Exported proteins and their role in Plasmodium infection

Loh, Han Ping

2019

Loh, H. P. (2019). Exported proteins and their role in Plasmodium infection. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/100418>

<https://doi.org/10.32657/10356/100418>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**EXPORTED PROTEINS AND
THEIR ROLE IN *PLASMODIUM*
INFECTION**

**LOH HAN PING
SCHOOL OF BIOLOGICAL SCIENCES**

2019

EXPORTED PROTEINS AND THEIR ROLE IN *PLASMODIUM* INFECTION

LOH HAN PING

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang
Technological University in partial fulfilment of
the requirement for the degree of Doctor of
Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

[Input Date Here]

07/06/2019

.....
Date

[Input Signature Here]

Loh Han Ping

.....
[Input Name Here]

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

[Input Date Here]

07/06/2019

.....
Date

[Input Supervisor Signature Here]

Prof Peter Preiser

.....
[Input Supervisor Name Here]

Authorship Attribution Statement

Please select one of the following; *delete as appropriate:

*(A) This thesis **does not** contain any materials from papers published in peer-reviewed journals or from papers accepted at conferences in which I am listed as an author.

~~*(B) This thesis contains material from [x number] paper(s) published in the following peer reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.~~

[Input Date Here]

07/06/2019

Date

[Input Signature Here]

Loh Han Ping

[Input Name Here]

Acknowledgements

I have been very fortunate to have many people helping me throughout my PhD studies in NTU and SIgN. I would like to offer my sincere gratitude to several people who have made my PhD journey such a wonderful experience. Firstly, I would like to thank my supervisors Prof. Peter Preiser and Prof. Laurent Renia for the guidance and support given to me during my PhD studies. Although both of my supervisors were very busy with their work schedule, they are always available for meet-ups if I have some pressing questions about the project that needed their advice. Prof Peter Preiser always have great plans for my project and propels me to think more about my own project. Prof Laurent Renia has always been my support with studies involving the immune system. His insights and knowledge in immunology has helped greatly in making my project goes smoothly.

I would also like to thank the members of Prof Preiser lab who had provided me with many advices and help during my stay in the lab. Dr Annie Gao is not only my neighbor in the office, she has always been my mentor and is always available to assist me when I encounter any problems in my projects. I would like to thank Dr Amy Huang and Dr Anthony Siau for teaching me the ropes on how to do mice work. Without the two of them my time working with mice would have been a disaster. I would also like to Dr Yam and Dr Regina Hoo for the many discussion sessions that we have throughout the years as these brings new insights into my project.

I would like to acknowledge the staff of the Animal Research Facility in LKC Clinical Science Building for the help given during the confusing times when the ARF is undergoing relocation.

Finally, I would like to thank my family; my parents and my sister for supporting my work and enduring the busy times when I was caught up working late in the lab and could not make it for family dinners. Your trust and constant encouragement has allowed me to persevere even when I face with setbacks in my studies.

Table of Contents

Acknowledgements	6
Table of Contents	8
List of figures	11
Abbreviations	13
Abstract	Error! Bookmark not defined.
1 Introduction	19
1.1 Overview of malaria	19
1.1.1 Malaria: the deadliest parasitic disease in the world	19
1.1.2 Symptoms of clinical malaria	20
1.3 Treatment and prevention	21
1.3.1 Vector control	22
1.3.2 Drugs	24
1.3.3 Vaccines	26
1.4 <i>Plasmodium</i> biology	35
1.4.1 The <i>Plasmodium</i> lifecycle	35
1.4.2 The <i>Plasmodium</i> ‘exportome’	37
1.5 Variant surface antigens	40
1.6 <i>Plasmodium</i> interspersed repeats	41
1.6.1 Structure of <i>pir</i> genes	42
1.6.2 Transcriptional control of <i>pirs</i>	42
1.6.3 Possible functions of <i>pirs</i>	44
1.7 Importance of antibodies against blood stage malaria	47
2 Objectives	48
Chapter 3 Materials and methods	51
Ethics statement	52
3.1 Materials	52
3.1.1 Animals	52
3.1.2 <i>Plasmodium yoelii</i> parasites	52
3.1.3 Chemicals, reagents and kits	53

3.2 Methods.....	56
3.2.1 Shortlisting of candidate proteins.....	56
3.2.2 Gene cloning	56
3.2.3 Protein expression and purification.....	57
3.2.4 Determination of protein concentration	59
3.2.5 Validation of proteins against serum from previous proteomic screen.....	60
3.2.6 Immunization regimen	60
3.2.7 ELISA	61
3.2.8 Mouse challenge and parasitemia assessment.....	61
3.2.9 Characterization of antibodies.....	62
3.2.10 Parasite transfection	63
3.2.11 Single cell cloning of transfected T2A-YIR parasites	64
3.2.12 Diagnostic PCR for checking of integrated Skip-peptide transgene.....	64
3.2.13 Isolation of trophozoite-stage total RNA from T2A-YIR transfected parasites.....	64
3.2.14 RNA sequencing and analysis.....	65
3.2.15 <i>In vivo</i> growth kinetic assay	66
3.2.16 <i>Ex vivo</i> growth kinetic assay	66
3.2.17 cDNA synthesis by reverse transcription	67
3.2.18 Quantitative Real-time PCR (qRT-PCR) Analysis for Expression of <i>yirs</i>	67
Chapter 4 Characterization of proteins identified in proteomic screen.....	69
4.1 Prior work done in the lab.....	70
4.1.1 Proteomic identification of antigens located on surface of infected RBC ...	70
4.1.2 Generation of immune sera	71
4.1.3 Characterization of antigens recognized by protective sera.....	72
4.2 Three novel proteins of unknown function were shortlisted.....	76
4.3 Generation of Halotag recombinant proteins	82
4.4 Purification of Halotag proteins	84
4.5 Generation of Histag recombinant proteins	85
4.6 Buffer optimization of Histag proteins	87
4.7 Purification of Histag proteins	89
4.7.1 Purification of MSP1-19	90
4.7.2 Purification of PY04147A and PY06306A.....	91
4.8 Recombinant proteins were recognized by protective immune sera.....	94

4.9 Immunization regime and antibody titer	95
4.10 Immunized mice produced antibodies specific to the proteins	96
4.11 Antibodies recognized full-length proteins in <i>P. yoelii</i> schizont lysate	97
4.12 Localization studies of proteins of interest in <i>P. yoelii</i>	98
4.12.1 Proteins were not located on the surface of infected erythrocytes	98
4.12.2 Proteins were located within individual merozoite	102
4.12.3 Proteins were not localized in the rhoptries of <i>P. yoelii</i> merozoites	104
4.13 Proteins did not confer protection against a lethal YM challenge	106
4.14 Discussion	110
Chapter 5 Endogenous expression of a single <i>yir</i> gene in <i>Plasmodium yoelii</i>	115
5.1 Expression of constitutively expressed <i>yir</i> gene in <i>P. yoelii</i>	116
5.1.1 Generation of <i>P. yoelii</i> constitutively expressing YIR parasite.	117
5.1.2 GFP tagged YIRs are expressed and exported	121
5.1.3 Single cell cloning and validation of clones.....	122
5.1.4 YIR-GFP is expressed in transfected parasites	123
5.1.6 <i>In vivo</i> growth phenotype.....	127
5.1.7 Growth defect in 0017400CA not due to loss of fitness of transfected cells	130
5.2 Discussion	132
Chapter 6 Evidence of co-regulation in <i>yir</i> expression	138
6.1 RNA preparation for RNA sequencing	139
6.2 Endogenous expression of a single <i>yir</i> gene results in the global downregulation of <i>yir</i> expression.....	141
6.3 Co-regulation of <i>yir</i> gene expression	143
6.4 Presence of motif in <i>yirs</i> under regulatory control.....	146
6.5 Discussion	151
7 Conclusion and future perspective	154
8 References	157
9 Appendix	168

List of figures

Figure 1: Pipeline of the malaria vaccine candidates currently in the pipeline (adapted from https://www.malariavaccine.org/projects).....	28
Figure 2: A schematic diagram depicting the CSP and the RTS/S vaccine	34
Figure 3: Lifecycle of <i>Plasmodium falciparum</i>	36
Figure 4: Schematic diagram of the organization of <i>pir</i> gene	42
Figure 5: Generation of sera for identification of antigens associated with protection in mouse malaria model	72
Figure 6: Proteomics approaches to identify antigens associated with protection	74
Figure 7: Alignment of protein sequence of PY04147 with orthologs in <i>P. falciparum</i> 3D7, <i>P. vivax</i> Sal-1 and <i>P. chabaudii</i>	79
Figure 8: Alignment of protein sequence of PY04372 with orthologs in <i>P. falciparum</i> 3D7, <i>P. vivax</i> Sal-1 and <i>P. chabaudii</i>	80
Figure 9: Alignment of protein sequence of PY06306 with orthologs in <i>P. falciparum</i> 3D7, <i>P. vivax</i> Sal-1 and <i>P. chabaudii</i>	81
Figure 10: Schematic diagram of the shortlisting process and region of peptide generation	82
Figure 11: Western blot images of Halotag induction	84
Figure 12: Coomassie Blue stained gel of different fractions obtained after Halotag purification	85
Figure 13: Recombinant protein expression in <i>E. coli</i> cells after induction	87
Figure 14: Western blot image of buffer optimization of proteins probed with anti-His antibodies	89
Figure 15: Coomassie Blue stained gels of MSP1-19 purification	91
Figure 16: Coomassie Blue stained gels of PY04372A and PY06306A purification	94
Figure 17: Recombinant proteins were probed with mice sera generated in previous screen	95
Figure 18: Immunization regime and amount of antigens injected	96
Figure 19: Western blot using polyclonal antibodies probed against <i>P. yoelii</i> extract	98
Figure 20: Antibodies generated in mice immunized by the different recombinant proteins recognized the native proteins expressed by <i>P. yoelii</i>	101
Figure 21: IFA showing co-staining of MSP1-19 and proteins of interest	103
Figure 22: Antibodies specific against the recombinant proteins were co-stained with MAEBL to check for co-localization in the rhoptries	105
Figure 23: Growth of <i>P. yoelii</i> YM parasites in mice immunized with different proteins and the corresponding survival curve	109
Figure 24: Plasmid map of skip peptide and schematic diagram of recombination	120
Figure 25: Live cell imaging of transfected PYYM_0017400 parasites at day 7 post transfection	121
Figure 26: Live cell imaging of HPXY T2A-GFP transfected parasites	122

