2008). *Pfmc-2TM*, the Maurer's cleft two-transmembane gene family contains 13 members. As their name indicates, they localize to the MCs but also the host surface and are clonally variant (Lavazec et al., 2007). The function of the 2TM superfamily in immune evasion and disease is still unclear.

B. Genetics of variant expression

1. Partitioning the genome: importance of the telomeres

The chromosomes of *Plasmodium falciparum* are partitioned like other eukaryotic genomes in internal regions and chromosome ends called telomeres. The telomeres and subtelomeres are regions of high instability, where deletions and recombination with other chromosomal ends occur. Chromosomal internal loci on the other hand are typically stable and enriched with genes involved in growth and maintenance i.e. ribosomal subunit, RNA helicase, and eukaryotic translation factor (Chen et al., 1998; Cortes et al., 2007; Freitas-Junior et al., 2000; Le Roch et al., 2003; Volkman et al., 2002). Telomeres are known for the presence of species-specific repetitive sequences that are ideally suited for recombination events between chromosomal ends. *Rep20* is a sequence of 21bp found at all 28 telomeres of *Plasmodium*, these sequences are thought to mediate the clustering of chromosome ends thus enhancing the chance of ectopic recombination (Freitas-Junior et al., 2000; O'Donnell et al., 2002). Recombination at the subtelomeres is a phenomenon that has occurred and is still ongoing in *var* genes. Such recombination leads to an accelerated generation of diversity with the shuffling within and between *var* domains, creating variants with potential novel cytoadhesive properties and immunogenicity (Kraemer et al., 2007). Hence, the generation of variants and their controlled transcription can be thought as a source of antigenic variation.

2. Transcriptional regulation

Plasmodium falciparum contains the canonical transcriptional machinery in common with other eukaryotes however; global transcriptional analysis has revealed highly coordinated gene expression during the intraerythrocytic cycle and gametocyte development (Bozdech et al., 2003). Waves of gene expression from time zero of invasion to egress of rupturing schizonts show that most of the genome is wired for sequential activation of gene pathways. However, a fraction of the genome shows patterns of differential expression in both clinical and culture-adapted isolates. The absence of chromosomal clustering of genes with similar expression patterns (across the lifecycle and in primary isolates) indicates that genes are independently regulated (Gardner et al., 2002). Such regulation at the transcriptional level requires transcriptional factors and other DNA-binding proteins as well as cis-acting elements at the 5'UTR. Only in the last few years have some strong candidates been found.

ApiAP2 transcriptional factors were recently discovered in *Plasmodium.* There are 28 copies of this gene in the genome and they are only found in Apicomplexa parasites like *T. gondii*, *C. parvum* and *T. annulata* (Iyer et al., 2008). ApiAP2s are expressed at different points in the intraerythrocytic life cycle thus making them likely players in various transcriptional events across different parasite stages (Balaji et al., 2005). So far, studies from two members of this gene family, PF14_0633 and PFF0200c, indicate that they bind specific palindromic promoter motif, TGCATGCA and GTGCAC respectively (De Silva et al., 2008). Surprisingly, PFF0200c, which contains two copies of the AP2 domain, is co-expressed with genes that harbor the motif in their 5'UTR, whereas the same is not true for PF14_0633, which only carries one AP2 domain (De Silva et al., 2008). This suggests that the Ap2 domain serves in the binding of *cis*-acting elements in the 5'UTR and probably also in the recruitment of other DNA-binding protein for gene regulation.

C. Epigenetic mechanisms of variant expression

Changes in gene expression can occur without alterations of the primary sequence of a gene. When these changes result in stably heritable phenotypes, they are defined as being under epigenetic control (Berger et al., 2009). The current understanding of such a phenomenon requires three components: an “epigenator”, transient environmental signals that initiate a response at the chromosomal level; an “epigenetic initiator”, factors that relegate signals to the chromosome (sometimes changing chromosome location in the nucleus); and “epigenetic maintainers”, factors which prolong the epigenetic effects beyond the lifespan of the affected cells (Berger et al., 2009).

In *P. falciparum*, epigenetic processes have been mostly described for the coordinated expression of *var* genes. It was shown that the observed monoallelic expression patterns were not due to changes at the DNA sequence nor the presence or absence of transcriptional factors when comparing a *var* gene at its active and silent states, but rather likely due to chromosome structure and promoter accessibility (Chen et al., 1998; Scherf et al., 1998). In fact, the expression or silencing of *var* genes and other variant antigens such as the *rifs* and *stevors* is associated with the chromatin structure of the region surrounding their promoters (Duraisingh et al., 2005; Freitas-Junior et al., 2005). Heterochromatin and euchromatin marks, or when the chromosomes are inaccessible or free to proteins involved in transcription, are many in eukaryotes and include histone posttranslational acetylation, methylation and ubiquitination (Cui and Miao, 2010). In *Plasmodium*, active *var* expression sites are associated with acetylation of histones H2 and H3, whereas silent loci with methylation of lysine 27 of histone 3 (H3K27) and tri-methylation of Lysine 9 of histone 3 (H3K9me3) (Chookajorn et al., 2007; Duraisingh et al., 2005; Freitas-Junior et al., 2005; Lopez-Rubio et al., 2007). It is believed that a yeast homologue of the sirtuin 2 protein, PfSir2, acts as a histone deacetylase and mediates the silencing of *var* promoters (Duraisingh et al., 2005; Freitas-Junior et al., 2005; Tonkin et al., 2009). This same protein acts in the process of heterochromatin spread starting at the telomeres spreading into internal regions of the chromosome by binding to telomere repeats *Rep20*, often found adjacent to *var* promoters (Freitas-Junior et al., 2005). In yeast, Sir2-mediated silencing spreads internally

by 3kb whereas in *Plasmodium* the region extends some 55kb (Freitas-Junior et al., 2005; Imai et al., 2000)! This spread is thought to control *var* gene expression as well as other genes located at the subtelomeres. In fact, the silent states of other variantly expressed genes, such as the *rifs* and *stevors* is associated with the heterochromatin mark H3K9me3 (Lopez-Rubio et al., 2009). In addition to histone posttranslational modifications, activation and repression of variant antigens is mediated by chromosomal repositioning to specific subnuclear regions. The data from *var* expression studies shows that while PfSir2 clusters (silent loci) are localized at the nuclear periphery in specific foci where chromatin is condensed forming the so called chromosome "bouquets", only the chromosome bearing the active *var* gene detaches from the PfSir2 bouquet to position itself at nuclear pores in an uncondensed, PfSir2-free state (Duraisingh et al., 2005; Freitas-Junior et al., 2005; Ralph et al., 2004).

Thus, one can see how variant expression can occur via epigenetic mechanisms at chromosome ends to regulate the expression of genes that play a role in antigenic variation. Interestingly, recent genetic studies have provided some evidence for co-regulation of *vars* and the three 2TM gene families. Specifically, the over-expression of a promoter from one family affects in a titratable manner the expression of genes from all four families, indicating that there are factors that bind regions common to these families (Howitt et al., 2009). It is still unclear how histone modification processes reach internal loci where for example some members of *var* genes, the UpsC type, are found. However, H3K9me3 marks have been found in some discrete intrachromosomal regions where members of multi-copy gene families also cluster (Lopez-Rubio et al., 2009; Salcedo-Amaya et al., 2009).

VI. Summary of Aims

This dissertation seeks to elucidate mechanisms of protein secretion in closely related parasites of the Apicomplexa phylum generally and in *Plasmodium falciparum* specifically, to develop methods to systematically identify putative antigens that function together in active disease. In the first chapter, by using comparative genomics and forward genetics, we aim to identify *P. falciparum*-like N-terminal export sequences together with the genes that harbor them in Apicomplexa parasites such as *Toxoplasma, Babesia* and *Cryptosporidium,* which like *Plasmodium spp*. reside in a parasitophorous vacuole for parts of their life cycle. Some of these exported genes are likely to play a major role in host cell remodeling (Maier et al., 2008; Plattner and Soldati-Favre, 2008).

In the next chapter, we focus on *P. falciparum* malaria to define gene co-expression modules in a severe malaria paradigm and explore their role in host-cell remodeling. Here, we work on two hypotheses: first, the functional genomics hypothesis that co-expressed genes function in the same biological pathway, second, that genes involved in severe disease are differentially expressed *in vivo* as opposed to *in vitro* culture conditions and can be systematically identified based on predicted export sequences. Many studies have looked at *P. falciparum* gene expression, using clinical isolates and culture-adapted parasites (wild-type or with perturbations), a small number have identified gene modules that function in invasion and growth, but none have expression data from clinical strains to identify gene modules relevant during disease (Date and Stoeckert, 2006; Hu et al., 2010). Finally, we apply analytical methods similar as those used at the genome level above on a small gene family called the *Plasmodium* Helical Intersperced Subtelomeric gene family a (*phista)* to characterize the role of *phista* genes in the pathogenic process of cytoadherence. Through these studies we aim to show the importance of antigenic secretion in the pathogenesis of *P. falciparum* parasites.

VII. References

Abbott, N.J., Ronnback, L., and Hansson, E. (2006). Astrocyte-endothelial interactions at the blood-brain barrier. Nature reviews. Neuroscience *7*, 41-53.

Adams, S., Brown, H., and Turner, G. (2002). Breaking down the blood-brain barrier: signaling a path to cerebral malaria? Trends in parasitology *18*, 360-366.

Adl, S.M., Simpson, A.G., Farmer, M.A., Andersen, R.A., Anderson, O.R., Barta, J.R., Bowser, S.S., Brugerolle, G., Fensome, R.A., Fredericq, S.*, et al.* (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. The Journal of eukaryotic microbiology *52*, 399-451.

Aikawa, M., Torii, M., Sjolander, A., Berzins, K., Perlmann, P., and Miller, L.H. (1990). Pf155/RESA antigen is localized in dense granules of Plasmodium falciparum merozoites. Experimental parasitology *71*, 326-329.

Alberts B, J.A., Lewis J, et al (2002). Molecular Biology of the Cell., 4th edn (New York: Garland Science).

Aley, S.B., Sherwood, J.A., and Howard, R.J. (1984). Knob-positive and knob-negative Plasmodium falciparum differ in expression of a strain-specific malarial antigen on the surface of infected erythrocytes. The Journal of experimental medicine *160*, 1585-1590.

Asada, M., Goto, Y., Yahata, K., Yokoyama, N., Kawai, S., Inoue, N., Kaneko, O., and Kawazu, S. (2012). Gliding Motility of Babesia bovis Merozoites Visualized by Time-Lapse Video Microscopy. PloS one *7*, e35227.

Avril, M., Tripathi, A.K., Brazier, A.J., Andisi, C., Janes, J.H., Soma, V.L., Sullivan, D.J., Jr., Bull, P.C., Stins, M.F., and Smith, J.D. (2012). A restricted subset of var genes mediates adherence of Plasmodium falciparum-infected erythrocytes to brain endothelial cells. Proceedings of the National Academy of Sciences of the United States of America *109*, E1782-1790.

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

Berendt, A.R., McDowall, A., Craig, A.G., Bates, P.A., Sternberg, M.J., Marsh, K., Newbold, C.I., and Hogg, N. (1992). The binding site on ICAM-1 for Plasmodium falciparum-infected erythrocytes overlaps, but is distinct from, the LFA-1-binding site. Cell *68*, 71-81.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes & development *23*, 781-783.

Bhattacharjee, S., Hiller, N.L., Liolios, K., Win, J., Kanneganti, T.D., Young, C., Kamoun, S., and Haldar, K. (2006). The malarial host-targeting signal is conserved in the Irish potato famine pathogen. PLoS pathogens *2*, e50.

Bhattacharjee, S., Speicher, K.D., Stahelin, R.V., Speicher, D.W., and Haldar, K. (2012a). PI(3)P-independent and -dependent pathways function together in a vacuolar translocation sequence to target malarial proteins to the host erythrocyte. Molecular and biochemical parasitology *185*, 106-113.

Bhattacharjee, S., Stahelin, R.V., Speicher, K.D., Speicher, D.W., and Haldar, K. (2012b). Endoplasmic Reticulum PI(3)P Lipid Binding Targets Malaria Proteins to the Host Cell. Cell *148*, 201-212.

Bishop, R., Musoke, A., Morzaria, S., Gardner, M., and Nene, V. (2004). Theileria: intracellular protozoan parasites of wild and domestic ruminants transmitted by ixodid ticks. Parasitology *129 Suppl*, S271-283.

Blythe, J.E., Yam, X.Y., Kuss, C., Bozdech, Z., Holder, A.A., Marsh, K., Langhorne, J., and Preiser, P.R. (2008). Plasmodium falciparum STEVOR proteins are highly expressed in patient isolates and located in the surface membranes of infected red blood cells and the apical tips of merozoites. Infection and immunity *76*, 3329-3336.

Boddey, J.A., Carvalho, T.G., Hodder, A.N., Sargeant, T.J., Sleebs, B.E., Marapana, D., Lopaticki, S., Nebl, T., and Cowman, A.F. (2013). Role of Plasmepsin V in Export of Diverse Protein Families from the Plasmodium falciparum Exportome. Traffic (Copenhagen, Denmark) *14*, 532-550.

Boddey, J.A., Hodder, A.N., Gunther, S., Gilson, P.R., Patsiouras, H., Kapp, E.A., Pearce, J.A., de Koning-Ward, T.F., Simpson, R.J., Crabb, B.S.*, et al.* (2010). An aspartyl protease directs malaria effector proteins to the host cell. Nature *463*, 627-631.

Bozdech, Z., Llinas, M., Pulliam, B.L., Wong, E.D., Zhu, J., and DeRisi, J.L. (2003). The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS biology *1*, E5.

Brayton, K.A., Lau, A.O., Herndon, D.R., Hannick, L., Kappmeyer, L.S., Berens, S.J., Bidwell, S.L., Brown, W.C., Crabtree, J., Fadrosh, D.*, et al.* (2007). Genome sequence of Babesia bovis and comparative analysis of apicomplexan hemoprotozoa. PLoS pathogens *3*, 1401-1413.

Brewster, D.R., Kwiatkowski, D., and White, N.J. (1990). Neurological sequelae of cerebral malaria in children. Lancet *336*, 1039-1043.

Brown, H., Hien, T.T., Day, N., Mai, N.T., Chuong, L.V., Chau, T.T., Loc, P.P., Phu, N.H., Bethell, D., Farrar, J.*, et al.* (1999). Evidence of blood-brain barrier dysfunction in human cerebral malaria. Neuropathology and applied neurobiology *25*, 331-340.

Brown, K.N., and Brown, I.N. (1965). Immunity to malaria: antigenic variation in chronic infections of Plasmodium knowlesi. Nature *208*, 1286-1288.

Bullen, H.E., Charnaud, S.C., Kalanon, M., Riglar, D.T., Dekiwadia, C., Kangwanrangsan, N., Torii, M., Tsuboi, T., Baum, J., Ralph, S.A.*, et al.* (2012). Biosynthesis, localization, and macromolecular arrangement of the Plasmodium falciparum translocon of exported proteins (PTEX). The Journal of biological chemistry *287*, 7871-7884.

Burleigh, B.A., and Woolsey, A.M. (2002). Cell signalling and Trypanosoma cruzi invasion. Cellular microbiology *4*, 701-711.

Carruthers, V.B., and Sibley, L.D. (1997). Sequential protein secretion from three distinct organelles of Toxoplasma gondii accompanies invasion of human fibroblasts. European journal of cell biology *73*, 114-123.

Chang, H.H., Falick, A.M., Carlton, P.M., Sedat, J.W., DeRisi, J.L., and Marletta, M.A. (2008). N-terminal processing of proteins exported by malaria parasites. Molecular and biochemical parasitology *160*, 107-115.

Chen, Q., Fernandez, V., Sundstrom, A., Schlichtherle, M., Datta, S., Hagblom, P., and Wahlgren, M. (1998). Developmental selection of var gene expression in Plasmodium falciparum. Nature *394*, 392-395.

Chen, X.M., O'Hara, S.P., Huang, B.Q., Nelson, J.B., Lin, J.J., Zhu, G., Ward, H.D., and LaRusso, N.F. (2004). Apical organelle discharge by Cryptosporidium parvum is temperature, cytoskeleton, and intracellular calcium dependent and required for host cell invasion. Infection and immunity *72*, 6806-6816.

Chookajorn, T., Dzikowski, R., Frank, M., Li, F., Jiwani, A.Z., Hartl, D.L., and Deitsch, K.W. (2007). Epigenetic memory at malaria virulence genes. Proceedings of the National Academy of Sciences of the United States of America *104*, 899-902.

Claessens, A., Adams, Y., Ghumra, A., Lindergard, G., Buchan, C.C., Andisi, C., Bull, P.C., Mok, S., Gupta, A.P., Wang, C.W.*, et al.* (2012). A subset of group A-like var genes encodes the malaria parasite ligands for binding to human brain endothelial cells. Proceedings of the National Academy of Sciences of the United States of America *109*, E1772-1781.

Cortes, A., Carret, C., Kaneko, O., Yim Lim, B.Y., Ivens, A., and Holder, A.A. (2007). Epigenetic silencing of Plasmodium falciparum genes linked to erythrocyte invasion. PLoS pathogens *3*, e107.

Cox, H.W. (1959). A study of relapse Plasmodium berghei infections isolated from white mice. Journal of immunology (Baltimore, Md. : 1950) *82*, 209-214.

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

Current, W.L., and Garcia, L.S. (1991). Cryptosporidiosis. Clinical microbiology reviews *4*, 325-358.

Cyrklaff, M., Sanchez, C.P., Kilian, N., Bisseye, C., Simpore, J., Frischknecht, F., and Lanzer, M. (2011). Hemoglobins S and C interfere with actin remodeling in Plasmodium falciparum-infected erythrocytes. Science (New York, N.Y.) *334*, 1283-1286.

Date, S.V., and Stoeckert, C.J., Jr. (2006). Computational modeling of the Plasmodium falciparum interactome reveals protein function on a genome-wide scale. Genome research *16*, 542-549.

de Koning-Ward, T.F., Gilson, P.R., Boddey, J.A., Rug, M., Smith, B.J., Papenfuss, A.T., Sanders, P.R., Lundie, R.J., Maier, A.G., Cowman, A.F.*, et al.* (2009). A newly discovered protein export machine in malaria parasites. Nature *459*, 945-949.

De Silva, E.K., Gehrke, A.R., Olszewski, K., Leon, I., Chahal, J.S., Bulyk, M.L., and Llinas, M. (2008). Specific DNA-binding by apicomplexan AP2 transcription factors. Proceedings of the National Academy of Sciences of the United States of America *105*, 8393-8398.

Decherf, G., Egee, S., Staines, H.M., Ellory, J.C., and Thomas, S.L. (2004). Anionic channels in malaria-infected human red blood cells. Blood cells, molecules & diseases *32*, 366-371.

DeRocher, A., Gilbert, B., Feagin, J.E., and Parsons, M. (2005). Dissection of brefeldin A-sensitive and -insensitive steps in apicoplast protein targeting. Journal of cell science *118*, 565-574.

DeRocher, A., Hagen, C.B., Froehlich, J.E., Feagin, J.E., and Parsons, M. (2000). Analysis of targeting sequences demonstrates that trafficking to the Toxoplasma gondii plastid branches off the secretory system. Journal of cell science *113 (Pt 22)*, 3969-3977.

Desai, S.A. (2012). Ion and nutrient uptake by malaria parasite-infected erythrocytes. Cellular microbiology *14*, 1003-1009.

Desai, S.A., Bezrukov, S.M., and Zimmerberg, J. (2000). A voltage-dependent channel involved in nutrient uptake by red blood cells infected with the malaria parasite. Nature *406*, 1001-1005.

Dorovini-Zis, K., Schmidt, K., Huynh, H., Fu, W., Whitten, R.O., Milner, D., Kamiza, S., Molyneux, M., and Taylor, T.E. (2011). The neuropathology of fatal cerebral malaria in malawian children. The American journal of pathology *178*, 2146-2158.

Dubremetz, J.F., Achbarou, A., Bermudes, D., and Joiner, K.A. (1993). Kinetics and pattern of organelle exocytosis during Toxoplasma gondii/host-cell interaction. Parasitology research *79*, 402-408.

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.

El Hajj, H., Lebrun, M., Arold, S.T., Vial, H., Labesse, G., and Dubremetz, J.F. (2007). ROP18 is a rhoptry kinase controlling the intracellular proliferation of Toxoplasma gondii. PLoS pathogens *3*, e14.

Esch, K.J., and Petersen, C.A. (2013). Transmission and epidemiology of zoonotic protozoal diseases of companion animals. Clinical microbiology reviews *26*, 58-85.

Fernandez, V., Hommel, M., Chen, Q., Hagblom, P., and Wahlgren, M. (1999). Small, clonally variant antigens expressed on the surface of the Plasmodium falciparum-infected erythrocyte are encoded by the rif gene family and are the target of human immune responses. The Journal of experimental medicine *190*, 1393-1404.

Fichera, M.E., and Roos, D.S. (1997). A plastid organelle as a drug target in apicomplexan parasites. Nature *390*, 407-409.

Foth, B.J., and McFadden, G.I. (2003). The apicoplast: a plastid in Plasmodium falciparum and other Apicomplexan parasites. International review of cytology *224*, 57-110.

Foth, B.J., Ralph, S.A., Tonkin, C.J., Struck, N.S., Fraunholz, M., Roos, D.S., Cowman, A.F., and McFadden, G.I. (2003). Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. Science (New York, N.Y.) *299*, 705-708.

Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature *407*, 1018-1022.

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.

Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., Carlton, J.M., Hall, N., Ren, Q., Paulsen, I.T., Pain, A., Berriman, M.*, et al.* (2005). Genome sequence of Theileria parva, a bovine pathogen that transforms lymphocytes. Science (New York, N.Y.) *309*, 134-137.

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.

Gehde, N., Hinrichs, C., Montilla, I., Charpian, S., Lingelbach, K., and Przyborski, J.M. (2009). Protein unfolding is an essential requirement for transport across the parasitophorous vacuolar membrane of Plasmodium falciparum. Molecular microbiology *71*, 613-628.

Gero, A.M., and O'Sullivan, W.J. (1990). Purines and pyrimidines in malarial parasites. Blood cells *16*, 467-484; discussion 485-498.

Gr**ü**ring, C., Heiber, A., Kruse, F., Flemming, S., Franci, G., Colombo, S.F., Fasana, E., Schoeler, H., Borgese, N., Stunnenberg, H.G.*, et al.* (2012b). Uncovering common principles in protein export of malaria parasites. Cell host & microbe *12*, 717-729.

Haase, S., Herrmann, S., Gr**ü**ring, C., Heiber, A., Jansen, P.W., Langer, C., Treeck, M., Cabrera, A., Bruns, C., Struck, N.S.*, et al.* (2009). Sequence requirements for the export of the Plasmodium falciparum Maurer's clefts protein REX2. Molecular microbiology *71*, 1003-1017.

Hakansson, S., Charron, A.J., and Sibley, L.D. (2001). Toxoplasma evacuoles: a two-step process of secretion and fusion forms the parasitophorous vacuole. The EMBO journal *20*, 3132-3144.

Harb, O.S., Chatterjee, B., Fraunholz, M.J., Crawford, M.J., Nishi, M., and Roos, D.S. (2004). Multiple functionally redundant signals mediate targeting to the apicoplast in the apicomplexan parasite Toxoplasma gondii. Eukaryotic cell *3*, 663-674.

Hasler, T., Handunnetti, S.M., Aguiar, J.C., van Schravendijk, M.R., Greenwood, B.M., Lallinger, G., Cegielski, P., and Howard, R.J. (1990). In vitro rosetting, cytoadherence, and microagglutination properties of Plasmodium falciparum-infected erythrocytes from Gambian and Tanzanian patients. Blood *76*, 1845-1852.

Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C., and Haldar, K. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science (New York, N.Y.) *306*, 1934-1937.

Homer, M.J., Aguilar-Delfin, I., Telford, S.R., 3rd, Krause, P.J., and Persing, D.H. (2000). Babesiosis. Clinical microbiology reviews *13*, 451-469.

Hommel, M., David, P.H., and Oligino, L.D. (1983). Surface alterations of erythrocytes in Plasmodium falciparum malaria. Antigenic variation, antigenic diversity, and the role of the spleen. The Journal of experimental medicine *157*, 1137-1148.

Horrocks, P., Pinches, R.A., Chakravorty, S.J., Papakrivos, J., Christodoulou, Z., Kyes, S.A., Urban, B.C., Ferguson, D.J., and Newbold, C.I. (2005). PfEMP1 expression is reduced on the surface of knobless Plasmodium falciparum infected erythrocytes. Journal of cell science *118*, 2507-2518.

Howitt, C.A., Wilinski, D., Llinas, M., Templeton, T.J., Dzikowski, R., and Deitsch, K.W. (2009). Clonally variant gene families in Plasmodium falciparum share a common activation factor. Molecular microbiology *73*, 1171-1185.

Hu, G., Cabrera, A., Kono, M., Mok, S., Chaal, B.K., Haase, S., Engelberg, K., Cheemadan, S., Spielmann, T., Preiser, P.R.*, et al.* (2010). Transcriptional profiling of growth perturbations of the human malaria parasite Plasmodium falciparum. Nature biotechnology *28*, 91-98.

Huynh, M.H., Opitz, C., Kwok, L.Y., Tomley, F.M., Carruthers, V.B., and Soldati, D. (2004). Trans-genera reconstitution and complementation of an adhesion complex in Toxoplasma gondii. Cellular microbiology *6*, 771-782.

Imai, S., Armstrong, C.M., Kaeberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. Nature *403*, 795-800.

Iyer, L.M., Anantharaman, V., Wolf, M.Y., and Aravind, L. (2008). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. International journal for parasitology *38*, 1-31.

Kalanon, M., Tonkin, C.J., and McFadden, G.I. (2009). Characterization of two putative protein translocation components in the apicoplast of Plasmodium falciparum. Eukaryotic cell *8*, 1146-1154.

Kappe, S., Bruderer, T., Gantt, S., Fujioka, H., Nussenzweig, V., and Menard, R. (1999). Conservation of a gliding motility and cell invasion machinery in Apicomplexan parasites. The Journal of cell biology *147*, 937-944.

Keeley, A., and Soldati, D. (2004). The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. Trends in cell biology *14*, 528-532.

Khattab, A., Bonow, I., Schreiber, N., Petter, M., Schmetz, C., and Klinkert, M.Q. (2008). Plasmodium falciparum variant STEVOR antigens are expressed in merozoites and possibly associated with erythrocyte invasion. Malaria journal *7*, 137.

Kohler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J., Palmer, J.D., and Roos, D.S. (1997). A plastid of probable green algal origin in Apicomplexan parasites. Science (New York, N.Y.) *275*, 1485-1489.

Kraemer, S.M., Kyes, S.A., Aggarwal, G., Springer, A.L., Nelson, S.O., Christodoulou, Z., Smith, L.M., Wang, W., Levin, E., Newbold, C.I.*, et al.* (2007). Patterns of gene recombination shape var gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates. BMC genomics *8*, 45.

Kraemer, S.M., and Smith, J.D. (2006). A family affair: var genes, PfEMP1 binding, and malaria disease. Curr Opin Microbiol *9*, 374-380.

Kyes, S.A., Rowe, J.A., Kriek, N., and Newbold, C.I. (1999). Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum.

Proceedings of the National Academy of Sciences of the United States of America *96*, 9333-9338.

Lauer, S., VanWye, J., Harrison, T., McManus, H., Samuel, B.U., Hiller, N.L., Mohandas, N., and Haldar, K. (2000). Vacuolar uptake of host components, and a role for cholesterol and sphingomyelin in malarial infection. The EMBO journal *19*, 3556-3564.

Lauer, S.A., Rathod, P.K., Ghori, N., and Haldar, K. (1997). A membrane network for nutrient import in red cells infected with the malaria parasite. Science (New York, N.Y.) *276*, 1122-1125.

Lavazec, C., Sanyal, S., and Templeton, T.J. (2006). Hypervariability within the Rifin, Stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. Nucleic acids research *34*, 6696-6707.

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

Lavstsen, T., Turner, L., Saguti, F., Magistrado, P., Rask, T.S., Jespersen, J.S., Wang, C.W., Berger, S.S., Baraka, V., Marquard, A.M.*, et al.* (2012). Plasmodium falciparum erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children. Proceedings of the National Academy of Sciences of the United States of America *109*, E1791-1800.

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 (New York, N.Y.) *301*, 1503-1508.

Leech, J.H., Barnwell, J.W., Miller, L.H., and Howard, R.J. (1984). Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. The Journal of experimental medicine *159*, 1567-1575.

Levine, N.D. (1988). Progress in taxonomy of the apicomplexan protozoa. The Journal of protozoology *35*, 518-520.

Lewallen, S., White, V.A., Whitten, R.O., Gardiner, J., Hoar, B., Lindley, J., Lochhead, J., McCormick, A., Wade, K., Tembo, M.*, et al.* (2000). Clinical-histopathological correlation of the abnormal retinal vessels in cerebral malaria. Archives of ophthalmology *118*, 924-928.

Lopez-Rubio, J.J., Gontijo, A.M., Nunes, M.C., Issar, N., Hernandez Rivas, R., and Scherf, A. (2007). 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Molecular microbiology *66*, 1296-1305.

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.

Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szestak, T., Chesson, J., Wu, Y., Hughes, K., Coppel, R.L.*, et al.* (2008). Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell *134*, 48-61.

Marechal, E., and Cesbron-Delauw, M.F. (2001). The apicoplast: a new member of the plastid family. Trends in plant science *6*, 200-205.

Marsh, K., Forster, D., Waruiru, C., Mwangi, I., Winstanley, M., Marsh, V., Newton, C., Winstanley, P., Warn, P., Peshu, N.*, et al.* (1995). Indicators of life-threatening malaria in African children. The New England journal of medicine *332*, 1399-1404.

Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004). Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science (New York, N.Y.) *306*, 1930-1933.

Martin, R.E., and Kirk, K. (2007). Transport of the essential nutrient isoleucine in human erythrocytes infected with the malaria parasite Plasmodium falciparum. Blood *109*, 2217-2224.

Milner, D.A., Jr., Pochet, N., Krupka, M., Williams, C., Seydel, K., Taylor, T.E., Van de Peer, Y., Regev, A., Wirth, D., Daily, J.P.*, et al.* (2012). Transcriptional Profiling of Plasmodium falciparum Parasites from Patients with Severe Malaria Identifies Distinct Low vs. High Parasitemic Clusters. PloS one *7*, e40739.

Mordue, D.G., Desai, N., Dustin, M., and Sibley, L.D. (1999). Invasion by Toxoplasma gondii establishes a moving junction that selectively excludes host cell plasma membrane proteins on the basis of their membrane anchoring. The Journal of experimental medicine *190*, 1783-1792.

Mosqueda, J., Olvera-Ramirez, A., Aguilar-Tipacamu, G., and Canto, G.J. (2012). Current advances in detection and treatment of babesiosis. Current medicinal chemistry *19*, 1504-1518.

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.

Newbold, C., Craig, A., Kyes, S., Rowe, A., Fernandez-Reyes, D., and Fagan, T. (1999). Cytoadherence, pathogenesis and the infected red cell surface in Plasmodium falciparum. International journal for parasitology *29*, 927-937.

Newton, C.R., Peshu, N., Kendall, B., Kirkham, F.J., Sowunmi, A., Waruiru, C., Mwangi, I., Murphy, S.A., and Marsh, K. (1994). Brain swelling and ischaemia in Kenyans with cerebral malaria. Archives of disease in childhood *70*, 281-287.

Nguitragool, W., Bokhari, A.A., Pillai, A.D., Rayavara, K., Sharma, P., Turpin, B., Aravind, L., and Desai, S.A. (2011). Malaria parasite clag3 genes determine channel-mediated nutrient uptake by infected red blood cells. Cell *145*, 665-677.

O'Donnell, R.A., Freitas-Junior, L.H., Preiser, P.R., Williamson, D.H., Duraisingh, M., McElwain, T.F., Scherf, A., Cowman, A.F., and Crabb, B.S. (2002). A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes. The EMBO journal *21*, 1231-1239.

Parsons, M., Karnataki, A., Feagin, J.E., and DeRocher, A. (2007). Protein trafficking to the apicoplast: deciphering the apicomplexan solution to secondary endosymbiosis. Eukaryotic cell *6*, 1081-1088.

Pei, X., Guo, X., Coppel, R., Bhattacharjee, S., Haldar, K., Gratzer, W., Mohandas, N., and An, X. (2007). The ring-infected erythrocyte surface antigen (RESA) of Plasmodium falciparum stabilizes spectrin tetramers and suppresses further invasion. Blood *110*, 1036-1042.

Plattner, F., and Soldati-Favre, D. (2008). Hijacking of host cellular functions by the Apicomplexa. Annual review of microbiology *62*, 471-487.

Ralph, S.A., Foth, B.J., Hall, N., and McFadden, G.I. (2004). Evolutionary pressures on apicoplast transit peptides. Molecular biology and evolution *21*, 2183-2194.

Riglar, D.T., Rogers, K.L., Hanssen, E., Turnbull, L., Bullen, H.E., Charnaud, S.C., Przyborski, J., Gilson, P.R., Whitchurch, C.B., Crabb, B.S.*, et al.* (2013). Spatial association with PTEX complexes defines regions for effector export into Plasmodium falciparum-infected erythrocytes. Nature communications *4*, 1415.

Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A., and Goldberg, D.E. (2010). Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. Nature *463*, 632-636.

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. Proceedings of the National Academy of Sciences of the United States of America *106*, 9655-9660.

Sargeant, T.J., Marti, M., Caler, E., Carlton, J.M., Simpson, K., Speed, T.P., and Cowman, A.F. (2006). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome biology *7*, R12.

Sato, S. (2011). The apicomplexan plastid and its evolution. Cellular and molecular life sciences : CMLS *68*, 1285-1296.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J., and Lanzer, M. (1998). 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 *17*, 5418-5426.

Schwab, J.C., Beckers, C.J., and Joiner, K.A. (1994). The parasitophorous vacuole membrane surrounding intracellular Toxoplasma gondii functions as a molecular sieve. Proceedings of the National Academy of Sciences of the United States of America *91*, 509-513.

Seeber, F., and Soldati-Favre, D. (2010). Metabolic pathways in the apicoplast of apicomplexa. International review of cell and molecular biology *281*, 161-228.

Seydel, K.B., Fox, L.L., Glover, S.J., Reeves, M.J., Pensulo, P., Muiruri, A., Mpakiza, A., Molyneux, M.E., and Taylor, T.E. (2012). Plasma concentrations of parasite histidine-rich protein 2 distinguish between retinopathy-positive and retinopathy-negative cerebral malaria in Malawian children. The Journal of infectious diseases *206*, 309-318.

Shikani, H.J., Freeman, B.D., Lisanti, M.P., Weiss, L.M., Tanowitz, H.B., and Desruisseaux, M.S. (2012). Cerebral malaria: we have come a long way. The American journal of pathology *181*, 1484-1492.

Sibley, L.D. (2010). How apicomplexan parasites move in and out of cells. Current opinion in biotechnology *21*, 592-598.

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., Pinches, R., Newbold, C.I., and Miller, L.H. (1995). Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell *82*, 101-110.

Spielmann, T., and Gilberger, T.W. (2010). Protein export in malaria parasites: do multiple export motifs add up to multiple export pathways? Trends in parasitology *26*, 6-10.

Spielmann, T., Hawthorne, P.L., Dixon, M.W., Hannemann, M., Klotz, K., Kemp, D.J., Klonis, N., Tilley, L., Trenholme, K.R., and Gardiner, D.L. (2006). A cluster of ring stage-specific genes linked to a locus implicated in cytoadherence in Plasmodium falciparum codes for PEXEL-negative and PEXEL-positive proteins exported into the host cell. Molecular biology of the cell *17*, 3613-3624.