Figure 27: Diagnostic PCR for checking of correct integrant	123
Figure 28: GFP fusion protein is expressed in transfected parasites	126
Figure 29: In vivo growth assay and mice survival after infection with wildtype YMmC or 0017400CA parasites	128
Figure 30: In vivo growth assay and mice survival after infection with wildtype YMmC or 0035300CA parasites	129
Figure 31: In vivo growth assay and mice survival after infection with wildtype YMmC or 0400700CA parasites	130
Figure 32: Ex vivo growth assay	131
Figure 33: Global yir expression is downregulated in PYYM_0017400CA cell lines	143
Figure 34: Expression of yir genes are co-regulated.....	146
Figure 35: ClustalW alignment of the last 100 amino acids of the 59 proteins detected in the RNA sequencing analysis.....	150

Abbreviations

ABSL2	Animal Biosafety Level 2
ACT	Artemisinin combination therapy
AIDS	Immunodeficiency syndrome
APS	Ammonium persulfate
BSA	Bovine serum albumin
CaNa ₂ EDTA	Calcium disodium versenate
Cas9	CRISPR-associated protein 9
CDS	Coding sequence
CFA	Complete Freund's Adjuvant
CHMI	Controlled human malaria infection
CRISPR	Clustered regularly interspaced short palindromic repeats
cRPMI	complete RPMI
DHA	Dihydroartemisinin
DMSO	Dimethyl Sulfoxide
Dpi	Days post infection
DV	Digestive vacuole
ELISA	Enzyme-linked immunosorbent assay
<i>E. coli</i>	<i>Escherichia coli</i>
EXP2	Exported protein 2
FACs	Fluorescence-activated cell sorting
gDNA	Genomic DNA
GOI	Gene of interest
HCC	Host cell cytoplasm
HSP	Heat shock protein
IACUC	Institutional Animal Care and Use Committee
InFA	Incomplete Freund's Adjuvant
i.p	Intraperitoneal
i.v	Intravenous
IPTG	Isopropyl β -D-1-thiogalactopyranoside

iRBC	Infected RBC
IRS	Indoor residual spraying
KCl	Potassium Chloride
KH ₂ PO ₄	Monopotassium phosphate
LLIN	Long-lasting insecticidal bed-nets
MSP1	Merozoite surface protein 1
NACLAR	National Advisory Committee for Laboratory Animal Research
Na ₂ HPO ₄	Disodium phosphate
ncRNA	Non-coding RNA
Ni-NTA	Nickel-Nitrilotriacetic acid
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate buffer saline
PBS-T	Phosphate buffer Saline-Tween 20
<i>P. berghei</i>	<i>Plasmodium berghei</i>
<i>P. brasilianum</i>	<i>Plasmodium brasilianum</i>
<i>P. chabaudi</i>	<i>Plasmodium chabaudi</i>
<i>p. cynomolgi</i>	<i>Plasmodium cynomolgi</i>
<i>P. falciparum</i>	<i>Plasmodium falciparum</i>
<i>P. knowlesi</i>	<i>Plasmodium knowlesi</i>
<i>P. inui</i>	<i>Plasmodium inui</i>
<i>P. malariae</i>	<i>Plasmodium malariae</i>
<i>P. ovale</i>	<i>Plasmodium ovale</i>
<i>P. schwetzi</i>	<i>Plasmodium schwetzi</i>
<i>P. simium</i>	<i>Plasmodium simium</i>
<i>P. vivax</i>	<i>Plasmodium vivax</i>
<i>P. yoelii</i>	<i>Plasmodium yoelii</i>
PFA	Paraformaldehyde
PfEMP1	<i>P. falciparum</i> erythrocyte membrane protein 1
<i>PfSPZ</i>	<i>P. falciparum</i> sporozoite
pI	Isoelectric point
PIESPs	Parasite Infected Erythrocyte Surface Proteins

PIRs	Plasmodium interspersed repeats
PLASMED	<i>Plasmodium</i> Lipophilic and Secondary structure Mediated Export Domain
PNEPs	PEXEL negative exported proteins
PTEX	<i>Plasmodium</i> Translocon of EXported proteins
RBC	Red blood cell
RIFINs	repetitive interspersed families of polypeptides
RMP	Rodent malaria parasite
RPMI	Roswell Park Memorial Institute
SDS	Sodium dodecyl sulfate
STEVOR	Subtelomeric variable open reading frame
TEMED	Tetramethylethylenediamine
Tris-HCL	TRIS Hydrochloride
TRX2	Thioredoxin 2
WHO	World Health Organization

Summary

Malaria is caused by the protozoan genus *Plasmodium* and infects about 200 – 300 million people annually. Due to the rise in drug resistance in malaria parasites, it is of utmost importance to develop an effective vaccine that can confer sterile protection in humans. Although extensive efforts to develop vaccines are underway, current vaccine strategy based on malaria antigens failed to confer substantial protection in humans.

The failure of current vaccine candidates to confer protection implies that the antigens associated with protection are still unknown. In order to try and address this problem, my project focuses on two approaches to study antigens and their impact on the immune system. The first is an untargeted approach method in which immune sera were screened to identify antigens which are highly associated with protection. The second method is a targeted approach focusing on the studying of the *Plasmodium* interspersed repeat (*pir*) multigene family using the rodent malaria parasite (RMP) *Plasmodium yoelii* as a model organism to understand the impact these proteins play in immunity. The *pir* multigene family was chosen for the targeted approach due to the variant nature of these protein which suggested that these proteins might be exposed to the immune system and therefore are under immune pressure.

From the untargeted approach, we intend to elucidate the mechanism of protective immunity by using a malaria mouse model to identify potential parasite antigens that can confer sterile protection in mice. A prior study in our lab had generated a library of parasite antigens that are recognized by the sera

from sterile protected mice. Based on this data, I have shortlisted three novel proteins which are highly associated with protection. I show that these three proteins are found within the parasite throughout the erythrocytic lifecycle and are present in the free merozoites as well. However, none of these three proteins could protect mice from the lethal *P. yoelii* YM infection. Interestingly many of the hits in the screen are parasite proteins found within the merozoites, suggesting that immunity against *P. yoelii* infection is directed towards the merozoite stages rather than the erythrocytic stage.

For the second part of the project, I aimed to study the *Plasmodium* interspersed repeat (PIR) multigene family and its role in malaria infection. The *pir* multigene family is the largest multigene family in *Plasmodium* however research on these proteins is hampered by the inability to study these genes in isolation. Using the rodent malaria parasite *P. yoelii* orthologs known as *yir*, I aim to constitutively activate a single *yir* gene to study its functional role. In this study, I had successfully adopted a skip peptide approach to constitutively activate a single member of the *yir* gene within the parasite.

RNAseq data showed that activation of a single *yir* caused a global silencing of the majority of *yirs* within the parasite population. I also observed that the activation of a single *yir* variant resulted in the upregulation of 4 additional *yirs* in the population. Analysis of the YIR protein sequence revealed that the YIRs which are co-regulated contain a PLASMED motif. The exact mechanism of how these PLASMED containing YIRs are co-regulation is currently unclear;

however, my results suggest that there are subsets of *yirs* in the parasite which are regulated independently.

1 Introduction

1.1 Overview of malaria

1.1.1 Malaria: the deadliest parasitic disease in the world

Malaria is a potentially life-threatening disease that is caused by parasites of the genus *Plasmodium*. Together with acquired immunodeficiency syndrome (AIDS) and tuberculosis, malaria remains one of the top 3 infectious diseases that exact a heavy toll on human life and health. According to the World Health Organization Malaria report 2017, it is estimated that approximately 216 million new cases of malaria occurred in 2016 that resulted in more than 445 thousand deaths. Children under the age of five are especially susceptible to malaria infection and accounted for more than 70% of malaria mortality. The number of malaria cases rose from 210 million in 2015 to 216 million in 2016. Although this rise is small, it implies that the significant reduction of malaria cases of 37% from 2000 to 2015 has slowed and is being reversed.

Malaria affects the wealth and wellbeing of individuals as well as the nation as a whole. In Africa, malaria is recognized as a disease of poverty and a cause of poverty with an estimated 40% of public health expenditure going into combatting or treating the disease (WHO, 2011). Six parasites species; *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*, *P. knowlesi* and the recently reported *P. cynomolgi* [1], have been reported to cause malaria in humans. Of these 6 species, *P. falciparum* causes the most severe form of malaria and accounts for approximately 95% of all malaria related deaths.

Non-human primate malaria parasites such as *P. knowlesi* and *P. cynomolgi* have been reported to be transmitted to humans living near forests where the natural host of the parasites live. Naturally acquired human infection of *P. knowlesi* were initially thought to be very rare until a large cohort study in Malaysia in 2004 revealed that more than half of the infections were in fact caused by *P. knowlesi* [2]. Other species of non-human primate malaria that can potentially be transmitted to humans are; *P. brasilianum* [3], *P. inui*, *P. schwetzi* [4] and *P. simium* [5].

1.1.2 Symptoms of clinical malaria

Early symptoms of malaria includes fever, chills, headache and vomiting and can be wrongly diagnosed as gastrointestinal infection [6]. In *P. falciparum* malaria, the onset of fever usually occurs towards the end of the incubation period (approximately 9 – 15 days) but a longer incubation period is also possible [7]. In non-immune individuals, malaria can progress rapidly to severe malaria unless appropriate treatment is started on time [6]. Severe malaria is associated with *P. falciparum* infections and includes cerebral malaria, acute renal failure, respiratory failure and anemia. As the progression to severe malaria can be rapid, malaria patients must be treated as soon as possible to promote survival.

P. ovale and *P. vivax* both cause periodic fever after the incubation period. The incubation time for *P. vivax* is 14 days while *P. ovale* is about 12 – 20 days. Although the disease caused by these two species is considered as mild and usually does not cause severe complications, there is evidence that showed that they can indeed cause severe complications despite the fact that overall parasite

burden is limited by the fact that these parasite species are restricted to reticulocytes. [8, 9]. *P. vivax* is known to form a quiescence uninucleate form in the liver known as hypnozoite, and these can result in relapses of the disease months or years after the initial infection [10, 11]. Although *P. ovale* is also a recurring malaria species, the existence of hypnozoites has never been biologically proven, and questions remained whether hypnozoites are really present in this species [12].

Depending on the strain of *P. malariae*, the incubation time can vary from 16 – 59 days [13, 14]. *P. malariae* does not cause relapses from the generation of hypnozoites but the blood stage can persist in the host for up to 70 years [15]. Although infection with *P. malariae* is usually not severe, it can in rare cases lead to edema and renal diseases [16].

1.3 Strategies to control malaria

Encouraged by the initial success of using dichloro-diphenyl-trichloroethane(DDT) against mosquitoes and chloroquine against malaria parasites, the WHO sets out a Global malaria eradication program(GMEP) in 1955. This plan was however short-lived and was largely abandoned 14 years later due to the realization that eradication in many areas was not possible with the current means available. The rise of resistance of mosquitoes to DDT and other insecticides and the emergence of drug resistant parasites was and remains a serious problem and malaria rates remain high in sub-Saharan Africa and South-East Asia. Due to the fast-evolving nature of the parasites, a combination of intervention strategies is thought to be the most effective way to control malaria.

Current methods to control malaria including vector control, drugs and vaccines, will be discussed below.

1.3.1 Vector control

The widespread use of DDT is the main reason for the reduction of malaria in endemic region from 1940 – 1970. One of the properties that makes DDT such an ideal agent for use against malaria is that it persists for a long time in the environment and thus does not need frequent application. However, this also mean that the hydrophobic DDT compound accumulated in the fatty tissues of animals and is only eliminated very slowly, resulting in the accumulation of DDT in predators at the top of the food chain including humans [17]. The half-life of dichlorodiphenyldichloroethylene (DDE), a major metabolite of DDT in the human body was estimated to be around 11 years [18].

The widespread use of DDT from 1940 – 1970 resulted in adverse impact on the environments, including mortality of birds due to DDT poisoning [19]. Furthermore, there was a detrimental effect on the reproductive system of birds in which egg shell thinning was observed in several species [20]. In contrast to the indiscriminate use of DDT in agricultural settings, the controlled use of DDT in indoor residual spraying (IRS) has minimal impact on the environment. For IRS, DDT is sprayed onto walls and surfaces in the house where the female *Anopheles* mosquitoes might land before or after a blood meal. In addition to killing mosquitoes that land on sprayed surfaces, DDT also has a excitorepellant effect, repelling mosquitoes from the sprayed vicinity [21]. Due to the

effectiveness of DDT in IRS for malaria vector control, DDT is still recommended by WHO for use in some African countries [22].

Other than IRS, the use of long-lasting insecticidal bed-nets (LLINs) has also contributed to the decline in malaria incidences. It is estimated that about 78% of malaria cases averted globally from 2000 to 2015 is due to the use of IRS and LLINs [23]. The use of LLINs is effective in Africa where the vector is mainly the *Anopheles gambiae* complex and *Anopheles funestus* and these mosquitoes feed at night. However, a recent study in Haiti found that LLINs are not as effective when the vector is different and thus the feeding habits are different as well [24]. The main issue with the use of LLINs is the rise of resistance to a class of insecticide known as Pyrethroid which are used in these bed nets [25]. To minimize the threat of LLINs becoming less effective, bi-treated bed nets or concurrent usage of IRS should be encouraged in endemic areas.

Other than the use of insecticides, genetic modification of the vectors is also increasingly being explored to complement the current vector control strategies. With the recent development in gene editing technologies such as clustered regularly interspaced short palindromic repeats/ CRISPR-associated protein 9 (CRISPR/Cas9) systems, it has become possible to perform gene editing on *Anopheles* mosquitoes. One such example is the CRISPR knockout of the fibrinogen-related protein 1(FREP1) gene in *Anopheles gambiae* which resulted in the mosquitoes being more resistant to infection by malaria parasites [26]. Another genetic approach is the alteration of the bacteria found normally in the mosquito to combat the parasites. By modifying a *Serratia* bacterium strain (AS1)

to secrete anti-plasmodial proteins, Wang et al. successfully inhibited the development of plasmodium parasites in these mosquitoes [27]. Although these genetic approaches look attractive, they have so far only been tested under laboratory settings and any environmental impact must be thoroughly evaluated before these genetically modified mosquitoes or bacteria can be released into the environment.

1.3.2 Drugs

Without an effective malaria vaccine, drugs remain one of the best treatment methods available to combat and control malaria. There are many drugs that are available to treat malaria and they are divided into 3 main classes of compounds: the quinolines, the antifolates and the artemisinin derivatives. However, for this thesis I will focus only on the use of the first antimalarial compound and its most important derivative; chloroquine, and artemisinin which is currently the first-line drug against malaria.

Quinine was the first compound with anti-malarial properties and was discovered in the 17th Century. The story of the discovery of quinine differs in South America and Europe, which involves either native Americans or the Spanish countess of Chinchon respectively [28]. Before 1820, quinine was extracted by grounding the dried bark of the cinchona tree into a fine powder, mixing it with a liquid and ingesting the mixture. In 1820, quinine was successfully extracted and purified from the bark of the cinchona tree and eventually replaced it as the standard treatment for malaria.

With rapid advances in synthetic drug discovery in the 1930s, chloroquine was synthesized and widely used against malaria infections around the world. Chloroquine was the first drug produced on a large scale for treatment and prevention of malaria infection. Chloroquine has activity against the blood stages of susceptible strains of *P. vivax* and *P. falciparum*, *P. malariae* and *P. ovale* [29]. The mode of action of chloroquine is not fully understood but is thought to be that the unprotonated form enters the digestive vacuole (DV) membrane and gets protonated within the DV [30]. The protonated form of chloroquine cannot pass through the membrane of the DV and accumulates in the DV of the iRBC and interferes with the detoxification mechanism, impairing the formation of hemozoin [31, 32].

The wide spread use of chloroquine started declining in the beginning of 1960 with the rise of resistance in *P. falciparum*. Despite the declining use, chloroquine remained in use in sub-Saharan regions for uncomplicated malaria until after 2000 [33]. Resistance against Chloroquine has been linked to multiple point mutations, particularly the K76T mutation in the *P. falciparum* chloroquine resistance transporter (*PfCRT*) [34]. This mutation resulted in a removal of a positive charge from the predicted substrate binding domain of *PfCRT*, thus allowing the protonated chloroquine to be transported out of the DV [35]. To date resistance against chloroquine has spread to many parts of the world where *falciparum* malaria is endemic [36].

Artemisinin combination therapy (ACT) is currently the first-line drug recommended by WHO for the treatment of *P. falciparum* infections. ACT is a

formulation comprising of the potent and fast acting artemisinin with other slower acting and longer half-life partner drugs such as mefloquine or lumefantrine and is taken orally for 3 days. The main reason for using two different drugs is to reduce the emergence of drug resistance and in areas that ACT is effective, the incidence of malaria related mortalities had dropped significantly [37].

Even with the use of ACT, resistance to artemisinin started to appear and was first reported in Cambodia in 2008 [38]. Since then, artemisinin resistant parasites have spread to other countries in the greater Mekong region [39]. The molecular basis of artemisinin resistance is linked to mutations in the kelch propeller region (K13) and K13 mutations have been used as a molecular marker for artemisinin resistance in *P. falciparum* [40]. If these artemisinin resistant parasites acquired resistance to the partner drugs, it can pose a threat to the control of elimination of malaria in this region and there is evidence that ACT failure is on the rise. The appearance of artemisinin resistance now poses a significant threat to the partner drug and resistance to K13 mutations in combination with two other mutations have been shown to be implicated in the declining efficacy of a 3-day dihydroartemisinin (DHA)-piperaquine treatment in Cambodia [41].

1.3.3 Vaccines

For the past 200 years, vaccines which protect against infectious diseases have saved many lives and resulted in the eradication of one of the deadliest diseases in human history, smallpox. Due to the ease of application and their effectiveness, vaccines are considered to be the most cost efficient and successful way to combat diseases. Malaria is a major health burden particularly in Africa thus an

effective vaccine against malaria would be a major advance in public health and is highly sought after by scientists around the world.

However, vaccines against parasites are significantly harder to develop due to the genetic complexity of the parasite (~ 5400 genes in *P. falciparum*). In fact, there was no vaccine available on the market to prevent parasitic diseases in humans until the approval of RTS,S-AS01 in 2015. The rationale that a malaria vaccine can be successfully developed is supported by the observation that people living in endemic areas developed immunity to malaria over time although this protection is short lived and not sterile. Malaria vaccine research has resulted in several potential candidates in the development pipeline, and the current progress of all these vaccine candidates are shown in Figure 1. Research on malaria vaccine can be broadly divided into three broad categories representing the distinct life cycle of the parasite within the human host. Pre-erythrocytic vaccines aim to prevent blood stage malaria by targeting the clinically silent liver stage. Blood stage vaccines aim to disrupt the parasites growth in the blood and transmission blocking vaccines which prevents mosquitoes from carrying the parasites from spreading them. Malaria vaccine design is based on two approaches; the whole parasite vaccine approach and the subunit vaccine approach.

Figure 1: Pipeline of the malaria vaccine candidates currently in the pipeline (adapted from <https://www.malariavaccine.org/projects>)

1.3.3.1 Whole parasite vaccine

Whole parasite vaccination is based on the concept of delivering attenuated parasites to induce protection against the sporozoites and liver stages thereby averting the pathogenic blood stage infection. Many strategies which attempt to arrest the infection before, during or shortly after the liver development stage had been developed. The 3 main strategies are Radiation Attenuated Sporozoites

(RAS), Genetically Attenuated Sporozoites (GAS) and Chemical Prophylaxis Sporozoites (CPS).

1.3.3.1.1 Radiation attenuated sporozoites (RAS) approach

Early works in the 1970s showed that sterile protection can be achieved with the RAS approach [42, 43]. Since then multiple studies have validated the high reproducibility of this vaccination model. Hoffman et al demonstrated that using the radiation attenuated sporozoites to inoculate volunteers had shown that a very high levels (>90%) of protective efficacy can be induced [44, 45]. However, to achieve such protective effects, the volunteers in the study were exposed to bites of >1000 irradiated mosquitoes. This requirement for >1000 mosquito bites to induce high level of efficacy is not practical for public health use thus there is a need to advance beyond the use of mosquitoes for administration of the sporozoites. Over the years, advancements in technology enable sporozoites to be extracted from the salivary glands aseptically, radiation attenuated, purified and cryopreserved. With this new approach, the Biotech company Sanaria reported the first malaria vaccine called the *Plasmodium falciparum* sporozoite (PfSPZ) vaccine that can achieve 100% protection in humans by using parasites isolated from mosquitoes [46]. In this study, all six individuals who received five doses of the vaccine intravenously were protected against a controlled human malaria infection (CHMI). One important point to note for this CHMI is that the challenge was done using homologous parasites. In addition, the number of volunteers in the study was small (n=6) therefore it might not be representative of the outcome when used on a larger scale.