Susomboon, P., Maneerat, Y., Dekumyoy, P., Kalambaheti, T., Iwagami, M., Komaki-Yasuda, K., Kawazu, S., Tangpukdee, N., Looareesuwan, S., and Kano, S. (2006). Down-regulation of tight junction mRNAs in human endothelial cells co-cultured with Plasmodium falciparum-infected erythrocytes. Parasitology international *55*, 107-112.

Suss-Toby, E., Zimmerberg, J., and Ward, G.E. (1996). Toxoplasma invasion: the parasitophorous vacuole is formed from host cell plasma membrane and pinches off via a fission pore. Proceedings of the National Academy of Sciences of the United States of America *93*, 8413-8418.

Taylor, S., Barragan, A., Su, C., Fux, B., Fentress, S.J., Tang, K., Beatty, W.L., Hajj, H.E., Jerome, M., Behnke, M.S*., et al.* (2006). A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen Toxoplasma gondii. Science (New York, N.Y.) *314*, 1776-1780.

Taylor, T.E., Fu, W.J., Carr, R.A., Whitten, R.O., Mueller, J.S., Fosiko, N.G., Lewallen, S., Liomba, N.G., and Molyneux, M.E. (2004). Differentiating the pathologies of cerebral malaria by postmortem parasite counts. Nature medicine *10*, 143-145.

Tilley, L., Sougrat, R., Lithgow, T., and Hanssen, E. (2008). The twists and turns of Maurer's cleft trafficking in P. falciparum-infected erythrocytes. Traffic (Copenhagen, Denmark) *9*, 187-197.

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.

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

Torii, M., Adams, J.H., Miller, L.H., and Aikawa, M. (1989). Release of merozoite dense granules during erythrocyte invasion by Plasmodium knowlesi. Infection and immunity *57*, 3230-3233.

Turner, G.D., Ly, V.C., Nguyen, T.H., Tran, T.H., Nguyen, H.P., Bethell, D., Wyllie, S., Louwrier, K., Fox, S.B., Gatter, K.C*., et al.* (1998). Systemic endothelial activation occurs in both mild and severe malaria. Correlating dermal microvascular endothelial cell phenotype and soluble cell adhesion molecules with disease severity. The American journal of pathology *152*, 1477-1487.

Turner, G.D., Morrison, H., Jones, M., Davis, T.M., Looareesuwan, S., Buley, I.D., Gatter, K.C., Newbold, C.I., Pukritayakamee, S., Nagachinta, B*., et al.* (1994). An immunohistochemical study

of the pathology of fatal malaria. Evidence for widespread endothelial activation and a potential role for intercellular adhesion molecule-1 in cerebral sequestration. The American journal of pathology *145*, 1057-1069.

Ueno, N., and Wilson, M.E. (2012). Receptor-mediated phagocytosis of Leishmania: implications for intracellular survival. Trends in parasitology *28*, 335-344.

van Dooren, G.G., Schwartzbach, S.D., Osafune, T., and McFadden, G.I. (2001). Translocation of proteins across the multiple membranes of complex plastids. Biochimica et biophysica acta *1541*, 34-53.

van Ooij, C., Tamez, P., Bhattacharjee, S., Hiller, N.L., Harrison, T., Liolios, K., Kooij, T., Ramesar, J., Balu, B., Adams, J.*, et al.* (2008). The malaria secretome: from algorithms to essential function in blood stage infection. PLoS pathogens *4*, e1000084.

Volkman, S.K., Hartl, D.L., Wirth, D.F., Nielsen, K.M., Choi, M., Batalov, S., Zhou, Y., Plouffe, D., Le Roch, K.G., Abagyan, R.*, et al.* (2002). Excess polymorphisms in genes for membrane proteins in Plasmodium falciparum. Science (New York, N.Y.) *298*, 216-218.

von Heijne, G., and Nishikawa, K. (1991). Chloroplast transit peptides. The perfect random coil? FEBS letters *278*, 1-3.

Waller, R.F., Keeling, P.J., Donald, R.G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A.F., Besra, G.S., Roos, D.S., and McFadden, G.I. (1998). Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America *95*, 12352-12357.

Wassmer, S.C., Lepolard, C., Traore, B., Pouvelle, B., Gysin, J., and Grau, G.E. (2004). Platelets reorient Plasmodium falciparum-infected erythrocyte cytoadhesion to activated endothelial cells. The Journal of infectious diseases *189*, 180-189.

Wetzel, D.M., Chen, L.A., Ruiz, F.A., Moreno, S.N., and Sibley, L.D. (2004). Calcium-mediated protein secretion potentiates motility in Toxoplasma gondii. Journal of cell science *117*, 5739-5748.

Whisson, S.C., Boevink, P.C., Moleleki, L., Avrova, A.O., Morales, J.G., Gilroy, E.M., Armstrong, M.R., Grouffaud, S., van West, P., Chapman, S.*, et al.* (2007). A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature *450*, 115-118.

White, V.A., Lewallen, S., Beare, N., Kayira, K., Carr, R.A., and Taylor, T.E. (2001). Correlation of retinal haemorrhages with brain haemorrhages in children dying of cerebral malaria in Malawi. Transactions of the Royal Society of Tropical Medicine and Hygiene *95*, 618-621.

WHO (2000). Severe falciparum malaria. World Health Organization, Communicable Diseases Cluster. Transactions of the Royal Society of Tropical Medicine and Hygiene *94 Suppl 1*, S1-90.

WHO (2012). World malaria report. (Geneva, Switzerland, WHO), p. 195.

Wickham, M.E., Culvenor, J.G., and Cowman, A.F. (2003). Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. The Journal of biological chemistry *278*, 37658-37663.

CHAPTER II

Shared elements of host-targeting pathways among apicomplexan parasites of differing life styles

Karell G. Pellé[1*], Rays H.Y. Jiang[1.2*], Byung-Ho Kang[3], David R. Allred[4] and Matthias Marti[1]

[1] Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA 02115, USA

[2] The Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA

[3]Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611, USA

[4] Department of Infectious Diseases and Pathology, University of Florida, Gainesville, FL 32611,USA

*These authors contributed equally to the work.

Correspondence to Matthias Marti (mmarti@hsph.harvard.edu), Department of Immunology and Infectious Diseases, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA. Tel +16174324126

CONTRIBUTIONS

This chapter was written as a manuscript for publication. Dr. Rays Jiang performed the computational analyses and generated figures for the corresponding results. Karell Pellé generated all transfection plasmids (*Cryptosporidium, Plasmodium* and *Babesia*) export domain, *P. falciparum* transgenic parasites, fluorescence microscopy imgaes and export quantification. Dr. David Allred generated *B. bovis* transgenic parasites and fluorescence microscopy images of these parasites with the help of Yu-Ping Xiao. Dr. Byung-Ho performed immune-electron microscopy experiments. The writing of the manuscript was a collective effort by all authors. Dr. Tobias Spielmann was instrumental in providing helpful discussion points and critical reading of the manuscript.

I. Introduction

The apicomplexan phylum consists of obligate intracellular parasites of a wide variety of vertebrate hosts. Apicomplexan parasites include *Cryptosporidium* and *Toxoplasma*, as well as the blood parasites *Babesia* and *Plasmodium.* The phylum is named after an apical organellar complex that specializes in secretion. This 'apical complex' includes specialized sub-cellular secretory structures such as the micronemes, rhoptries and dense granules that sequentially release their contents during egress, host cell attachment and invasion. To achieve replication and evasion of host immune surveillance, these parasites modulate the host cellular environment by secretion of proteins into the host cell. *Plasmodium* and *Cryptosporidium* parasites are permanently engulfed by a parasitophorous vacuole membrane (PVM), making this a particularly complicated challenge, whereas the structure disappears very early in the *Babesia* intraerythrocytic cycle and in *Theileria* (Asada et al., 2012; Lingelbach and Joiner, 1998).

The human malaria parasite, *Plasmodium falciparum*, is by far the most intensively studied system among Apicomplexa with regard to protein secretion and export into the host cell. Signal requirements and trafficking routes for exported proteins have been characterized in some detail, whereas export pathways in other genera remain unknown. In *Plasmodium*, a bipartite signal consisting of a classical N-terminal eukaryotic signal sequence followed by a novel pentameric motif with the consensus sequence RxLxE/Q/D has been defined as a functional domain required for the export of parasite proteins into the host cell (Hiller et al., 2004; Marti et al., 2004). The motif has been termed PEXEL (*Plasmodium* Export Element (Marti et al., 2004)) or HTS (Host Targeting Signal (Hiller et al., 2004)). Subsequent identification of a parasite-derived putative translocon complex in the PVM (de Koning-Ward et al., 2009), as well as phosphatidylinositol-3-phosphate (PI3P) recognition (Bhattacharjee et al., 2012b) by the motif and proteolytic cleavage via a parasite protease (Boddey et al., 2010; Russo et al., 2010) at the parasite endoplasmic reticulum, demonstrate the existence of a genuine signal-mediated protein export pathway in *Plasmodium*. The presence of a bipartite signal for protein secretion and export also facilitated systematic characterization, classification and comparative analysis of the parasite "exportome" (Sargeant et al., 2006).

Using *Plasmodium* as a template we aimed to identify signals required for protein export in other apicomplexan parasites. For this purpose we performed a comparative *in silico* analysis of protein secretion and export across ten apicomplexan parasite species, representing five genera differing significantly in lifestyles. Specifically, we analyzed the available genome sequences of *Babesia bovis, Theileria parva, Toxoplasma gondii* and *Cryptosporidium parvum*. This analysis predicted the presence of PEXEL-like motifs in a subset of *Babesia* and *Cryptosporidium* proteins. We demonstrate that core motifs from both *Babesia* and *Cryptosporidium* are required for proper protein export when expressed in transgenic *Plasmodium* parasites. Moreover, we show in transgenic *Babesia* parasites that a complete export domain, including the core motif and flanking regions is required for proper regulated protein trafficking and export via the spherical body, demonstrating presence of a genuine signal-mediated export pathway in *Babesia*. Based on these findings a model is presented suggesting step-wise evolution of protein export pathways in the apicomplexan lineage.

II. Results

A. Definition of apicomplexan secretomes

As a first step for the identification of putative signals involved in protein export in Apicomplexa, we defined the set of proteins predicted or known to enter the secretory pathway (i.e., the secretome) in the species where annotated genome sequences were available. The secretome was estimated based on the presence of a known or predicted N-terminal signal peptide. Components of the core secretory machinery were removed from further analysis by information inferred from subcellular localizations of homologous proteins in the SwissProt database. The genomes analyzed included *T. parva*, *B. bovis*, *T. gondii* and *C. parvum*, and for comparison the genomes of six *Plasmodium* species, *P. falciparum*, *P. vivax*, *P. knowlesi*, *P. chabaudi*, *P. yoelii* and *P. berghei,* were included. The entire set of secreted proteins was clustered into families using the Markov-based method TribeMCL (Enright et al., 2002). With this method 331 families

comprised of secreted proteins with more than five members each were identified, with the majority being in *Plasmodium* parasites. The largest families of secreted proteins are part of the PIR superfamily of exported *Plasmodium* antigens (Janssen et al., 2004). Large non-*Plasmodium* families include the variant erythrocyte surface antigen (VESA) family in *Babesia* (O'Connor et al., 1997), a family of exported *Theileria* antigens with FAINT (Frequently associated with *Theileria*) domains (Schmuckli-Maurer et al., 2009) and a large family of SAG-related parasite surface antigens termed SRS in *Toxoplasma* (Jung et al., 2004).

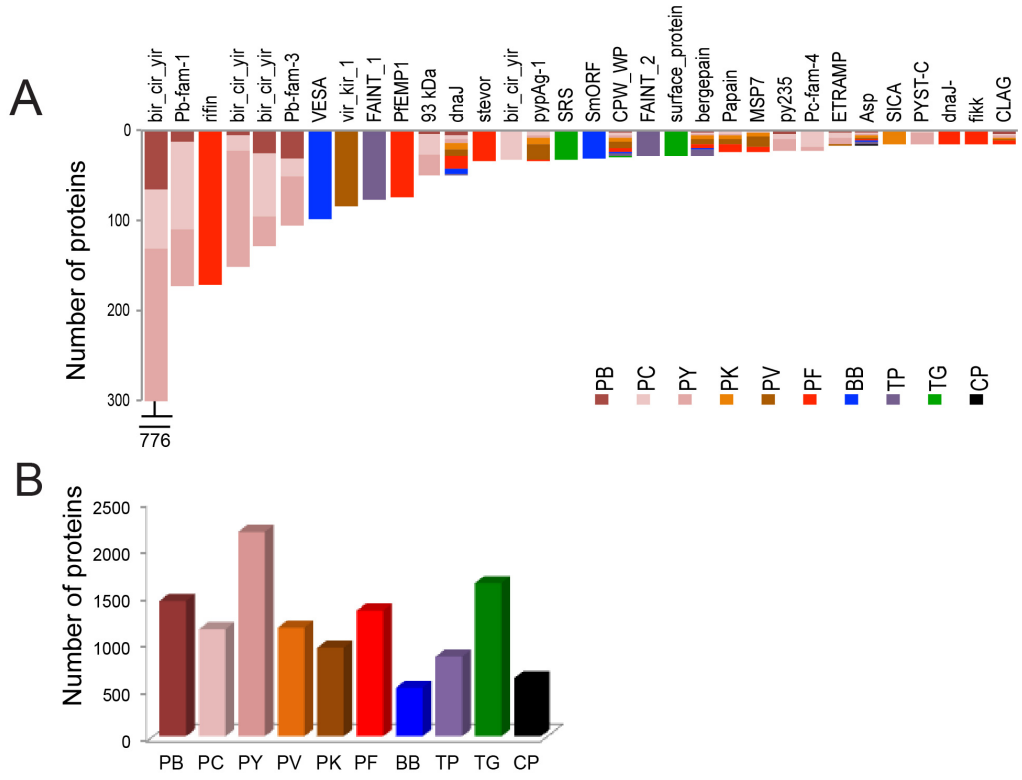


Figure 2- 1: **Protein secretion and trafficking in Apicomplexa**. **A.** Diversity of secreted protein families across Apicomplexa. Families are ranked by the number of paralogs and colored by species. **B.** Abundance of secreted proteins across Apicomplexa. Total number of secreted proteins per species.

A. Identification of a PEXEL-like motif in *Babesia* and *Cryptosporidium*

To identify sequence motifs or domains potentially involved in protein export we first scanned the apicomplexan secretomes for enrichment of any conserved sequence downstream of the signal sequence and within the first 150 amino acids from the N-terminus. This analysis only yielded the *Plasmodium* PEXEL motif. Next we specifically searched the secretomes for enrichment of PEXEL or PEXEL-like motifs within the same protein region downstream of the N-terminal signal sequence. Specifically, we quantified the enrichment of motifs by comparing the occurrence with the average frequency in the same protein sets after permutation. This analysis demonstrated the presence of a PEXEL-like motif, Rx(x)L in *Babesia* that occurs more than a hundred times above permutated background ($p<0.01$). A similar motif search and permutation test also showed enrichment for a PEXEL motif, RxLxE in *Cryptosporidium* ($p<0.01$). Importantly, both *Cryptosporidium* and *Babesia* motifs are positionally conserved in the N-terminus of the protein and downstream of the signal sequence, as is the case for the RxLxE/D/Q motif in *Plasmodium*, suggesting analogous function. In contrast, we did not identify any enrichment of a positionally conserved motif in *Theileria* or *Toxoplasma* (Figure 2-2A).

Our secretome analysis suggested that *Cryptosporidium* lacks large family expansion in secreted proteins (Figure 2-1). Likewise, the *Cryptosporidium* PEXEL-like motif RxLxE is present in only 15 proteins within the secretome (Figure 2-2C). Most gene families in *Cryptosporidium* have unknown function except the genes encoding Mucin, Insulinase and oocyst wall proteins. The few proteins with a predicted PEXEL-like motif are scattered across several small *Cryptosporidium* families previously classified as SKSR, FLGN and MEDLE. As no efficient system for reporter expression in intracellular *Cryptosporidium* parasites yet exists, we tested whether the PEXEL-like motif of two *Cryptosporidium* proteins (a protein containing a DnaJ domain and a hypothetical protein) can mediate secretion of a YFP reporter into the host RBC in *Plasmodium*. Indeed, in both cases the YFP fusion protein was exported into the host cell (Figure 2-2B). These initial experiments suggest functionality of the *Cryptosporidium* motif in a heterologous context. Systematic evaluation of this putative signal is required to determine whether it indeed represents part of a novel *Cryptosporidium* export pathway.

Like *Plasmodium*, *Babesia* infects RBCs, and both lineages show expansion of large membrane-bound proteins transported to the host cell surface. These are the variant PfEMP1 antigens in *P. falciparum* and the variant VESA antigens in *Babesia* (Leech et al., 1984; Allred et al., 1993). The proteins of these 2 families also share structural similarities in that they lack an N-terminal signal sequence but contain a single internal transmembrane domain. In both cases orientation and mode of translocation into the ER remain unclear. After VESA1, smORF is the second largest protein family in *B. bovis* (Brayton et al., 2007). Although smORF genes are interspersed among *ves* gene clusters in the genome, no evidence for their transcriptional co-regulation with *ves* genes has been found (Brayton et al., 2007; Ferreri et al., 2012). Their extreme variability suggests involvement in antigenic variation, but the subcellular localization of smORF proteins has yet to be critically determined (Brayton et al., 2007; Ferreri et al., 2012). SBP2 is the third largest protein family in *B. bovis*, localizing first to spherical bodies before being associated with the cytoplasmic side of the RBC membrane (Dowling et al., 1996).

We identified a positionally constrained PEXEL-like motif in multiple paralogs of the smORF and the SBP2 families, as well as several additional proteins, but not in VESA (Figure 2-2C). Additional PEXEL containing proteins include the secreted spherical body protein, SBP1, and a family of putative membrane proteins of which SBP1 is a member (Morahan et al., 2011; Brayton et al., 2007). Additional single copy proteins that are related to neither smORF nor SBP2 include the two secreted spherical body proteins, SBP3 and SBP4 (Struck et al., 2008; Terkawi et al., 2011), and a few hypothetical proteins. Altogether, our analysis demonstrated that all known secreted spherical body proteins and a significant number of as yet uncharacterized proteins harbor a PEXEL-like motif in *B. bovis*. To ask whether there are conserved properties within the region of the PEXEL motif, the charge and hydrophobicity pattern of the motif and its flanking regions were analyzed. A computational method was designed to calculate the average charge and hydropathy patterns from a group of sequences so that positive charge, negative charge, and hydropathy characteristics can be assessed simultaneously. Sequence alignments with the N-termini of smORF and the SBP2 cluster, respectively, revealed that in both cases the motif is flanked by consistently patterned stretches of positively charged, negatively charged, and

hydrophobic amino acids (Figure 2-2C). Therefore, we propose that it is more appropriate to consider the *Babesia* PEXEL-like motif as part of a larger, definable “export domain” downstream of the signal sequence.

Figure 2- 2: Identification and analysis of putative export motifs in *Babesia* and *Cryptosporidium*. A. Occurrence of conserved domains in the N-terminus of *Babesia* and *Cryptosporidium* proteins. Left panel: *Babesia* and *Cryptosporidium* but not *Toxoplasma* and *Theileria* show enrichment of an Rx(x)L and an RxLxE motif, respectively, when compared to a permuted data set. The occurrence of the conserved motifs and the permuted set is plotted. *P. vivax* serves as a positive control. Right Panel: Plotting the amino acid position of the putative motifs demonstrates positional conservation in *Babesia* and *Cryptosporidium* but not in *Theileria* and *Toxoplasma*. **B. The *Cryptosporidium* PEXEL-like motif.** Top panel: Protein families with PEXEL-like motif. *Cryptosporidium* does not show significant expansion of secreted proteins, less than 20 proteins are predicted to have an RxLxE motif. Middle panel: Sequence and charge composition of the *Cryptosporidium* PEXEL-like motif. Bottom panel: Expression of the *Cryptosporidium* export motif in *P. falciparum*. The N-termini of the CGD3 (a DnaJ homolog) and CGD2 (a hypothetical protein), respectively, were expressed as YFP fusions in transgenic *P. falciparum* parasites. Protein export into the host cell was monitored across the asexual parasite cycle. The left panel shows representative images from early (ring), developing (trophozoite) and late (schizont) stage parasites. The right panel shows quantification of reporter localization during development. At ring stages, the signal is confined to the parasites, while the reporter is increasingly secreted into the host cell in trophozoite stages and the majority localizes to the host cell in mature schizont stages. **C. The *Babesia* PEXEL-like motif.** Top panel: Protein families with PEXEL-like motif. *Babesia* has significantly expanded the smORF and SBP2 families. Bottom panel: Sequence and charge composition of the *Babesia* PEXEL-like motif. Shown are a sequence alignment of representative proteins, as well as a logo and a charge composition plot of the smORF family and SBP2 protein cluster.

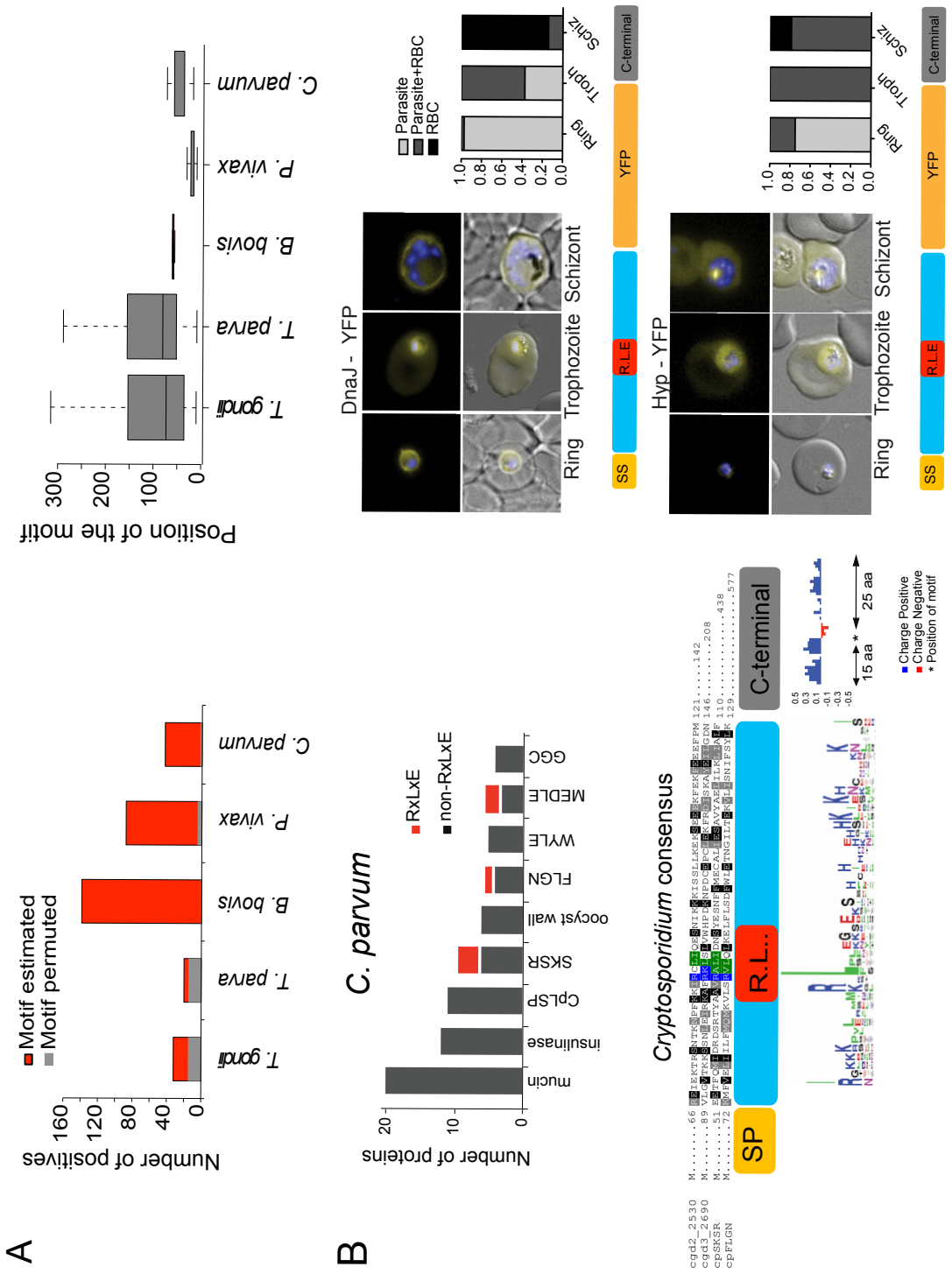


Figure 2- 2 continued: Identification and analysis of putative export motifs in *Babesia* and *Cryptosporidium*.

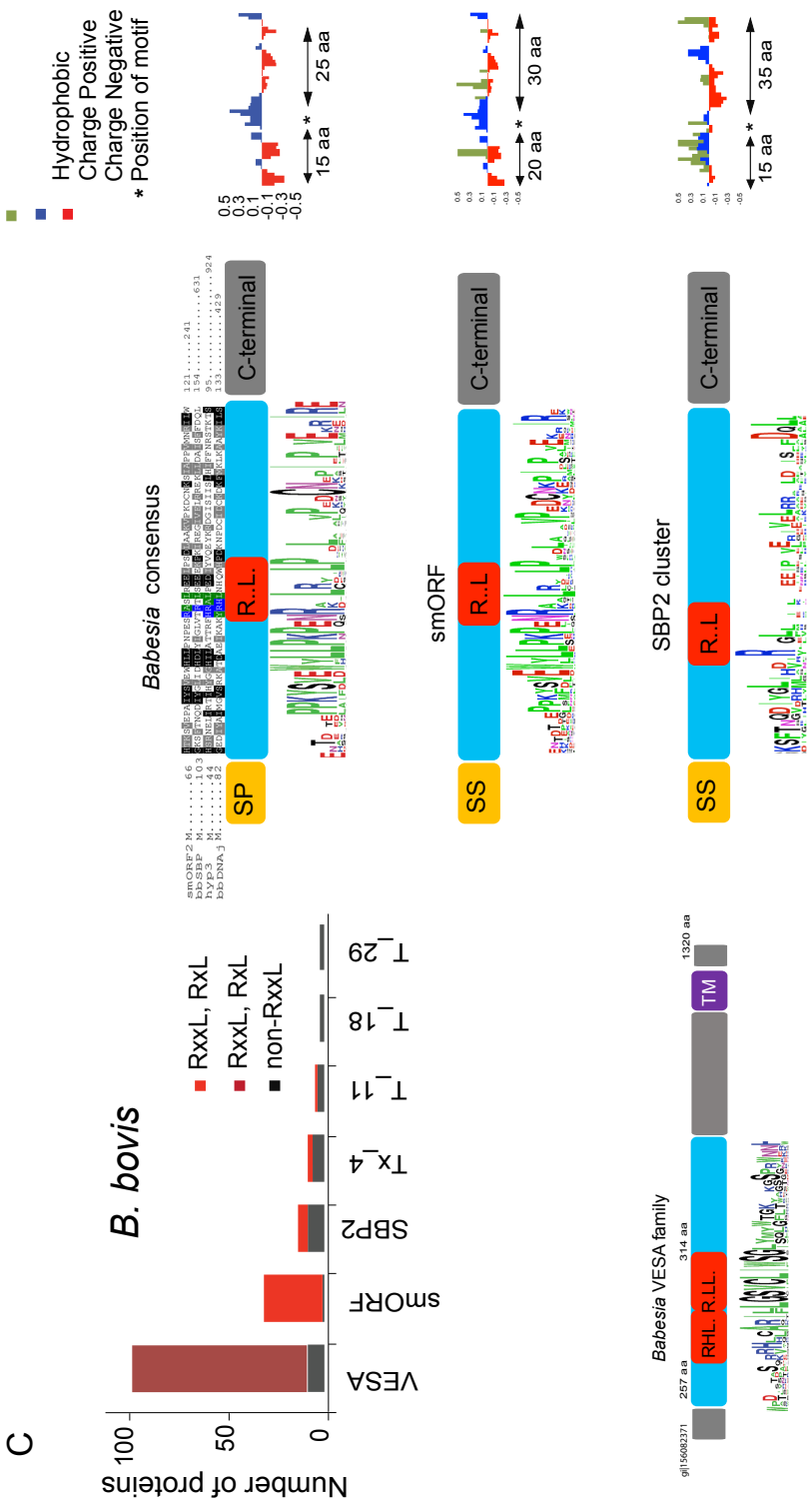


Figure 2- 2 continued: Identification and analysis of putative export motifs in *Babesia* and *Cryptosporidium*.

B. Characterization of a *Babesia* host-targeting pathway

To determine whether the putative PEXEL-like motif is involved in protein export into the host cell we fused N-terminal portions of *Babesia* proteins to YFP and localized the reporter in *B. bovis* parasites. Direct observation of live parasites was not informative, as GFP fails to fluoresce within the bovine erythrocyte when exported by *B. bovis* (not shown); reasons for this are not clear. Therefore, an immunofluorescence assay (IFA) using anti-GFP antibody was employed to localize fusion proteins (Figure 2-3A). Control experiments were also performed to confirm specificity of GFP labeling (Figure S1A). Transgenic parasites expressing native N-termini, including the entire "export domain" of SBP2 and two smORF paralogs fused to YFP, showed strong reporter labeling in the parasite in early parasite stages (Figure 2-3A). Later during development, the reporter was increasingly present in the host cell but appeared to be almost solely clustered in punctate structures. The localization of the YFP fusions was further defined at the ultrastructural level, by immunoelectron microscopy (immunoEM) experiments, using GFP antibodies and asynchronous *B. bovis* parasites (Figure 2-3B). In IFA experiments, few intensely fluorescent spots were observed within the parasites (Figure 2-3A white arrows). By immunoEM most of the fusion protein was correspondingly observed within large vesicular organelles in the meront stage (Figure 2-3B white arrows, and inserts I and II). Based on the morphologic characteristics of a location between the nucleus and apical end with its rhoptries and micronemes, electron density similar to that of the infected RBC cytoplasm, homogeneous density, spherical shape, and approximately 100 - 300 nm diameter, we conclude that these are a babesial organelle referred to as "spherical bodies" (Terkawi et al., 2011). In some cases the reporter was also found in small vesicles at the parasite periphery, probably en route to the host cell (Figure 2-3B, insert II black arrows). In later stages, gold particles were localized to the infected RBC membrane and sometimes to vesicles within the host RBC (Figure 2-3B, inserts III and IV white arrowheads). In the absence of spherical bodies, trophozoites showed host cell membrane labeling only, whereas a non-transfected parasite control was completely negative for GFP antibody staining (Supplementary Figure S1B) (Dowling et al., 1996). Altogether, these

experiments indicate that the YFP fusion proteins are targeted to and accumulate within spherical bodies before being exported into the host cell and to the erythrocyte plasma membrane during intracellular development of the parasite.

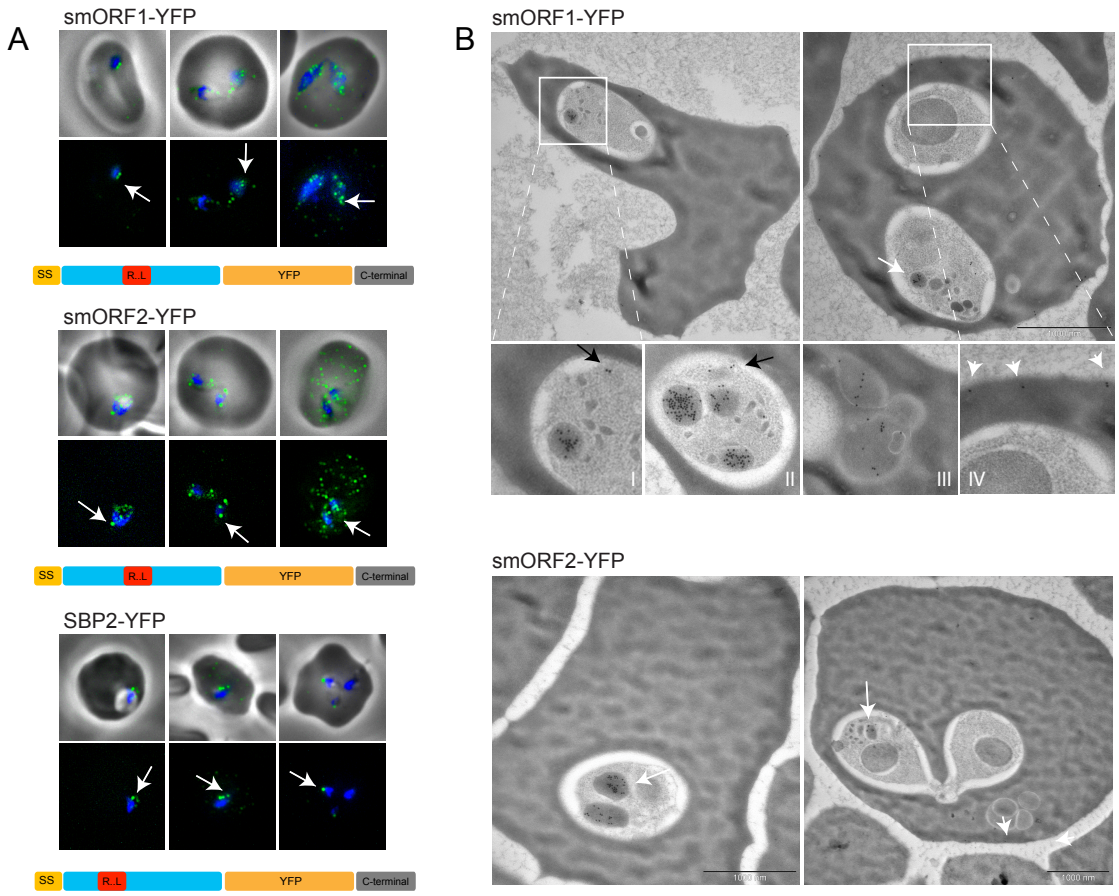


Figure 2- 3: Analysis of smORF-YFP and SBP2-YFP expression in transgenic *Babesia* parasites. A. IFA analysis of reporter expression. In all experiments, representative Images were captured across the *Babesia* RBC cycle. IFA demonstrates that both smORF-YFP fusions as well as SBP2-YFP show reporter labeling in the parasite and in multiple punctuate structures in the host cell and close to the surface membrane. Early meront stages show only parasite labeling, mainly at the apical end, while later stages show increasing labeling in the host cell. **B. Immunoelectron microscopic analysis of smORF1-YFP and smORF2-YFP distribution**. Staining with GFP antibodies shows major concentration of the reporter in apical organelles reminiscent of spherical bodies (white arrows). Later stages also show clusters of gold particles at the iRBC surface membrane (white arrowheads). Inserts show magnified sections of parasites with putative spherical body labeling (inserts I and II), as well as parasite vesicles (BII), as well internal RBC membrane (insert III) and RBC surface membrane labeling (insert IV).