A follow-up trial in 2017 focused on the question of a heterologous parasite challenge and reduced number of vaccine doses [47]. This study enrolled 31 healthy volunteers with the volunteers allocated in the ‘vaccine group’ administered with 3 doses of *PfSPZ* vaccine at eight weeks interval. 19 weeks after the final immunization, 9 out of 14 (64%) of the vaccinated volunteers showed no sign of malaria after CHMI with homologous parasites. To evaluate the heterologous protective effect of the vaccine, 6 of the protected volunteers were exposed to mosquitoes infected with another strain of *P. falciparum* 33 weeks after the final immunization. In this group, 5 out of 6 (83%) of the volunteers were protected against malaria infection. The researchers found that the vaccine activated CD8⁺ T cell response which is strain-transcending and induced antibody responses in all vaccinated individuals.

Despite the success of *PfSPZ* vaccine, the widespread use of the vaccine is currently hampered by challenges in the distribution and administration of the vaccine. The need for up to three intravenous injection of the vaccine over a period of a few months in order to achieve protection might be difficult due to low compliance of patients. Another hurdle to overcome is that the vaccine is currently dissected manually in a sterile laboratory setting but this can be solved if there is a way to increase production using robotics. Another challenge is that the vaccine must be kept cryopreserved after extraction, a situation that might be hard to fulfil in some endemic areas such as Africa. Efficacy of the vaccine in clinical trials will need to be assessed to determine if such laborious method is feasible for large scale production of the vaccine.

1.3.3.1.2 Genetically attenuated sporozoites (GAP) approach

GAS approaches aim to genetically modify the malaria parasites such that it cannot complete its development in the liver stage. This strategy overcomes the need of fine tuning as part of the cGMP process. A *P. yoelii* GAS model that features a deletion of P52 and P36 showed complete liver arrest and confers sterile protection [48]. This model was adapted in a human challenge study in which 6 volunteers were administered the (P52-/P36-) GAS [49]. However, after immunization with GAP from 200 mosquito bites, 1 volunteer developed clinical malaria. Analysis of the parasites revealed that the parasites still retains the (P52-/P36-) deletions showing that these genetically altered parasites can somehow complete the liver stage when introduced at high doses.

The application of GAS method relies on the disruption of the development of this parasite during the liver stage and this must be complete as otherwise even a single infected hepatocyte can cause a full-blown blood stage infection. More research must be conducted in selecting the gene to be deleted to prevent breakthrough blood-stage infections.

1.3.1.1.3 Chemical Prophylaxis Sporozoites (CPS) approach

Recent experiments have shown that injection of wildtype sporozoites under the simultaneous application of a chemical prophylaxis is a good way to induce protective immunity. Several drugs were tested using this approach, but chloroquine is the most widely used. Chloroquine kills the blood stage parasites but has no effect on the liver stages. Parasites in the CPS approach are able to

complete liver stage development but are rapidly killed upon released into the blood stream.

Initial experiments on mice injected with whole sporozoites under the cover of chloroquine treatment lead to high level of protection against subsequent sporozoite challenges [50]. Similar studies on humans have shown that CPS can also be used on humans with similar efficiencies [51]. Interestingly in this study, when challenged with a different *P. falciparum* strain one year after CPS, the majority of the volunteers were still protected against the heterologous challenge. This implies that CPS can induce protection lasting for more than a year which is important for the development of whole sporozoite vaccines. It was found that the development of the parasite in the liver is necessary for the protective responses observed as addition of primaquine (a drug that kills the intra hepatic stages) reduces the protective efficacy induced by CPS.

Other than chloroquine, many other anti-malarial drugs have also been studied. A recent study using mefloquine in CPS demonstrated that its efficacy is comparable to that when chloroquine was used [52].

1.3.3.2 Subunit vaccine

Whole parasite vaccine strategies deliver a wide range of antigens and acts like a multiepitope vaccine, leading to high level of protection. In contrast, none of the subunit vaccines tested so far have been able to induce the high level of protective immunity seen in humans vaccinated with whole parasites. Although whole parasite vaccines can induce high levels of protection, the need for precise

irradiation and maintenance of consistent quality for cGMP makes them a challenge for large scale production.

Considering these difficulties, most of the efforts in malaria vaccine research is focused on the development of subunit vaccines. Subunit vaccines have the advantage of being cheaper to produce, easier to store and administer as compared to whole parasite vaccines. A number of vaccine candidates have progressed to clinical trials but none has yet achieved sufficient protection against malaria infection. The most effective malaria subunit vaccine to date is the RTS,S/AS01 which is a hybrid protein particle formulated in a multi-component adjuvant named AS01.

1.3.3.2.1 RTS,S/AS01

RTS,S/AS01 is a malaria subunit vaccine targeting the pre-erythrocytic stages and is currently the most advanced vaccine candidate in the development pipeline. As of 2015, it is the only candidate that is licensed for production as a malaria vaccine by the European Medicines Agency (EMA). RTS,S development is based on the observation that circumsporozoite protein (CSP) alone is able to protect against malaria [53]. CSP is a predominant surface antigen on the sporozoites and is expressed during the sporozoite and liver stages. The vaccine was engineered from the central repeat region and the T-cell epitope of the CSP and fused to hepatitis B surface antigen (HBsAg) [54]. The fusion protein when mixed with native HBsAg will self-assemble into a virus-like particle exposing the fusion protein on the surface. A schematic diagram of the RTS,S vaccine is shown in Figure 2. The AS01 adjuvant system for RTS,S is liposome-

based, comprising of 3-O-desacyl-4'-monophosphoryl lipid A, the saponin QS-21 and a toll-like receptor 4 ligand [55] .

Figure 2: A schematic diagram depicting the CSP and the RTS/S vaccine[56]. The RTS,S vaccine contains the last 16 NANP repeats of the CSP proteins and the entire C terminus.

Despite being the most advanced vaccine candidate in the pipeline, the efficacy of the vaccine in clinical trials are disappointing. In phase 2a trial, RTS,S protected ~50% of subjects against CHMI 2 to 3 weeks after the last immunization, but the protection drops to 22% 5 months after the last immunization [57]. Phase 3 trial conducted on African infants 6 to 12 weeks of age showed that RTS,S/AS01 is able to reduce rates of clinical and severe malaria acquired over a 12 month period by 31.3% and 36.6% respectively [58]. High levels of anti-CSP IgG and CD4⁺ T cell responses were detected in individuals protected after immunization, suggesting a role for both humoral and cell-mediated immunity [57].

1.4 *Plasmodium* biology

1.4.1 The *Plasmodium* lifecycle

The *Plasmodium* parasite has a complex life cycle which involves both a mosquito vector and a vertebrate host (Figure 3). The asexual cycle in humans begins with an infected *Anopheles* mosquito depositing sporozoites under the skin of the host during a blood meal. These sporozoites transverse the skin and enter the blood vessels which allows them to migrate to the liver. In the liver the sporozoites invade hepatocytes and undergo several rounds of asexual replication to form thousands of liver merozoites. Work using *P. yoelii* showed that the merozoites then bud from the infected hepatocyte as merozoites and subsequently burst in the lung capillaries to invade circulating erythrocytes [59]. Within the erythrocyte, the parasite developed through distinct phases of its asexual lifecycle namely the ring, trophozoite and schizont stages. The schizont stage is characterized by numerous rounds of cell divisions, giving rise to new merozoites. Egress of the schizonts releases these merozoites into the blood stream to infect new erythrocytes, thus continuing the intra-erythrocytic cycle. In *P. falciparum*, one intra-erythrocytic cycle takes about 48 hours and a ruptured schizont releases about 16 to 32 merozoites into the blood stream. Clinical manifestation of malaria is caused by this 48-hour cyclic proliferation of asexual parasites within the RBC.

While most of the parasites undergo asexual development in the host, some parasites developed into sexual forms known as gametocytes. These gametocytes circulate in the peripheral blood and allow for the transmission of parasites to mosquitoes during subsequent bites. When the mosquito picks up these

gametocytes, fertilization occurs within the mosquito midgut and result in the development of ookinetes. These ookinetes then transverse the midgut epithelial cells to the basal lamina where they develop into oocysts. Matured oocyst releases sporozoites which travel to the salivary glands of the mosquito and from there are injected into a human host when the mosquito feeds, thereby continuing the cycle.

Figure 3: Lifecycle of *Plasmodium falciparum*. Adapted from Sauerwein et al, 2011 [60]

1.4.2 The Plasmodium ‘exportome’

Plasmodium parasites have evolved to grow in terminally differentiated erythrocytes, which are anucleated and have very little biosynthetic capacity. Due to the lack of biosynthetic capacity, many of the nutrients needed for parasite development need to be imported into the host from the plasma. To do this, the parasite exports hundreds of proteins to remodel the host cell. These includes the addition of channels which permits the uptake of nutrients and removal of metabolic waste [61, 62].

While growing within the erythrocytes, the parasite must be able to avoid detection and clearance by the host immune system. One major site of clearance of infected erythrocytes is the spleen and export of PfEMP1 onto the surface of erythrocytes assist in cytoadherence and avoids splenic clearance of the parasites [63]. The sequestration of parasites in vascular endothelial cells leads to the pathogenesis of severe malaria such as cerebral and placental malaria [64, 65].

Due to the importance of these exported proteins in parasite survival and virulence, they have become the focus of scientific research. The importance of these exported proteins was highlighted when a study into these exported proteins revealed that at least 25% of these proteins are essential to the parasite for survival, thus making them valuable targets for drug intervention [66].

1.4.2.1 Targeting motif and signals in malaria export

Proteins that are exported and enter the secretory pathway must cross the parasite plasma membrane (PPM) and the parasitophorous vacuole membrane (PVM) to reach the host cell cytoplasm (HCC) and thus require the presence of special

export signals. A hydrophobic transmembrane domain usually located at the NH₂-terminus of exported proteins serves as a signal peptide for the insertion into the endoplasmic reticulum (ER) and is required for trafficking proteins into the parasitophorous vacuole (PV). However, to cross the PVM extra signals are required and most exported proteins contain a simple motif known as Plasmodium export element (PEXEL) or host targeting (HT) signal, with the consensus sequence RxLxE/Q/D that is located approximately 20 amino acids downstream of the signal peptide [67, 68]. The discovery of PEXEL has broadly expanded the knowledge on how protein export is carried out in *Plasmodium* and was used to predict the exportome of *P. falciparum* [67]. The presence of these targeting motifs also suggested that these exported proteins must pass through a central portal to enter the host cell cytoplasm.