To determine whether the PEXEL-like motif is required for targeting to spherical bodies and/or protein secretion into the host cell we generated a modified N-terminus for the smORF1 fusion protein, replacing the PEXEL-like motif with Alanines (RxxL to AAAA YFP fusion). In analogy to *Plasmodium*, we hypothesized that absence of the motif would result in default secretion into the host cell. IFA showed significantly reduced labeling in the parasite and enhanced labeling of the iRBC, suggesting different export kinetics or efficiency compared to wild-type smORF1 (Figure 2-4A). ImmunoEM microscopy demonstrated that the reporter localized to spherical bodies but was often observed in smaller vesicles in the parasite (Figure 4A, inserts II-IV). Frequently, clusters of gold particles were found at the periphery of the parasite and at the parasite host interface (Figure 2-4A, insert DII: black arrows). Together these data suggest that a significant proportion of the reporter is targeted to the host cell, either directly through an alternative export route, or through the spherical bodies, but with reduced retention. To further investigate this phenotype we generated an additional series of mutants where we fused only the signal sequence (SS) to YFP (Figure 2-4B). In the transgenic lines expressing SS-YFP fusions of smORF1, smORF2 and SBP2 spherical body localization of the reporter was nearly abrogated whereas localization to the red blood cell compartment appeared to be essentially quantitative by IFA. These data suggest that, in the absence of any organellar retention signal secretion into the host cell represents the default pathway in *Babesia*. The experiments also suggest contribution by export domain sequences flanking the PEXEL-like motif in spherical body targeting as the smORF1 Alanine mutant shows an intermediate phenotype rather than complete rerouting of the reporter to the default pathway.

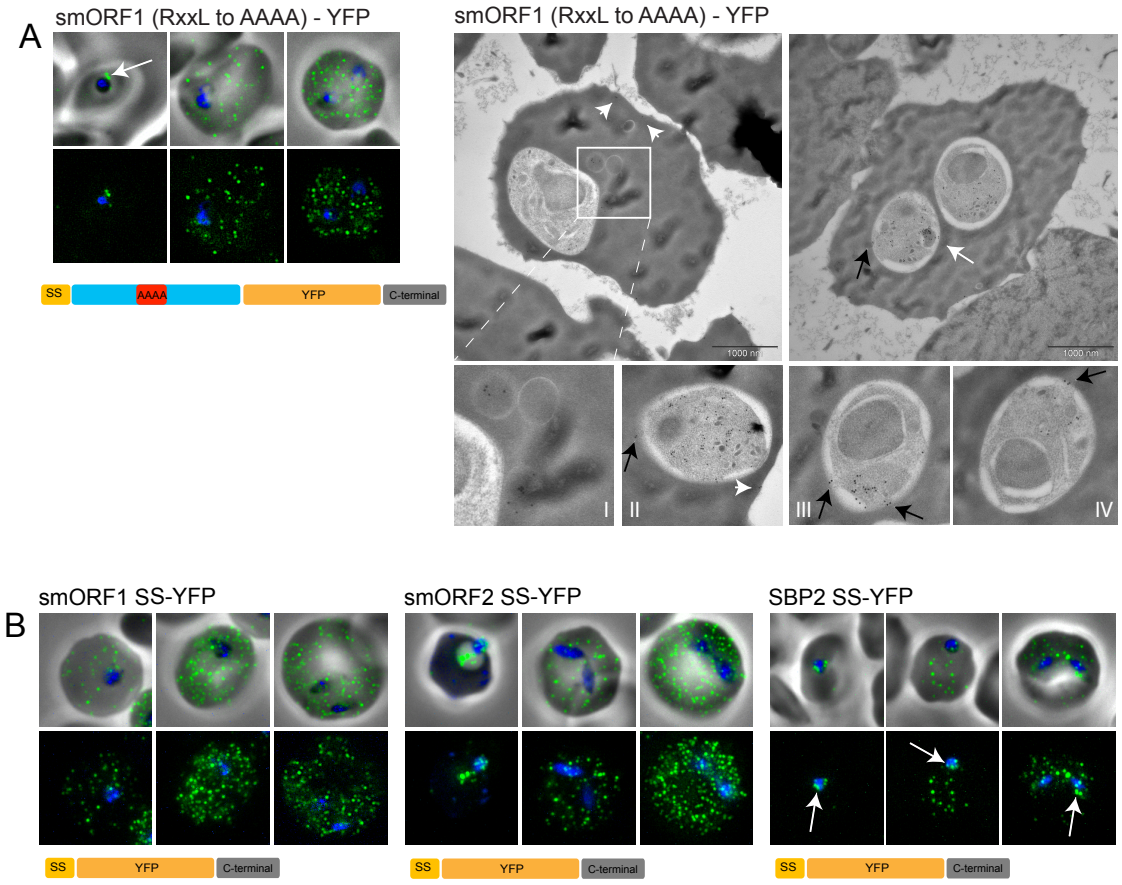


Figure 2- 4: Analysis of transgenic *Babesia* parasites expressing mutant N-termini fused to YFP. **A. RxxL to AAAA YFP fusion of smORF1**. IFA analysis suggests major reporter distribution in punctuate structures in the iRBC, whereas little to no labeling is observed in parasites. Likewise immunoelectron microscopy shows significantly reduced GFP labeling in spherical bodies (white arrows) and strong localization to RBC membranes (white arrow heads). Importantly, parasites show multiple clusters of gold particles outside of spherical bodies, some of them are at the host parasite interface (inserts BII - BIV: black arrows). **B. SS-YFP fusions of smORF proteins and SBP2.** IFA shows strong reporter labeling in the host cell in all three lines and complete absence of parasite localization, even in early parasite stages.

C. The *Babesia* PEXEL-like motif exports a reporter in transgenic *Plasmodium falciparum*

The experiments in transgenic *Babesia* suggested that a PEXEL-like motif is required for efficient targeting of the reporter to spherical bodies, which are not described in *Plasmodium*. To determine whether this motif can functionally complement the *Plasmodium* PEXEL motif, we expressed a series of fusion proteins in *P. falciparum* parasites. For this purpose we first expressed YFP fusions of the native SBP2 and two smORF N-termini, including the PEXEL-like motifs, in transgenic *P. falciparum*. IFA localization and quantification thereof demonstrated in all three cases that the native *Babesia* N-terminus is sufficient for protein export in this heterologous context (Figure 2-5A). Co-localization experiments with the Maurer's cleft marker Skeleton-binding protein 1 (SBP1) revealed that during later stages the smORF2-YFP fusion was targeted to these parasite-induced membrane structures within the host RBC. In order to determine whether the PEXEL-like motif itself is required for export in *P. falciparum*, we generated transgenic *P. falciparum* parasites expressing alanine mutants of the smORF1 and smORF2 core motif (RxxL to AAAA). Both mutant smORF -YFP fusions showed a localization pattern suggesting accumulation at the level of PV or PVM (Figure 2-5B). Together these data demonstrate that the *Babesia* export domain with the intact core PEXEL-like motif can functionally complement the *Plasmodium* export pathway.

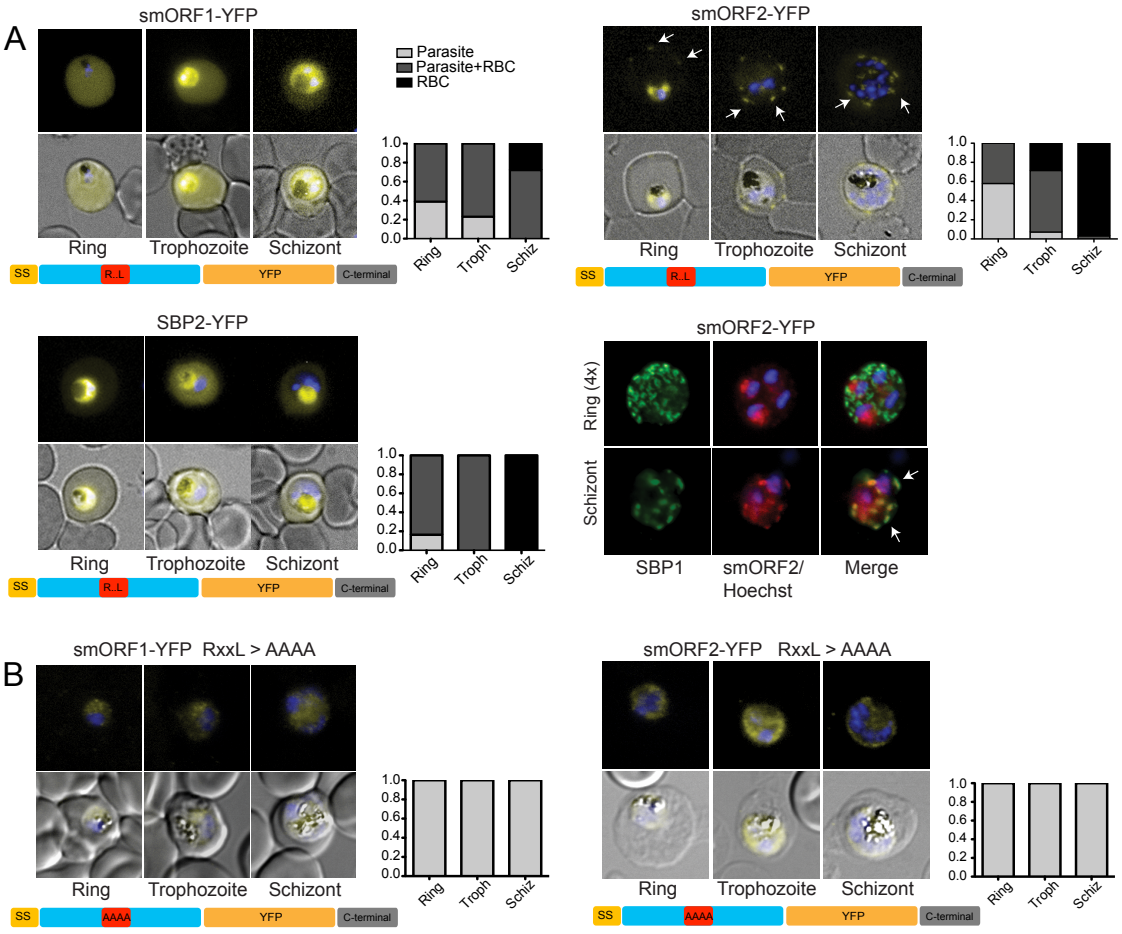


Figure 2- 5: Experimental validation of the *Babesia* PEXEL-like export motif function by heterologous expression in *P. falciparum*. A. Native N-termini fused to YFP. smORF-YFP and SBP2-YFP fusions are efficiently exported into the host cell. The smORF2-YFP fusion localizes to punctuate structures; IFA shows co-localization with the Maurer's cleft marker SBP1. **B. Mutant N-termini fused to YFP**. RxxL to AAAA mutant lines of smORF1-YFP and smORF2-YFP show a complete block of protein export. In all lines except for the mutants, reporter localization was quantified across asexual development by analyzing > 50 cells per line.

III. Discussion

The identification and functional characterization of a signal-mediated export pathway in *Plasmodium* has finally provided a mechanistic explanation for the secretion of multiple parasite antigens into the host RBC. Here we present first evidence of a novel host-targeting pathway in *Babesia* and its implications for the evolution of protein secretion in apicomplexan parasites.

Plasmodium proteins targeted to the host cell harbor a PEXEL motif (or a functionally equivalent motif in the case of a few PEXEL-negative exported proteins) for translocation across the PVM (Boddey et al., 2013; Hiller et al., 2004; Marti et al., 2004; Gruring et al., 2012). In the absence of any additional signal, reporter proteins fused to an N-terminal signal sequence are targeted to the PV, suggesting that secretion to the PV is the default pathway in *Plasmodium* (Marti et al., 2004; Waller et al., 2000; Wickham et al., 2001). How does selective secretion into the host cell work in other Apicomplexa, including those lacking a permanent PVM such as *Theileria* and *Babesia*? To answer this question we attempted to identify and functionally investigate sequence motifs downstream of the signal sequence of secreted proteins in apicomplexan parasites *Cryptosporidium parvum*, *Toxoplasma gondii*, *Theileria parva* and *Babesia bovis*. *In silico* identification and analysis of the secretomes from these four species revealed the presence of a positionally conserved PEXEL-like motif in a subset of *Cryptosporidium* and *Babesia* proteins, but not in *Theileria* and *Toxoplasma*. This was unexpected as *Babesia* is more closely related to *Theileria* than to *Plasmodium* (Cornillot et al., 2012; Kappmeyer et al., 2012). Notably, a recent *in silico* analysis of *Babesia* proteins did not identify any conserved sequence motifs in exported proteins, indicating the difficulty of finding such sequences (Gohil et al., 2013). Thus, a lack of PEXEL-like sequences in *Theileria* does not necessarily imply a lack of protein export to the host cell.

Both *Babesia* and *Plasmodium* invade and replicate within RBCs. However, a fundamental difference in the biology of these two parasites is that the membrane surrounding the parasite, the PVM, is maintained throughout the developmental cycle in *Plasmodium* but disappears within minutes after invasion in *Babesia* (Asada et al., 2012). Clearly, both similarities

and differences may be expected in the export of proteins to the host because of this difference. Analogous to the Maurer's clefts and knob structures in *Plasmodium*, membranous structures presumed to be of parasite origin are found within the RBC cytoplasm, and structures functionally analogous to the "knobs" formed by *P. falciparum* may be formed on the host cell surface by *B. bovis* (Hutchings et al., 2007; Rudzinska and Trager, 1977; Wright, 1972, 1973) (c.f., Figures 2-3B and 2-4A). We have identified in 59 *B. bovis* proteins a positionally conserved PEXEL-like motif preceded by a signal sequence, including all the known secreted spherical body proteins, SBP1, SBP2, SBP3 and SBP4. A closely related motif was also found in VESA1 polypeptides, which lack an N-terminal signal sequence, but the motif was not conserved within the same region and possesses a significant hydrophobic patch (Figure 2-2C). Functional analysis in transgenic *B. bovis* parasites demonstrated that neither the hydrophobic flanks nor the core PEXEL-like motif is required for targeting of a reporter to spherical bodies or subsequent secretion into the host cell. However, both appear to be necessary for retention within spherical bodies for properly regulated export. Deletion of the entire region including the PEXEL-like motif resulted in apparent direct secretion into the host cell without significant accumulation in spherical bodies, whereas mutation of the motif with retention of wild type flanking sequences showed an intermediate phenotype. The direct secretion is analogous to the default secretion phenotype into the PV that has previously been described in *Plasmodium*. Although VESA1 polypeptides contain an analogous domain and motif, they lack an N-terminal signal sequence, and it remains to be determined whether there is any overlap between export of VESA1 polypeptides and the spherical body pathway.

A number of exported proteins in apicomplexan parasites are synthesized then stored, later being injected into the host cell during the initial invasion process. Examples are injection of rhoptry kinases during *Toxoplasma* invasion, and export of dense granule proteins such as RESA upon invasion in *P. falciparum* (Peixoto et al., 2010; Bokhari et al., 2008; Rug et al., 2004). Based on their characteristics *Babesia* spherical bodies may be functionally equivalent to dense granules in certain aspects. However, in contrast to dense granules they can release at least some of their contents throughout the entire intracellular cycle and not only during or immediately

following invasion. In particular, SBP1 and SBP4 appear to be primarily secreted into the host cell much later, after division of the parasite (Morahan et al., 2011; Terkawi et al., 2011). Our experiments with chimeric fusion proteins similarly demonstrate increased reporter secretion in later stages of the cycle (Figure 2-3), and significantly reduced retention and accumulation within the spherical body with mutation of the PEXEL-like motif or deletion of the entire export domain (Figure 2-4). These observations demonstrate the existence of a signal-mediated, regulated export pathway via spherical bodies in *Babesia*. Differential dynamics of secretion suggest that spherical bodies facilitate the selective, regulated release of cargo, and perhaps protein maturation. Future studies are required to determine the function of the core motif and the flanking regions in spherical body targeting and export, as well as the mechanism of protein discharge into the host cell and processes of parasite-induced host cell remodeling.

In addition to *Babesia* our initial secretome analysis also identified a PEXEL-like motif in *Cryptosporidium*. This parasite is at the base of the apicomplexan tree and probably is most closely related to Gregarines (Marti et al., 2003). Upon invasion of epithelial cells, *Cryptosporidium* develops as a trophozoite within an extracytoplasmic vacuole that is separated from the host cytoplasm by a membrane. Electron dense regions where parasite and vacuolar membrane attach have been suggested to represent transport channels for the secretion of parasite material into the host cell, but secretion of specific parasite antigens after invasion has not been formally demonstrated (Huang et al., 2004). It remains to be determined whether the putative export motif in *Cryptosporidium* proteins targets them through such a channel, or whether the signal is required for targeting of proteins to secretory organelles before export into the host cell during the process of invasion. Our experiments in *P. falciparum* suggest that N-termini of *Cryptosporidium* proteins possessing a PEXEL-like motif and its flanking regions are sufficient for export of a reporter into the host cell. Recent successes in establishing a transfection protocol in *Cryptosporidium* will facilitate direct functional analyses in that parasite (Li et al., 2009).

Functionality of the *Babesia* export motif in *Plasmodium* indicates that common principles of protein export are likely to hold in the two apicomplexan lineages, despite the key difference

with regard to the PVM. Apart from the requirement of a related core motif for protein export both systems appear to depend on the presence of flanking regions with specific, consistent characteristics. Our experiments demonstrate that these regions are necessary in *Babesia* for efficient retention within the spherical body for regulated release. In *Plasmodium*, requirement for both upstream and downstream flanking sequences in protein export has been demonstrated as well (Bhattacharjee et al., 2006; Janssen et al., 2004; Knuepfer et al., 2005). A recent study also identified a non-random distribution of hydrophobicity pattern in the PEXEL flanking sequences of Rifin, Stevor and PfEMP1 (Hiss et al., 2008). We further investigated the regions flanking the *Plasmodium* PEXEL using a sliding window approach. Specifically, we constructed a set of HMMs using different lengths of flanking sequence to search for positive signals in exported proteins harboring a core PEXEL motif, as defined previously (Sargeant et al., 2006). The window size and window position with the highest number of hits is suggestive of conserved sequence information in the investigated region. The largest number of positive domains was recovered when the flanking regions were defined as 40 amino acid residues in length, with 15 amino acids upstream towards the signal sequence and 25 amino acids downstream of the core motif (Figure 2-6A). Non-random sequence distribution in the flanking regions is reflected by local charge bias, with positive charges occurring upstream and negative charges occurring in the immediate downstream sequences flanking the PEXEL motif. This charge pattern appears to be a general rule in *Plasmodium* proteins with a PEXEL core motif, including the membrane-bound Rifin family, but it does not apply to PfEMP1. Moreover, this charge distribution is similar to the one in the SBP2 and SmORF protein domains in *Babesia*, whereas the VESA1 polypeptide domains are dominated by hydrophobic amino acids. The inexact positioning of the PEXEL or PEXEL-like motif relative to the charged patches would suggest that it is the existence of the patches that is crucial, perhaps as a part of an electron motive mechanism driving polypeptide transfer across the membrane. In analogy to *Babesia*, it may be more appropriate to think of the entire sequence module in *Plasmodium* as a functional “export domain”. In *B. bovis,* proteins with a PEXEL-like motif are grouped into three clusters, smORF, SBP2, and VESA1, suggesting expansion of each family from a small number of founding ancestral exported proteins. Similarly, in *Plasmodium* all

the exported protein families occur in a species- or lineage- specific manner, with the exception of a large cluster of short soluble proteins previously described as PHIST (*Plasmodium* Helical Interspersed Subtelomeric (Sargeant et al., 2006)), and a second large cluster represented by the PIR superfamily of membrane-bound proteins (Janssen et al., 2004) (Figure 2-6B).

We propose herein a stepwise model for the evolution of regulated protein export into the host cell among Apicomplexa. We propose that this capacity began with the initial acquisition of the core motif and subsequent adaptation of export pathways within their respective host environment. In this model the ancestral function of an export motif (and specific flanking regions) in Apicomplexa was to target proteins to secretory organelles. This hypothesis is supported by recent identification of a PEXEL-like motif in a few *Toxoplasma* proteins required for protein targeting to the PVM via dense granules (Hsiao et al., 2013). While Toxoplasma does not appear to translocate proteins across the PVM, our data in *Babesia* support the idea that a specific signal is required for efficient organelle trafficking and retention, even in the absence of a permanent PVM, as proteins would otherwise be secreted into the host cell by default. The presence of a permanent PVM in *Plasmodium* has selected for the evolution of at least one, and possibly more, translocon-mediated pathways in order to facilitate efficient protein export into the host cell. Our heterologous expression experiments in *P. falciparum* suggest that a key step in export via a translocon remains functionally interchangeable with sorting to an apical organelle in other Apicomplexa. In *Plasmodium*, most exported proteins do not appear to undergo organellar staging, the parasite instead relying on transcriptional regulation to control the timing of protein export during development. In *B. bovis*, which has a much less well-defined cell cycle, regulation of export may occur primarily through organellar staging and timed release (Franssen et al., 2003). If the ancestral state were to degrade the PVM after invasion, as is the case in *Babesia* and *Theileria*, the question arises as to the added benefit of retaining and expanding it, as occurs in *Plasmodium*. Lack of a PVM may facilitate more efficient soluble nutrient uptake, whereas the presence of a PVM may allow selective import of host products such as hemoglobin, but necessitates the establishment of highly selective machinery for protein export and perhaps more extensive transcriptional regulation.

Figure 2- 6: **Revisiting the *Plasmodium* PEXEL and the evolution of protein export in Apicomplexa. A. Analysis of PEXEL flanking sequences defines an "export domain**. Top panel: Defining a *Plasmodium* export domain. A sliding window HMM screen was applied to test enrichment in variable sequence sizes flanking the PEXEL motif. The exportome as defined previously based on presence of a PEXEL motif (Sargeant et al., 2006) was used as an input. Bottom panel: Sequence and charge composition of the export domain. The *Plasmodium spp* logo plot represents all proteins with predicted PEXEL domain across *Plasmodium*. The basic residues are colored blue, acidic residues red, and hydrophobic green. Protein charge profiles are calculated with a sliding window approach with the EMBOSS program protein charge analysis. The average charge distribution of export domains is plotted. **B. Occurrence of exported protein families across *Plasmodium*.** Plasmodium families encoding secreted proteins (with > 5 paralogs) were used as input. Only the exported PHIST and PIR super families have members across all *Plasmodium* species, apart from the merozoite surface protein families MSP1 and MSP7, and the PVM family etramp. **C. A model on the evolution of protein export in a common ancestor of *Plasmodium* and *Babesia***. The ancestral apicomplexan export pathway targets proteins to specific secretory organelles. In *Toxplasma*, proteins with a PEXEL-like motif are targeted to the PVM via dense granules (Hsiao et al., 2013), while they are exported into the host cell in *Babesia* and *Plasmodium*. In *Plasmodium*, the core machinery for motif recognition and trafficking to secretory organelles is also functional for protein targeting to the host cell via a dedicated translocon in the PVM.

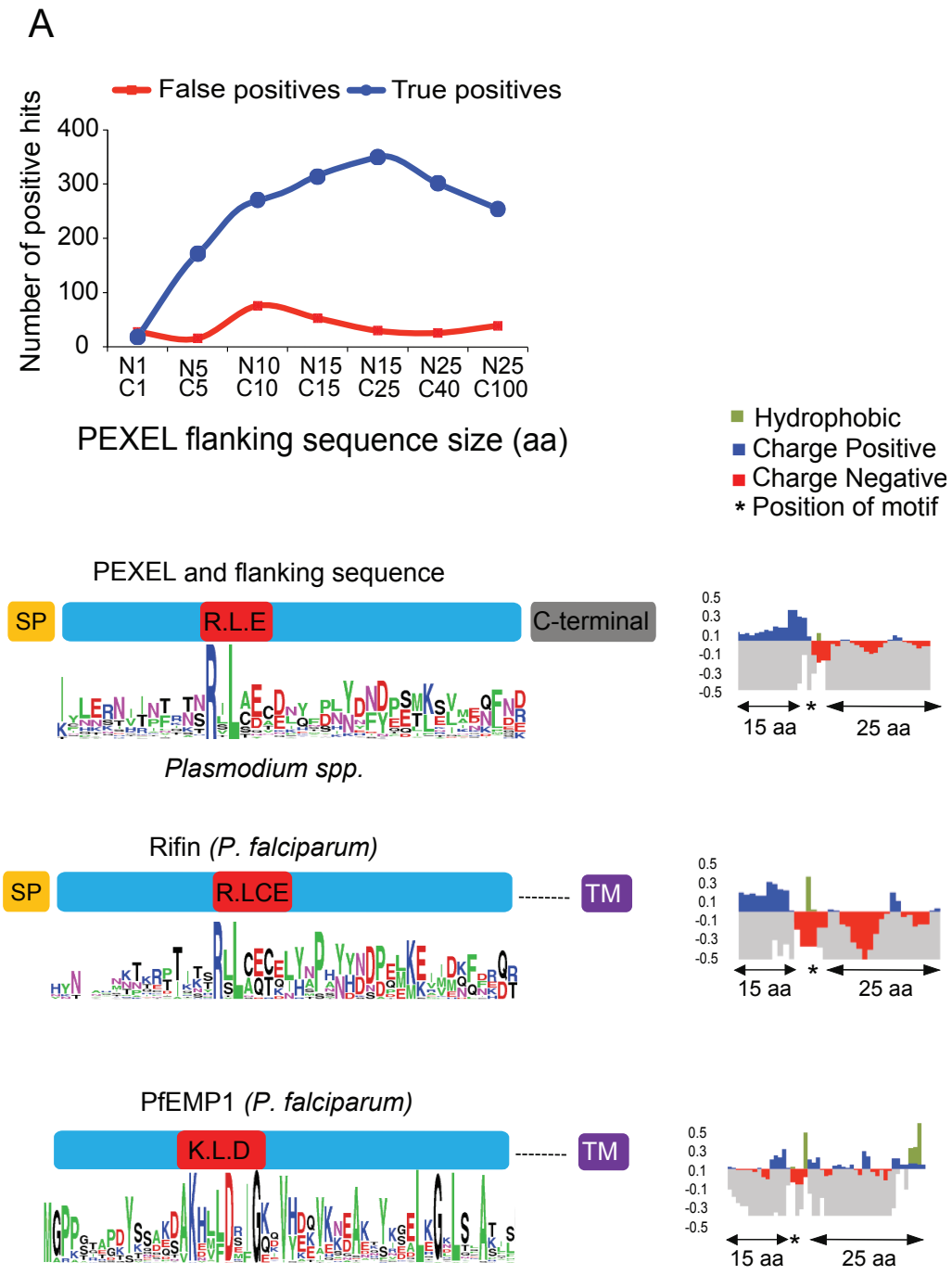


Figure 2- 6: continued: Revisiting the *Plasmodium* PEXEL and the evolution of protein export in Apicomplexa.

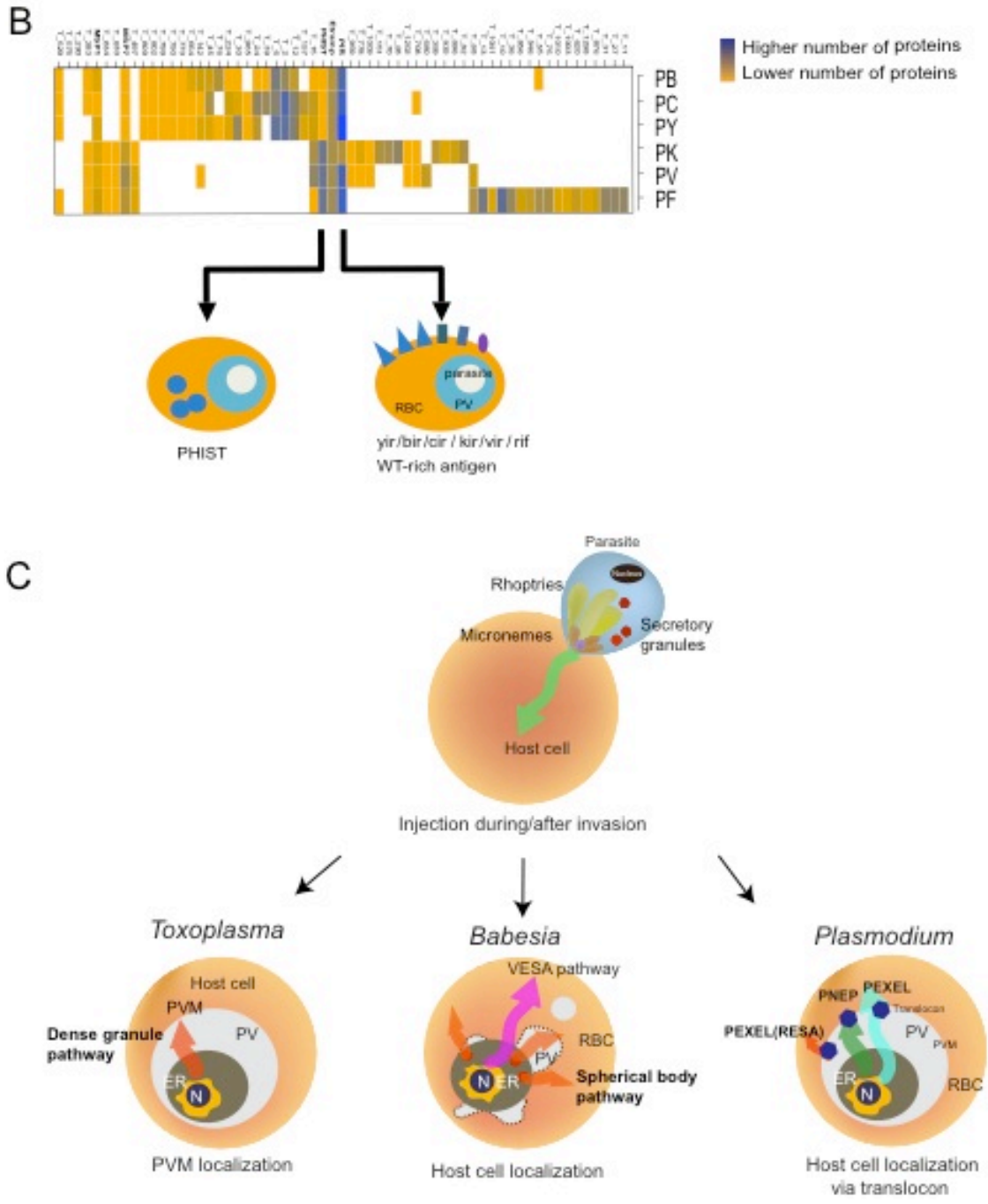


Figure 2- 6 continued: Revisiting the *Plasmodium* PEXEL and the evolution of protein export in Apicomplexa.

IV. Materials and Methods

Plasmids

All *P. falciparum* expression plasmids are based on pHHK(+)YFP (Marti et al., 2004). Sequences encoding target N-termini were amplified from genomic DNA of *B. bovis* strain C9.1 (for *SBP2* and 2 *smORF* genes) and from *C. parvum* Iowa II isolate oocysts (for cgd2_2530 and cgd3_2690), using Phusion DNA polymerase (New England BioLabs, Ipswich MA). Amplicons were digested with NotI and AvrII and cloned into pHHK(+)YFP, replacing the sequence encoding the KAHRP N-terminus.

The *Babesia* expression vector is also based on pHHK(+)YFP. However, the hDHFR expression cassette was removed with HindIII, and the *PfHSP86* promoter driving the trafficking cassette was replaced with the bidirectional 434 bp intergenic region derived from the locus of active *ves* transcription of *B. bovis* (Levy and Ristic, 1980). After replacement, this promoter drives expression of the trafficking cassette in one direction and in the other a blasticidin-S-deaminase (*bsd*) cassette (including the *B. bovis* b-tubulin terminator region). This expression vector is termed pBab-Exp. *Babesia* sequences encoding N-terminal trafficking signals were inserted upstream of the YFP gene using the NotI and AvrII sites, replacing sequence encoding the KAHRP N-terminus.

In vitro culture and transfection of *P. falciparum* and *B. bovis* parasites

The *P. falciparum* 3D7 reference line was used in this study. Parasites were cultivated under standard conditions in O+ human erythrocytes in RPMI 1640 supplemented with human serum to a final concentration of 10%, and gassed with 5% CO_2 and 0.5% O_2 in N_2 at 37°C as described (Trager and Jensen, 1976). Parasites were transfected with 50-100 µg of plasmid DNA by electroporation (0.310 kV, 950 mF), and transgenic parasites selected by addition of 5µM WR99210, as described (Wu et al., 1996).

The *B. bovis* CE11 strain (O'Connor and Allred, 2000) was cultured, using the microaerophilous stationary phase system (Levy and Ristic, 1980), under 5% O_2 and 5% CO_2 in

N_2 at 37°C, as described elsewhere (Allred et al., 1993). Parasites were grown in single-donor bovine erythrocytes, in medium supplemented with 40% (v/v) bovine serum from the same donor. For transfection, cultures at 5 - 10% parasitized erythrocytes were collected, washed into 1x Cytomix buffer (Allred et al., 1993), and prepared as 50% packed cell volume suspension in Cytomix buffer. The cell suspensions were transfected with 50-55 mg of plasmid DNA by electroporation 1.25 kV, 25 mF, and 200 Ω resistance) Cells were collected into complete *Babesia* culture medium (Vega et al., 1986) at 2.5% hematocrit (normal bovine erythrocytes). After 8 hours, blasticidin-S was added to a final concentration of 25 mM, and parasites were maintained under drug selection thereafter. Parasites typically emerged after approximately 1 week.

Microscopy

Transgenic *P. falciparum* parasites were analyzed live by epifluorescence microscopy, using an inverted microscope with differential interference contrast (DIC) (Zeiss; Jena, Germany). For analysis sorbitol synchronized parasites were washed by centrifugation and nuclei stained using Hoechst 33342 in PBS for 5 min at 37ºC prior to mounting on glass slides. For each transgenic *P. falciparum* line, 30-50 cells from three stages (ring, trophozoite and schizont) were captured for subsequent quantification of protein export dynamics. For co-localization of the smORF2 mutant with SBP1, parasites were analyzed by immunofluorescence analysis using fixed and permeabilized cells as described (Tonkin et al., 2009), and polyclonal rabbit anti-SBP1 antibodies (Blisnick et al., 2000) at a dilution of 1: 500 were used.

Transfected *B. bovis* parasites were washed two times with VYM buffer (Vega et al., 1986) by centrifugation, then resuspended and observed live. Alternatively, parasitized erythrocytes were processed for fixed-cell immunofluorescence as published (Tonkin et al., 2004), with minor modifications: fixative included 0.012% of glutaraldehyde and cells were permeabilized with 0.01% NP40. Rabbit polyclonal antiserum to GFP was used as primary antibody and goat anti-rabbit IgG-Alexa488 as secondary antibody. Each line was developmentally asynchronous when processed. Cells were observed with an Olympus BX50

microscope equipped for epifluorescence, phase contrast or DIC imaging, using a 100x, NA 1.3 immersion oil objective. From 30-50 images were captured of each transfected line. Immunofluorescence images were processed by Wiener Filter pre-conditioned Landweber constrained iterative deconvolution, using a theoretical point-spread function. The "Parallel Iterative Deconvolution" plug-in for ImageJ (P. Wendykier, Wolfram Research, Inc.; https://sites.google.com/site/piotrwendykier/software/deconvolution/paralleliterativedeconvolution) was used to perform the mathematical manipulations.

Immunoelectron microscopy

Bovine red blood cells infected with *B. bovis* were concentrated, frozen by high-pressure freezing, and freeze-substituted for 24 h at -80ºC with anhydrous acetone containing 0.25% (v/v) glutaraldehyde and 0.1% uranyl acetate. Samples were slowly warmed to -20ºC over a 48 h period, and washed 4 times with 100% ethanol. Cells were infiltrated at -20ºC with LR White resin (Ted Pella, Redding, CA) in steps of 33%, 66%, and 100%, over a 24 h period. Following two additional changes of 100% resin, samples were cured in BEEM capsules at 60ºC for 24 h. Samples were sectioned, and immunolabeling and electron microscopic imaging were performed as described (Kang, 2010).