Plasmodium Translocon of EXported proteins (PTEX) complex is the nexus of protein export located on the PVM and is present in all mammalian *Plasmodium* parasites [69]. The complex is comprised of essential core components; Heat shock protein 101 (HSP101), which is a ClpB-like AAA+ ATPase component, PTEX150 and exported protein 2 (EXP2) [70, 71]. In *P. berghei*, two other proteins HSP88 and thioredoxin 2 (TRX2) were shown to be non-essential and possibly fulfil auxiliary roles in the PTEX complex [70, 71].

Although the actual mechanism of how PEXEL mediates protein export is still unknown, it has been reported that the PEXEL motif is able to bind to phosphatidylinositol-3-phosphate (PI3P) in the endoplasmic reticulum (ER) [72] before being cleaved by plasmepsin V in the ER to reveal the export signal

(xE/Q/D) at the N terminus of cargo proteins [73, 74]. Both the binding of PI3P to PEXEL and subsequent cleavage by plasmepsin V is thought to be determining factors for protein export but the actual export mechanism is still unknown. Currently, it is still unclear whether the last residue of the PEXEL motif that remains after plasmepsin V cleavage is important for specificity as there is no apparent export information after the residue.

The presence of PEXEL negative exported proteins (PNEPs) suggested that the predicted exportome identified based on the presence of PEXEL is incomplete. Two such proteins, Skeletal binding protein 1 (SBP1) [75] and membrane associated histidine rich protein (MAHRP) [76, 77] have been identified and are both localized to the Maurer's cleft. Another PNEP is PfEMP1 which is trafficked into the erythrocyte and subsequently displayed on the erythrocyte surface, where it binds to host surface receptors on the microvasculature leading to malaria associated complications [78]. PNEPs are not cleaved by plasmepsin V but are exported into the host cell cytoplasm through PTEX [79, 80].

The use of PEXEL for protein export is not unique in *P. falciparum*, these motifs are also found in other malaria species infecting humans *P. vivax* and *P. knowlesi*, and the rodent malaria species *P. chabaudi*, *P. berghei*, and *P. yoelii*. However in other *Plasmodium* species, there seems to be a significantly smaller proportion of exported proteins which are PEXEL proteins [81]. This suggests that PNEPs play a bigger role in protein export in these parasites. Members of the Plasmodium interspersed repeats (PIR) superfamily lacks the classical PEXEL motif, with the exception of repetitive interspersed families of polypeptides

(RIFINs) and subtelomeric variable open reading frame (STEVAR) in *P. falciparum* and a subset of VIRs in *P. vivax* [67, 82, 83]. Through sequence alignment and secondary structure prediction a new non-PEXEL signal termed PLASMED was identified [84]. The characterization of PLASMED suggested that the secondary structure of the protein might play a role in export however, there is no data of secondary structure mediated export in other *Plasmodium* species as of now and it seems that PLASMED is unique to rodent malaria parasites.

1.5 Variant surface antigens

One of the major hurdles in the development of a malaria vaccine is the variable nature of infected erythrocyte surface proteins. These variant surface antigens (VSAs) are located on the surface of the parasites or infected erythrocytes and are encoded by large multigene families. Multigene families are a group of proteins that are thought to share a common ancestor and are derived from each other by gene duplication and divergence. These genes often sit in the subtelomeric region of chromosomes and are constantly undergoing recombination thus leading to high sequence variability [85]. Most of our understanding of the function of VSA is based on the large number of studies carried out on PfEMP1 of *Plasmodium falciparum*, with only in the last few years some new data starting to shed light on the important biological roles of the STEVAR and RIFIN multigene family of this parasite. Unlike PfEMP1 which is unique to *P. falciparum*, small variant surface antigens (sVSA), broadly ranging in size from about 30 – 40 kDa, are found in all *Plasmodium spp* [86-89]. They include *stevor*

and *rif* along with the *pir* (*Plasmodium* interspersed repeats) families that are found in all *plasmodium* spp. including the human parasite *P. vivax* and the zoonotic simian parasite *P. knowlesi* along with the rodent parasites *P. chabaudi*, *P. berghei* and *P. yoelii*. Proteins encoded by the *P. falciparum* *stevor* and *rifin* multigene families are thought to be distantly related to the *pir* multigene family of rodent and other non-human primate and human malaria parasites [88, 90]. The *var* gene family in *P. falciparum* is by far the most well studied VSAs and has been shown to be involved in sequestration and immune evasion. In this report we will be focusing on the largest multigene family known as *Plasmodium* interspersed repeats, which is found in all *Plasmodium* species affecting humans, primates and rodents [88].

1.6 *Plasmodium* interspersed repeats

The *pir* multigene family is present in all *Plasmodium* genomes sequenced so far and is the largest multigene family in *Plasmodium*. Members of this family include *stevor* (~35) and *rif* (~200) in *P. falciparum*; *vir* (~350), *kir* (~68), *cyir* (~256) in human/ primate malaria *P. vivax*, *P. knowlesi* and *P. cynomolgi* respectively; and *yir* (~800), *cir* (~200), *bir* (~180) in rodent malaria *P. yoelii*, *P. chabaudi* and *P. berghei* respectively [88, 90, 91]. Despite being discovered more than a decade ago, not much is known about the role *pirs* play in parasite biology. Being the largest multigene family in *Plasmodium*, the large proportion of genome dedicated to *pirs* supports the importance of these proteins in the biology of the parasites.

1.6.1 Structure of *pir* genes

Most *pir* genes share a similar structure: a short first exon, long second exon containing conserved cysteine residues, a predicted trans-membrane domain, and a highly conserved final exon [92] (Figure 4). Most members of the PIR family are PEXEL negative proteins, with the exception of RIFINs and STEVOR in *P. falciparum* and YIR D in *P. vivax* [82], and are arranged and transcribed in a centromere to telomere direction [93]. *In-silico* analysis of the large *yir* family in *P. yoelii* showed that approximately 56% of the *yirs* can be classified into five distinct phylogenetic groups while the rest are too diverse to be assigned into any groups [94]. The five groups of *yir* genes differs in their chromosomal location, presence of signal peptide and alternative splicing events.

Figure 4: Schematic diagram of the organization of *pir* gene. Representation of a typical *pir* genes showing a 3-exon structure. Transmembrane regions are shown in hatched box. (Image adapted from Portillo et al. [92])

1.6.2 Transcriptional control of *pirs*

Studies using microarray on *P. yoelii yir* genes showed that almost 50% of the *yir* genes are transcribed in low amounts during an infection [95]. The transcription of *yirs* has also been found to switch rapidly during infection, with cloned daughter cells expressing a different *yir* repertoire from the parent as early

as 10 cycles *in vivo* [95]. Similar results were also obtained from *P. chabaudii* and *P. vivax* in which 40% and 59% of the *pir* gene repertoire was expressed in the parasite population during parasite infection respectively [96, 97]. Early studies on *P. vivax* showed that unlike *var* genes which only express one variant at any one time, the *vir* genes are not clonally expressed which means that a single parasite clone can express more than one *vir* at a time [98]. Similar results were found for *yirs* in which a single parasite expresses only 1 – 3 *yir* genes, suggesting that the transcription of *yirs* are under tight regulation in individual parasite [95]. In addition, antigenic switching has also been described in laboratory previously for the *cir* genes in *P. chabaudii* [99]. Taken together, these results suggest that the *pir* genes are under tight regulation even though they are the largest multigene family in the *Plasmodium* species, however their precise role in the parasite biology is still unknown.

Differential expression of *pirs* have been shown to be involved in modulating the virulence of *Plasmodium* infection. Serial passage of parasites in rodents had been demonstrated to cause an increase in the virulence of the parasites [100, 101] and passage of the infected blood through mosquito ‘resets’ the parasite to its original characteristics [102, 103]. Spence et al demonstrated that vector transmission of serially passaged *P. chabaudii* infected blood modified the expression of approximately 10% of the genome with majority of the regulated genes coding for exported proteins [104]. Of these regulated genes, the *cir* multigene family was found to be intensely regulated with 57% of the *cir* genes upregulated after mosquito passage. They showed that serial passaging resulted in the expression of a dominant *cir* within the population and passaging through

mosquitoes removes the *cir* dominance and instead promotes the general upregulation of other *cirs* within the parasite population. The immune profile in mice infected with mosquito passaged blood has enhanced expansion of activated CD8 α^+ and CD8 α^- dendritic cells which in turn increased the magnitude of the effector CD4 $^+$ T cell response. This suggests that increasing the repertoire of *cirs* within the parasite population resulted in a stronger host immune response. Their results clearly showed that mosquito transmission can alter gene expression of *pirs* which in turn modifies the host immune response to the infection.

1.6.3 Possible functions of *pirs*

Although the function of *pir* genes are not well defined, it has been previously proposed that different subsets within the *pir* multigene family exhibit different functions. For example, in *P. yoelii* and *P. berghei*, a stage specific role was assumed since not all *pir* genes are transcribed equally when different life cycle stages were compared [105, 106]. In *P. chabaudii*, the in silico classification of *cirs* into subfamily A and B resulted in the identification of a group of long *cirs* which are characterized by the large insertions and additional transmembrane domains, which might also be interpreted as hints for distinct functions of these proteins [107].

The observation that expression of *pirs* is modulated by the immune system [105], coupled with their localization on iRBC surface, support their role as an immune evasion mechanism employed by the parasites. This was also demonstrated in *P. vivax* where *virs* are found to be differentially transcribed in different patient samples indicating possible roles in immune evasion [97]. Microarray analysis of

P. yoelii infection [95] and *P. vivax ex vivo* culture [97] showed that a large repertoire of *pir* genes can be expressed in a population of parasites. The presence of many different PIRs in a population allows for variants to be maintained, thereby acting as a form of immune evasion by overwhelming the immune system with many targets for antibody generation. For this to be effective, the antibodies generated by the body against these *pir* repertoire must not be cross-reactive. The finding that the specificity of antibodies generated against different VIR subsets reacts specifically with a certain population of vivax infected reticulocytes from the same patient somewhat supports that little cross-reactivity is generated against non-specific VIR.

Although *P. yoelii* do not sequester significantly and all forms of the parasites can be detected in peripheral blood, *P. chabaudi* has been shown to both sequester on endothelial cells and form rosettes with uninfected RBCs [108]. Although *P. berghei* is also known to sequester, deletion studies of BIRs ruled out the role of BIRs in mediating sequestration [109]. In the case of *P. vivax*, expression of a subset of *virs* in transgenic vir-expressing *P. falciparum* lines has shown that some variants of VIRs can mediate binding to ICAM-1 on endothelial cells [110]. In *P. vivax* infected patients, *virs* have been speculated to be involved in spleen specific cytoadherence and maintenance of chronic infection [98].