Annotation of secreted proteins in Apicomplexa

The total annotated protein set of the 10 apicomplexan parasites, *Theileria parva*, *Cryptosporidium parvum*, *Toxoplasma gondii*, *Babesia bovis*, *Plasmodium berghei*, *Plasmodium yoelii*, *Plasmodium chabaudi*, *Plasmodium knowlesi*, *Plasmodium vivax* and *Plasmodium. Falciparum,* was retrieved from PlasmoDB and subjected to SignalP3.0 prediction (Bendtsen et al., 2004). Positive candidates were predicted either by the NN (Neural Network) or the HMM (Hidden Markov Model) method, with scores larger than 0.8. In addition, they were required to have a signal sequence cleavage site between 10 and 30 amino acids from the N-terminus, or one or two internal transmembranes in case a signal sequence was not predicted. Because all known secreted proteins in Apicomplexa have less than three transmembrane domains,

membrane proteins with three or more transmembrane domains such as transporters were separated from the data set. Proteins targeted to mitochondria or residing in the ER were removed based on IntroPro hits and/or Swissprot protein homology (BLASTP e value < 1e-5).

The entire set of secreted proteins was clustered into families using the Markov-based method TribeMCL (Enright et al., 2002). With this method 331 families with more than 5 members each were identified (the largest ones are shown in Figure 1A), the vast majority of them in *Plasmodium* parasites.

ACKNOWLEDGEMENTS

We would like to thank Dr. Honorine Ward (Tufts University School of Medicine, Boston), for providing Cryptosporidium parvum genomic DNA, Dr. W. Clay Smith (University of Florida) for polyclonal GFP antiserum, and Dr. Brian Cooke (Monash University, Melbourne) for polyclonal PfSBP1 antibodies.

V. References

Allred, D.R., Hines, S.A., and Ahrens, K.P. (1993). Isolate-specific parasite antigens of the Babesia bovis-infected erythrocyte surface. Molecular and biochemical parasitology *60*, 121-132.

Asada, M., Goto, Y., Yahata, K., Yokoyama, N., Kawai, S., Inoue, N., Kaneko, O., and Kawazu, S. (2012). Gliding Motility of Babesia bovis Merozoites Visualized by Time-Lapse Video Microscopy. PloS one *7*, e35227.

Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. J Mol Biol *340*, 783-795.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes & development *23*, 781-783.

Bhattacharjee, S., Hiller, N.L., Liolios, K., Win, J., Kanneganti, T.D., Young, C., Kamoun, S., and Haldar, K. (2006). The malarial host-targeting signal is conserved in the Irish potato famine pathogen. PLoS pathogens *2*, e50.

Bhattacharjee, S., Stahelin, R.V., Speicher, K.D., Speicher, D.W., and Haldar, K. (2012). Endoplasmic Reticulum PI(3)P Lipid Binding Targets Malaria Proteins to the Host Cell. Cell *148*, 201-212.

Blisnick, T., Morales Betoulle, M.E., Barale, J.C., Uzureau, P., Berry, L., Desroses, S., Fujioka, H., Mattei, D., and Braun Breton, C. (2000). Pfsbp1, a Maurer's cleft Plasmodium falciparum protein, is associated with the erythrocyte skeleton. Molecular and biochemical parasitology *111*, 107-121.

Boddey, J.A., Carvalho, T.G., Hodder, A.N., Sargeant, T.J., Sleebs, B.E., Marapana, D., Lopaticki, S., Nebl, T., and Cowman, A.F. (2013). Role of Plasmepsin V in Export of Diverse Protein Families from the Plasmodium falciparum Exportome. Traffic.

Boddey, J.A., Hodder, A.N., Gunther, S., Gilson, P.R., Patsiouras, H., Kapp, E.A., Pearce, J.A., de Koning-Ward, T.F., Simpson, R.J., Crabb, B.S.*, et al.* (2010). An aspartyl protease directs malaria effector proteins to the host cell. Nature *463*, 627-631.

Bokhari, A.A., Solomon, T., and Desai, S.A. (2008). Two distinct mechanisms of transport through the plasmodial surface anion channel. The Journal of membrane biology *226*, 27-34.

Brayton, K.A., Lau, A.O., Herndon, D.R., Hannick, L., Kappmeyer, L.S., Berens, S.J., Bidwell, S.L., Brown, W.C., Crabtree, J., Fadrosh, D.*, et al.* (2007). Genome sequence of Babesia bovis and comparative analysis of apicomplexan hemoprotozoa. PLoS pathogens *3*, 1401-1413.

Cornillot, E., Hadj-Kaddour, K., Dassouli, A., Noel, B., Ranwez, V., Vacherie, B., Augagneur, Y., Bres, V., Duclos, A., Randazzo, S.*, et al.* (2012). Sequencing of the smallest Apicomplexan genome from the human pathogen Babesia microti. Nucleic acids research *40*, 9102-9114.

de Koning-Ward, T.F., Gilson, P.R., Boddey, J.A., Rug, M., Smith, B.J., Papenfuss, A.T., Sanders, P.R., Lundie, R.J., Maier, A.G., Cowman, A.F.*, et al.* (2009). A newly discovered protein export machine in malaria parasites. Nature *459*, 945-949.

Dowling, S.C., Perryman, L.E., and Jasmer, D.P. (1996). A Babesia bovis 225-kilodalton spherical-body protein: localization to the cytoplasmic face of infected erythrocytes after merozoite invasion. Infection and immunity *64*, 2618-2626.

Enright, A.J., Van Dongen, S., and Ouzounis, C.A. (2002). An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res *30*, 1575-1584.

Ferreri, L.M., Brayton, K.A., Sondgeroth, K.S., Lau, A.O., Suarez, C.E., and McElwain, T.F. (2012). Expression and strain variation of the novel "small open reading frame" (smorf) multigene family in Babesia bovis. International journal for parasitology *42*, 131-138.

Franssen, F.F., Gaffar, F.R., Yatsuda, A.P., and de Vries, E. (2003). Characterisation of erythrocyte invasion by Babesia bovis merozoites efficiently released from their host cell after high-voltage pulsing. Microbes and infection / Institut Pasteur *5*, 365-372.

Gohil, S., Kats, L.M., Seemann, T., Fernandez, K.M., Siddiqui, G., and Cooke, B.M. (2013). Bioinformatic prediction of the exportome of Babesia bovis and identification of novel proteins in parasite-infected red blood cells. International journal for parasitology.

Gruring, C., Heiber, A., Kruse, F., Flemming, S., Franci, G., Colombo, S.F., Fasana, E., Schoeler, H., Borgese, N., Stunnenberg, H.G.*, et al.* (2012). Uncovering common principles in protein export of malaria parasites. Cell host & microbe *12*, 717-729.

Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C., and Haldar, K. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science (New York, N.Y.) *306*, 1934-1937.

Hiss, J.A., Przyborski, J.M., Schwarte, F., Lingelbach, K., and Schneider, G. (2008). The Plasmodium export element revisited. PloS one *3*, e1560.

Hsiao, C.H., Luisa Hiller, N., Haldar, K., and Knoll, L.J. (2013). A HT/PEXEL Motif in Toxoplasma Dense Granule Proteins is a Signal for Protein Cleavage but not Export into the Host Cell. Traffic.

Huang, B.Q., Chen, X.M., and LaRusso, N.F. (2004). Cryptosporidium parvum attachment to and internalization by human biliary epithelia in vitro: a morphologic study. The Journal of parasitology *90*, 212-221.

Hutchings, C.L., Li, A., Fernandez, K.M., Fletcher, T., Jackson, L.A., Molloy, J.B., Jorgensen, W.K., Lim, C.T., and Cooke, B.M. (2007). New insights into the altered adhesive and mechanical properties of red blood cells parasitized by Babesia bovis. Molecular microbiology *65*, 1092-1105.

Janssen, C.S., Phillips, R.S., Turner, C.M., and Barrett, M.P. (2004). Plasmodium interspersed repeats: the major multigene superfamily of malaria parasites. Nucleic acids research *32*, 5712-5720.

Jung, C., Lee, C.Y., and Grigg, M.E. (2004). The SRS superfamily of Toxoplasma surface proteins. International journal for parasitology *34*, 285-296.

Kang, B.H. (2010). Electron microscopy and high-pressure freezing of Arabidopsis. Methods in cell biology *96*, 259-283.

Kappmeyer, L.S., Thiagarajan, M., Herndon, D.R., Ramsay, J.D., Caler, E., Djikeng, A., Gillespie, J.J., Lau, A.O., Roalson, E.H., Silva, J.C*., et al.* (2012). Comparative genomic analysis and phylogenetic position of Theileria equi. BMC genomics *13*, 603.

Knuepfer, E., Rug, M., and Cowman, A.F. (2005). Function of the plasmodium export element can be blocked by green fluorescent protein. Molecular and biochemical parasitology *142*, 258-262.

Leech, J.H., Barnwell, J.W., Miller, L.H., and Howard, R.J. (1984). Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. The Journal of experimental medicine *159*, 1567-1575.

Levy, M.G., and Ristic, M. (1980). Babesia bovis: continuous cultivation in a microaerophilous stationary phase culture. Science (New York, N.Y.) *207*, 1218-1220.

Li, W., Zhang, N., Liang, X., Li, J., Gong, P., Yu, X., Ma, G., Ryan, U.M., and Zhang, X. (2009). Transient transfection of Cryptosporidium parvum using green fluorescent protein (GFP) as a marker. Molecular and biochemical parasitology *168*, 143-148.

Lingelbach, K., and Joiner, K.A. (1998). The parasitophorous vacuole membrane surrounding Plasmodium and Toxoplasma: an unusual compartment in infected cells. Journal of cell science *111 (Pt 11)*, 1467-1475.

Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004). Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science (New York, N.Y.) *306*, 1930-1933.

Marti, M., Li, Y., Schraner, E.M., Wild, P., Kohler, P., and Hehl, A.B. (2003). The secretory apparatus of an ancient eukaryote: protein sorting to separate export pathways occurs before formation of transient Golgi-like compartments. Molecular biology of the cell *14*, 1433-1447.

Morahan, B.J., Strobel, C., Hasan, U., Czesny, B., Mantel, P.Y., Marti, M., Eksi, S., and Williamson, K.C. (2011). Functional analysis of the exported type IV HSP40 protein PfGECO in Plasmodium falciparum gametocytes. Eukaryotic cell *10*, 1492-1503.

O'Connor, R.M., and Allred, D.R. (2000). Selection of Babesia bovis-infected erythrocytes for adhesion to endothelial cells coselects for altered variant erythrocyte surface antigen isoforms. Journal of immunology *164*, 2037-2045.

O'Connor, R.M., Lane, T.J., Stroup, S.E., and Allred, D.R. (1997). Characterization of a variant erythrocyte surface antigen (VESA1) expressed by Babesia bovis during antigenic variation. Molecular and biochemical parasitology *89*, 259-270.

Peixoto, L., Chen, F., Harb, O.S., Davis, P.H., Beiting, D.P., Brownback, C.S., Ouloguem, D., and Roos, D.S. (2010). Integrative genomic approaches highlight a family of parasite-specific kinases that regulate host responses. Cell host & microbe *8*, 208-218.

Rudzinska, M.A., and Trager, W. (1977). Formation of merozoites in intraerythrocytic Babesia microti: an ultrastructural study. Canadian journal of zoology *55*, 928-938.

Rug, M., Wickham, M.E., Foley, M., Cowman, A.F., and Tilley, L. (2004). Correct promoter control is needed for trafficking of the ring-infected erythrocyte surface antigen to the host cytosol in transfected malaria parasites. Infection and immunity *72*, 6095-6105.

Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A., and Goldberg, D.E. (2010). Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. Nature *463*, 632-636.

Sargeant, T.J., Marti, M., Caler, E., Carlton, J.M., Simpson, K., Speed, T.P., and Cowman, A.F. (2006). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome biology *7*, R12.

Schmuckli-Maurer, J., Casanova, C., Schmied, S., Affentranger, S., Parvanova, I., Kang'a, S., Nene, V., Katzer, F., McKeever, D., Muller, J.*, et al.* (2009). Expression analysis of the Theileria parva subtelomere-encoded variable secreted protein gene family. PLoS ONE *4*, e4839.

Struck, N.S., Herrmann, S., Schmuck-Barkmann, I., de Souza Dias, S., Haase, S., Cabrera, A.L., Treeck, M., Bruns, C., Langer, C., Cowman, A.F.*, et al.* (2008). Spatial dissection of the cis- and trans-Golgi compartments in the malaria parasite Plasmodium falciparum. Molecular microbiology *67*, 1320-1330.

Terkawi, M.A., Seuseu, F.J., Eko-Wibowo, P., Huyen, N.X., Minoda, Y., AbouLaila, M., Kawai, S., Yokoyama, N., Xuan, X., and Igarashi, I. (2011). Secretion of a new spherical body protein of Babesia bovis into the cytoplasm of infected erythrocytes. Molecular and biochemical parasitology *178*, 40-45.

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.

Tonkin, C.J., van Dooren, G.G., Spurck, T.P., Struck, N.S., Good, R.T., Handman, E., Cowman, A.F., and McFadden, G.I. (2004). Localization of organellar proteins in Plasmodium falciparum using a novel set of transfection vectors and a new immunofluorescence fixation method. Molecular and biochemical parasitology *137*, 13-21.

Trager, W., and Jensen, J.B. (1976). Human malaria parasites in continuous culture. Science (New York, N.Y.) *193*, 673-675.

Vega, C.A., Buening, G.M., Rodriguez, S.D., and Carson, C.A. (1986). Concentration and enzyme content of in vitro-cultured Babesia bigemina-infected erythrocytes. The Journal of protozoology *33*, 514-518.

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

Wickham, M.E., Rug, M., Ralph, S.A., Klonis, N., McFadden, G.I., Tilley, L., and Cowman, A.F. (2001). Trafficking and assembly of the cytoadherence complex in Plasmodium falciparum-infected human erythrocytes. The EMBO journal *20*, 5636-5649.

Wright, I.G. (1972). An electron microscopic study of intravascular agglutination in the cerebral cortex due to Babesia argentina infection. International journal for parasitology *2*, 209-215.

Wright, I.G. (1973). Ultrastructural changes in Babesia argentina-infected erythrocytes in kidney capillaries. The Journal of parasitology *59*, 735-736.

Wu, Y., Kirkman, L.A., and Wellems, T.E. (1996). Transformation of Plasmodium falciparum malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. Proceedings of the National Academy of Sciences of the United States of America *93*, 1130-1134.

CHAPTER III

Transcriptional profiling of genes at the host-parasite interface identifies expression modules involved in malaria pathogenesis

CONTRIBUTIONS

We would first like to thank the families and the children who are enrolled at the Malaria Research Ward in the Queen Elizabeth Central Hospital in Blantyre Malawi for the study of pediatric cerebral malaria, without whom the work in this chapter would not have been possible.

This work was done in collaboration with Dr. Curtis Huttenhower, Dr. Terrie Taylor, Dr. Danny Milner, and Dr. Karl Seydel. The computational analyses (generation of gene co-expression modules and GSEA analysis) were performed by Vagheesh Narasimhan and Kevin Oh and the remaining by Karell Pellé. The chapter was written by Karell Pellé with helpful discussion and input by Dr. Curtis Huttenhower, Dr. Matthias Marti, Dr. Pierre-Yves Mantel and Vagheesh Narasimhan.

I. Introduction

Malaria is a disease that has claimed the life of over 1.2 million people in 2010 alone (Murray et al., 2012). This disease is mainly caused by *Plasmodium falciparum*, a protozoan parasite transmitted by Anopheles mosquitoes. Although a preventable and treatable disease, the growing emergence of drug resistance against first line chemotherapies and insecticides, and the lack of an effective vaccine make this disease a global concern (WHO, 2012). In particular, the complexity of this parasite's life cycle and the gap in knowledge of a great proportion of its genome render identifying key regulators of pathways involved in pathogenesis a challenge.

Plasmodium falciparum has a genome of ~5300 genes. Over 60% of the genome remains uncharacterized in part due to the absence of phylogenetic information (Gardner et al., 2002). However, genes implicated in pathogenesis have been identified and act primarily during the erythrocytic stages of infection, in particular in trophozoites and schizonts (Beeson et al., 2002). These two stages are rare in human peripheral blood as they sequester in the microvasculature of tissues where they cause pathologies. In *falciparum* malaria, pathology (and by extension clinical outcome) is dependent on multiple factors including parasite pathogenesis, host polymorphisms and immune status (Miller et al., 2002). The parasitic processes that contribute to pathogenesis include invasion, host-cell remodeling, cytoadherence and immune evasion. In disease, invasion can be linked to haemolytic anemia and impaired erythropoiesis via the release of malaria toxins from lysis of infected erythocytes in the blood (Miller et al., 2002).

Host cell remodeling includes increased host membrane rigidity through the incorporation of parasite factors, associated with blood vessel occlusion and tissue hypoxia (Glenister et al., 2002; Hosseini and Feng, 2012). Other changes involve the appearance of membrane protrusions called "knobs", which act as platforms for the presentation of membrane-anchored variant surface antigens called *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) (Aikawa et al., 1996; Nagao et al., 2000). PfEMP1 mediate cytoadherence to host cell receptors on the endothelium of the microvasculature. This results in the sequestration of parasites in various host tissues where they elicit local immune responses and pathological sequelae leading in some instances to the severe diseases of cerebral malaria (CM) in

immunologically naïve children or pregnancy-associated malaria (PAM) in pregnant women (Fried and Duffy, 1996; Turner et al., 1998). Cytoadherence is mediated by mutual exclusive expression of variants of PfEMP1 surface proteins encoded by *var* genes (Scherf et al., 1998; Su et al., 1995). *Var* genes are classified by promoter sequence type and their chromosomal location in 5 different Ups group: UpsA, B, C, D and E (Lavstsen et al., 2003). While the UpsE *var*, *var2csa*, is linked to PAM via PfEMP1 binding to chondroitin sulfate A (CSA) in the placenta, UpsA *var*s are linked to CM via intercellular adhesion molecule 1 (ICAM-1) binding in the brain (Avril et al., 2012; Claessens et al., 2012; Duffy et al., 2006; Jensen et al., 2004; Lavstsen et al., 2012). The switch in PfEMP1 surface expression allows for both sequestration, which enables parasites to grow and multiply while escaping spleen-dependent killing and immune evasion (Miller et al., 2002; Smith et al., 1995).

Many parasite factors are known to modify the red blood cell membrane. These factors are secreted during invasion or exported in the host cell during growth and form protein-protein interaction with components of the RBC cytoskeleton (Maier et al., 2009). Some of these factors include the main knob protein, knob associated histidine protein (KAHRP), which binds to alpha-spectrin repeats, actin and ankyrin (Kilejian et al., 1991; Magowan et al., 2000; Pei et al., 2005). KAHRP also binds PfEMP1 and KAHRP$^-$ parasites lack cytoadherence ability (Crabb et al., 1997; Waller et al., 1999). The mature parasite-infected erythrocyte surface antigen (MESA) binds protein 4.1 and the absence of this interaction results in parasite growth defects (Waller et al., 2002). The ring-infected surface antigen (RESA) released during invasion binds protein 4.1R and provides iRBC resistance to febrile temperatures (Da Silva et al., 1994; Pei et al., 2007a; Silva et al., 2005). *P. falciparum* erythrocyte membrane protein 3 (PfEMP3) through its interaction with spectrin at a site near actin-protein 4.1R junctions alters the deformability properties of the iRBC membrane and the disruption of the encoding loci results in the loss of cytoadherence (Pei et al., 2007b; Waterkeyn et al., 2000). The identification of these host-remodeling proteins required a gene-by-gene targeted approach using some challenging tools, forward and reverse genetics and mass spectrometry. Recently, in an attempt to systematically identify proteins with similar function, Maier and colleagues selected a set of genes based several criteria including the

presence or absence of a *Plasmodium* Export Element (PEXEL) motif which targets proteins to the host cell and transcriptional expression during the intra-erythrocytic lifecycle followed by functional validation (Hiller et al., 2004; Maier et al., 2008; Marti et al., 2004). Over 80 genes were knocked-out and 17 had a phenotype related to cytoadherence, knob-formation or iRBC membrane rigidity (Maier et al., 2008). The majority of these hits (76%) are genes containing a PEXEL motif indicating the importance of exported proteins in host cell remodeling, cytoadherence and consequently pathogenesis.

The functional characterization of single genes in their involvement in pathogenesis is a laborious process. Furthermore, understanding the relationship of these genes to one another and their role in clinical outcome is challenging. However, previously identification of important parasite biological pathways using whole transcriptional profiling has been achieved. In these studies, significant changes in the expression of certain genes in wild type parasites or in response to drug or heat shock showed that predicting gene function with robustness is feasible (Date and Stoeckert, 2006; Hu et al., 2010; Oakley et al., 2007). These studies integrated *in silico* and experimental data to generate global and functional relationships between gene products. Date and Stoeckert for example in the goal to determine *P. falciparum* gene function integrated phylogenetic, expression, Rosetta stone linkage and mass spectrometry data into a model to compute a protein interaction network (2006). This generated the *P. falciparum* “interactome” (Date and Stoeckert, 2006). They identify sub-networks, or hubs in which uncharacterized genes can be functionally linked to a pathway by association. Hu and colleagues on the other hand take a more targeted approach (2010). In order to identify the full set of genes involved in invasion, or the “invadome”, they subject parasites to different growth inhibiting compounds and create a interaction network by integration of phylogenetic, expression, domain-domain interaction prediction and yeast-2-hybrid data (Hu et al., 2010).

The approach used in these studies however may not be suitable for the identification of genes involved in processes important *in vivo*, either modulated *in vivo* or even only occurring *in vivo*, such as pathogenesis. These genes are likely modulated by factors in the host that are absent in culture conditions (i.e. humoral immunity). In fact, gene products involved in antigenic

variation can shift expression pattern during cultivation. This phenomenon was observed in *var* genes during culture-adaptation where an overall downregulation of all *vars* occured just after 10 days (Peters et al., 2007). Additionally, surface antigens involved in antigenic variations in *P. knowlesi*, such as the schizont-infected cell agglutination (SICA) fail to be detected in agglutination experiments when passaged in splenectomised monkeys (Barnwell et al., 1983). This is due to the lack of a host immune response, which in spleen intact monkeys would kill all circulating infected cells. Thus, the use of expression data generated from parasites sampled from malaria-stricken patients to identify pathways such as the invadome is crucial in particular if there is evidence of gene expression changes under different environment or shifts in acting proteins within pathways. Here, we aim to define global and local functional relationships of genes involved pathogenesis. We use the *P. falciparum* exportome as a representative list of genes encoding proteins at the host-parasite interface (Boddey et al., 2013; Sargeant et al., 2006). Using large-scale parasite *in vivo* expression data we are able to show that genes encoding proteins at the host-parasite interface are co-expressed to form functional modules containing known parasite virulence genes.

II. Results

A. Transcriptional profiling of genes at the host-parasite interface

We first sought out to organize the exportome into functional modules of co-expressed genes using integration of malaria patient microarray transcriptional data. We selected 129 microarrays of peripheral blood samples from patients with malaria from different African geographic and transmission areas, Sénégal and Malawi (see Methods). These samples were collected at time of admission and therefore are dominantly ring stage parasites. After normalization, all microarrays were integrated. We then computed normalized Pearson correlations (z-score) of the expression of all gene pairs in the transcriptome to generate a global functional network focused on *in vivo* co-expression.

Within this global functional network, we focused on a subgraph containing genes from the exportome. We identified 25 exportome genes that are constitutively expressed (high expression across most patients) and 98 genes that are variably expressed (ranging in expression across patients). We then clustered this exportome subgraph into 10 constitutive and 10 variable modules, which we call “Cons” and “Variable” respectively (Figure 3-1). These 20 modules have different network properties, ranging from gene family composition, delta z score and epigenetic marks of regulation

B. Properties of constitutive and variable export modules

The export modules range in size (n= [1-35]) and connectivity (Δ= [0.383-2.397]). Connectivity (or delta z-score) here is defined as the within module average z-score (average of the correlations between all gene pairs within the module) minus the average z-score of gene pairs between module genes and the transcriptome. It reflects the significance of the connections within a module as opposed to background. Most export modules, both Constitutive and Variable, contain 1-10 genes. Two outliers, Variable modules 6 and 8, contain 21 and 35 genes and have high connectivity, Δ= 1.958 and 1.819 respectively (Figure 3-1B). The export modules are mainly composed of genes from the *phist, fikk, stevor, rif* and *hyp* families, with a few single-member families such as *kahrp* and *garp* and other unannotated genes. Interestingly, *rifins* and *stevors* are predominant in variable modules in particular V6 and V8, whereas members of the other gene families are found in both constitutive and variable export modules (Figure 3-1C). One export module Cons 6, is composed of 5 exportome genes and has an overall connectivity of Δ= 1.779 (Figure 3-1B,D). Members of this module include two genes which are directly involved in PfEMP1 export to the iRBC surface and abolishment of PfEMP1-mediated cytoadherence, MAL7P1.172 and PF13_0076. These genes are 2 of the 6 hits identified in a genetic deletion screen of over 80 genes for cytoadherence (Maier et al., 2008). Another gene of Cons 6 is *rex3*, whose locus is found on a segment of chromosome 9, which is linked to cytoadherence (Day et al., 1993) (Figure 3-1B,D). We therefore confirm this module as a cytoadherence module.

To also identify genes that are not predicted at the host-parasite interface (i.e. not exported) but could potentially be functional with export modules, we extended the edges between export module nodes and nodes in the *P. falciparum* transcriptome within the *in vivo* integrated network. In doing so, we identified a module linked to invasion (Figure 3-1B,D). Merozoite surface protein 4, MSP4, belongs to Variable module 5 (Δ= 0.383) whose other members include non-exported genes such as glideosome-associated protein 50 (GAP50), a known component of the glideosome complex which powers merozoite gliding motility, actin, actin-depolymerizing factor, calcium-dependent protein kinase 1 (CDPK1) and genes involved in energy break-down such as an inorganic pyrophosphatase and P-type ATPase (Baum et al., 2006). Interestingly, these genes or their paralogs are also present in the invadome identified by Hu and colleagues after growth perturbations on *in vitro* parasite strains (Hu et al., 2010).

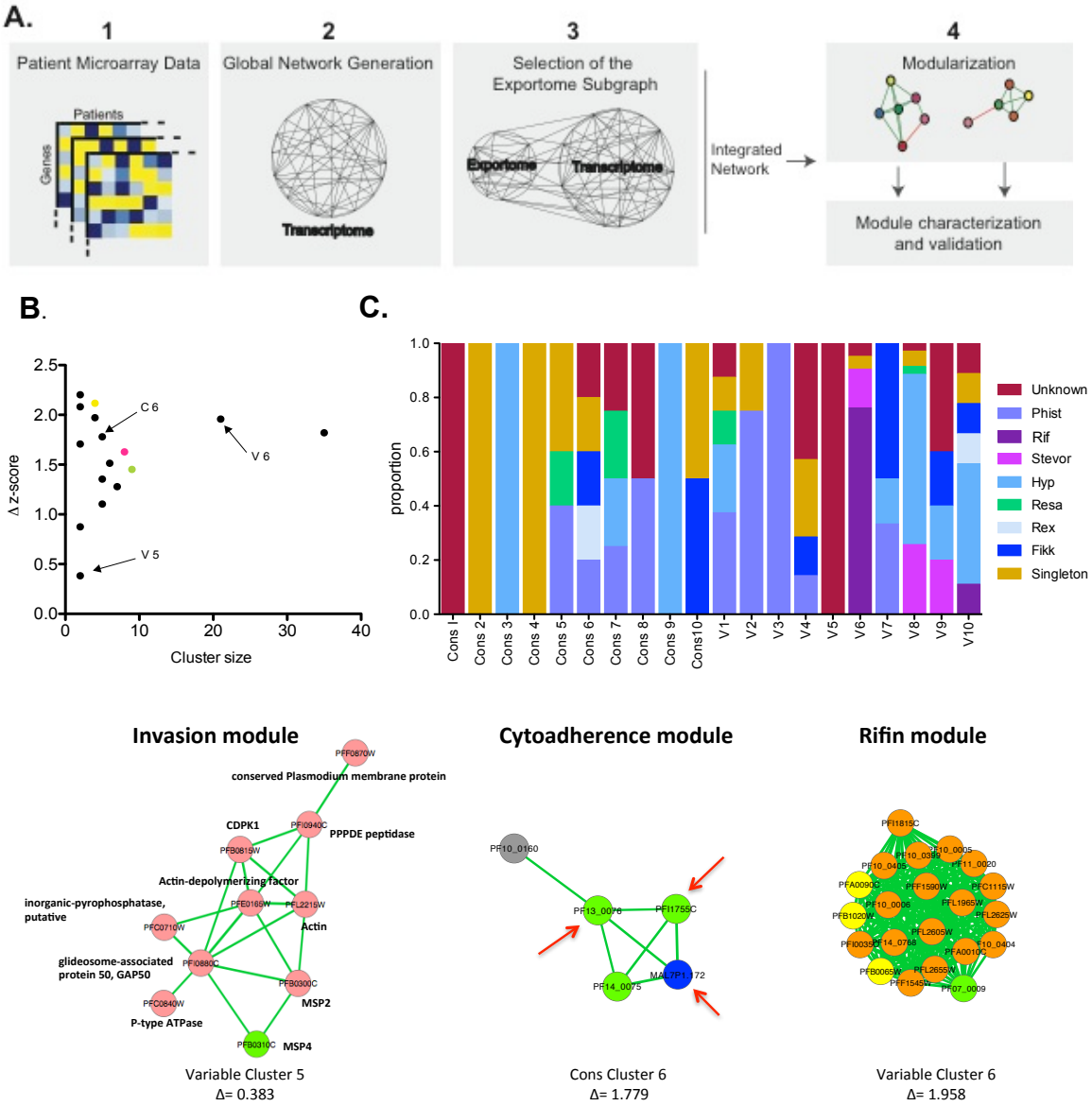


Figure 3- 1: **Generation of export modules and their characteristics. A.** Schematic summary of the methodology to generate expression modules of genes at the host-parasite interface **B.** Delta z-score vs. module size. Variable modules 1, 2, and 10, which will be used later on for validation are shown in pink, gold and green respectively. **C.** Family composition of constitutive and variable export modules. **D**. Examples of modules include a module involved in cytoadherence, "cytoadherence module" (red arrows indicate genes involved in cytoadherence (Day et al., 1993; Maier et al., 2008)); gliding motility, "invasion module"; and a module involved in antigenic variation, "rifin module". These three modules are indicated by arrows on the delta z-score graph in B.

C. Module validation by qRT-PCR in cerebral malaria patients

We then sought to validate variable gene modules by quantitative real time-PCR using peripheral blood samples from Malawian children diagnosed with CM (n=30). This is a different cohort than the one used for microarrays. These children were enrolled in 2010. They are part of the same ongoing study of pediatric malaria in Malawi. Three variable export modules were selected for validation based on the following criteria: a) high delta z-score, b) 3D7 *in vitro* expression of some gene members, c) role in cytoadherence, remodeling of the iRBC plasma membrane, or surface expression and d) unknown function. Variable export module 1 included *phista* (*PF14_0752*), *resa-2* (*PF11_0512), phistb* (*PFB0900c*), pf70 (*PF10_0025*) and *phistc* (*PF10_0021*); variable export module 2 included the surface protein *piesp2* (*PFE0060w*) and *phistb* (*PFB0080c*); and variable export module 10 included *rex4* (*PFI1760w*) and *hyp11* (*PFI1750c*) (Figure 3-2A). Genes within these three modules were strongly correlated (p-values< 2E-4) (Figure 3-2B). We did notice that V2 and V10 were correlated to each other and both were anti-correlated to V1, although this fell below significance (Figure 3-2B).

To assess whether variable export modules are co-expressed with *var* genes, we computed correlations between genes from variable modules 1, 2, 10 and *var* genes. Although UpsA *var* genes are dominant in this cohort (p<0.0001) as expected for parasites from cerebral malaria patients, we did not see any significant correlation between Ups groups and *var* genes (Figure 3-3).

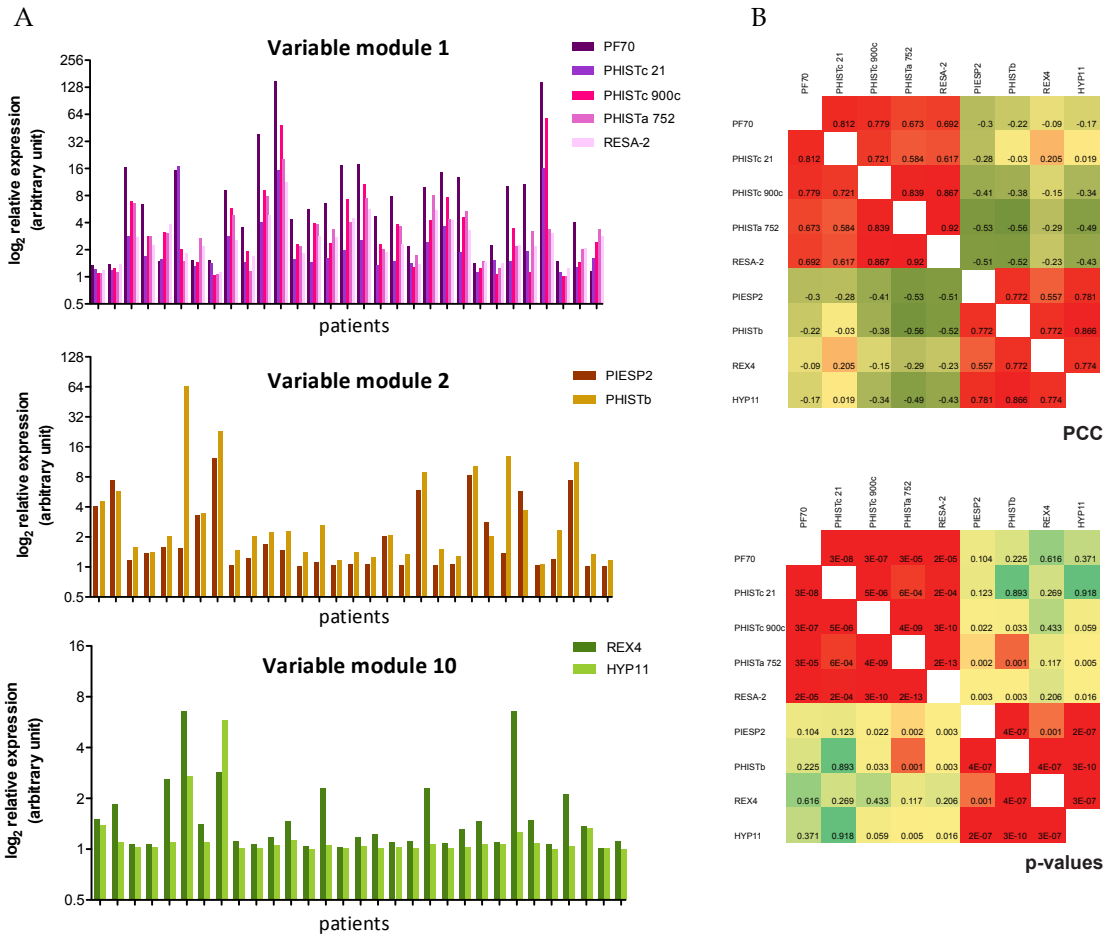


Figure 3- 2: **Transcriptional co-expression of variable export modules in *P. falciparum* parasites from CM patients isolates. A.** Transcriptional expression of candidate genes in V1, V2 and V10 was determined in 30 CM isolates (x-axis) by qRT-PCR. Each bar corresponds to the relative expression of module representative genes for each patient. The expression is shown relative to the endogenous marker, seryl-t-RNA synthetase. **B**. Pearson correlation of all candidate gene pairs shows co-expression within each module and between V2 and V10 genes. The top matrix shows Pearson correlations between all gene pairs. The bottom matric shows corresponding p-values. Red signifies higher correlation/p-value.