Localization of PIRs within iRBC changes over time during parasite development and this has been shown in both *P. vivax* [110] and *P. chabaudi* [108]. Surface localization of PIRs had been shown in *P. falciparum* [111, 112] and *P. vivax* [92, 110] but no direct evidence of surface localization has been

shown for rodent PIRs. In *P. falciparum*, RIFIN has been shown to be expressed on the surface of infected RBCs and mediate rosetting via binding to the blood group A antigen [113, 114]. In addition, a subset of RIFINs have been found to bind to inhibitory receptors on the surface of immune cells to achieve immune modulation during intraerythrocytic development [115]. STEVOR has been demonstrated to bind to glycophorin C and mediate rosetting and to possibly facilitate merozoite invasion [111].

Together these results demonstrate that members of the *pir* superfamily have multiple functions critical for parasite survival that at this stage we do not fully understand. A summary of the known information of *pir* genes across the *Plasmodium* species can be found in Table 1.

	<i>P. falciparum</i> (RIFIN, STEVOR)	<i>P. vivax</i> (VIR)	<i>P. berghei</i> (BIR)	<i>P. chabaudii</i> (CIR)	<i>P. yoelii</i> (YIR)
Percentage of pir gene transcribed in population	-	59% [97]	-	40% [96]	50% [95]
Allelic exclusion	Stevor – No [116] RIFIN- No [117]	No [98, 118]	-	-	No [95]
Sequestration	Yes, STEVOR [111]	Yes, indirect evidence [110]	Yes, but not mediated by BIRs [109]	Yes [108]	-
Subcellular localization	STEVOR-irbc surface and MC [111, 112] RIFIN-irBC surface and MC [112]	irBC surface [92, 110]	Exported [119]	Internal and exported [108]	Internal and exported [84, 120]
Interaction with immune cells	RIFIN – binds to LILRB1 or LAIR1 [115]	-	-	-	-

Table 1: Summary of known information about *pirs* in the different *Plasmodium* species. - refers to information which are not available.

1.7 Importance of antibodies against blood stage malaria

The importance of antibodies in protection against malaria was first demonstrated by Cohen and colleagues in 1961 in which passive transfer of antibodies from malaria-immune adults reduced fever and parasitemia in children acutely ill with malaria [121]. Early studies have shown that antibody response to malaria is

linked to the exposure to blood stage parasites rather than sporozoites [122]. Antibody mediated protection can also be seen in mice model of malaria, although the degree of protection varies depending on the parasite-host combination used [123]. Passive transfer of serum from mice immunized with blood stage antigens confers protection to naïve mice against blood stage parasite challenge [124]. Antibodies can protect against blood stage parasites by opsonization [125], complement mediated lysis [126] or direct blocking of merozoite invasion [127]. This is dependent on the antibody subclass induced and is partially determined by the type of adjuvants used [128].

Despite the discovery by Cohen et al. that passive transfer of serum from malaria immune adults can reduce clinical malaria symptoms in children, we still do not know which antigens are responsible for protection and this makes the development of vaccines challenging. Therefore, it is of utmost importance for the malaria community to establish the identity of antigens that might be immunological correlates of protection. This will bring us to our objective which we aim to use a targeted and untargeted approach to study exported proteins and their functions in iRBCs.

2 Objectives

Malaria parasites export many proteins into the host cell cytoplasm to modify the host for their own survival [61]. Some of these exported proteins eventually get displayed on the surface of infected erythrocytes and play an important role in disease pathogenesis [64, 65]. Although there are many different screens done in different laboratories to try and identify these parasite infected erythrocyte

surface proteins (PIESPs), many of these proteins are not yet validated and their functions are still unknown [129, 130].

Since the parasite spends much of its time within the erythrocyte, we hypothesized that surface proteins might play an important role in immunity due to the extensive time that these proteins are exposed to the immune system. To increase our understanding of PIESPs, we used an untargeted approach in which proteomics and bioinformatics will be utilized to identify a library of proteins which are potentially expressed on infected erythrocyte surface.

The second part of my work is a targeted study to characterize members of the *pir* multigene family, which were shown to be exported but is still relatively poorly characterized. The main reason for focusing on a multigene family in our study is to address the inability of researchers to study these genes in isolation. Although the sheer numbers of these genes in the parasite genome likely makes them unsuitable as vaccine candidates, the fact that the parasite invest so much resources to maintain such a high number of *pir* genes within its genome implies that these genes must perform vital functions. The evidence that *pils* have been implicated in immune modulation and evasion makes them an interesting target to study. I aim to develop a set of tools which allows for the dissection of the functions of these genes and open up the research into this multigene family. Specifically, I aim to activate a single *yir* within the parasite and characterize the growth and localization of these proteins.

To achieve our specific aims, I utilized *P. yoelii*, a rodent malaria model to address the following objectives:

Proteomics approach to identify protective antigens

1. Identification of novel antigens on surface of infected erythrocytes
2. Evaluation of the protective capabilities of identified proteins
3. Characterization of proteins using molecular methods

Targeted characterization of *yir* multigene family

1. Activation of a constitutively expressed *yir* in *P. yoelii*
2. Functional characterization of *yir*
3. Determination of allelic exclusion in *yir*

Chapter 3

Materials and methods

Ethics statement

This study was carried out in strict accordance with the recommendations of the National Advisory Committee for Laboratory Animal Research (NACLAR) guidelines under the Animal & Birds (Care and use of Animals for Scientific Purposes) rules of Singapore. The protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of Nanyang Technological University (Approval number: ARF SBS/NIE-A0379). All efforts were made to minimize suffering of animals used in this study.

3.1 Materials

3.1.1 Animals

Wildtype male BALB/c mice (age 6– 8 weeks) were obtained from Invivos Pte Ltd. Mice were kept in Animal Biosafety Level 2 (ABSL2) condition at Nanyang Technological University Animal Research Facility.

3.1.2 *Plasmodium yoelii* parasites

Wild type *P. yoelii* parasites were obtained from Malaria Research and Reference Reagent Resource Center (MR4) and transgenic *P. yoelii* parasites were generated in our laboratory. Parasite lines were maintained in wild type BALB/c mice, infected erythrocytes were cryopreserved and kept in liquid nitrogen for long term storage.

3.1.3 Chemicals, reagents and kits

Table 2. Reagents used in *in vivo* studies

Name	Active compound	Brand	Dosage	Purpose
Valabarb	Pentobarbitone sodium	Jurox	40-60 mg/kg (in saline)	Euthanasia
Isoflurane, USP	Isoflurane	HSB Veterinary Supply Inc.	Inhalation	Inhaling anaesthesia
Pyrimethamine	Pyrimethamine	MP biomedical	10mg/kg (in DMSO for i.p); 70µg/ml (in drinking water)	Drug for killing parasites without DHFR
Heparin	Heparin	Sigma-Aldrich	0.2mg/mouse (in PBS)	Anticoagulant
Freunds complete adjuvant	-	Sigma-Aldrich	Intramuscular injection mixed with protein at 50:50 ratio	Adjuvant for immunization
Freunds incomplete adjuvant	-	Sigma-Aldrich	Intramuscular injection mixed with protein at 50:50 ratio	Adjuvant for immunization

Table 3. Other chemicals used in experiments

Name	Brand
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich
Histodenz™ nonionic density gradient medium	Sigma-Aldrich
RPMI media 1640	Thermo Fisher Scientific
Fetal Bovine Serum	Thermo Fisher Scientific
Penicillin/ streptomycin	Sigma-Aldrich
Giemsa Stain	Sigma-Aldrich
1 kb DNA ladder	Thermo Fisher Scientific
Agarose	Vivantis
B-Per Bacterial Protein Extraction Reagent	Thermo Fisher Scientific
Laemmli buffer	Sigma Aldrich
ExactPro Broad Range (10-245kDa) Prestained Protein Ladder	1 st base

PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa	Thermo Fisher Scientific
Protein marker	1 st base
10% SDS solution	Bio-rad
Bis-acrylamide (37.5% w/v)	Bio-rad
Tetramethylethylenediamine (TEMED)	Bio-rad
Ammonium Persulphate (APS)	Sigma-Aldrich
Acetic acid	Merck Millipore
Coomassie Brilliant Blue	AppliChem
Bovine Serum Albumin (BSA)	Sigma-Aldrich
Sodium Chloride (NaCl)	Merck Millipore
TRIS Hydrochloride (Tris-HCl)	Sigma-Aldrich
Calcium disodium versenate (CaNa ₂ EDTA)	Sigma-Aldrich
Potassium Chloride (KCl)	Merck Millipore
Disodium phosphate (Na ₂ HPO ₄)	Sigma-Aldrich
Monopotassium phosphate (KH ₂ PO ₄)	Sigma-Aldrich
Hydrochloric Acid (HCl)	Merck Millipore
Ethanol	Merck Millipore
Methanol	Merck Millipore
Sodium Hydroxide	Merck Millipore
Triton-X 100	USB biochemicals
TWEEN [®] 20	Amresco Inc.
Pierce Protease Inhibitor Tablets	Pierce
Trizol Reagent	Invitrogen
RNaseZap	Invitrogen
Imidazole	Sigma-Aldrich
Nickel-NTA agarose	Qiagen
Paraformaldehyde (PFA)	Sigma-Aldrich
Percoll	Sigma-Aldrich

Table 4. Kits used in experiments

Name	Brand	Usage
Basic Parasite nucleofector solution kit II	Lonza	Parasite Transfection
Restriction enzymes	New England Biolabs	PCR reaction
Restriction enzymes (fast digest)	Thermo Fisher Scientific	PCR reaction
KOD Neo plus DNA Polymerase	Toyobo	PCR reaction
GoTaq[®] Green Master Mix	Promega	Colony PCR
FavorPrep Plasmid Extraction Mini Kit	Favorgen	Plasmid purification
NucleoBond [®] Xtra Midi Kit	Macherey-Nagel	Plasmid purification

HaloTag® Protein Purification System	Promega	Protein purification
NucleoSpin Blood®	Macherey-Nagel	DNA extraction from parasites
RNeasy Mini Kit	Qiagen	RNA purification

Table 5. Buffers and Solutions

Name	Composition
Incomplete RPMI (iRPMI) for parasite culture	Dissolve 16.2g RPMI powder and 2.3g sodium bicarbonate in 1L distilled water. Filter through 0.2µm membrane to sterilize.
complete RPMI (cRPMI) for parasite culture	iRPMI supplemented with 20% FBS and 30µg/ml Gentamycin
Histodenz buffer pH 7.5	5 mmol/L Tris HCl, 3mmol/L KCL, 0.3mmol/L CaNa ₂ EDTA
Histodenz solution	Dissolve 27.6g Histodenz powder in 100ml Histodenz buffer, autoclave for 20 mins and store in 4°C
10X PBS buffer	Dissolve 8g NaCl, 0.2g KCl, 1.44g Na ₂ HPO ₄ and 0.24g KH ₂ PO ₄ in 0.8L distilled water. Adjust to pH 7.4 with HCl and top up to 1L.
FACS buffer	1X PBS with 1% BSA
Saline	Dissolve 0.9g NaCl in 100ml distilled water, filter through 0.2µm membrane to sterilize.