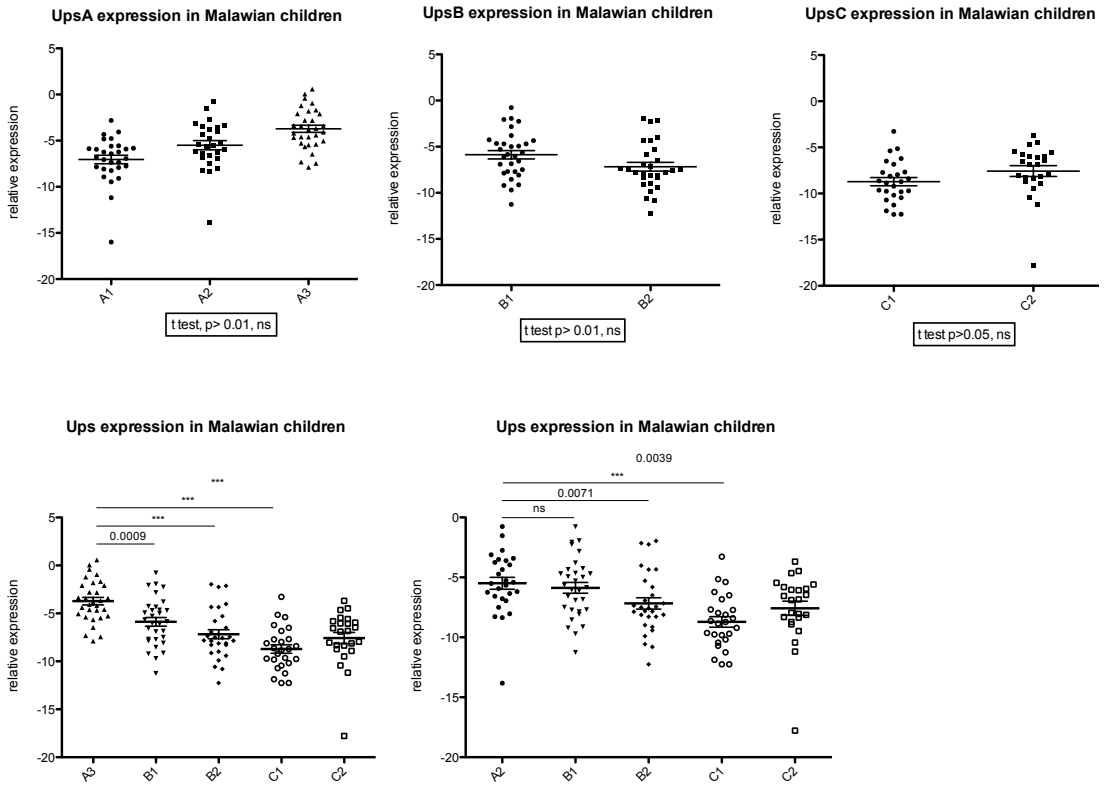


Figure 3- 3: ***var*** **transcript abundance by Ups group in parasites from CM patients**. The transcript abundance of Ups A, B and C was determined in parasites from children with cerebral malaria collected at admission. Different primer sets were used for each Ups group. Abundance is represented as log2 relative expression to seryl-t-RNA synthetase. In each graph, each data point represents one sample, the mean is shown with a bar. A student t-test was used to determine differences in the means. *** $p<0.0001$. No significance was reached when comparing UpsA using the A1 primer set to UpsB or C.

D. Module validation by qRT-PCR in an *in vitro* model of cerebral malaria

Previous studies show that parasites selected on human brain microvascular endothelial cells (HBMEC) upregulate the expression of a subset of UpsA-like *var* genes, which are also dominant in circulating parasites from children with CM (Avril et al., 2012; Claessens et al., 2012; Lavstsen et al., 2012). These HBMEC-selected parasites can also bind other human tissues such as pulmonary and cardiac microvascular endothelial cells, HPMEC and HCMEC respectively (Avril et al., 2012). Parasites selected on human tissues can therefore be used as an i*n vitro* model of CM. Thus, we surveyed the expression of variable modules in an *in vitro* model of CM using ring stage parasites that were panned on HBMEC and further selected on HPMEC (Figure 3-4A). We also used two cell lines, R29 and PF13 that were independently selected for rosetting. These rosetters upregulate UpsA *var* genes (Janes et al., 2011). These parasite lines were kindly provided by Dr. Smith (Seattle Biomedical Research Institute). Here, variable modules overall correlated less than in CM patients, especially variable module 10 (V1 p-values>0.002; V2 p-value= 0.004; V10 p-value=0.125) (Figure 3-4B). Of notice, two genes in V1, *phista (PF14_0752*) and *resa-2* (*PF11_0512*), did not correlate with the rest of genes in this export module (p-values> 0.235) suggesting that these two genes may be differentially regulated from the rest of the module. The fact that we could not reconstitute with high significance variable export modules in panned lines indicates that they are modulated *in vivo.* This also confirms the significant expression correlation that exists between these genes in their respective modules as defined by using *in vivo* co-expression data.

To understand the lack of correlation between genes in these 3 variable export modules and *var* genes, we also checked for *var* transcript abundance by Ups group. However, no Ups group (A, B or C) is dominant in these selected lines (Figure 3-5). We suspect that this lack of selection for the dominance of one *var* group is due to culturing as described by others previously and this also has similar effect on the selection of variable export modules *in vitro* (Peters et al., 2007).

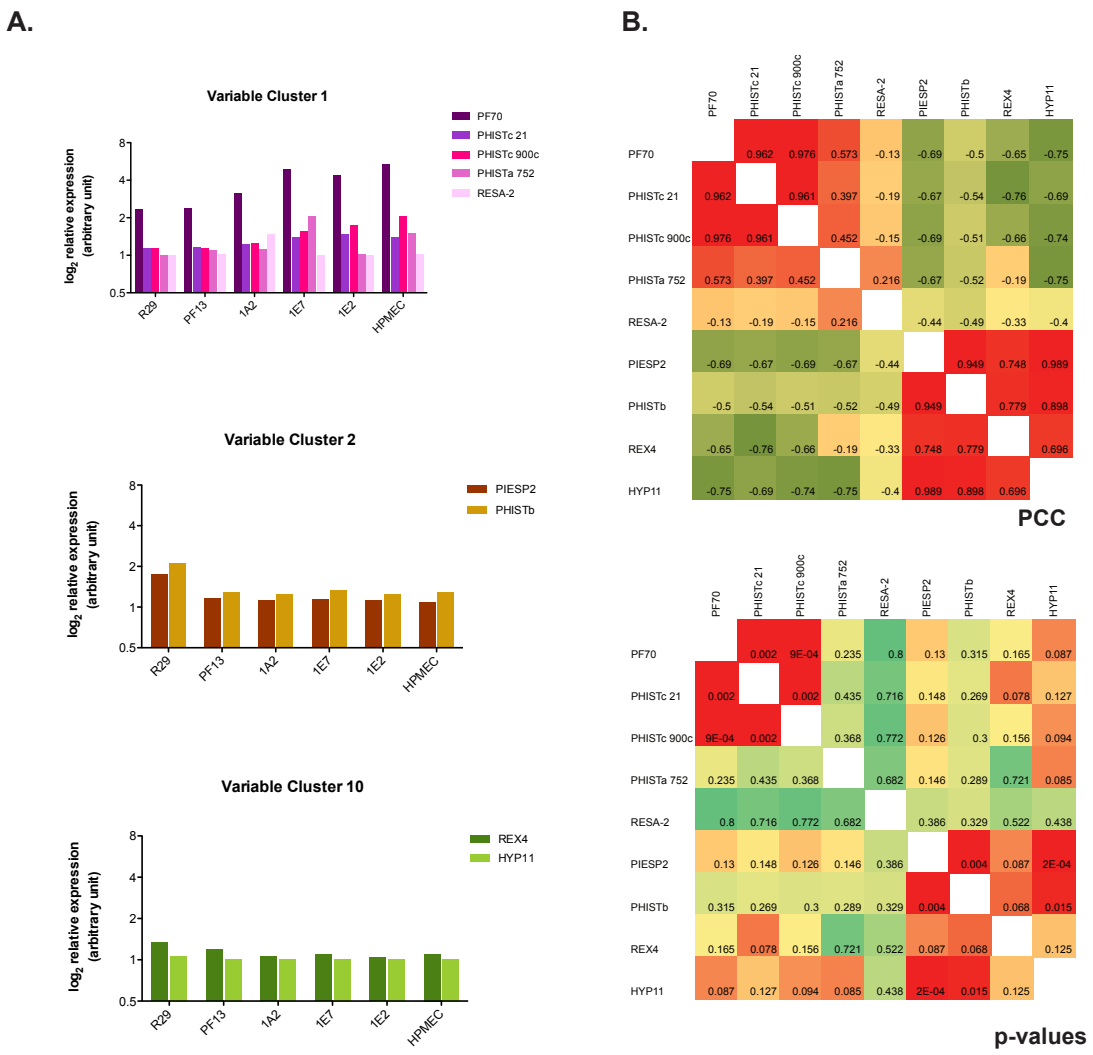


Figure 3- 4: Transcriptional co-expression of variable export modules 1, 2 and 10 in *P. falciparum* parasites with cytoadhesive properties. **A**. Transcriptional expression of candidate genes in V1, V2 and V10 was determined in 6 parasite lines: 4 IT4 cytoadherent lines and two rosetting lines R29 and PF13 (x-axis). The IT4 clones 1A2, 1E7, 1E2 were selected on HBMEC (brain) and a fourth independent HBMEC-adherent clone re-selected on HPMEC (lung). The expression of each gene is shown relative to the endogenous marker seryl-t-RNA synthetase. **B**. Pearson correlation of all candidate gene pairs was performed. The top matrix shows Pearson correlations between all gene pairs. The bottom matrix shows corresponding p-values. Red signifies higher correlation/p-value. Of notice, two genes in V1 do not significantly correlate with the rest of the module.

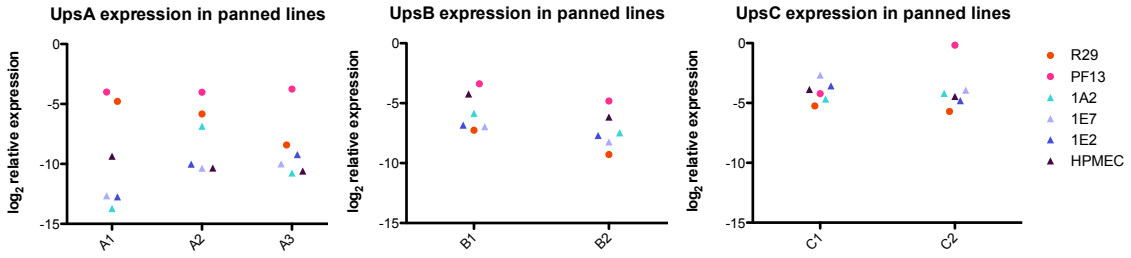


Figure 3- 5: ***Var*** **transcript abundance by Ups group in cytoadherence-selected parasites**. The transcript abundance of Ups A, B and C was determined in parasites selected for binding to different host cells. R29 and PF13 are rosetters (circles on graph). The IT4 clones 1A2, 1E7, 1E2 were selected on HBMEC and a fourth independent HBMEC-adherent clone re-selected on HPMEC (lung) (triangles on graph). Different primer sets were used for each Ups group. Abundance is represented as relative expression to an endogenous marker seryl-t-RNA synthetase.

E. A protein in variable export module I is exported to the host cell surface

To further validate gene-coexpression modules, we sought to determine the subcellular localization of a protein encoded by a gene in variable module I. *PF10_0025* is subtelomeric and maximally expressed very early in ring stages and encodes a protein called PF70 for its predicted molecular weight (Le Roch et al., 2004; Ma et al., 1996). This protein has a signal sequence followed by a PEXEL motif and therefore predicted to localize in the host cytosol (Marti et al., 2004). Also, the genetic deletion of PF70 in a cytoadherent parasite line, CS2, which is able to bind chondroitin sulfate A (CSA) through the UpsE *var* gene, *var2csa*, results in a loss of cytoadherence and surface expression of the PfEMP1, suggesting the essential role of PF70 in PfEMP1 trafficking to the surface (Maier et al., 2008). One other gene within this variable module 1 is also linked to PfEMP1. This is the *phistc* pf10_0021 which was showed to bind the intracellular domain of PfEMP1 and (Mayer et al., 2012). Therefore, we performed IFA experiments in 3D7 asexual parasites to determine if PF70 is exported in the host environment and whether it co-localizes with known virulence factors at the surface, PfEMP3 and KHARP. Both PfEMP3 and KAHRP bind host cytoskeleton proteins and their deletion results in a loss of cytoadherence (Pei et al., 2007; Waterkeyn et al., 2000; Kilejan et al., 1991; Magowan et al., 2000; Pei et al., 2005; Crabb et al., 1997; Waller et al., 1999). In early stages PF70 is readily detectable in the parasite but can also be seen as punctate structures at the surface of the red blood cell (Figure 3-6). The protein forms larger aggregates in schizont stages. Although PF70 does not fully localize with PfEMP3 or KAHRP, it seems associated with the iRBC membrane in foci deprived of PfEMP3 and KAHRP (Figure 3-6). It is possible that these proteins do not co-localize because the context in which co-localization and perhaps interaction occurs is altered. In fact, these IFAs were performed in unselected, long-term *in vitro* cultured 3D7 parasites. It may be that only under the right *in vivo* stimulus, proteins encoded by genes within variable export module 1 are highly co-expressed and present in the host environment to mediate certain functions (i.e. cytoadherence) lacking under unselected *in vitro* conditions. These stimuli as seen in the previous section are present *in vivo* but lack in culture-adapted parasites that have lost *var*

gene dominance (Figure 3- 2-5). Host cell expression, membership in a variable module and its effect on PfEMP1 trafficking increases the role of PF70 in host cell remodeling in particular cytoadherence. However, further work is needed to elucidate the function of this protein.

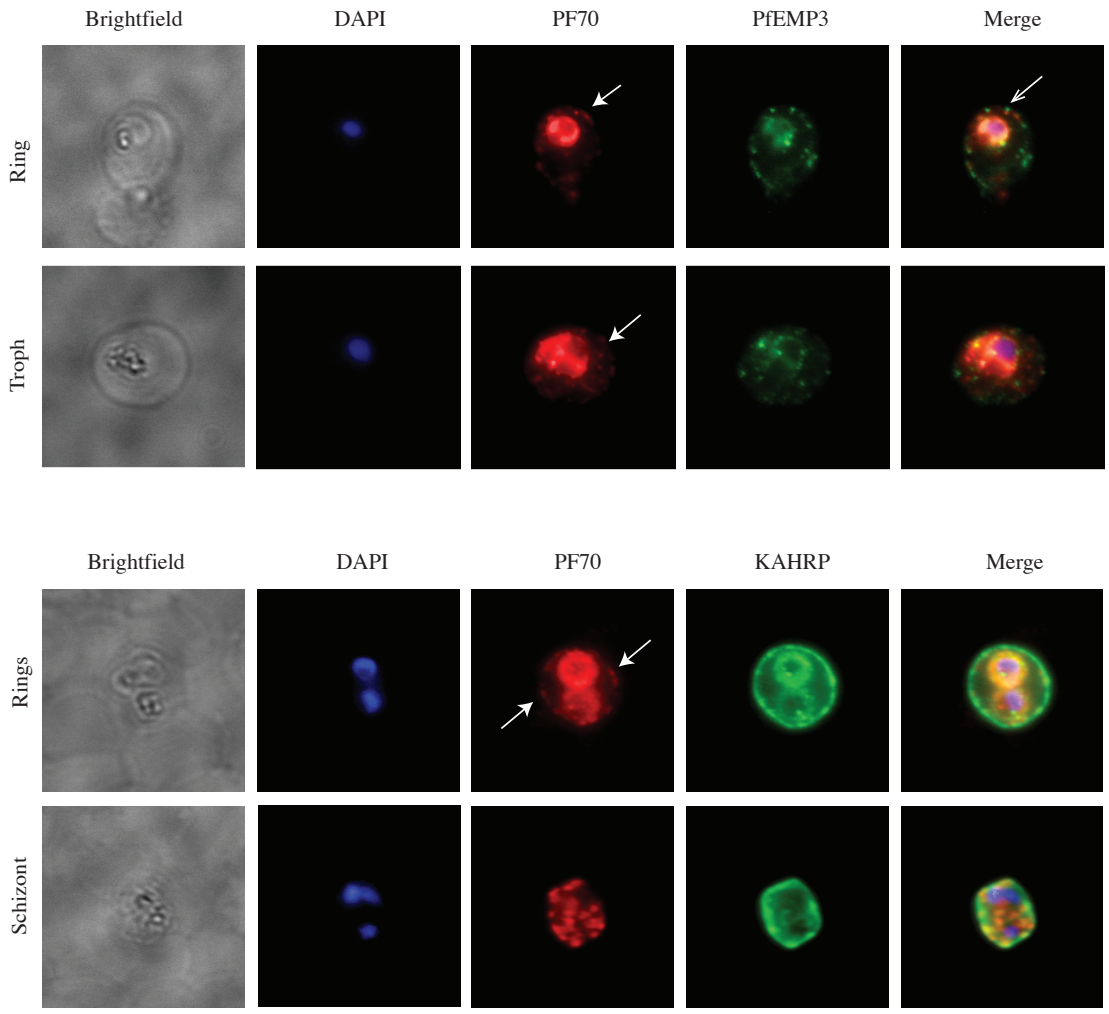


Figure 3- 6: Subcellular localization of PF70 in *P. falciparum* infected erythrocytes. Anti-PF70 antibodies were used to detect PF70 in asexual *P. falciparum* parasites. PF70 signal (red) was enriched in the parasite in ring stages as well as in the host cell as seen by punctate structures at the erythrocyte periphery and as large aggregates throughout the host in schizonts (white arrows). PF70 co-localized with PfEMP3 (top panel) and KAHRP (bottom panel) in the parasite but was only closely associated at the periphery of the host cell (open arrow). DAPI staining labels parasite nuclei.

F. Epigenetic marks of regulation are enriched in variable export modules

In *P. falciparum*, genes that display variant expression *in vitro* include those encoding proteins that are exported to the erythrocyte surface including *var* and genes involved in invasion. These genes are mostly located at the subtelomeres i.e. *var, rifs* and *stevors*, and are sometimes found in tandem (Gardner et al., 2002). Furthermore, variant expression is often associated with heterochromatin marks, where transcriptionally silent loci are enriched in H3K9me3 (Flueck et al., 2009; Lopez-Rubio et al., 2009; Salcedo-Amaya et al., 2009). In higher eukaryotes, H3K9me3 is a feature of constitutive heterochromatin (Trojer and Reinberg, 2007). In *P. falciparum* however, heterochromatin is a facultative state. Silent loci marked by H3K9me3 can be activated upon different environmental cues, giving rise to variegated gene expression.

To understand the variable and constitutive expression of exportome genes of export modules, we gathered four different types of information. First, if genes are subtelomeric; this information was retrieved for the database PlasmodDB. Second, if H3K9me3 marks are present at transcriptionally silent loci. This inforrmation can be found in two published studies: Lopez-Rubio et al., 2009 and Salcedo-Amaya et al., 2009. Third, if heterochromatin protein 1 (HP1) binds H3K9me3-enriched regions, published by Flueck et al., 2009. And lastly, whether genes show clonal variant expression in four different *P. falciparum* strains and thus take membership in the variantome published by Rovira-Graells et al., 2012 (Table 3-1). Strikingly, genes from Cons export modules do not show any enrichment for epigenetic marks of regulation nor clonal variation, although they are located at the subtelomeres. Conversely, a great proportion of genes from variable export modules are enriched for H3K9me3 and HP1 binding in their promoter regions during silent states, are part of the variantome and are located at the subtelomeres. Variable modules as opposed to Cons modules appear to be transcriptionally regulated by epigenetic mechanisms.

There are a few exceptions. One gene in Cons export modules is subtelomeric, undergoes clonal variation *in vitro* and has marks of epigenetic regulation (Table 3-1). This gene, PF13_0010, belongs to Cons 5 (n=6) and encodes a glycophorin-binding protein potentially

involved in invasion (PF13_0010). Within Cons 5, it is the least connected to the module, with an average within module z-score of 0.232 as compared to the module's overall average z-score of 0.942 (Figure 3-1). Because of the strong lack in epigenetic marks in Cons export modules PF13_0010 under stricter module definition may be classified in a variable module. We also noticed that only two genes in variable export module 1 are enriched for epigenetic marks of expression regulation. These two genes, *phista* (PF14_0752) and *resa-2* (PF11_0512) were used in the validation of variable module 1. They showed lower correlation with the rest of the export module in parasites panned on human brain and pulmonary endothelial cells and in rosetting parasites (Figure 3-4B). However, export modules were defined based on *in vivo* gene co-expression. The loss in correlation may be due to the environmental conditions in which parasites are assayed (host versus tissue culture system). We suggest that the mechanisms of regulation of the expression of certain genes may be different but the overall outcome *in vivo* is transcriptional co-expression. It is also possible that the heterochromatin marks that are detected in culture-adapted parasites do not cover the full gene set that it repressed by similar mechanisms in patients with active disease.

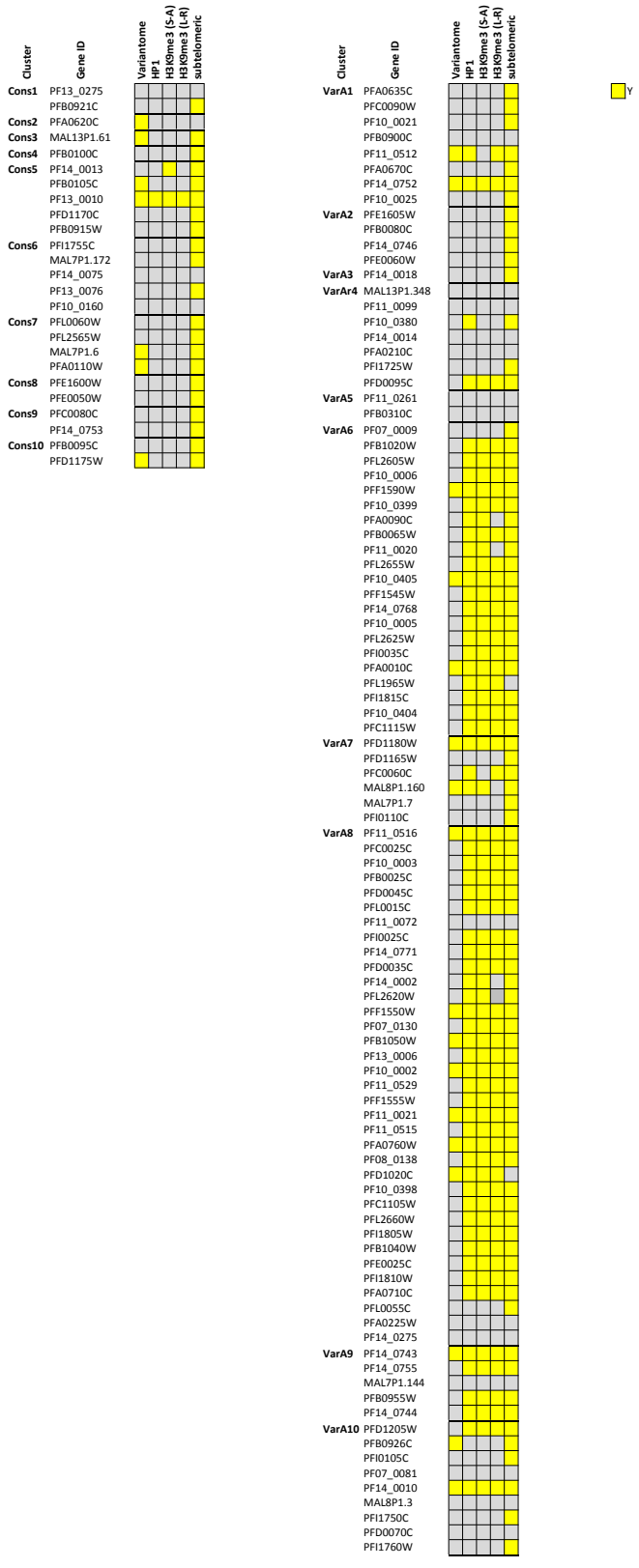


Table 3- 1: Variable export modules are associated with epigenetic marks of expression regulation. The table was generated using published data from 5 sources: Variantome (Rovira-Graells et al., 2012); HP1 binding at H3K9me3 enriched loci (Flueck et al., 2009); H3K9me3 enrichment at transcriptionally silent loci based on ChiP-chip experiments ("S-A" Salcedo-Amaya et al., 2009 and "L-R" Lopez-Rubio et al., 2009); Subtelomeric chromosomal location defined as a loci within 150kB of chromosome ends (plasmoDB.org).

G. Functional enrichment of export modules include "Protein Binding" and "Protein Folding"

To evaluate the biological roles and pathway memberships of genes within the functional modules that are not exported, we retrieved GO functions from the open access database, PlasmoDB (Aurrecoechea et al., 2009). We find that a total of 51 genes are associated with export modules and that these genes are redundant in both Cons and Variable export modules. The most enriched GO function for genes co-expressed with export modules fall in the categories of "Protein binding" and "Protein folding" (12.8%), whereas genes that are strongly anti-correlated in "Transcription" (11.5%) and "Translation" (11.2%). Surprisingly, 16% of associated genes do not have any GO function (Figure 3-7). This suggests that export modules are co-expressed with other genes that play a role in protein complex formation and in assuring proper protein conformation. In fact, some of the genes in the "Protein folding" class contain genes with DnaJ domain. The "Protein binding" class are genes with WD-repeat and the armadillo "ARM" domains. These domains have previously been shown to mediate protein complexing by acting as scaffolds for protein-protein interaction (Li and Roberts, 2001; Tewari et al., 2010).

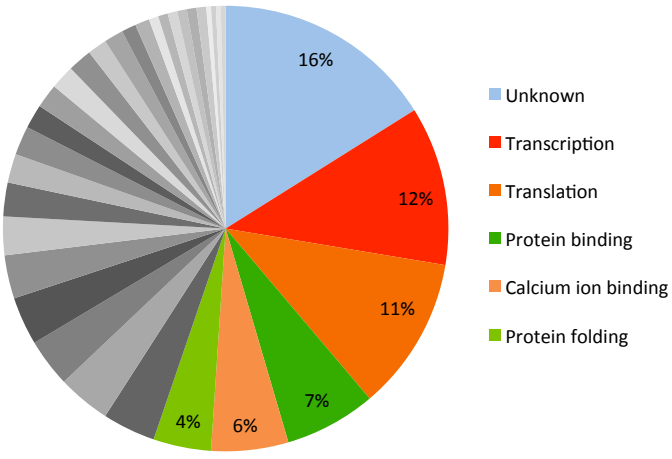


Figure 3- 7: Representation of GO functions of export module gene that do not belong to the exportome. Non-exportome genes mostly lack GO function assignment. Most GO functions fall in Transcription & Translation (anticorrelation, red slices), Protein Binding and Folding (positive correlation, green slices). Shown are the 6 top GO functions.

H. Variable export modules are associated with pathology in CM patients

We then wanted to elucidate the role of export modules in cerebral malaria using a children cohort in Malawi. A GSEA analysis shows that only variable modules are associated with clinical phenotypes (Nom p-value< 0.055; FDR q-value< 0.25) (Table 3-2). V4 was associated with "sleeping under a bednet", a parameter that is directly linked to the likelihood of getting bitten by a vector, as well as fever. "Respiratory distress", "spleen size" and "liver size" are also enriched with different variable modules. These clinical phenotypes are reflective of organ pathologies often associated with local accumulation of parasites via cytoadherence of iRBCs in the microvasculature of these tissues (Mackintosh et al., 2004). Of note, "respiratory distress" was enriched with five different modules, as compared to other phenotypes associated with only one or two, mainly composed of members of the *fikk, phist*, and *hyp* multi-gene families. Respiratory distress needs to be interpreted with great caution as it is often a secondary effect during malaria infections and can be caused by other pathogens (Taylor et al., 2012).

The enrichment of only variable export modules with different organ pathologies in a severely diseased cohort with CM, suggests that these multi-family gene modules interact dynamically with the host environment.

Phenotype	Cluster	Nom p-value*	FDR q-value ϕ	Top 4 enriched genes			
BEDNETSLP	Variable 4	0.0096	0.1012	PFA0210C	PF14_0014 (HYP10)		
FEVER	Variable 4	0.0208	0.1299	PFA0210C	PFD0095C	*PF14_0014*	
LIVER	Variable 8	0.0020	0.1677	*PF08_0138*	*PF10_0398*	*PF10_0003*	*PF14_0771*
RESPDISTRES	Variable 10	< 10^{-5}	0.0012	*PFD0070C*	PFI1760W (REX4)	PFB0926C (HYP2)	
	Variable 4	0.0212	0.1673	PF11_0099 (HSP)	PFA0210C	PF10_0380 (FIKK)	MAL8P1.3 (HYP9)
	Variable 2	0.0247	0.1690	PF14_0746 (PHISTb)	PFE0060W (PIESP2)	PFB0080C (PHISTb)	PFD0095C (PHISTb)
	Variable 1	0.0533	0.2273	PFC0090W	PF10_0021 (PHISTc)	PF10_0025 (HYP16)	PF11_0512 (RESA-2)
SPLEEN	Variable 9	0.0263	0.2498	*PFB0955W*	PF14_0755 (HYP15)	MAL7P1.144 (FIKK)	PF14_0744 (PHISTa)

Table 3- 2: Variable export modules are associated with disease in a Malawian cohort of CM. GSEA analysis was performed on expression data from a cohort of children with CM in Malawi (n=59). Listed are modules significantly associated with any phenotype and the top 4 genes of the export modules, which are significantly associated with the phenotype (*Nom p-values <0.055; ϕFDR q-values <0.25). Clusters that are differentially expressed in this cohort are associated with disease phenotype whereas no constitutively expressed clusters meet the threshold for significance

III. Discussion

Genes at the host-parasite interface are likely mediating interactions with the host to allow parasite survival, which encompasses immune evasion and parasite sequestration deep in the tissues. We hypothesize that these genes are likely modulated by components from the host, which are themselves dynamic (i.e. immune response, receptor type and availability). In this study we have identified two kinds of genes at the host-parasite interface. One group is highly expressed across patients with malaria and the other variably expressed. The first group is likely a core group of genes required for cytoadherence and immune evasion. Some of these genes are well known and include *KAHRP*, *MAL7P1.172* and *RESA*. The second group, because of its mosaic transcriptional expression across patients may be significant in antigenic variation.

Export gene modules are associated with more than one pathogenic process. We identified a variable export module involved in invasion, in particular the glideosome complex. Other variable modules include large *rifin* and *stevor* clusters. In fact, there is evidence that *rifins* and *stevors* are variantly expressed (Fernandez et al., 1999; Kyes et al., 1999; Lavazec et al., 2007). RIFINs and STEVORs are both surface-expressed (Blythe et al., 2008; Petter et al., 2008). While RIFINs are targeted by the immune system and lead to antigenic variation, the incorporation of STEVORs in the RBC membrane decreases membrane deformability (shear modulus) and thus the capacity of iRBC to migrate through capillaries (Fernandez et al., 1999; Sanyal et al., 2012). STEVORs and RIFINs have also been localized at the merozoite apical tip making it a possibility that variable export modules containing members of these gene families are also involved in invasion (Blythe et al., 2008; Khattab et al., 2008). Other variable export modules that we identify are heterogeneous. They are composed of *phist, fikk, hyp*, and other single gene families. There is growing evidence supporting the role of these gene families in pathogenic processes. For instance a subset of *phistb* contain a DnaJ domain, a cardinal characteristic of HSP40 chaperones, and are thought to mediate translocation of parasite proteins across the iRBC cytoplasm, however their function with HSP70 remains to be elucidated (Sargeant et al., 2006). Additionally, a PHISTC domain binds the intracellular domain of PfEMP1 (Mayer et al., 2012). The deletion of *MAL7P1.172*, a PHISTC, abrogates PfEMP1 presentation at

the surface (Maier et al., 2008). Further, FIKK members are differentially expressed *in vitro* and modulate cytoskeleton proteins by phosphorylation events in trophozoite and schizont stages (Nunes et al., 2007). They do not change the adhesive capabilities of parasites but they cause considerable decrease in iRBC membrane rigidity (Nunes et al., 2010).

Although we did not find a correlation between variable export modules by qRT-PCR in children with CM and parasite lines that were panned on HBMEC and HPMEC or rosetting parasites, we do find common features of epigenetic regulation between variable modules and *var* genes. Recent studies have attempted to understand the genetic mechanisms regulating *var* genes expression and other multi-gene families located at the subtelomeres. In one particular study, the authors show that the induction of *var* promoters (UpsA, B, and C) and a head-to-head rifA1 promoter are able to silence endogenous *var* transcription, yet other promoters of gene families known to undergo variant expression (*phistb, stevor, pfmc-2tm*) do not share this competence. In fact, non-*var* promoters do not modulate the expression of genes within their own respective families (Witmer et al., 2012). This suggests that control of singular *var* expression is an exclusive privilege of *var* promoters and that there is no regulatory cross talk between *var* and non-*var* gene families. Perhaps this could also apply to our results to explain the lack of correlation between the different modules and *vars*. Furthermore, the observations made by this group were done in a 3D7 *P. falciparum* background, unselected for binding. Whether cytoadherence cues, or other environmental cues, interfere with the genetic regulation of these families is unknown. The fact remains that *var, phist, rif, stevor, hyp* and many other genes are variantly expressed *in vivo* and share common structural features, this points towards epigenetic mechanisms of transcriptional regulation. This agrees with a study where the *P. falciparum* silent information regulator 2 (PfSir2) paralogs (Sir2A and SirB, histone deacetylases involved in transcriptional silencing) coordinate the expression of virulence genes and antigenic variation (Tonkin et al., 2009). Additionally, the expression of invasion genes and genes involved in nutrient transport is also under epigenetic control (Cortes et al., 2007; Nguitragool et al., 2011). This suggests that the parasite adapts to changes in the host environment by modulating the

expression of genes at the host-parasite interface by taking advantage of the plasticity and reversibility of epigenetic mechanisms of transcriptional regulation.

In other systems, epigenetic responses have been associated with hypoxia. Reduction in oxygen levels, or hypoxia, can occur under normal or pathological conditions. Under hypoxic conditions, cells adapt through angiogenesis, erythropoiesis, glucose transport and are able to switch from an aerobic to a glycolytic metabolic state (Singh et al., 2012). These adaptive responses have been linked to transcriptional changes through the activity of a family of transcription factors Hypoxia Inducible Factor (HIF) (Wang and Semenza, 1993). HIFs form heterodimers. While one subunit (β) is constitutively expressed in the cell, the other (α) is oxygen-sensitive and when hydroxylated is targeted to the proteasome for degradation (Sutter et al., 2000). Under hypoxic conditions, HIF dimers form and bind specific DNA sequences termed "hypoxia response elements" (HREs), leading to the upregulation of genes bearing HREs. HIFs also interact with histone modifying enzymes including histone deacetylases (HDACs) and are thought to mediate HDAC recruitment in chromosomal regions where transcriptional repression will occur (Kato et al., 2004). Overall, this mechanism is a quick adaptation to hypoxia-mediated tissue damage (Singh et al., 2012). Interestingly, hypoxia is common in falciparum malaria in particular in cerebral malaria. The effect of hypoxia on *P. falciparum* gene expression is unknown. However, whole-genome transcriptional studies have reported the existence of parasites from clinical samples under various metabolic states (Daily et al., 2007; Milner et al., 2012). Three metabolic clusters were found: one related to glycolysis, a second to respiration and a third cluster was associated with environmental stress (heat shock, oxidative and osmotic stress) (Daily et al., 2007). The first cluster was also linked to starvation (Daily et al., 2007). It is therefore possible that hypoxia, although likely contributed by *P. falciparum* in the host through vasoconstriction, like nutrient deprivation is another stress which is tolerated by *Plasmodium* through gene expression modulation. Across multiple samples, this reveals itself as the variable gene expression we detected in variable export modules. A large number of variable export modules associated with respiratory distress (Table 3-2) and it is still not clear how these genes would be influenced by limited oxygen level. They are predicted to localize in the host cell,

potentially even at the surface, and could be involved in sensing and monitoring changes in oxygen levels in the host environment.