Table 6 Stains and antibodies used for FACS

Name	Brand
Hoechst 33342 nuclei acid stain	Thermo Fisher Scientific
Dihydroethidium	Sigma-Aldrich
Anti-mouse CD45 APC-eFluor 780 Clone: 30-F11	eBiosciences Inc

Table 7. Softwares for data analysis

Name	Developer
ImageJ	Wayne Rasband
FlowJo V10	Treestar Inc, Ashland, OR
Trimmomatic	Anthony Bolger
Kallisto	Pachter lab

3.2 Methods

3.2.1 Shortlisting of candidate proteins

A total of 202 proteins were identified in previous screen and these were assigned a score of 1 to 10, with 10 being the protein which is most likely to be associated with protection and possibly located on the surface of the infected RBC. In order to select for proteins that have a higher association with protection, 24 proteins with a score of ≥ 8 were selected for further analysis. The following selection criteria were used to further shortlist the 24 proteins; Presence of transmembrane domain, presence of signal peptide, presence of orthologs in *P. falciparum*, not found within the RBC. Under these selection criteria, 3 hypothetical proteins were shortlisted for downstream work. Refer to Chapter 4.1 for a detailed writeup of the selection procedure.

3.2.2 Gene cloning

3.2.2.1 HaloTag vector

Selected sequences of the 4 genes were obtained from PlasmoDB and primers were designed to amplify the coding region of the genes (primer list), omitting N terminal signal peptide and predicted transmembrane domain. The PCR-amplified inserts were ligated into a pH6HTN His6HaloTag T7 Vector (Promega), between the SacI and NotI sites. Ligations were transformed into

DH5 α cells (Invitrogen) via heat shock treatment and transformants were selected on agar plate supplemented with 100 μ g/ml Ampicillin. Resulting clones were verified by sequencing. DNA from the correct clone was subsequently transformed into one step KRX competent cells (Promega) for protein expression.

3.2.2.2 His-tag vector

pET24a vector (Novagen) was used as the expression vector and PCR amplified inserts were ligated in between the NheI and XhoI sites (primer list). Ligated plasmids were transformed into DH5 α cells (Invitrogen) and positive clones were verified by sequencing. Correct plasmid clones were transformed into BL21(DE3) cells (NEB) for protein expression.

3.2.2.3 Skip peptide vector

P. yoelii knock-in constructs were generated by modifying the existing ePL plasmid which was published previously for use in *P. yoelii* episomal transfection [84] to create an integration plasmid T2A-ePL. DNA corresponding to the last 500 base-pair of PYYM_0017400 gene was generated by chemical gene synthesis (Genscript). We replaced the portion of the ePL which contains (eGFP tag, *P. berghei* EF1 constitutive promoter, the 3' UTR region of *P. berghei* DHFR/TS gene and the selectable marker *T. gondii* DHFR gene) with an eGFP-T2A-TgDHFR construct. Ligated products were transformed into XL10-gold cells and plated on Ampicillin plates. All positive clones were validated by sequencing.

3.2.3 Protein expression and purification

3.2.3.1 Halotag proteins

Induction was done by seeding a starter culture at 1:100 dilution and adding 0.02M Rhamnose (Promega) when OD₆₀₀ reaches 0.4 – 0.5. Induction was carried out overnight for 20 hours at 37°C. Cells were pelleted and lysed using B-Per (Thermo Fisher), following manufacturer's instructions. Presence of proteins after induction was confirmed by running an aliquot of cell lysate in an SDS-PAGE and staining with Halotag TMR Ligand (Promega). Lysed bacteria cells were pelleted and the supernatant was incubated with equilibrated HaloLink™ Resin (Promega) for 1 hour with end to end mixing. Purification was done as per manufacturer instruction with minor changes made to the TEV removal step. Instead of 1 step TEV removal, TEV was removed by 2 rounds of 15 minutes incubation at room temperature with Hislink™ resin (Promega). Changes made to the buffers used in the purification can be found in the appendix (Table i).

3.2.3.2 Histag proteins

Protein induction was done by incubating BL21(DE3) cells at 37°C until OD₆₀₀ reaches 0.6 – 0.8. Induction was carried out by adding 1mM IPTG and incubating for 4 hours at 37°C. Bacteria cell pellet was collected by centrifugation at 5000 rpm for 10 minutes. Bacteria pellet was resuspended in purification buffer and sonicated for 20 minutes (45 amplitude, 6 sec on, 6 sec off). Presence of recombinant proteins were confirmed by running an immunoblot with an aliquot of cell lysate and probing with Anti-His antibodies (BioLegend). Lysed bacteria

cells were centrifuged at 13000g for 20 minutes and the pellet was collected for insoluble protein purification.

Lysis buffer (purification buffer + 8M Urea) was added to the pellet, resuspended and incubated at 37°C for 2 hours on a rotator. The suspension was centrifuged at 13000g for 20 minutes and the supernatant containing the dissolved proteins was collected. Ni-NTA agarose resin (nickel-Nitrilotriacetic acid, Qiagen) was added to the supernatant and incubated at 37°C for 1 hour on a rotator. Resin was loaded onto column and allow the solution to flow through. The resin was washed once with 5 column volume of buffer and bound proteins were eluted with increasing concentration of imidazole. 30ul of each fraction was collected for gel electrophoresis to determine which fraction to pool. Buffer exchange was done by placing pooled protein fractions into a 10,000 MW cutoff dialysis tube (Thermos) and dialyze with PBS overnight at 4°C with slow stirring.

3.2.4 Determination of protein concentration

Protein concentration of all purified proteins was determined by comparing the intensity of Coomassie blue-stained bands on SDS-PAGE gels. Purified proteins and serial dilutions of purified bovine serum albumin (BSA) were subjected to electrophoresis on an SDS-PAGE gel. The gel was stained with Coomassie blue and destained until the bands can be seen clearly. Intensity of the sample and BSA bands were quantified with ImageJ V.1.48 (National Institutes of Health). A standard curve was obtained from the intensity reading of the BSA dilution and were compared to the intensity of the protein sample to get the amount of

protein loaded. The concentration of the proteins was determined by dividing the amount of protein by the volume loaded.

3.2.5 Validation of proteins against serum from previous proteomic screen

1µg of purified proteins were loaded into each well and electrophoresis was carried out until the dye front exits the gel. Proteins were transferred onto a 0.2µm nitrocellulose membrane (Bio-Rad) by semi-wet transfer. Membrane blocking was carried out for 1 hour with 5% skim milk diluted with PBST and subsequently washed 3x with PBST for 5 minutes. Primary antibody incubation was done overnight at 4°C by incubating membrane with 4 immune sera generated previously by our lab (immune sera dilution 1:1000). Membrane was washed 3x with PBST and incubated with goat anti mouse IgG HRP secondary antibody (BioLegend) at a dilution of 1:5000 dilution for 1 hour. Final washing step was carried out 3x for 10 minutes with PBST. Bands were visualized by adding SuperSignal chemiluminescence substrate (Thermos) to the membrane and exposed to X-Ray film.

3.2.6 Immunization regimen

Groups of 5 male Balb/C mice (8 weeks old) were injected with 50µg of purified proteins at 2 weeks interval. First immunization was done via subcutaneous route with recombinant protein emulsified in Complete Freund's adjuvant (CFA, Sigma) at 1:1 ratio. Subsequent booster immunizations were done at 2 weeks interval with recombinant proteins emulsified in incomplete Freund's adjuvant (InFA, Sigma) at 1:1 ratio. Sera were collected from immunized mice 1 week

after the 3rd booster shot for ELISA to check for the presence of protein specific antibodies.

3.2.7 ELISA

96-well flat-bottomed Nunc Maxisorp plates (Thermo Scientific) were coated with 200ng of peptide diluted in coating buffer (eBiosciences) overnight at 4°C. Insoluble proteins were first dissolved in 2% SDS-PBS before diluting to 0.02% for coating overnight at room temperature. The plates were washed 3x with PBS-T and blocked with 3% BSA-PBS-T. Serum from mice were added in triplicates and incubated at room temperature for 2 hours (Dilutions of antibodies tested range from 1:4000 to 1:128000). After washing 3x with PBS-T, plates were incubated with horseradish-conjugated goat anti-mouse IgG (BioLegend) at a dilution of 1:5000. Enzyme reaction was developed by adding tetramethylbenzidine (TMB, eBiosciences) and stopped with 2N H₂SO₄. The OD was measured by using an ELISA microplate (Biorad) reader subtracting 540nm reading from 450nm readings. Background determination was determined from wells without serum and was subtracted from the mean of each sample. Results were expressed as index of reactivity (IR = OD₄₅₀ values of test sample divided by value of cutoff) Cutoff values were set at 3 standard deviations above the mean of sera from mice injected with adjuvant + PBS. Samples were considered positive if the IR value is > 1.

3.2.8 Mouse challenge and parasitemia assessment

Prior to infection, a Histodenz gradient was performed to extract schizont stage parasites. Schizonts were counted on hemocytometer and mice were

subsequently challenged with 10^3 parasitized RBC via intravenous injection. Parasitemia was measured daily from day 4 onwards using flowcytometry. Parasites were stained with 1 μ M Hoechst 33342, 5 μ g/ml dihydroethidium and anti CD45 APC/Cy7 (1:600) in a volume of 100 μ l for 20 minutes at room temperature. The samples were acquired on BD Fortessa (BD Biosciences) using the UV laser (305 nm) to detect Hoechst 33342, the blue laser (488 nm) for dihydroethidium, and the red laser (633nm) for APC/Cy7. Mice with 0% parasitemia reading on the FACS were double checked with thin blood smears. Parasitemia was monitored after day 4 post infection by counting at least 1000 erythrocytes from blood smears stained with Giemsa.

3.2.9 Characterization of antibodies

3.2.9.1 Immunoblotting for proteins in *P. yoelii* schizont lysate

Schizont samples purified using histodenz gradient were lysed in sample loading dye and subjected to SDS-PAGE. Antibodies obtained from the immunization were used to probe the membrane.