All in all, we have defined co-expression modules composed of known and yet described genes encoding proteins at the host-parasite interface. These modules hold much information as to the different pathways that are involved in parasite viability during active disease. Some of these pathways may even be completely novel as we recall that 60% of the genome remains unannotated. As we learn more about *P. falciparum* biology and expand our toolkit to study this parasite as close to its natural environment in the human host, the functional players and their relationship to disease will become more clear.

IV. Subjects and Methods

Ethics

This study was approved by the institutional review boards of the University of Malawi College of Medicine, Albert Einstein College of medicine and the Brigham and Women's Hospital. The cohort of this study comprises of children admitted during 2010 in the pediatric unit of the Malaria Research Ward (MRW) located in the Queen Elizabeth Central Hospital in Blantyre, Malawi. Pediatric patients were enrolled in an observational study of malaria pathogenesis upon receipt of written informed consent during admission at MRW. Diagnostic criteria, clinical management, laboratory investigations and treatment protocols have been previously described (Taylor et al., 2004). All patients in this study meet the clinical definition of cerebral malaria.

Sample Handling

After written informed consent was received, 3mLs of peripheral blood from children meeting the clinical definition of cerebral malaria were collected at time of admission. These samples were rapidly preserved in Tri Reagent, shaken and frozen at -80C. Samples were shipped from MRW to HSPH in liquid nitrogen and thawed at room temperature prior to RNA extraction. Thirty samples were processed for RNA extraction.

Patient Microarray Expression Data

Multiple microarray data sets (a total of 129) were utilized to study co-expression networks of genes at the host-parasite interphase in the context of active malaria disease. These microarrays were performed using peripheral blood material of patients with malaria at time of admission. Patients vary in malaria disease severity, from asymptomatic malaria to severe cerebral malaria, age and geographic location. These include two senegalese cohorts of mild to severe malaria, a Sénégal 2005 cohort published in Daily et al., 2005 and a Sénégal 2008 cohort (unpublished but accessed through the Broad Institute); and a Malawi 2009 cohort of severe malaria (severe anemia and cerebral malaria) published in Milner et al., 2012.

Identification of gene co-expression modules during disease

To identify gene networks potentially modulating disease course in patients with malaria, we first determined which genes predicted at the host-parasite interphase, or the exportome, are differentially expressed or not across patients. We used the exportome list from Boddey et al., 2013. For this, a rank sum score was generated for all genes within the exportome across every microarray without normalization. The rows were sorted from high to low rank sum score. The top 8% genes were defined as constitutively expressed. The bottom 10%, "Other", were defined as genes with no –to- low expression overall and were removed from the "ranked" exportome list.

We then constructed gene networks within the ranked exportome to identify genes that are co-expressed and extended these networks to the entire transcriptome to identify other co-regulated genes. Essentially, for each microarray we computed a co-expression matrix using the Spleipnir tool "Distancer" to generate z-score between all gene pairs. Z-scores are based on the Fisher transformation of Pearson Correlations. Each co-expression matrix was then separately normalized by centering at 0 and diving each value by its standard deviation. We then integrated all matrices by averaging to generate an integrated global network where gene co-expression greater than a z-score corresponding to alpha=0.01. Subnetworks containing genes from the ranked exportome (constitutive and variantly expressed genes) were retrieved and each group subsequently clustered into non-overlapping expression clusters (10 containing constitutively expressed genes and 10 for variantly expressed genes). The clusters were then visualized on Cytoscape, where each node corresponds to a gene, and each edge to the z-score corresponding to a gene pair from the integrated network.

Gene Set Enrichment Analysis

A GSEA was performed to assess the correlation between the expression of genes within each module and clinical phenotype. For this purpose *P. falciparum* Gene Ontology leaf annotations were obtained from PlasmoDB and propagated into the ontology using the Sleipnir functional genomics library. Gene sets containing <2 genes were removed, and overlapping gene sets were merged (combining the top 10% of gene set pairs ranked by the fraction of shared genes relative

to the total size of the smaller set). These gene sets were finally combined with the *P. falciparum* pathways from KEGG.

The microarrays used (a final total of 56) were performed using venous blood sampled from Malawian cohort of children enrolled in a 2009 study on cerebral malaria and matching clinical phenotype. The cohort is described in (Milner et al., 2012). Two sets of analysis were performed. First to test for the significance of each gene and second to test each cluster of genes previously identified against the phenotype. For each phenotype extreme outliers were removed and all phenotype gene pairs were tested using Spearman rank correlation for continuously valued features and the Mann–Whitney U test (for 2-valued factors) or the Kruskal-Wallis rank sum test (for multivalued factors) for discrete factors. Significance was assessed using an empirical bootstrap P value over 5000 randomizations (nonparametric tests and bootstrapping were chosen to avoid effects from remaining experimental outliers). P values were adjusted for multiple hypothesis testing, using Benjamini Hochberg false discovery rate (FDR) q values and used to examine associations between individual genes and the phenotypes.

In order to examine the effect and enrichment of gene pathways on phenotype, the normalized test statistic from each of these gene-phenotype tests was used as input in a GSEA analysis to look for enrichment of these functional modules. Further, the clusters obtained were used themselves as pathways to look for enrichment of these clusters with the phenotypes of interest. The enrichment of both the resulting pathway sets was then assessed using GSEA over 1000 bootstraps. A significant association was found when the normalized p-values <0.055 and FDR q-values <0.25.

***Plasmodium falciparum* cytoadherence-selected lines**

IT4 clone R29 and 3D7 clone PF13 rosetting parasites were generated by selection with mAB and rosetting phenotype confirmed (Janes et al., 2011). The cytoadherence-selected parasites were generated by binding ItG parasites first on ICAM-1 then on human brain microvascular endothelial cells (HBMEC), subsequently cloned, expanded in culture and finally re-measured for binding. These ItG/HBMEC selected clones are 1E2, 1A2 and 1E7. One ItG/HBMEC clone was

additionally selected on human pulmonary microvasculature endothelial cells, the HPMEC line. A detailed description of the generation of these lines can be found in the methods section of Avril et al., 2012. All six parasite lines were cultured in supplemented RPMI mediums as described (Trager and Jensen, 1976). Ring-stage parasites were obtained by two rounds of synchronization using 5% sorbitol solution. Parasites were pelleted and frozen in Tri Reagent prior to RNA extraction.

Validation by quantitative Real Time-PCR and association to *var*

Primer pairs targeting genes from Variant cluster 1, 2 and 10 were optimized on gDNA isolated from *P. falciparum* 3D7 parasites. *Var* degenerate primers used in this study are the Rottmann primer set previously published (Rottmann et al., 2006). RNA was extracted from patient peripheral blood samples, cytoadherence-selected and rosetting lines conserved in Tri Reagent. Essentially, RNA was separated by an initial chloroform extraction. The RNA layer was then processed using the RNeasy mini kit (Qiagen) and then subjected to a Dnase digest treatment. The quality of the RNA was determined on a 1% agarose RNA denaturing gel (formaldehyde) and by Nanodrop. cDNA synthesis was then performed using the SuperScript Reverse Transcriptase III (Invitrogen). qRT-PCR was run on Applied Biosystem qPCR machines using SYBR green. Reactions were run in triplicate for each gene and gene expression was normalized to the marker seryl-t-RNA synthetase.

Immunofluorescence Assay

Approximately 500µL of cells from a mixed stage culture were pelleted by centrifugation at 2000rpm for 2min, and then pellets were washed twice in 1X PBS. Parasite pellets were fixed in 4% paraformaldehyde/ 0.0075% gluteraldeyde for 30min at room temperature in suspension. Pellets were washed in 1X PBS and permeabilized using 0.1% Triton X-100 for 10min at room temperature. Parasites were then washed in 1X PBS and cells were blocked in 3% w/v bovine serum albumin (BSA) in PBS for 1hr. Primary antibodies were than added in their respective dilutions and cells were incubated for 4hrs: PF70 rabbit polyclonal antibodies (Ma et al., 1996),

gifted by Dr. Kironde, anti KAHRP mouse monoclonal antibodies (clone 2G12) a gift from Dr. Diane Taylor and anti-PfEMP3 monoclonal mouse antibody. Cells were washed for 5 min in 3% w/v BSA three times, then blocked with secondary antibodies (in BSA) for 2 hrs (Invitrogen's goat anti-rabbit alexa 594 (1:750) and goat anti-mouse alexa 488 (1:1000). After 3 washes in 1X PBS, cells were resuspended in a small volume of 1X PBS and mounted on microscope glass slides using VectaShield containing 4',6-diamidino-2-phenylindole (DAPI) nuclear dye. Parasites were visualized on a traditional fluorescence microscope. Images were edited in Adobe Photoshop.

Statistical tests for clinical parameters

Statistical analyses of gene expression by qRT-PCR association to clinical parameters was done in GraphPad Prism.

ACKNOWLEDGEMENTS

We would like to thank Dr. Joseph Smith for providing *P. falciparum* lines panned on human brain microvascular endothelial cells, human pulmonary microvascular endothelial cells, and rosetting parasite lines, Dr. Diane Taylor for antibodies against KAHRP and Dr. Fred Kironde for the PF70 antibodies.

V. References

Aikawa, M., Kamanura, K., Shiraishi, S., Matsumoto, Y., Arwati, H., Torii, M., Ito, Y., Takeuchi, T., and Tandler, B. (1996). Membrane knobs of unfixed Plasmodium falciparum infected erythrocytes: new findings as revealed by atomic force microscopy and surface potential spectroscopy. Experimental parasitology *84*, 339-343.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S.*, et al.* (2009). PlasmoDB: a functional genomic database for malaria parasites. Nucleic acids research *37*, D539-543.

Avril, M., Tripathi, A.K., Brazier, A.J., Andisi, C., Janes, J.H., Soma, V.L., Sullivan, D.J., Jr., Bull, P.C., Stins, M.F., and Smith, J.D. (2012). A restricted subset of var genes mediates adherence of Plasmodium falciparum-infected erythrocytes to brain endothelial cells. Proceedings of the National Academy of Sciences of the United States of America *109*, E1782-1790.

Barnwell, J.W., Howard, R.J., Coon, H.G., and Miller, L.H. (1983). Splenic requirement for antigenic variation and expression of the variant antigen on the erythrocyte membrane in cloned Plasmodium knowlesi malaria. Infection and immunity *40*, 985-994.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.W., Green, J.L., Holder, A.A., and Cowman, A.F. (2006). A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. The Journal of biological chemistry *281*, 5197-5208.

Beeson, J.G., Amin, N., Kanjala, M., and Rogerson, S.J. (2002). Selective accumulation of mature asexual stages of Plasmodium falciparum-infected erythrocytes in the placenta. Infection and immunity *70*, 5412-5415.

Blythe, J.E., Yam, X.Y., Kuss, C., Bozdech, Z., Holder, A.A., Marsh, K., Langhorne, J., and Preiser, P.R. (2008). Plasmodium falciparum STEVOR proteins are highly expressed in patient isolates and located in the surface membranes of infected red blood cells and the apical tips of merozoites. Infection and immunity *76*, 3329-3336.

Boddey, J.A., Carvalho, T.G., Hodder, A.N., Sargeant, T.J., Sleebs, B.E., Marapana, D., Lopaticki, S., Nebl, T., and Cowman, A.F. (2013). Role of Plasmepsin V in Export of Diverse Protein Families from the Plasmodium falciparum Exportome. Traffic (Copenhagen, Denmark) *14*, 532-550.

Claessens, A., Adams, Y., Ghumra, A., Lindergard, G., Buchan, C.C., Andisi, C., Bull, P.C., Mok, S., Gupta, A.P., Wang, C.W.*, et al.* (2012). A subset of group A-like var genes encodes the malaria parasite ligands for binding to human brain endothelial cells. Proceedings of the National Academy of Sciences of the United States of America *109*, E1772-1781.

Cortes, A., Carret, C., Kaneko, O., Yim Lim, B.Y., Ivens, A., and Holder, A.A. (2007). Epigenetic silencing of Plasmodium falciparum genes linked to erythrocyte invasion. PLoS pathogens *3*, e107.

Crabb, B.S., Cooke, B.M., Reeder, J.C., Waller, R.F., Caruana, S.R., Davern, K.M., Wickham, M.E., Brown, G.V., Coppel, R.L., and Cowman, A.F. (1997). Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress. Cell *89*, 287-296.

Da Silva, E., Foley, M., Dluzewski, A.R., Murray, L.J., Anders, R.F., and Tilley, L. (1994). The Plasmodium falciparum protein RESA interacts with the erythrocyte cytoskeleton and modifies erythrocyte thermal stability. Molecular and biochemical parasitology *66*, 59-69.

Daily, J.P., Le Roch, K.G., Sarr, O., Ndiaye, D., Lukens, A., Zhou, Y., Ndir, O., Mboup, S., Sultan, A., Winzeler, E.A.*, et al.* (2005). In vivo transcriptome of Plasmodium falciparum reveals overexpression of transcripts that encode surface proteins. The Journal of infectious diseases *191*, 1196-1203

Daily, J.P., Scanfeld, D., Pochet, N., Le Roch, K., Plouffe, D., Kamal, M., Sarr, O., Mboup, S., Ndir, O., Wypij, D.*, et al.* (2007). Distinct physiological states of Plasmodium falciparum in malaria-infected patients. Nature *450*, 1091-1095.

Date, S.V., and Stoeckert, C.J., Jr. (2006). Computational modeling of the Plasmodium falciparum interactome reveals protein function on a genome-wide scale. Genome research *16*, 542-549.

Day, K.P., Karamalis, F., Thompson, J., Barnes, D.A., Peterson, C., Brown, H., Brown, G.V., and Kemp, D.J. (1993). Genes necessary for expression of a virulence determinant and for transmission of Plasmodium falciparum are located on a 0.3-megabase region of chromosome 9. Proceedings of the National Academy of Sciences of the United States of America *90*, 8292-8296.

Duffy, M.F., Maier, A.G., Byrne, T.J., Marty, A.J., Elliott, S.R., O'Neill, M.T., Payne, P.D., Rogerson, S.J., Cowman, A.F., Crabb, B.S.*, et al.* (2006). VAR2CSA is the principal ligand for chondroitin sulfate A in two allogeneic isolates of Plasmodium falciparum. Molecular and biochemical parasitology *148*, 117-124.

Fernandez, V., Hommel, M., Chen, Q., Hagblom, P., and Wahlgren, M. (1999). Small, clonally variant antigens expressed on the surface of the Plasmodium falciparum-infected erythrocyte are encoded by the rif gene family and are the target of human immune responses. The Journal of experimental medicine *190*, 1393-1404.

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 pathogens *5*, e1000569.

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

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.

Glenister, F.K., Coppel, R.L., Cowman, A.F., Mohandas, N., and Cooke, B.M. (2002). Contribution of parasite proteins to altered mechanical properties of malaria-infected red blood cells. Blood *99*, 1060-1063.

Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C., and Haldar, K. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science (New York, N.Y.) *306*, 1934-1937.

Hosseini, S.M., and Feng, J.J. (2012). How malaria parasites reduce the deformability of infected red blood cells. Biophysical journal *103*, 1-10.

Hu, G., Cabrera, A., Kono, M., Mok, S., Chaal, B.K., Haase, S., Engelberg, K., Cheemadan, S., Spielmann, T., Preiser, P.R.*, et al.* (2010). Transcriptional profiling of growth perturbations of the human malaria parasite Plasmodium falciparum. Nature biotechnology *28*, 91-98.

Janes, J.H., Wang, C.P., Levin-Edens, E., Vigan-Womas, I., Guillotte, M., Melcher, M., Mercereau-Puijalon, O., and Smith, J.D. (2011). Investigating the host binding signature on the Plasmodium falciparum PfEMP1 protein family. PLoS pathogens *7*, e1002032.

Jensen, A.T., Magistrado, P., Sharp, S., Joergensen, L., Lavstsen, T., Chiucchiuini, A., Salanti, A., Vestergaard, L.S., Lusingu, J.P., Hermsen, R.*, et al.* (2004). Plasmodium falciparum associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A var genes. The Journal of experimental medicine *199*, 1179-1190.

Kato, H., Tamamizu-Kato, S., and Shibasaki, F. (2004). Histone deacetylase 7 associates with hypoxia-inducible factor 1alpha and increases transcriptional activity. The Journal of biological chemistry *279*, 41966-41974.

Khattab, A., Bonow, I., Schreiber, N., Petter, M., Schmetz, C., and Klinkert, M.Q. (2008). Plasmodium falciparum variant STEVOR antigens are expressed in merozoites and possibly associated with erythrocyte invasion. Malaria journal *7*, 137.

Kilejian, A., Rashid, M.A., Aikawa, M., Aji, T., and Yang, Y.F. (1991). Selective association of a fragment of the knob protein with spectrin, actin and the red cell membrane. Molecular and biochemical parasitology *44*, 175-181.

Kyes, S.A., Rowe, J.A., Kriek, N., and Newbold, C.I. (1999). Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America *96*, 9333-9338.

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

Lavstsen, T., Salanti, A., Jensen, A.T., Arnot, D.E., and Theander, T.G. (2003). Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions. Malaria journal *2*, 27.

Lavstsen, T., Turner, L., Saguti, F., Magistrado, P., Rask, T.S., Jespersen, J.S., Wang, C.W., Berger, S.S., Baraka, V., Marquard, A.M.*, et al.* (2012). Plasmodium falciparum erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children. Proceedings of the National Academy of Sciences of the United States of America *109*, E1791-1800.

Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson, K.C., Holder, A.A., Carucci, D.J.*, et al.* (2004). Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. Genome research *14*, 2308-2318.

Li, D., and Roberts, R. (2001). WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. Cellular and molecular life sciences : CMLS *58*, 2085-2097.

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.

Ma, H.W., Ray, P., Dhanda, V., Das, P.K., Paliwal, S., Sahoo, N., Patra, K.P., Das, L.K., Singh, B., and Kironde, F.A. (1996). A novel 70-kDa Triton X-114-soluble antigen of Plasmodium falciparum that contains interspecies-conserved epitopes. Experimental parasitology *83*, 322-334.

Mackintosh, C.L., Beeson, J.G., and Marsh, K. (2004). Clinical features and pathogenesis of severe malaria. Trends in parasitology *20*, 597-603.

Magowan, C., Nunomura, W., Waller, K.L., Yeung, J., Liang, J., Van Dort, H., Low, P.S., Coppel, R.L., and Mohandas, N. (2000). Plasmodium falciparum histidine-rich protein 1 associates with the band 3 binding domain of ankyrin in the infected red cell membrane. Biochimica et biophysica acta *1502*, 461-470.

Maier, A.G., Cooke, B.M., Cowman, A.F., and Tilley, L. (2009). Malaria parasite proteins that remodel the host erythrocyte. Nature reviews. Microbiology *7*, 341-354.

Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szestak, T., Chesson, J., Wu, Y., Hughes, K., Coppel, R.L.*, et al.* (2008). Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell *134*, 48-61.

Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004). Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science (New York, N.Y.) *306*, 1930-1933.

Mayer, C., Slater, L., Erat, M.C., Konrat, R., and Vakonakis, I. (2012). Structural analysis of the Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. The Journal of biological chemistry *287*, 7182-7189.

Miller, L.H., Baruch, D.I., Marsh, K., and Doumbo, O.K. (2002). The pathogenic basis of malaria. Nature *415*, 673-679.

Milner, D.A., Jr., Pochet, N., Krupka, M., Williams, C., Seydel, K., Taylor, T.E., Van de Peer, Y., Regev, A., Wirth, D., Daily, J.P.*, et al.* (2012). Transcriptional Profiling of Plasmodium falciparum Parasites from Patients with Severe Malaria Identifies Distinct Low vs. High Parasitemic Clusters. PloS one *7*, e40739.

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.

Nagao, E., Kaneko, O., and Dvorak, J.A. (2000). Plasmodium falciparum-infected erythrocytes: qualitative and quantitative analyses of parasite-induced knobs by atomic force microscopy. Journal of structural biology *130*, 34-44.

Nguitragool, W., Bokhari, A.A., Pillai, A.D., Rayavara, K., Sharma, P., Turpin, B., Aravind, L., and Desai, S.A. (2011). Malaria parasite clag3 genes determine channel-mediated nutrient uptake by infected red blood cells. Cell *145*, 665-677.

Nunes, M.C., Goldring, J.P., Doerig, C., and Scherf, A. (2007). A novel protein kinase family in Plasmodium falciparum is differentially transcribed and secreted to various cellular compartments of the host cell. Molecular microbiology *63*, 391-403.

Nunes, M.C., Okada, M., Scheidig-Benatar, C., Cooke, B.M., and Scherf, A. (2010). Plasmodium falciparum FIKK kinase members target distinct components of the erythrocyte membrane. PloS one *5*, e11747.

Oakley, M.S., Kumar, S., Anantharaman, V., Zheng, H., Mahajan, B., Haynes, J.D., Moch, J.K., Fairhurst, R., McCutchan, T.F., and Aravind, L. (2007). Molecular factors and biochemical pathways induced by febrile temperature in intraerythrocytic Plasmodium falciparum parasites. Infection and immunity *75*, 2012-2025.

Pei, X., An, X., Guo, X., Tarnawski, M., Coppel, R., and Mohandas, N. (2005). Structural and functional studies of interaction between Plasmodium falciparum knob-associated histidine-rich protein (KAHRP) and erythrocyte spectrin. The Journal of biological chemistry *280*, 31166-31171.

Pei, X., Guo, X., Coppel, R., Bhattacharjee, S., Haldar, K., Gratzer, W., Mohandas, N., and An, X. (2007a). The ring-infected erythrocyte surface antigen (RESA) of Plasmodium falciparum stabilizes spectrin tetramers and suppresses further invasion. Blood *110*, 1036-1042.

Pei, X., Guo, X., Coppel, R., Mohandas, N., and An, X. (2007b). Plasmodium falciparum erythrocyte membrane protein 3 (PfEMP3) destabilizes erythrocyte membrane skeleton. The Journal of biological chemistry *282*, 26754-26758.

Peters, J.M., Fowler, E.V., Krause, D.R., Cheng, Q., and Gatton, M.L. (2007). Differential changes in Plasmodium falciparum var transcription during adaptation to culture. The Journal of infectious diseases *195*, 748-755.

Petter, M., Bonow, I., and Klinkert, M.Q. (2008). Diverse expression patterns of subgroups of the rif multigene family during Plasmodium falciparum gametocytogenesis. PloS one *3*, e3779.

Rottmann, M., Lavstsen, T., Mugasa, J.P., Kaestli, M., Jensen, A.T., Muller, D., Theander, T., and Beck, H.P. (2006). Differential expression of var gene groups is associated with morbidity caused by Plasmodium falciparum infection in Tanzanian children. Infection and immunity *74*, 3904-3911.

Rovira-Graells, N., Gupta, A.P., Planet, E., Crowley, V.M., Mok, S., Ribas de Pouplana, L., Preiser, P.R., Bozdech, Z., and Cortes, A. (2012). Transcriptional variation in the malaria parasite Plasmodium falciparum. Genome research *22*, 925-938.

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. Proceedings of the National Academy of Sciences of the United States of America *106*, 9655-9660.

Sanyal, S., Egee, S., Bouyer, G., Perrot, S., Safeukui, I., Bischoff, E., Buffet, P., Deitsch, K.W., Mercereau-Puijalon, O., David, P.H.*, et al.* (2012). Plasmodium falciparum STEVOR proteins impact erythrocyte mechanical properties. Blood *119*, e1-8.

Sargeant, T.J., Marti, M., Caler, E., Carlton, J.M., Simpson, K., Speed, T.P., and Cowman, A.F. (2006). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome biology *7*, R12.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J., and Lanzer, M. (1998). 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 *17*, 5418-5426.

Silva, M.D., Cooke, B.M., Guillotte, M., Buckingham, D.W., Sauzet, J.P., Le Scanf, C., Contamin, H., David, P., Mercereau-Puijalon, O., and Bonnefoy, S. (2005). A role for the Plasmodium falciparum RESA protein in resistance against heat shock demonstrated using gene disruption. Molecular microbiology *56*, 990-1003.

Singh, N., Sharma, G., and Mishra, V. (2012). Hypoxia inducible factor-1: its potential role in cerebral ischemia. Cellular and molecular neurobiology *32*, 491-507.

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., Pinches, R., Newbold, C.I., and Miller, L.H. (1995). Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell *82*, 101-110.

Su, X.Z., Heatwole, V.M., Wertheimer, S.P., Guinet, F., Herrfeldt, J.A., Peterson, D.S., Ravetch, J.A., and Wellems, T.E. (1995). The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. Cell *82*, 89-100.

Sutter, C.H., Laughner, E., and Semenza, G.L. (2000). Hypoxia-inducible factor 1alpha protein expression is controlled by oxygen-regulated ubiquitination that is disrupted by deletions and missense mutations. Proceedings of the National Academy of Sciences of the United States of America *97*, 4748-4753.

Taylor, T.E., Fu, W.J., Carr, R.A., Whitten, R.O., Mueller, J.S., Fosiko, N.G., Lewallen, S., Liomba, N.G., and Molyneux, M.E. (2004). Differentiating the pathologies of cerebral malaria by postmortem parasite counts. Nature medicine *10*, 143-145.

Taylor, W.R., Hanson, J., Turner, G.D., White, N.J., and Dondorp, A.M. (2012). Respiratory manifestations of malaria. Chest *142*, 492-505.

Tewari, R., Bailes, E., Bunting, K.A., and Coates, J.C. (2010). Armadillo-repeat protein functions: questions for little creatures. Trends in cell biology *20*, 470-481.

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.

Trager, W., and Jensen, J.B. (1976). Human malaria parasites in continuous culture. Science (New York, N.Y.) *193*, 673-675.

Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? Molecular cell *28*, 1-13.

Turner, G.D., Ly, V.C., Nguyen, T.H., Tran, T.H., Nguyen, H.P., Bethell, D., Wyllie, S., Louwrier, K., Fox, S.B., Gatter, K.C.*, et al.* (1998). Systemic endothelial activation occurs in both mild and severe malaria. Correlating dermal microvascular endothelial cell phenotype and soluble cell adhesion molecules with disease severity. The American journal of pathology *152*, 1477-1487.

Waller, K.L., Cooke, B.M., Nunomura, W., Mohandas, N., and Coppel, R.L. (1999). Mapping the binding domains involved in the interaction between the Plasmodium falciparum knob-associated histidine-rich protein (KAHRP) and the cytoadherence ligand P. falciparum erythrocyte membrane protein 1 (PfEMP1). The Journal of biological chemistry *274*, 23808-23813.

Waller, K.L., Nunomura, W., Cooke, B.M., Mohandas, N., and Coppel, R.L. (2002). Mapping the domains of the cytoadherence ligand Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) that bind to the knob-associated histidine-rich protein (KAHRP). Molecular and biochemical parasitology *119*, 125-129.

Wang, G.L., and Semenza, G.L. (1993). General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia. Proceedings of the National Academy of Sciences of the United States of America *90*, 4304-4308.

Waterkeyn, J.G., Wickham, M.E., Davern, K.M., Cooke, B.M., Coppel, R.L., Reeder, J.C., Culvenor, J.G., Waller, R.F., and Cowman, A.F. (2000). Targeted mutagenesis of Plasmodium falciparum erythrocyte membrane protein 3 (PfEMP3) disrupts cytoadherence of malaria-infected red blood cells. The EMBO journal *19*, 2813-2823.

WHO (2012). World malaria report. (Geneva, Switzerland, WHO), p. 195.

Witmer, K., Schmid, C.D., Brancucci, N.M., Luah, Y.H., Preiser, P.R., Bozdech, Z., and Voss, T.S. (2012). Analysis of subtelomeric virulence gene families in Plasmodium falciparum by comparative transcriptional profiling. Molecular microbiology *84*, 243-259.

CHAPTER IV

Characterization of a novel gene family in *Plasmodium falciparum*, PHISTa

CONTRIBUTIONS

This chapter was written in its entirety by Karell Pellé. All experiments and analyses were performed by Karell Pellé.

I. Introduction

Antigenic variation is defined as the controlled clonal expression switch of antigenic epitopes that leads to phenotypic changes that allows the cell to subvert immune recognition. It has been described in pathogenic bacteria and protozoa. These antigens are primarily surface-expressed and typically members of large polymorphic multi-gene families (Kyes et al., 2001). By switching from one antigenic determinant to another, *Plasmodium falciparum* is able to escape the immune system and establish chronic infection (Biggs et al., 1991; Roberts et al., 1992; Smith et al., 1995). *Plasmodium falciparum* antigenic variation occurs in two pathogenic processes where parasite proteins are exposed to the host, invasion and cytoadherence. In terms of invasion, the erythrocyte binding antigens (EBAs) have been shown to switch expression from one member to another, whereas cytoadherence has been linked to *var* gene expression switching (Cortes et al., 2007; Scherf et al., 1998). Other variable surface antigens (VSA) have been identified (*rif, stevors* and *pfmc-2tms*) but their role in antigenic variation remains to be confirmed.

Cytoadherence is an interactive mechanism between parasitized RBCs and host cells. It is mediated by the binding of a family of adhesins expressed on the surface of iRBCs, *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) to host cell ligands in the microvasculature (Leech et al., 1984). PfEMP1 proteins are encoded by *var* genes (Su et al., 1995). Cytoadherence leads to the sequestration of parasitized RBC aggregates in the microvasculature leading to severe infections such as cerebral malaria (CM) and pregnancy-associated malaria (PAM) (MacPherson et al., 1985; Pongponratn et al., 1991). The presentation of PfEMP1 on the iRBC surface is still not fully understood, but it relies on accessory proteins and iRBC remodeling as shown in a genetic screen for parasite genes involved in cytoadherence and host cell rigidity (Maier et al., 2008). This screen identified genes that are directly involved in the translocation of PfEMP1 from the parasite to the iRBC plasma membrane (PM). Some of these genes belong to a large gene family for which to date no function have been assigned, the *phist.*

Plasmodium **helical interspersed subtelomeric** (PHIST) multi-gene family is composed of 3 conserved families (a, b and c) of predicted soluble proteins that have radiated in

primate *Plasmodium* parasites from a single rodent PHIST ancestor (Sargeant et al., 2006). PHIST proteins contain an N-terminal signal sequence followed by an export motif (PEXEL) and a well-conserved C-terminal region that contains the protein's core domain composed of 4 conserved alpha helices (~150 amino acids). The PHIST domain along with the Duffy-Binding-Like (DBL) domain found in PfEMP1 molecules are unique and exclusive to *Plasmodium.* Both domains contain well-conserved tryptophan and cysteine residues (Sargeant et al., 2006; Templeton and Deitsch, 2005). In PfEMP1, the DBL domain provides the adhesin differential affinity for host cell receptors and is highly polymorphic, whereas the PHIST function remains unknown (Buffet et al., 1999; Rowe et al., 1997; Smith et al., 2000). *Phist* genes are primarily localized at the subtelomeres of all chromosomes and show transcriptional peaks in late schizonts and ring stages (Le Roch et al., 2003; Sargeant et al., 2006). Some *phist* members also show expression variation *in vitro* and this seems to be due to epigenetic mechanisms (Flueck et al., 2009; Lopez-Rubio et al., 2009; Rovira-Graells et al., 2012; Salcedo-Amaya et al., 2009). The three *phist* families have expanded differentially in *Plasmodia* species and several recent investigations seem to point to a chaperone function for this family.

The PHISTC family is composed of 16 paralogs that have radiated in *P. falciparum*, *P. vivax and P. knowlesi.* They show the most sequence length variability in the amino-terminal region following the PHIST domain (Sargeant et al., 2006). In terms of function, PHISTC were shown to associate with the red blood cell membrane. They interact with the RBC cytoskeletal protein ankyrin and localize to knobs of placental malaria parasites (LaCount et al., 2005; Lasonder et al., 2002). PHISTC also interact with PfEMP1. One PHISTC domain binds the intracellular domain of PfEMP1 suggesting a role in the translocation of PfEMP1 to the surface of the iRBC (Mayer et al., 2012). Evidently, two *phistc* genetic deletions in CS2, a *P. falciparum* strain that confers adhesion to chondroitin sulfate A (CSA) via the stable expression of PfEMP1 encoded by the *var* gene *var2csa,* led to a loss of cytoadherence and re-localization of PfEMP1 from the surface to the MCs (Maier et al., 2008; Salanti et al., 2004).

The PHISTB family is composed of 23 paralogs of which 7 have an additional type III DnaJ domain at the carboxy-terminal end of the PHIST domain (Sargeant et al., 2006). The DnaJ domain is found in heat shock protein 40-HSP40, that regulate Adenosine triphosphate (ATP) hydrolysis function of HSP70. Whereas type I and II can bind substrate and regulate HSP70 function, the function of type III is still unclear (Walsh et al., 2004). Typically, the HSP40-HSP70 partnership functions in the control of protein homeostasis including folding and unfolding and protein trafficking (Walsh et al., 2004). In *P. falciparum*, cytosolic HSP40 functionally interacts with cytosolic *falciparum* HSP70 as well as human HSP70 (huHSP70) (Botha et al., 2011). Recently, Grover and colleagues identified a parasite HSP70 (PfHSP70) and showed that the protein localizes in punctate structures in the host cytosol including MCs (Grover et al., 2013). In another study, huHSP70 was identified in KAHRP fractions and was released from complexes by addition of ATP (Banumathy et al., 2002). These findings suggest that parasite exported HSP40s partner up with PfHSP70 to chaperone parasite proteins across the PVM into the host cell and potentially proteins involved in pathogenesis as PfEMP1 transit through MCs before being trafficked to the surface; and that partnership between parasite HSP40 and hHSP70 may play a role in knob formation (Maier et al., 2008). So far PHISTB-DnaJ have been localized at the surface of iRBC: PFB0090c localizes to knobs with PfEMP1, KHARP and PfEMP3 (Acharya et al., 2007) while the vaccine candidate RESA (PFA0110w) is integrated in the host cytoskeleton causing increased membrane rigidity and heat shock resistance during febrile episodes (Acharya et al., 2007; Da Silva et al., 1994; Mills et al., 2007; Silva et al., 2005). Additionally, the genetic deletion of some PHISTB-DnaJ was not achievable suggesting the essential nature of these proteins for parasite survival *in vitro.*

The PHISTA family is composed of 25 members. They are the shortest PHIST proteins as some consist of a recessed signal sequence, PEXEL motif and the PHISTA domain (Sargeant et al., 2006). Ph*ista* genes are unique in the sense that they are the only *phist* family restricted to the *P. falciparum* lineage. The selective expansion of this family in *P. falciparum* may have occurred to support the function of a virulence process exclusive to this *Plasmodium* species—the only human species with cytoadherence capacity—and responsible for *falciparum* malaria

disease severity (Maier et al., 2009). Similarly, *var* have no known orthologs and are involved in cytoadherence, rosetting and antigenic variation (Kyes et al., 2001; Smith et al., 1995; Rowe et al.,1997).

Whole transcriptional studies have been widely used to predict gene function in particular when a controllable perturbation is applied or when different parasite sources are used. In such studies, *phista* transcripts can be detected in parasites drawn from patient venous blood and their expression varies across patients. In particular, one *phista* gene, PF14_0752 is upregulated *in vivo* as well as in parasites selected for binding to CD36 and human brain endothelial cells (HBECs) (Claessens et al., 2012; Daily et al., 2005; Mok et al., 2007). In HBEC-selected parasites, only 15 non-VSAs are upregulated with the dominant *var* and *PF14_0752* shows the highest induction (Claessens et al., 2012). This was extremely surprising and strengthens the role of PF14_0752 in iRBCs adhesion in the brain microvasculature, and potentially CM. Overall, these findings suggest that *phista* expression may be modulated by factors that vary from patient to patient and/or under different parasite adhesive states. In fact, the differential expression of *phista* genes is governed by epigenetic processes and these allow inheritable yet reversible transcriptional expression. These processes include histone 3 lysine 9 trimethylation (H3K9me3) enrichment in the 5'UTR of transcriptionally silent loci, enrichment of heterochromatin protein I (HPI) bound H3K9me3 regions, and expression variation across isogenic clones (and not copy number variation) (Flueck et al., 2009; Lopez-Rubio et al., 2009; Rovira-Graells et al., 2012; Salcedo-Amaya et al., 2009). Interestingly, heterochromatin at transcriptionally silent loci is also found in subtelomeric genes of other multi-copy families including *var, rif* and *stevor,* all involved in host cell remodeling (ref. Chapter I).

Lastly, PHISTA may play a role in sexual differentiation. Two *phista* genes *PF14_0744* and *PF14_0748* are transcriptionally upregulated and their products enriched at the onset of gametocytogenesis (Eksi et al., 2005; Silvestrini et al., 2005; Silvestrini et al., 2010). These genes are localized either in the gametocyte PV or the host cell cytosol, indicating that they are specifically recruited for the remodeling of the iRBC for sexual differentiation. The promoter region of these two genes was also shown to drive the expression of a fluorescent marker in

young gametocytes, stages I and II (Buchholz et al., 2011). Of importance, young gametocytes (stage I-IV) are absent in the peripheral blood of diseased individuals and are thought to sequester in poorly defined sites in the body where they mature to stage V gametocytes. These mature stages are detectable in the peripheral blood. Although it does not seem like sequestration of sexual stages is mediated by cytoadherence, exported proteins such as PHISTA PF14_0744 and PF14_0748 are still of great importance as they may be involved in the initial formation of sexual stages (Silvestrini et al., 2012).

Overall, the PHIST family has not been fully functionally described: the PHISTC domain forms stable interactions with PfEMP1 and mediates the adhesin's trafficking to the iRBC PM; PHISTB-DnaJ provide heat shock resistance during febrile episodes and potentially partner up with HSP70 in proteostasis maintenance during trafficking across the PVM and to MCs; and PHISTA play an as yet understood role in host cell remodeling during sexual differentiation and cytoadherence. It is clear that this family contributes to many events where they may function as chaperones of proteins encoded by genes that under a similar transcriptional program.

In this chapter, we are particularly interested in PHISTAs. This family has radiated solely in *P*. *falciparum* and is involved in two major *P. falciparum* processes that have serious ramifications for malaria eradication and vaccine development: cytoadherence and transmission. PHISTA have been repeatedly associated with cytoadherence, a process, which, *in vivo* causes pathology and the severe diseases of CM and PAM, and sexual differentiation, the latter which ultimately leads to transmission. However, the amount of information known about PHISTA is incomplete: only a few genes are represented on microarrays of the published field studies and the subcellular localization of the PHISTA domain is unknown. In this study, we examine the composition and the dynamics of *phista* expression in long-term culture-adapted parasites as well as in parasites selected for binding and parasites from CM patients. We hypothesize that *phista* gene expression varies *in vitro* while specific members are upregulated in cytoadherent parasites and in parasites from CM patients.

II. Results

A. Absence of *phista* gene selection in culture-adapted parasites

During culture adaptation parasite surface antigens (on the merozoite surface and the iRBC) are no longer under immune selection, cytokines are absent, nutrients are continuously provided by replenishing culture-medium and temperature is regulated. Essentially, parasites are in a controlled and stable environment and at optimal growth conditions. However, in this minimal environment parasites undergo major and often irreversible changes. These include chromosomal deletions detected by pulse field gel electrophoresis resulting in the loss of knobs, cytoadherence and gametocytogenesis (Biggs et al., 1989; Day et al., 1993). Adaptation to culture conditions, in particular the removal of parasites from circulation where they are continuously cleared by the spleen also leads to a loss of antigenic variation. A similar effect was observed in *P. knowlesi* experiments using parasites passaged in splenectomized and spleen-intact Macaque monkeys (Barnwell et al., 1983). At the gene level, transcription is also massively modulated. This was well described for *var* genes where the dominance of one *var* in primary isolates was abrogated and the overall *var* transcript abundance decreased with time in culture (Peters et al., 2007). Therefore, we first examined the expression of *phista* genes in long-term (>2month) culture-adapted parasites to determine transcriptional profile in an environment lacking selection.

The expression of 13 *phista* genes was measured in 15 parasites strains, including reference strains of different genetic backgrounds such as 3D7, IT4 clone FCR3 and NF54 and *ex vivo* parasites from Senegalese patients that have been cultured well over 3 months. Here, we did not notice the dominance of one or a subset of *phista* genes (Figure 4-2A). This suggests that under *in vitro* conditions there is no selection for the expression of any one *phista* gene. To further confirm this result, we turned to a clonal tree composed of a parental *ex vivo* line TH137.07 and four derived clones. Here, the expression profile of each clone is divergent from that of the parental line. In fact, each clone has a different subset of *phista* genes with either

increased or decreased levels of expression, indicating that *phista* genes are clonally variant *in vitro* (Figure 4-1).

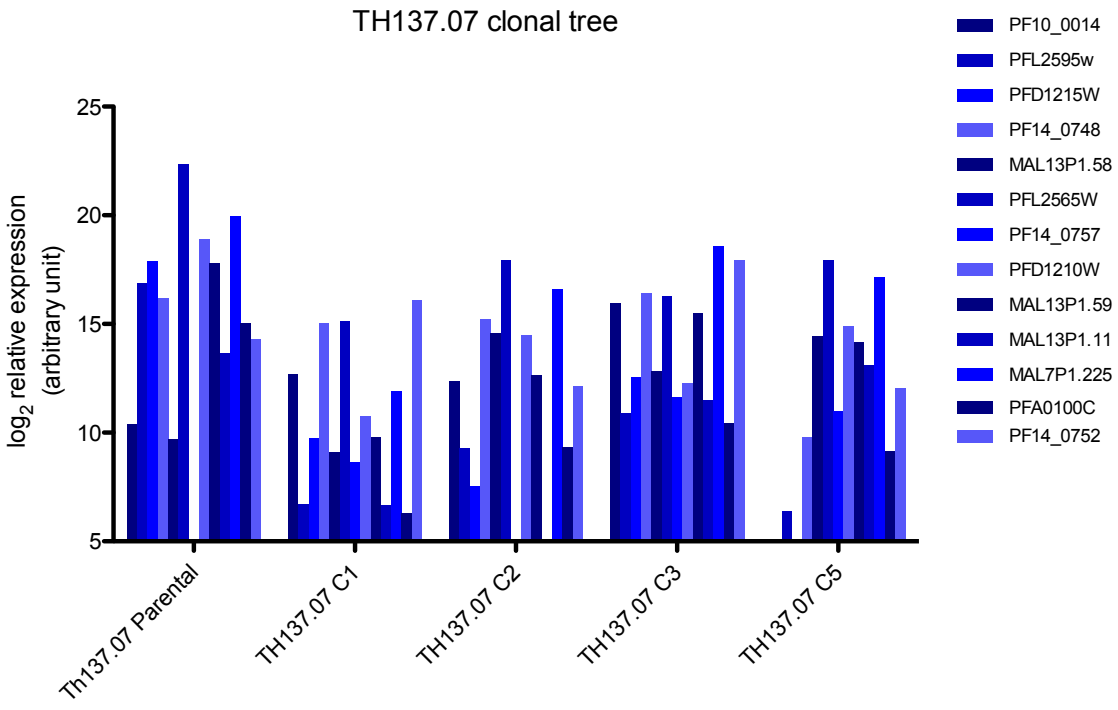


Figure 4- 1: **Lack of selection of *phista* genes culture-adapted parasite clones**. The transcriptional expression of 13 *phista* genes in a parent and four derived clones was assayed in ring stage parasites by qRT-PCR. Clones and parent cell lines show differences in their repertoire of *phista* expressed. Gene expression is normalized to skeleton binding protein 1 (SBP1) (a normalizer for ring stage parasites) by the delta Ct method.

B. *Phista* gene expression is modulated in cytoadherence-switched clones and parasites from children with cerebral malaria

The *phist* family has been linked to cytoadherence and a specific *phista* member; *PF14_0752* is upregulated during PAM and parasites selected for binding to various host cell receptors (Claessens et al., 2012; Mok et al., 2007). Therefore, we surveyed the expression of *phista* genes in parasites with adhesive properties using either, parasites experimentally selected for binding to specific host receptors and thus expressing one dominant *var* gene, or parasites directly sampled from the peripheral blood of Malawian children with CM without *ex vivo* culturing. Parasites selected for binding, or "binders" include IT4 parasites selected on human brain microvascular endothelial cells (HBMEC), HBMEC-selected parasite cross-adherent to human pulmonary endothelial cells (HPMEC) and parasites selected to form rosettes. For *in vivo* parasites, our guiding hypothesis is two-fold: first, parasites circulating in the peripheral blood (ring stages) are the same as those adhering to the endothelium of various tissue sites including the brain (trophozoites and schizonts). In fact, a study on a similar cohort in Malawi, determined that parasites in the peripheral blood of children with clinical CM were genetically identical as those found sequestered in tissues at autopsy (Milner et al., 2012). Second, they express one dominant *var* gene, which is translated into the PfEMP1 providing adhesion. We determined in the previous Chapter that parasites from the peripheral blood from this Malawian children cohort express UpsA type *var* genes at significantly higher levels than other *vars* and similar conclusions were reached in other cohorts of Tanzanian children with CM (Lavstsen et al., 2012).

In both binders and parasites from CM patients 3 *phista* genes are dominant: *PF10_0014, PFL2565w* and *PF14_0752* have the highest mean expression (Figure 4-3A). The same observation can be made when looking at CM parasites and binders separately (Figure 4-4). When comparing the expression of *phista* in these two groups with *in vitro* adapted parasites, the three dominant *phista* have comparable mean expression levels as *phista in vitro* and the unselected *phista* genes are expressed at lower levels. This suggests that *in vitro, phista* are all

upregulated and comparatively expressed whereas in CM parasites and binders, unselected *phista* genes are downregulated (Figure 4-3).

C. A subset of *phista* genes are variantly expressed

We showed that a specific subset of *phista* genes is highly expressed in parasites with cytoadherent properties as well as parasites from CM patients. *In vitro,* this selection is absent and genes are expressed at comparable levels. To understand if members of this gene family are coordinately expressed, i.e. if a subset of genes has similar expression patterns across isolates, we first determined their variable nature and then whether they are co-expressed. To determine the transcriptional behavior of *phista* genes, we compared the average gene rank of each member across isolates to define which genes that are constitutively highly expressed (top 20% of *phista* with high expression across most isolates), variantly expressed (variable expression across isolates). We then computed Pearson correlations of the expression of all gene pairs to determine co-expression patterns.

In the first part of the analysis we found that in both groups (culture-adapted parasites and cytoadherent and CM parasites) one *phista* gene *PFL2565w* is constitutively expressed and *PFL2595w* is the most variable gene (Figure 4-2B and 4-3B). In culture-adapted parasites two other genes *PF14_0748* and *MAL7P1.225* are also constitutively expressed. However, in CM and cytoadherent parasites these two genes fall within the ranks of variantly expressed genes and are replaced by *PF14_0752* and *PF10_0014* (Figure 4-2B and -3B). In terms of co-expression, the expression of *phista* genes is more correlated in CM and cytoadherent parasites than in culture-adapted parasites (corrected p-values > 2.3E-04) (Figure 4-2C & 3C). Three genes do not seem to correlate with the rest of the gene family, including genes that are constitutively expressed and/or dominant, *PF10_0014* and *PF14_0752*, and two variantly expressed genes *MAL7P1.225* and *PFD1215w* (Figure 4-3C). These genes are unique in that they are not co-regulated with any other *phista* genes analyzed in this study.

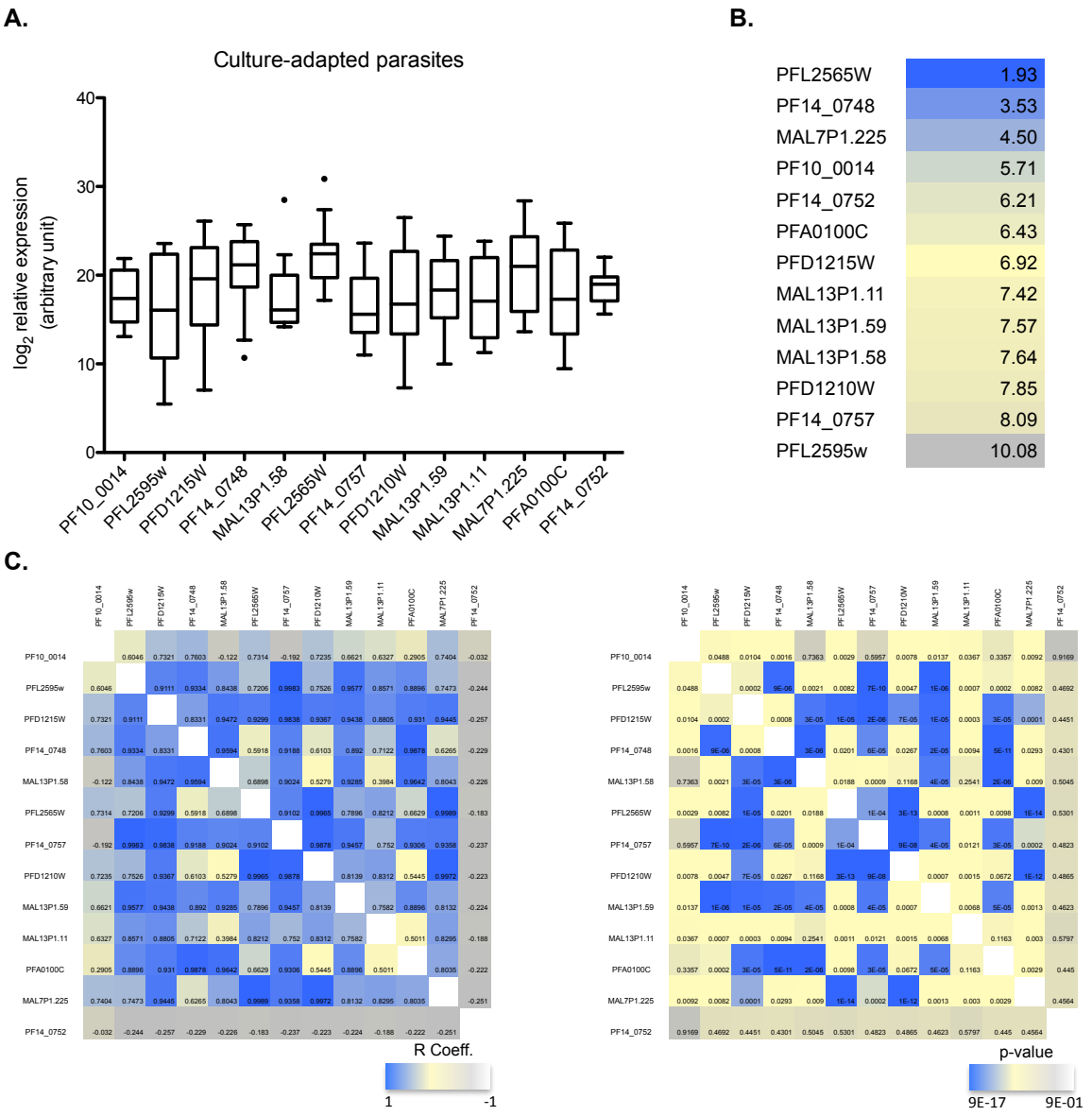


Figure 4- 2: **Expression of *phista* genes in unselected and culture-adapted parasites**. **A. Expression of 13 *phista* genes in culture-adapted parasites by qRT-PCR (n=15).** Shown is expression relative to the ring-stage marker SBP1 as box-and-whiskers plots (box line= mean, Tukey whiskers for the interquartile distance and outliers as dots). **B. Constitutive vs. variantly expressed *phista* genes.** High ranked genes correspond to constitutively highly expressed genes (blue) and variantly expressed genes are shown in yellow grade whereas genes with constitutive low expression levels are in grey **C. *Phista* co-expression** Correlation coefficient of *phista* genes and corresponding p-values (significance is set at the corrected p-values 2E-04, yellow to blue shading).

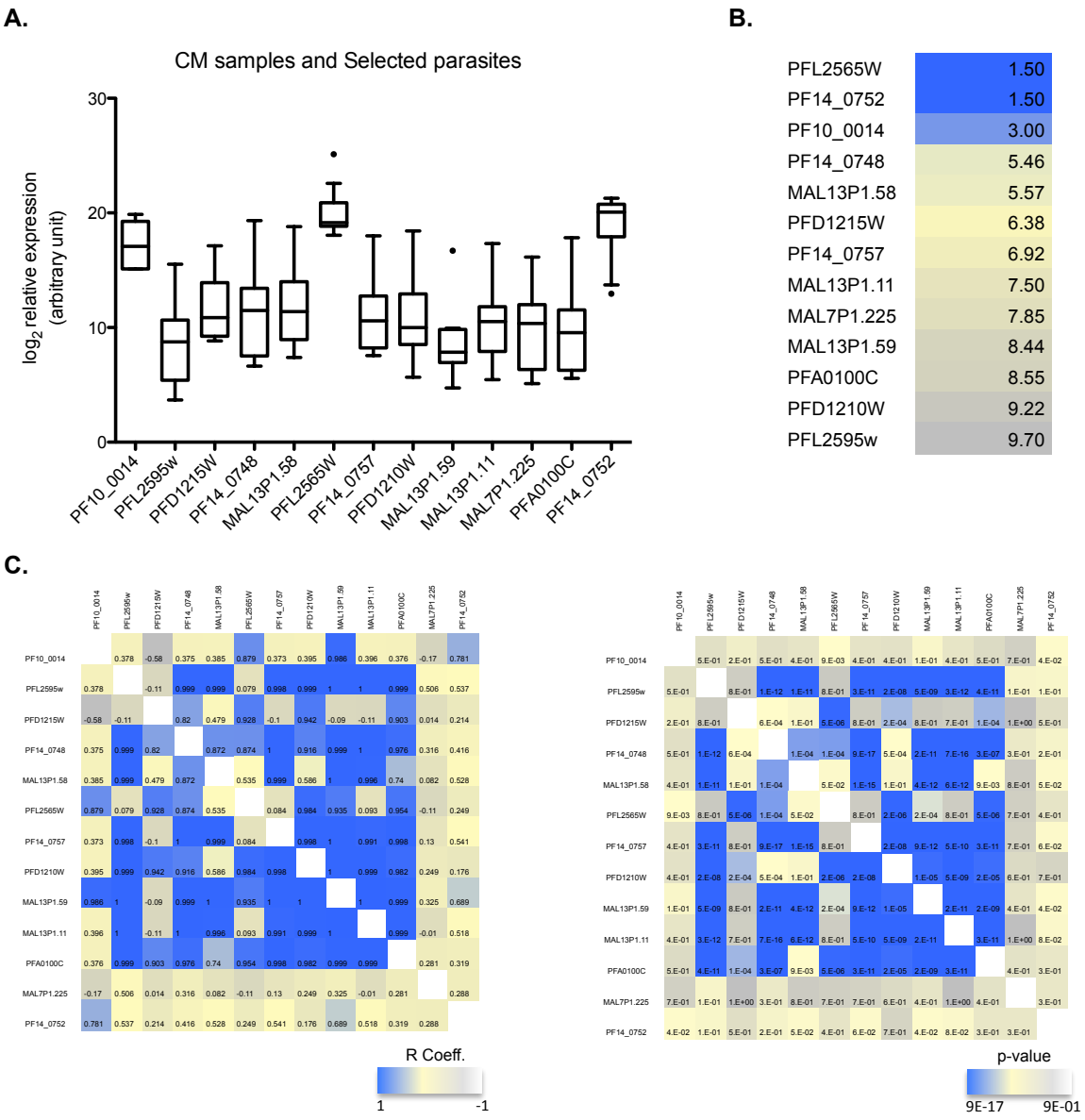


Figure 4- 3: **Expression of *phista* genes in cytoadherence-selected and cerebral malaria parasites. A. Expression of 13 *phista* genes in cytoadherence-selected (HBMEC n=3, HPMEC n=1, rosetters n=2) and CM parasites (n=8) by qRT-PCR.** Shown is expression relative to the ring-stage marker SBP1 as box-and-whiskers plots (box line= mean, Tukey whiskers for the interquartile distance and outliers as dots) **B. Constitutive vs. variantly expressed *phista* genes**. High ranked genes correspond to constitutively highly expressed genes (blue) and variantly expressed genes are shown in yellow grade whereas genes with constitutive low expression levels are in grey C**. *Phista* co-expression**. Correlation coefficient of *phista* genes and corresponding p-values (significance is set at the corrected p-values 2E-04, yellow to blue shade) are shown in gene-gene matrices.

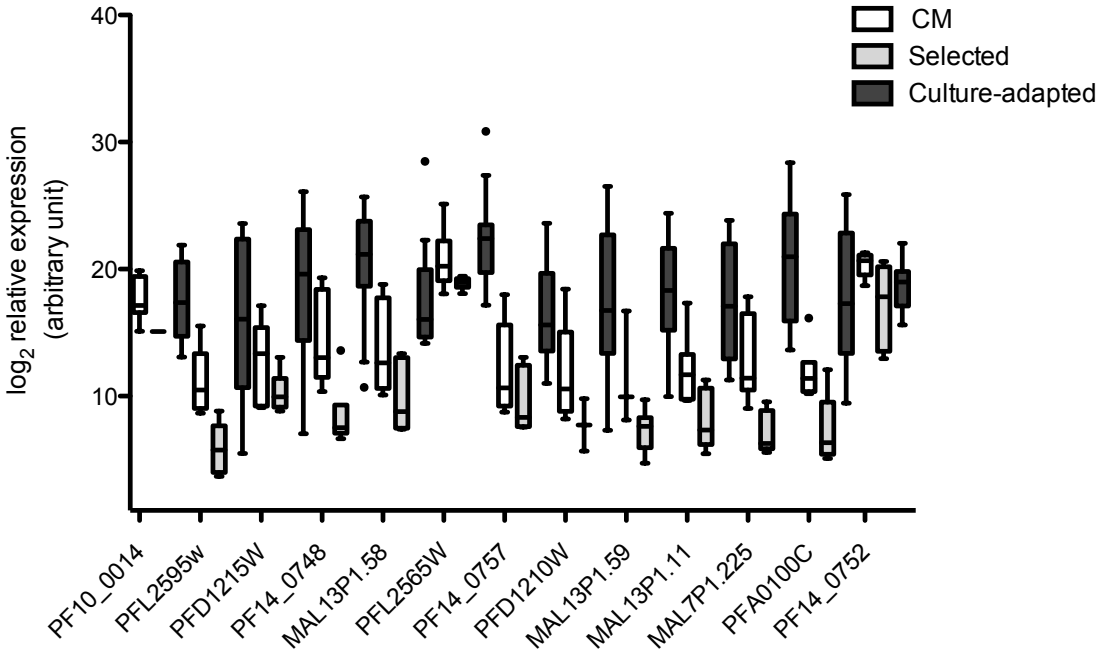


Figure 4- 4: ***Phista*** **gene expression in CM, cytoadherence-selected and culture-adapted parasites**. Shown is the mean expression of 13 *phista* genes in parasites from CM patients (white bars), parasites selected for binding (light grey bars) and unselected/ culture-adapted parasites (dark grey bars). Shown is *phista* gene expression relative to the ring-stage marker SBP1 as box-and-whiskers plots (box line= mean, Tukey whiskers for the interquartile distance and outliers as dots).

D. The expression of some *phista* may be under epigenetic regulation

As previously mentioned, multi-gene families located at the subtelomeres show unique transcriptional profile in *P. falciparum*. These encode exported proteins such as *var, rif, stevor, pfmc2tm* and *fikk* as well as proteins predicted to be exported including *phists.* The variant expression of these genes *in vitro* is consistent with epigenetic regulation as seen with the heterochromatin mark, H3K9me3, at their transcriptionally silent loci, binding of H3K9me3 by HP1 and a strong association between these marks and variant expression in isogenic clones (Flueck et al., 2009; Lopez-Rubio et al., 2009; Rovira-Graells et al., 2012; Salcedo-Amaya et al., 2009). In all 4 studies, some *phista* genes were included in ChiP-chip arrays and microarrays. Therefore, we collected information from all these studies to assess if the observed transcriptional behavior of *phista* genes determined in this study is consistent with what has been recently published. Although there are still information gaps within the *phista* family, there is a general consensus between the four studies (Figure 4-5, left part of matrix). When comparing our transcriptional data from culture-adapted parasites and CM and cytoadherent parasites, we also see some agreement. For example, *PFL2565w* is not enriched with heterochromatin marks nor is it a member of the Variantome (Figure 4-5). Agreeably, we also determined that this gene is constitutively expressed across all parasites assayed (Figure 4-2B and -3B). There are however some discrepancies. For *PF10_0014* and *PF14_0752* this can be explained by the change in expression in CM and cytoadherent parasites due to selection, for *MAL7P1.225* currently there is too little published data to draw any conclusion. Also, our finding of *PF14_0748* as being constitutively expressed in culture-adapted parasites maybe biased to our arbitrary threshold. (Figure 4-5).

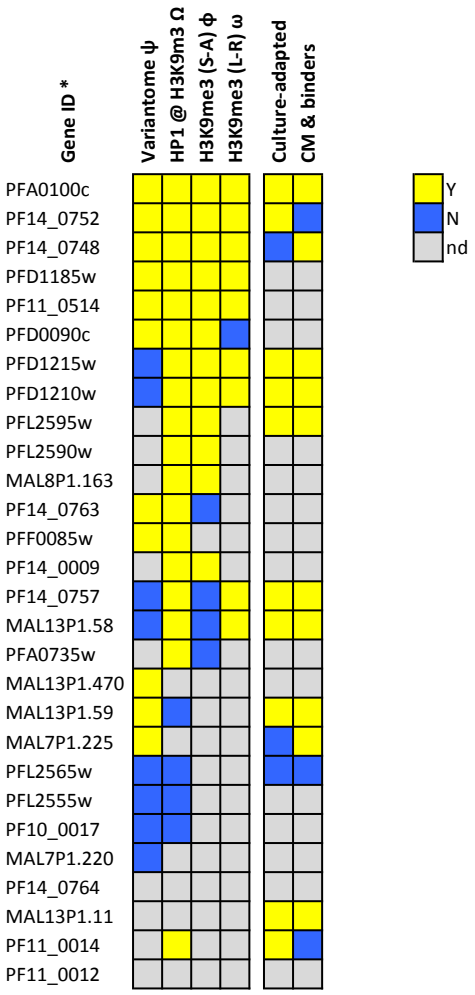


Figure 4- 5: Distribution of variant expression in the *phista* gene family. *Phista* genes that are present in the *P. falciparum* variantome and/or positive for heterochromatin marks when transcriptionally silent (H3K9me3 bound by HP1, H3K9me3 enrichment at silent loci) are shown in the left side of the matrix. *Phista* genes that are variantly expressed (Y) or constitutively expressed in culture-adapted parasites or CM parasite and binders are shown in the right side of the matrix (N). nd, non-determined.* *Phista* family was defined in Sargeant et al., 2006; $^{\Psi}$ Variantome (Rovira-Graells et al., 2012); $^{\Omega}$ HP1@H3K9me3 (Flueck et al., 2009); H3K9me3 enrichment based on ChIP-chip experiments: $^{\varphi}$ Salcedo-Amaya et al., 2009 and $^{\omega}$ Lopez-Rubio et al., 2009.

E. Subcellular localization of PHISTA domains in developing parasites

The subcellular localization of PHISTA proteins in iRBC remains unknown. PHISTA are relatively short proteins (~150aa), composed of a signal sequence and a PEXEL motif directly upstream of the PHIST domain. They lack TM domains and are therefore not predicted to anchor within membranes. To assess the subcellular localization of PHISTA domain, we generated 3 transgenic parasite lines expressing HA-tagged PHISTA domains under a HSP86 promoter for constitutive expression across the parasite development. We picked 3 *phista* genes with divergent transcriptional behavior: *PFL2565w* is constitutively expressed, lacks heterochromatin marks and is not part of the variantome (Figure 4-2B, -3B, -5); *PFD1185w* is variantly expressed and under epigenetic regulation based on recently published data (Figure 5); and *PF10_0014* which is selected in cytoadherent parasites (Figure 4-2B & -2C). PHISTA domains can be detected in the parasite in early ring stages (Figure 6). The signal is then localized in the iRBC cytoplasm in trophozoite and schizont stages leaving the PV completely devoid of signal. Additionally, PHISTA domains do not seem to aggregate in any structures such as the MCs as seen by the absence of punctate signal and co-localization with the MC-resident protein SBP1.

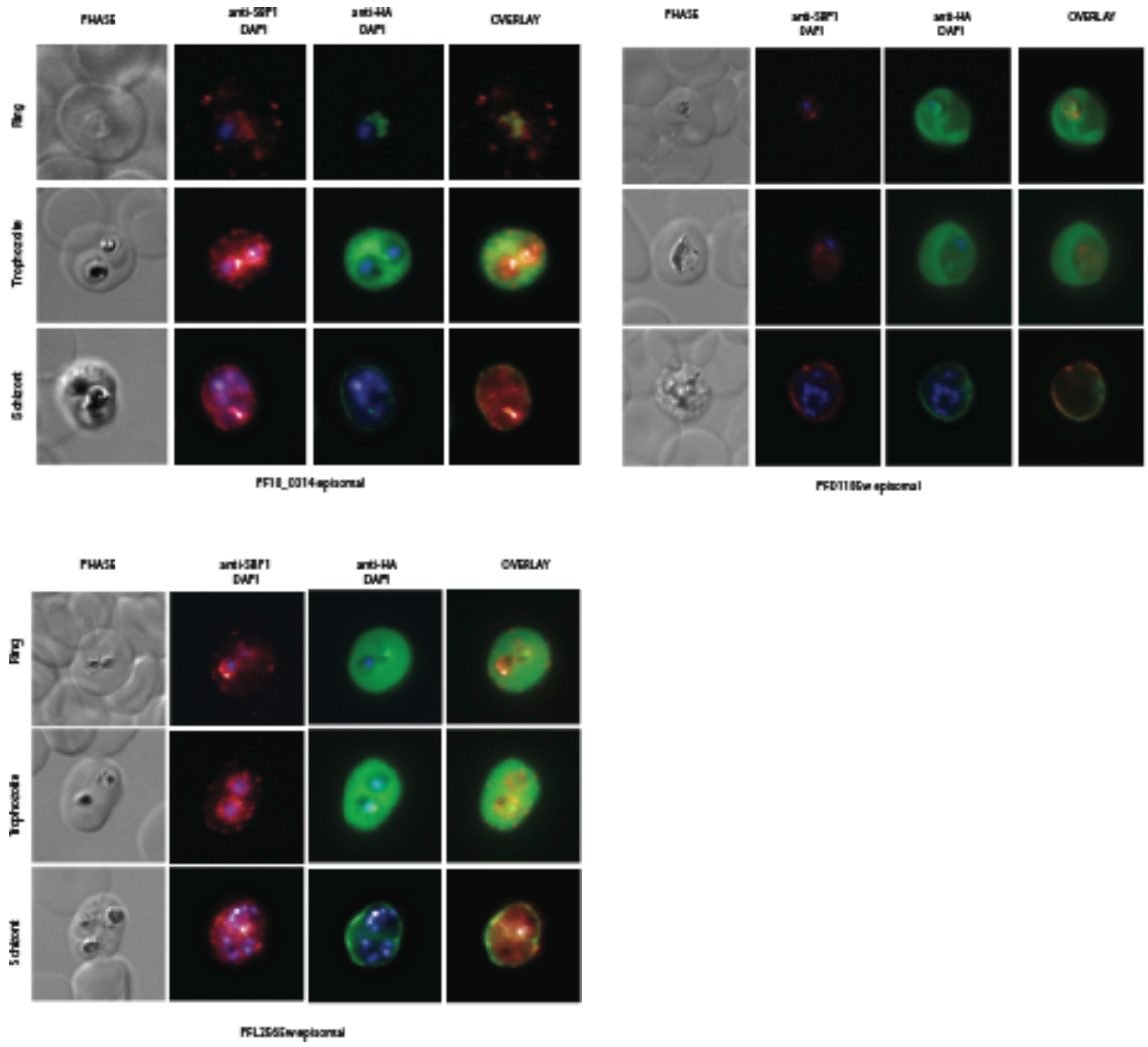


Figure 4- 6: Subcellular localization of three PHISTA domains during blood stage development. Episomal expression of three HA-tagged PHISTA domains in P2G12 *P. falciparum* parasites, PF10_0014, PFD1185w and PFL2565w. Localization of the reporter was assayed by immunofluorescence using anti-HA antibodies and co-localization with SBP1, a MC-resident protein using anti-SBP1 antibodies.

III. Discussion

In this work, I have determined that the *phista* family is composed of genes that have different transcriptional behavior *in vitro* as compared to parasites from the peripheral blood of children with CM and parasites that have been selected on human tissue or selected to form rosettes. In particular, a subset of *phista* genes is overexpressed in parasites with adhesive phenotypes.

At the expression level *phista* genes can be classified into two groups: constitutively or variantly expressed. Whereas a large subset is variantly expressed in culture-adapted, CM and selected parasites, a small number shows less expression variation. This pattern of expression is different from *var* genes where one *var* is dominantly expressed at any given time, yet similar to what has been observed for *rif, stevor* and *fikk* genes where multiple transcripts are highly expressed at any given time (Fernandez et al., 1999; Kyes et al., 1999; Lavazec et al., 2007; Nunes et al., 2007; Petter et al., 2007).

Like the aforementioned gene families, *phista genes* are also subtelomeric, and the variation in expression of some members is regulated by reversible modifications to chromatin and not by copy number variation (Flueck et al., 2009; Lopez-Rubio et al., 2009; Rovira-Graells et al., 2012; Salcedo-Amaya et al., 2009) (Figure 4-5). Therefore, the expression of all these families may be controlled by epigenetic processes that allow the parasite to respond rapidly to changes in the environment such as fluctuations in nutrient sources, changes in host cell ligand and antibody recognition and neutralization. In fact for each of these environmental changes, variant gene expression has been observed as a response by the parasite. The cytoadherence linked asexual genes (*clags)* variant expression has been linked to solute transport *eba*-140 variant expression to host cell invasion and *var* gene switching to immune evasion (Biggs et al., 1991; Cortes et al., 2007; Nguitragool et al., 2011). Here we postulate that *phista* expression is also modulated by cytoadherence.

In cytoadherence-selected parasites, three *phista* genes are highly expressed as compared to the rest of the family. These same genes are also dominant in CM samples, suggesting that a subset of *phista* genes is preferentially expressed in parasites that have

adhesive phenotypes. Although parasites from CM patients were not directly sampled from affected tissues, including the brain, as this would include great risks to the patients (children under five years of age were included in this study), we suspect that the parasites circulating in the peripheral blood have the same transcriptional profile as those emerging from sequestration in the microvasculature of damaged tissues. These parasites may present a transcriptional profile reflective of cytoadhesive parasites. However, it is still not known whether CM parasites sequestered in different tissues (as observed in autopsy studies) express different adhesion ligands or are able to cross-adhere to different tissues (Seydel et al., 2006). *In vitro*, current evidence seems to point towards cross-adhesion. The selection of parasites on HBMEC leads to the upregulation of UpsA-type *var* genes. These parasites can also adhere to pulmonary, dermal and cardiac microvascular endothelial cells (Avril et al., 2012).

It is not clear what role *phista* genes play in cytoadherence, however so far members of this family along with a handful of other genes have been transcriptionally linked to this process. In a recent study, the *phista* gene *PF14_0752*, membrane-associated histidine rich protein 1 (*mahrp1*), *P. falciparum* erythrocyte membrane protein 3 (*pfemp3*), *resa-1* and 3 other uncharacterized genes with a PEXEL motif, showed the highest transcriptional upregulation after selection for binding to HBMECs (Claessens et al., 2012). In another study, *PF14_0752* was also dramatically upregulated after selection for binding to CD36 (Mok et al., 2007). These genes all have transcriptional peaks in ring stages and could be translated before the chosen PfEMP1 is presented on the surface of trophozoite stages (Le Roch et al., 2003). Also, the proteins encoded by these genes are either involved in PfEMP1 trafficking or in altering the make-up and mechanical properties of the iRBC membrane (Da Silva et al., 1994; Glenister et al., 2002; Mills et al., 2007; Silva et al., 2005; Spycher et al., 2008). Since PHISTA domains do not localize to membrane structures as determined by IFA experiments, they may act as accessory proteins in the translocation events of PfEMP1 or other proteins involved in cytoadherence.

In summary, we have demonstrated that *phista* genes (which are highly underrepresented on microarray platforms), show differential transcriptional behavior *in vitro*

compared to cytoadherent parasites. Their unique presence in *falciparum* species and the fact that they encode soluble exported proteins ties them to cytoadherence.

IV. Materials and Methods

Phista primers

Primers for 13 *phista* genes were designed to anneal to non-polymorphic regions using published sequence diversity data from field isolates (PlasmoDB), optimized for qRT-PCR by using for *P. falciparum* 3D7 genomic DNA.

Culture-adapted parasites strains

Long-term culture-adapted parasites used in this study were NF54, 3D7, 3D7 clone P2G12, HB3, FCR3, FCR3/CS2 and W2Mef (W2 strain resistant for mefloquine). We also used a subset of clinical isolates from Gambia, Pf2004 and from three regions of Sénégal Thies (TH), Vélingara (V) and Pikine (P) over the span of 2004-2008. These parasites have been culture-adapted, frozen and thawed repeatedly. These included TH232.02, V56.04, V35.04, TH231.08, P06.08, Th137.07 and 4 clones generated by serial dilution by Dr. Amy Bei. We also used a set of parasites that were selected for binding to human tissues, kind gifts from Dr. Joe Smith. These included IT4 parasites selected on human brain microvascular endothelial cells (HBMEC) clones 1E2, 1A2 and 1E7; human pulmonary microvascular endothelial cells (HPMEC), the HPMEC clone. The generation and cloning of these parasites is described in Avril et al., 2012. We also used rosetting parasites, R29 and PF13. These were independently selected for rosetting using monoclonal antibodies on IT4 clone R29 and 3D7 clone PF13 (described in (Janes et al., 2011)).

Parasites from children with cerebral malaria

Peripheral blood samples from 8 children patients diagnosed with cerebral malaria were used. These eight samples are from the Malawian cohort described in Chapter III (see Chapter III "Subjects and Methods").

Parasite culture conditions

All *P. falciparum* lines used in this study were cultivated under standard conditions in O+ human erythrocytes, diluted to 4% hematocrit, and in culture medium composed of RPMI-1640 (Sigma-Aldrich) supplemented with 0.18% sodium bicarbonate and human serum to a final concentration of 10%. Cultures were kept in an atmosphere of 5% CO2 and 0.5% O2 in N2 at 37°C as previously described (Trager and Jensen, 1976).

Quantitative RT-PCR of parasites strains

To isolate ring-stage RNA samples, parasites were synchronized using 5% D-sorbitol solution. The presence of pure ring stages was confirmed by Giemsa smears after which five mls of culture was washed in 1X PBS and Trizol added to parasite pellets. Immediately, pellets were vortexed until homogenous and directly frozen at -80°C. To isolate RNA, Trizol-pellets were thawed at 37°C and equal amount of chloroform was added to extract the RNA layer. The RNA layer was then purified using the RNeasy mini kit (Qiagen). RNA was then submitted to a Dnase treatment (Turbo DNase, Invitrogen) to remove any traces of DNA. The purity and quantity of RNA was then determined using Nanodop system and cDNA was synthesized using SuperScript RT III (Invitrogen). A first qPCR run was performed to the “minus reverse transcriptase” control to test for any contaminating DNA species. If amplification occurred, the cDNA reaction was not used in further assays. For qRT-PCR, 1ng of cDNA was used per reaction. PHISTa gene expression was normalized to the ring-stage gene, encoding skeleton binding protein 1 (SBP1) by the delta Ct method. All qRT-PCR were run on Applied BioSystems machines.

Generation of transgenic parasites

All *P. falciparum* expression plasmids are based on the pHHK(+) vector (*P. falciparum* HSP86 promoter/ KAHRP N-terminus (signal sequence and export domain)/ KAHRP domain/ 3xHA-2xTy tag). Sequences encoding target PHISTA domain were amplified from genomic DNA of *P. falciparum* strain 3D7, using Phusion DNA polymerase (New England BioLabs, Ipswich MA). Amplicons were digested with the restriction enzymes AvrII and XhoI (New England BioLabs,

Ipswich MA) and ligated into pHHK(+) downstream of KAHRP export domain and upstream of a triple HA double TY tag, replacing the KAHRP N-domain. Maxipreps (Origene) of each plasmid were generated after transformation and selection of PMC103 electro-competent *Escherichia coli* stocks. P2G12 infected erythrocytes were transfected with 50-100 µg of purified plasmid DNA by electroporation using a BioRad GenePulser set at 0.310kV and 950uF. Transfected parasitized erythrocytes were put back in supplemented medium with 200-300µL of fresh RBCs. Drug selection was initiated by addition of 5nM WR99210 9-24hrs after transfection.

Indirect Immunofluorescence Assays

To generate synchronous PHISTa expressing iRBCs, transgenic parasites were grown to 5% parasitemia and synchronized for ring stages using 5% D-sorbitol. Parasites at ring, trophozoite and schizont stages were collected for intra-cellular localization by indirect immunofluorescence assays. Approximately 500µL of cells were pelleted by centrifugation at 2000rpm for 2min, and then pellets were washed twice in 1X PBS. Parasite pellets were fixed in 4% paraformaldehyde/ 0.0075% gluteraldeyde for 30min at room temperature in suspension. Pellets were washed in 1X PBS and permeabilized using 0.1% Triton X-100 for 10min at room temperature. Parasites were then washed in 1X PBS and cells were blocked in 3% w/v bovine serum albumin (BSA) in PBS for 1hr. Primary antibodies were than added in their respective dilutions and cells were incubated for 4hrs: Roche Applied Science rat anti-HA 1:1000 (clone 3F10), and rabbit anti-SBP1 1:500. Cells were washed for 5 min in 3% w/v BSA three times, then blocked with secondary antibodies (in BSA) for 2 hrs (Invitrogen's goat anti-rabbit alexa 594 (1:750) and goat anti-rat alexa 488 (1:1000). After 3 washes in 1X PBS, cells were resuspended in a small volume of 1X PBS and mounted on microscope glass slides using VectaShield containing 4',6-diamidino-2-phenylindole (DAPI) nuclear dye. Parasites were visualized on a traditional fluorescence microscope. Images were edited in Adobe Photoshop.

V. References

Acharya, P., Kumar, R., and Tatu, U. (2007). Chaperoning a cellular upheaval in malaria: heat shock proteins in Plasmodium falciparum. Molecular and biochemical parasitology *153*, 85-94.

Avril, M., Tripathi, A.K., Brazier, A.J., Andisi, C., Janes, J.H., Soma, V.L., Sullivan, D.J., Jr., Bull, P.C., Stins, M.F., and Smith, J.D. (2012). A restricted subset of var genes mediates adherence of Plasmodium falciparum-infected erythrocytes to brain endothelial cells. Proceedings of the National Academy of Sciences of the United States of America *109*, E1782-1790.

Banumathy, G., Singh, V., and Tatu, U. (2002). Host chaperones are recruited in membrane-bound complexes by Plasmodium falciparum. The Journal of biological chemistry *277*, 3902-3912.

Barnwell, J.W., Howard, R.J., Coon, H.G., and Miller, L.H. (1983). Splenic requirement for antigenic variation and expression of the variant antigen on the erythrocyte membrane in cloned Plasmodium knowlesi malaria. Infection and immunity *40*, 985-994.

Biggs, B.A., Gooze, L., Wycherley, K., Wollish, W., Southwell, B., Leech, J.H., and Brown, G.V. (1991). Antigenic variation in Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America *88*, 9171-9174.

Biggs, B.A., Kemp, D.J., and Brown, G.V. (1989). Subtelomeric chromosome deletions in field isolates of Plasmodium falciparum and their relationship to loss of cytoadherence in vitro. Proceedings of the National Academy of Sciences of the United States of America *86*, 2428-2432.

Botha, M., Chiang, A.N., Needham, P.G., Stephens, L.L., Hoppe, H.C., Kulzer, S., Przyborski, J.M., Lingelbach, K., Wipf, P., Brodsky, J.L.*, et al.* (2011). Plasmodium falciparum encodes a single cytosolic type I Hsp40 that functionally interacts with Hsp70 and is upregulated by heat shock. Cell stress & chaperones *16*, 389-401.

Buchholz, K., Burke, T.A., Williamson, K.C., Wiegand, R.C., Wirth, D.F., and Marti, M. (2011). A high-throughput screen targeting malaria transmission stages opens new avenues for drug development. The Journal of infectious diseases *203*, 1445-1453.

Buffet, P.A., Gamain, B., Scheidig, C., Baruch, D., Smith, J.D., Hernandez-Rivas, R., Pouvelle, B., Oishi, S., Fujii, N., Fusai, T.*, et al.* (1999). Plasmodium falciparum domain mediating adhesion to chondroitin sulfate A: a receptor for human placental infection. Proceedings of the National Academy of Sciences of the United States of America *96*, 12743-12748.

Claessens, A., Adams, Y., Ghumra, A., Lindergard, G., Buchan, C.C., Andisi, C., Bull, P.C., Mok, S., Gupta, A.P., Wang, C.W.*, et al.* (2012). A subset of group A-like var genes encodes the malaria parasite ligands for binding to human brain endothelial cells. Proceedings of the National Academy of Sciences of the United States of America *109*, E1772-1781.

Cortes, A., Carret, C., Kaneko, O., Yim Lim, B.Y., Ivens, A., and Holder, A.A. (2007). Epigenetic silencing of Plasmodium falciparum genes linked to erythrocyte invasion. PLoS pathogens *3*, e107.

Da Silva, E., Foley, M., Dluzewski, A.R., Murray, L.J., Anders, R.F., and Tilley, L. (1994). The Plasmodium falciparum protein RESA interacts with the erythrocyte cytoskeleton and modifies erythrocyte thermal stability. Molecular and biochemical parasitology *66*, 59-69.

Daily, J.P., Le Roch, K.G., Sarr, O., Ndiaye, D., Lukens, A., Zhou, Y., Ndir, O., Mboup, S., Sultan, A., Winzeler, E.A.*, et al.* (2005). In vivo transcriptome of Plasmodium falciparum reveals overexpression of transcripts that encode surface proteins. The Journal of infectious diseases *191*, 1196-1203.

Day, K.P., Karamalis, F., Thompson, J., Barnes, D.A., Peterson, C., Brown, H., Brown, G.V., and Kemp, D.J. (1993). Genes necessary for expression of a virulence determinant and for transmission of Plasmodium falciparum are located on a 0.3-megabase region of chromosome 9. Proceedings of the National Academy of Sciences of the United States of America *90*, 8292-8296.

Eksi, S., Haile, Y., Furuya, T., Ma, L., Su, X., and Williamson, K.C. (2005). Identification of a subtelomeric gene family expressed during the asexual-sexual stage transition in Plasmodium falciparum. Molecular and biochemical parasitology *143*, 90-99.

Fernandez, V., Hommel, M., Chen, Q., Hagblom, P., and Wahlgren, M. (1999). Small, clonally variant antigens expressed on the surface of the Plasmodium falciparum-infected erythrocyte are encoded by the rif gene family and are the target of human immune responses. The Journal of experimental medicine *190*, 1393-1404.

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 pathogens *5*, e1000569.

Glenister, F.K., Coppel, R.L., Cowman, A.F., Mohandas, N., and Cooke, B.M. (2002). Contribution of parasite proteins to altered mechanical properties of malaria-infected red blood cells. Blood *99*, 1060-1063.

Grover, M., Chaubey, S., Ranade, S., and Tatu, U. (2013). Identification of an exported heat shock protein 70 in Plasmodium falciparum. Parasite *20*, 2.

Janes, J.H., Wang, C.P., Levin-Edens, E., Vigan-Womas, I., Guillotte, M., Melcher, M., Mercereau-Puijalon, O., and Smith, J.D. (2011). Investigating the host binding signature on the Plasmodium falciparum PfEMP1 protein family. PLoS pathogens *7*, e1002032.

Kyes, S., Horrocks, P., and Newbold, C. (2001). Antigenic variation at the infected red cell surface in malaria. Annual review of microbiology *55*, 673-707.

Kyes, S.A., Rowe, J.A., Kriek, N., and Newbold, C.I. (1999). Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America *96*, 9333-9338.

LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld, L.W., Ota, I., Sahasrabudhe, S., Kurschner, C.*, et al.* (2005). A protein interaction network of the malaria parasite Plasmodium falciparum. Nature *438*, 103-107.

Lasonder, E., Ishihama, Y., Andersen, J.S., Vermunt, A.M., Pain, A., Sauerwein, R.W., Eling, W.M., Hall, N., Waters, A.P., Stunnenberg, H.G.*, et al.* (2002). Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. Nature *419*, 537-542.

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

Lavstsen, T., Turner, L., Saguti, F., Magistrado, P., Rask, T.S., Jespersen, J.S., Wang, C.W., Berger, S.S., Baraka, V., Marquard, A.M.*, et al.* (2012). Plasmodium falciparum erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children. Proceedings of the National Academy of Sciences of the United States of America *109*, E1791-1800.

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 (New York, N.Y.) *301*, 1503-1508.

Leech, J.H., Barnwell, J.W., Miller, L.H., and Howard, R.J. (1984). Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. The Journal of experimental medicine *159*, 1567-1575.

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.

MacPherson, G.G., Warrell, M.J., White, N.J., Looareesuwan, S., and Warrell, D.A. (1985). Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. The American journal of pathology *119*, 385-401.

Maier, A.G., Cooke, B.M., Cowman, A.F., and Tilley, L. (2009). Malaria parasite proteins that remodel the host erythrocyte. Nature reviews. Microbiology *7*, 341-354.

Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szestak, T., Chesson, J., Wu, Y., Hughes, K., Coppel, R.L.*, et al.* (2008). Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell *134*, 48-61.

Mayer, C., Slater, L., Erat, M.C., Konrat, R., and Vakonakis, I. (2012). Structural analysis of the Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. The Journal of biological chemistry *287*, 7182-7189.

Mills, J.P., Diez-Silva, M., Quinn, D.J., Dao, M., Lang, M.J., Tan, K.S., Lim, C.T., Milon, G., David, P.H., Mercereau-Puijalon, O.*, et al.* (2007). Effect of plasmodial RESA protein on deformability of human red blood cells harboring Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America *104*, 9213-9217.

Milner, D.A., Jr., Vareta, J., Valim, C., Montgomery, J., Daniels, R.F., Volkman, S.K., Neafsey, D.E., Park, D.J., Schaffner, S.F., Mahesh, N.C.*, et al.* (2012). Human cerebral malaria and Plasmodium falciparum genotypes in Malawi. Malaria journal *11*, 35.

Mok, B.W., Ribacke, U., Winter, G., Yip, B.H., Tan, C.S., Fernandez, V., Chen, Q., Nilsson, P., and Wahlgren, M. (2007). Comparative transcriptomal analysis of isogenic Plasmodium falciparum clones of distinct antigenic and adhesive phenotypes. Molecular and biochemical parasitology *151*, 184-192.

Nguitragool, W., Bokhari, A.A., Pillai, A.D., Rayavara, K., Sharma, P., Turpin, B., Aravind, L., and Desai, S.A. (2011). Malaria parasite clag3 genes determine channel-mediated nutrient uptake by infected red blood cells. Cell *145*, 665-677.

Nunes, M.C., Goldring, J.P., Doerig, C., and Scherf, A. (2007). A novel protein kinase family in Plasmodium falciparum is differentially transcribed and secreted to various cellular compartments of the host cell. Molecular microbiology *63*, 391-403.

Peters, J.M., Fowler, E.V., Krause, D.R., Cheng, Q., and Gatton, M.L. (2007). Differential changes in Plasmodium falciparum var transcription during adaptation to culture. The Journal of infectious diseases *195*, 748-755.

Petter, M., Haeggstrom, M., Khattab, A., Fernandez, V., Klinkert, M.Q., and Wahlgren, M. (2007). Variant proteins of the Plasmodium falciparum RIFIN family show distinct subcellular localization and developmental expression patterns. Molecular and biochemical parasitology *156*, 51-61.

Pongponratn, E., Riganti, M., Punpoowong, B., and Aikawa, M. (1991). Microvascular sequestration of parasitized erythrocytes in human falciparum malaria: a pathological study. The American journal of tropical medicine and hygiene *44*, 168-175.

Roberts, D.J., Craig, A.G., Berendt, A.R., Pinches, R., Nash, G., Marsh, K., and Newbold, C.I. (1992). Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature *357*, 689-692.

Rovira-Graells, N., Gupta, A.P., Planet, E., Crowley, V.M., Mok, S., Ribas de Pouplana, L., Preiser, P.R., Bozdech, Z., and Cortes, A. (2012). Transcriptional variation in the malaria parasite Plasmodium falciparum. Genome research *22*, 925-938.

Rowe, J.A., Moulds, J.M., Newbold, C.I., and Miller, L.H. (1997). P. falciparum rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. Nature *388*, 292-295.

Salanti, A., Dahlback, M., Turner, L., Nielsen, M.A., Barfod, L., Magistrado, P., Jensen, A.T., Lavstsen, T., Ofori, M.F., Marsh, K.*, et al.* (2004). Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. The Journal of experimental medicine *200*, 1197-1203.

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. Proceedings of the National Academy of Sciences of the United States of America *106*, 9655-9660.

Sargeant, T.J., Marti, M., Caler, E., Carlton, J.M., Simpson, K., Speed, T.P., and Cowman, A.F. (2006). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome biology *7*, R12.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J., and Lanzer, M. (1998). 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 *17*, 5418-5426.

Seydel, K.B., Milner, D.A., Jr., Kamiza, S.B., Molyneux, M.E., and Taylor, T.E. (2006). The distribution and intensity of parasite sequestration in comatose Malawian children. The Journal of infectious diseases *194*, 208-205.

Silva, M.D., Cooke, B.M., Guillotte, M., Buckingham, D.W., Sauzet, J.P., Le Scanf, C., Contamin, H., David, P., Mercereau-Puijalon, O., and Bonnefoy, S. (2005). A role for the Plasmodium falciparum RESA protein in resistance against heat shock demonstrated using gene disruption. Molecular microbiology *56*, 990-1003.

Silvestrini, F., Bozdech, Z., Lanfrancotti, A., Di Giulio, E., Bultrini, E., Picci, L., Derisi, J.L., Pizzi, E., and Alano, P. (2005). Genome-wide identification of genes upregulated at the onset of gametocytogenesis in Plasmodium falciparum. Molecular and biochemical parasitology *143*, 100-110.

Silvestrini, F., Lasonder, E., Olivieri, A., Camarda, G., van Schaijk, B., Sanchez, M., Younis Younis, S., Sauerwein, R., and Alano, P. (2010). Protein export marks the early phase of gametocytogenesis of the human malaria parasite Plasmodium falciparum. Molecular & cellular proteomics : MCP *9*, 1437-1448.

Silvestrini, F., Tiburcio, M., Bertuccini, L., and Alano, P. (2012). Differential adhesive properties of sequestered asexual and sexual stages of Plasmodium falciparum on human endothelial cells are tissue independent. PloS one *7*, e31567.

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., Pinches, R., Newbold, C.I., and Miller, L.H. (1995). Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell *82*, 101-110.

Smith, J.D., Subramanian, G., Gamain, B., Baruch, D.I., and Miller, L.H. (2000). Classification of adhesive domains in the Plasmodium falciparum erythrocyte membrane protein 1 family. Molecular and biochemical parasitology *110*, 293-310.

Spycher, C., Rug, M., Pachlatko, E., Hanssen, E., Ferguson, D., Cowman, A.F., Tilley, L., and Beck, H.P. (2008). The Maurer's cleft protein MAHRP1 is essential for trafficking of PfEMP1 to the surface of Plasmodium falciparum-infected erythrocytes. Molecular microbiology *68*, 1300-1314.

Su, X.Z., Heatwole, V.M., Wertheimer, S.P., Guinet, F., Herrfeldt, J.A., Peterson, D.S., Ravetch, J.A., and Wellems, T.E. (1995). The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. Cell *82*, 89-100.

Templeton, T.J., and Deitsch, K.W. (2005). Targeting malaria parasite proteins to the erythrocyte. Trends in parasitology *21*, 399-402.

Trager, W., and Jensen, J.B. (1976). Human malaria parasites in continuous culture. Science (New York, N.Y.) *193*, 673-675.

Walsh, P., Bursac, D., Law, Y.C., Cyr, D., and Lithgow, T. (2004). The J-protein family: modulating protein assembly, disassembly and translocation. EMBO reports *5*, 567-571.

CHAPTER V

Concluding Remarks

CONTRIBUTIONS

This chapter was written by Karell Pellé.

I. Summary and General Discussion

In this dissertation, I aimed to answer different questions that are all tied to one bigger problem: how do intracellular parasites manipulate the host cell in order to establish infections? There are many dimensions to this problem and a priori to solve this problem there are a few concepts that need to be explored. First, it requires a detailed understanding of the mechanisms that occur at the cellular level that allow parasites to establish residence in a host cell, successfully. Second, the establishment of infection is a systemic process that requires parasites to propagate in different cells, tissues and organs and this hints to coordinated parasitic processes that either allow, or restrict, such spread. Third, infections are a complex balance between parasite survival processes and host defenses that result in a gradation of symptoms and disease severity. This points to parasite pathogenic features that can be fine-tuned in response to host defense mechanisms and factors that fluctuate in the host. In this work, we approached this problem by addressing these various points.

In Chapter II, I investigate the sequence requirements for protein export in Apicomplexa parasites of similar life styles. In fact, for these parasites an intracellular lifestyle is not facultative. As developed in the introductory Chapter, Apicomplexan parasites such as *Toxoplasma* and *Plasmodium* are devoid of *de novo* synthesis of certain amino acids and thus require host cells for growth and replication. Apicomplexa parasites all reside in a parasitophorous vacuole after invading the host cell. We find that *Plasmodium, Babesia* and *Cryptosporidium* share similar export domains. These domains are all functional in *Plasmodium* whereas in *Babesia* the domain acts as a dense granule (secretory organelle) targeting sequence. We propose that the functionality of this domain has evolved to serve different tasks depending on the type of "residence" the parasite establishes in its host cell. In fact, *Babesia* only remains in a PV for a few minutes after invasion whereas *Plasmodium* parasite remain in the PV until egress and utilize the domain for export during its development in the PV.

The proteins that are secreted from the parasite into the host, through secretory organelles during invasion or during development in the PV, are either potentially essential for parasite viability and/or involved in processes at the host-parasite interface such as host cell remodeling. Host cell

remodeling comprises all the parasite-mediated processes that alter the make-up of the infected host cell. These processes can be thought of as means for the parasite to establish successful host colonization, intracellular growth and replication, which require nutrient and energy acquisition, immune evasion and transmission.

Many of the players in host cell remodeling have been localized outside the parasite's plasma membrane and are often members of multi-gene families. Gene family expansion creates functional redundancy even across life cycle stages. *P. falciparum* multi-gene families that harbor an export domain are commonly localized at the subtelomeres, which are regions of the chromosomes highly prone to recombination events (Sargeant et al., 2006). At the gene family level, this creates diversity, which can potentially alter the function of certain family members. On top of diversity driven by recombination, known subtelomeric genes (i.e. 2TM gene families) encoding surface-expressed proteins show regions of high nucleotide diversity or single nucleotide polymorphisms (SNPs) (Lavazec et al., 2006; Volkman et al., 2002). These are RBC anchored membrane protein exposed to antibody recognition and may be therefore targeted by the immune system and in the case of PfEMP1 and RIFINS has lead to the evolution of antigenic variation. Furthermore, subtelomeric genes have been shown to undergo variant expression through epigenetic mechanisms, where changes in heterochromatin state of gene loci either silences or activates gene transcriptional expression (Flueck et al., 2009; Lopez-Rubio et al., 2009; Rovira-Graells et al., 2012; Salcedo-Amaya et al., 2009; Tonkin et al., 2009). Genes involved in host cell remodeling thus have all the features that could allow the parasite to adapt and tolerate environmental changes, within a host, or between hosts.

In Chapter III, we take advantage of this insight in order to systematically identify and describe these genes. Not only do we identify genes at the host-parasite interface, we do so in their natural state using microarray transcriptional expression data of parasites directly sampled from patients with *P. falciparum* malaria. We discover gene co-expression modules, groups of genes, which through co-expression across clinical isolates, are hypothesized to function in the same pathway. We are able to validate some of these modules through the association of known genes in particular pathways such as invasion and cytoadherence. These genes modules are rich

in information. They directly link genes to disease. In fact, we observed that variable modules (groups of genes that co-vary together across patients) are associated with pathology whereas constitutive modules do not. Additionally, the loci of genes from these variable modules are enriched in heterochromatin marks when transcriptionally silent. This highly suggests that the variable expression of these genes is a result of parasite responses to changes in the host environment. These responses could range from evasion of host humoral immunity or adaptation to nutrient and even oxygen deprivation, temperature variation and changes in host receptor type and availability. It remains to be determined how these genes and their product actually influence disease. Additionally, genes in constitutive export modules encoding exported proteins are potentially interesting for disease control purposes. Not mentioned in Chapter III, we identified an export module containing the vaccine candidate *resa.* This gene forms a small module containing 3 other *phista* genes, including *phista* PFL2565w, which was determined in Chapter IV as a constitutively highly expressed gene across all samples and selectively upregulated in CM patients. This partially confirms the definition of this resa module as a constitutive module and suggests that it includes other genes that could be added to the malaria vaccine candidate selection pipeline in particular if they meet the criteria for immunogenicity.

Finally, in Chapter IV, we focus on one gene family: the *phista* multi-gene family. This gene family is a perfect candidate to validate the use of expression data (as done in Chapter III) in order to study gene function in host cell remodeling. In Chapter II, the *phist* superfamily was described as an ancestral plasmodial gene family as it is the only family found across the entire genus that harbors an export domain. In Chapter III, *phists* were found in a number of modules both constitutive and variant in nature. The growing amount of data on *phists* suggests that they are involved in two processes, cytoadherence and host cell remodeling during sexual development, and may function as chaperones. We determine that *phista* genes are variantly expressed in unselected culture-adapted parasites and that a few members are dominantly expressed in parasites with binding phenotypes and parasites from children with cerebral malaria. One of these *phista* genes is *PF14_0752* also present in variable module 1 (Chapter III). Therefore it is possible that this module is involved in cytoadherence, specifically the

translocation of PfEMP1 to the surface as the genetic deletion of another member of this module PF11_0512, abrogates PfEMP1 export (Maier et al., 2008).

II. Future Perspectives

In this dissertation, we have identified single genes as well as gene modules at the host-parasite interface, which potentially function in host cell remodeling. Additional work remains to validate and elucidate their function. For this to be achieved, functional studies using sophisticated tools such as forward genetics (i.e. protein overexpression) or reverse genetics (i.e. genetic deletion) in combination with phenotypic read-out related to cytoadherence, host cell rigidity, resistance to heat shock, growth among other assays could be performed *in vitro*. However, there are limitations with such studies. First, generating transgenic *P. falciparum* parasites is technically difficult due to the A/T richness of its genome, which makes cloning complicated and proper plasmid integration events rare (Gardner et al., 2002; Maier et al., 2008). Second, *in vitro* adaptation often leads to genetic mutations which may be as severe as deletions of large chromosome fragment, therefore one has to always keep in mind the fragility of parasites in culture and experimental error that can arise from it (Day et al., 1993; Ribacke et al., 2007). Third, gene networks identified *in vivo* although relevant to disease may be different when parasites are studied under controlled culture conditions. As described in Chapter III and IV, studying parasites *in vitro* essentially removes any selection that may have caused the phenotype initially. However, some tools have allowed the reproduction of such phenotypes *in vitro*, for example panning experiments on soluble receptors or human endothelial cells allows for the enrichment of parasite subpopulations able to cytoadhere.

The work in this dissertation is focused on asexual stages of the intra-erythrocytic life cycle. However, host cell remodeling also occurs in other stages. For example, a few *phista* genes are potential major players in sexual development. Within this family, we have linked a few members to cytoadherence in asexual stages. As mentioned earlier, gene family expansion may serve a role in diversifying function across parasite stages. Therefore, *phista* genes may also be

involved in cytoadherence in gametocytes. This suggests that host cell remodeling can also be studied in other parasite stages and that the function of genes involved in this process may be redundant. In terms of drug and vaccine development this has importance as a single compound or a single epitope can target multiple parasite stages. In particular for a vaccine, a single epitope could block disease by targeting asexual stages in the erythrocytic cycle (when pathology occurs) and block transmission by targeting gametocytes in the blood and potentially in the mosquito if an antibody response has been elicited. Ultimately there is great potential to increase efficacy of the tool used when multiple stages are targeted.

III. References

Day, K.P., Karamalis, F., Thompson, J., Barnes, D.A., Peterson, C., Brown, H., Brown, G.V., and Kemp, D.J. (1993). Genes necessary for expression of a virulence determinant and for transmission of Plasmodium falciparum are located on a 0.3-megabase region of chromosome 9. Proceedings of the National Academy of Sciences of the United States of America *90*, 8292-8296.

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 pathogens *5*, e1000569.

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.

Lavazec, C., Sanyal, S., and Templeton, T.J. (2006). Hypervariability within the Rifin, Stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. Nucleic acids research *34*, 6696-6707.

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.

Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szestak, T., Chesson, J., Wu, Y., Hughes, K., Coppel, R.L.*, et al.* (2008). Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell *134*, 48-61.

Ribacke, U., Mok, B.W., Wirta, V., Normark, J., Lundeberg, J., Kironde, F., Egwang, T.G., Nilsson, P., and Wahlgren, M. (2007). Genome wide gene amplifications and deletions in Plasmodium falciparum. Molecular and biochemical parasitology *155*, 33-44.

Rovira-Graells, N., Gupta, A.P., Planet, E., Crowley, V.M., Mok, S., Ribas de Pouplana, L., Preiser, P.R., Bozdech, Z., and Cortes, A. (2012). Transcriptional variation in the malaria parasite Plasmodium falciparum. Genome research *22*, 925-938.

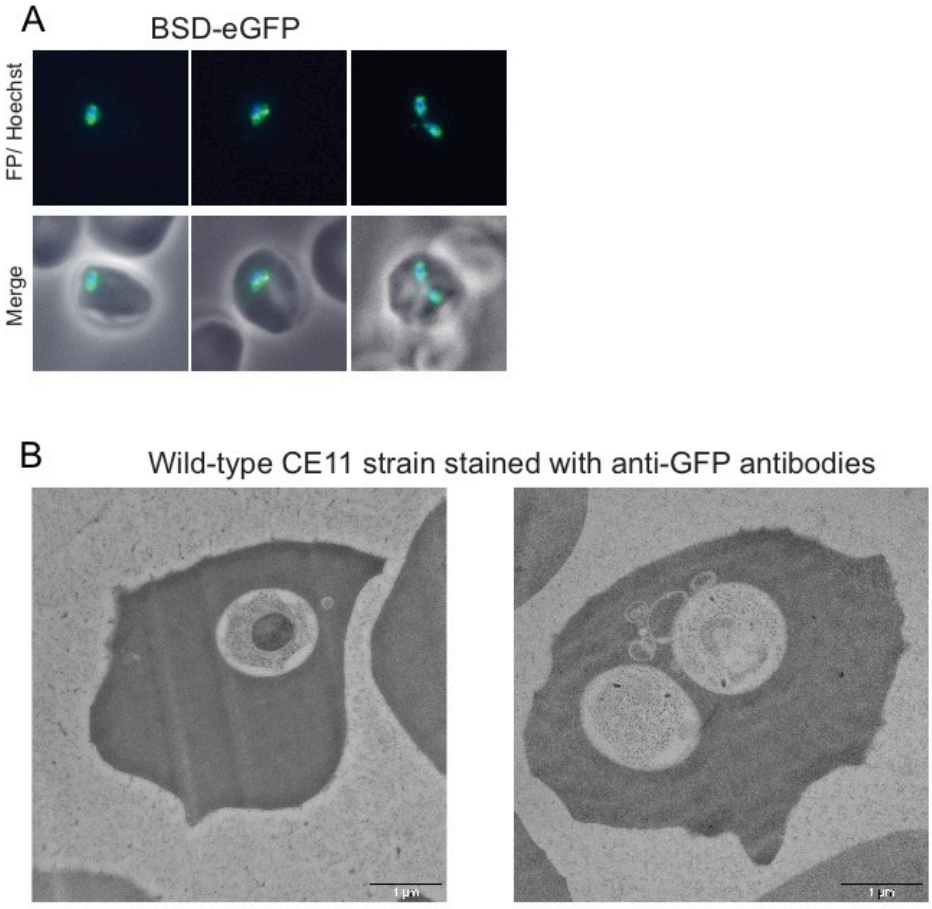
Salcedo-Amaya, A.M., van Driel, M.A., Alako, B.T., Trelle, M.B., van den Elzen, A.M., Cohen, A.M., Janssen-Megens, E.M., van de Vegte-Bolmer, M., Selzer, R.R., Iniguez, A.L.*, et al.* (2009). Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America *106*, 9655-9660.

Sargeant, T.J., Marti, M., Caler, E., Carlton, J.M., Simpson, K., Speed, T.P., and Cowman, A.F. (2006). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome biology *7*, R12.

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.

Volkman, S.K., Hartl, D.L., Wirth, D.F., Nielsen, K.M., Choi, M., Batalov, S., Zhou, Y., Plouffe, D., Le Roch, K.G., Abagyan, R.*, et al.* (2002). Excess polymorphisms in genes for membrane proteins in Plasmodium falciparum. Science (New York, N.Y.) *298*, 216-218.

ANNEX



Supplemental figure S1: Control experiments and smORF2-YFP ultrastructural distribution in *B. bovis*. A. IFA with the BSD-eGFP parasite line. This line expresses a BSD-eGFP fusion protein in the parasite cytoplasm. Both IFA and live imaging do not show any labeling outside of the parasite. **B. Immunoelectron microscopy with CE11 parasites.** Wild type parasites do not show any reactivity when stained with anti-GFP antibodies. Altogether these control experiments demonstrate that the signal obtained by IFA and immunoelectron microscopy is specific.