3.2.9.2 Immunofluorescence (IF) assay

P. yoelii 17X 1.1 infected erythrocytes were obtained by tail bleed from BALB/c mice and washed 3x with incomplete RPMI. Packed blood cells were diluted with FBS and thin blood smears were made and stored in -20°C until use for IF assay. Slides were fixed with cold acetone/methanol (9:1 dilution) for 10 minutes at 4°C. Blocking was done by incubating slides in blocking buffer (3% BSA diluted in PBS) for 1 hour on a rocker at room temperature. Slides were washed 3x with PBS for 5 mins. Mouse serum (1:1000) was added and incubated for 2 hours at

37°C, overnight at 4°C and 2 hours at room temperature. Slides were washed 3x with PBS for 5 mins and incubated with DyLight 594 – Goat anti mouse IgG secondary antibody (Jackson Immuno Research Laboratories) (1:800) in the dark for 1 hour at room temperature. Slides were washed and incubated with Hoescht stain for 10 minutes. 2 drops of Vectorshield mounting media (Vector Laboratories) was added to slide and covered with coverslip. Slides were visualized with fluorescence microscope (Nikon)

3.2.10 Parasite transfection

Mice infected with wildtype *P. yoelii* YM or *P. yoelii* 17X 1.1 parasites were used as parasite donor. Mice were euthanized with Valbarb by i.p. injection when the parasitemia reaches 15%-60%. Infected blood was collected through cardiac puncture with syringe containing heparin solution. Parasitized blood was washed with iRPMI and centrifuged at 2100 rpm with brake 0 for 5 minutes to remove serum and buffy coat. The RBC pellet was passed through a 50%/ 60% Histodenz gradient to isolate schizonts. The schizonts were collected and washed once in iRPMI and were allowed to mature with slow shaking in 37°C for 3 hours. Matured schizonts were collected by centrifuging at 2100rpm for 5 mins.

3.2.10.1 Skip peptide transfection

The T2A-YIR vector were linearized with EcoRV and concentrated with ethanol precipitation prior to transfection. 12µg of linearized T2A-YIR plasmid were used in each transfection. Transfection was carried out by mixing the plasmids with 100ul Nucleofactor solution (Lonza) and electroporated with Amaxa Nucleofactor II (Lonza) using program U-33. Transfected parasites were injected

intravenously into mice and transfectants were selected with a 3 day on and off regime of pyrimethamine (MD laboratories) injections.

3.2.11 Single cell cloning of transfected T2A-YIR parasites

T2A-YIR transfected *P. yoelii* infected blood were obtained from BALB/c mice and passed through a 50%/ 60% Histodenz gradient to isolate schizonts. Schizont density were counted using hemocytometer and cells were diluted to obtain a cell density of 1 cell per 100ul. 100ul of cell suspension was injected via i.v. into each mouse. The mice were injected with pyrimethamine (MD laboratories) every day for the first 3 days and were maintained subsequently with pyrimethamine added to drinking water. Genomic DNA was extracted from parasites after cloning and diagnostic PCR were conducted and sequenced to check for correct genomic integration.

3.2.12 Diagnostic PCR for checking of integrated Skip-peptide transgene

Forward PCR primers were designed to bind within the first 200 base pair of the gene of interest. Reverse primer was designed to bind in the GFP region of the transgene (Primer list). DNA was isolated from parasites when the parasitemia reaches 30% and genomic DNA was extracted with NM Blood kit.

3.2.13 Isolation of trophozoite-stage total RNA from T2A-YIR transfected parasites

Mice were inoculated intravenously with parasites in the ring stages and harvested 32 - 34 hours post infection to get trophozoite enriched parasites. Mice were sacrificed and whole blood was collected and passed through a non-woven fabric (NWF) filter to remove leucocytes. The filtrate was then centrifuged at

2200 rpm for 5 mins (acc5, dec0) and washed 2 times with iRPMI. 200ul of erythrocyte pellet was added to 2ml Trizol (prewarmed to 37°C), inverted a few times to get rid of cell clumps and incubated at 37°C for 10 minutes. The Trizol-erythrocyte mixture was then frozen in dry ice and kept in -80°C until ready for RNA extraction. RNA extraction was carried out as stated in Z. Bozdech et al [131]. The extracted RNA was purified using RNA cleanup kit following manufacturer's protocol (Qiagen) and stored in -80°C until further analysis.

3.2.14 RNA sequencing and analysis

Triplicates of each RNA samples were outsourced to SCELSE for RNA sequencing. RNA library preparation was performed according to Illumina's TruSeq Stranded mRNA protocol with 1,000ng of total RNA as input. The PCR amplification step, which selectively enriches for library fragments that have adapters ligated on both ends, was performed according to the manufacturer's recommendation but the number of amplification cycles was reduced to 12. Each library was barcoded with one of Illumina's TruSeq LT RNA barcodes to enable library pooling for sequencing.

Finished libraries were quantitated using Promega's QuantiFluor dsDNA assay and the average library size was determined on an Agilent Tapestation 4200. Library concentrations were then normalized to 4nM and validated by qPCR on a QuantStudio-3 real-time PCR system (Applied Biosystems), using the Kapa library quantification kit for Illumina platforms (Kapa Biosystems). DNA libraries were then pooled at equimolar concentrations and sequenced on an Illumina MiSeq sequencer at a read-length of 150bp paired-end.

Removal of Illumina adapters was made with the use of Trimmomatic [132] and Kallisto was used to create an index file from 17X 1.1 CDS (extracted from Plasmodb on 28 May 2018) for quantification. We performed the RNAseq analysis using Kallisto to map the reads against the annotated PY17X1.1 index file with bootstrap value of 100 and thread 4 to obtain the transcript per million (TPM) count. Differential gene analysis was done by processing the TPM from Kallisto with DEseq2 software using the *tximport* pipeline [133].

3.2.15 *In vivo* growth kinetic assay

Blood containing T2A-YIR transfected parasites and YM-mCherry parasites were harvested from mice via cardiac puncture and schizonts were isolated via a 50%/ 60% Histodenz gradient. Cells were diluted to obtain a cell density of 100cells/ 100ul. 100ul of cell suspension was injected into each mouse via IV injection. Mice were provided with drinking water supplemented with pyrimethamine throughout the whole duration of the experiment. Mice were bleed daily after day 4 post infection to check parasitemia via flow cytometry and thick blood smear. Once the parasitemia starts to decline in recovering mice, blood smears were made every 3 days until the mice is cleared of the infection.

3.2.16 *Ex vivo* growth kinetic assay

The ex vivo kinetic assay was adopted from Chang et al [134]. Briefly, Blood containing T2A-YIR transfected parasites and YM-mCherry parasites at about 10% parasitemia were harvested from mice via cardiac puncture and passed through a NWF filter to remove leucocytes. The blood was passed through a 65% percoll gradient to isolated rings and uninfected erythrocytes. The parasites were

then diluted to obtain 1 – 3% parasitemia and 2% hematocrit. Cells were seeded at 100ul per well in a 96 well plate (NUNC) and kept in 37°C for 24 hours. For parasitemia check, Cells were sampled every 3 hours for parasitemia growth using flow cytometry and thick blood smear. For parasitemia check via flow cytometry, cells were incubated in the dark with 5 µg/ml DHE and 8 µm Hoechst 33342 for 30 minutes at 37°C. After incubation, 400ul of ice cold PBS was added.

3.2.17 cDNA synthesis by reverse transcription

Reverse transcription of RNA was performed with SuperScript III reverse transcriptase. The first master mix was done as follows: 1µg RNA sample, 2µl Oligo (dT): pdN9: TS-Oligo 25µm:25µm:50µm primers (10mM, Invitrogen) and top up to 8µl with DEPC water. Incubate Mastermix 1 in PCR at 65°C for 5 minutes and chilled on ice for 2 minutes. Prepare mastermix 2 when mastermix 1 was being incubated. Mastermix 2 contains: 4µl 5X First Strand Buffer (Invitrogen), 2µl Dithiothreitol (DTT, 0.1M, Invitrogen), 1µl RNaseOUT (Thermo Scientific) and 3µl DEPC water. Add 11µl Mastermix 2 into mastermix 1 and incubate at room temperature for 2 minutes. After incubation, add in 1µl SuperScript III reverse transcriptase (200U/µl, Invitrogen) and mix gently. Incubate the mixture of mastermix 1 and mastermix 2 at 42°C for 50 minutes and 70°C for 15 minutes. cDNA was stored at -20°C.

3.2.18 Quantitative Real-time PCR (qRT-PCR) Analysis for Expression of *yirs*

qRT-PCR was done by preparing master mix 1 which contains: 10µl SYBR Green qPCR mastermix (Invitrogen), 0.2µl forward primer (10µM), 0.2µl reverse

primer (10 μ m), 9.6 μ l DEPC water (Invitrogen). The 19 μ l of mastermix was added to 1 μ l of cDNA that was pre-aliquoted into the 0.2ml 96- well plate (Thermo Scientific). The wells were sealed with MicroAmp adhesive film (Thermo Scientific), centrifuged at 500rpm for 3 minutes and placed into 7500 real-time pcr system (Applied Biosystem). The following protocol was used:

50°C2 min 1 cycle
95°C2 min 1 cycle
95°C 3 sec 40 cycles
60°C30 sec 40 cycles

Chapter 4

Characterization of proteins identified in proteomic screen

Previous proteomics screening and analysis has generated a list of proteins associated with protection prior to me joining the lab as a PhD student [135]. I have included these works in Section 4.1 and they served as background information for the actual work that was done by me from section 4.2 onwards.

4.1 Prior work done in the lab

4.1.1 Proteomic identification of antigens located on surface of infected RBC

Malaria proteins exported into the host cell and subsequently expressed on the surface of infected erythrocytes are attractive targets for vaccine development. This is because the parasite spends most of its life cycle within the erythrocyte and the immune system has the most time to react to the parasites at this stage. In addition, parasite infected erythrocyte surface proteins (PIESPs) are exposed to the immune cells and can be targeted by antibodies [136].

Although PIESPs are attractive targets, many of these such as PfEMP1, RIFIN and STEVOR belong to multigene families, are highly polymorphic and contributes to antigenic variations and immune evasion [137, 138]. This makes them unsuitable to be used as vaccine candidates because the parasite can easily switch expression to another form of the protein. Although protective immune response can develop against malaria, most subunit vaccines have not been shown to induce protection with substantial efficacy. This shows that the antigens that are responsible for protection have yet to be identified or protection against malaria requires a complex combination of responses. Since protection against blood stage parasites involved humoral immunity, our lab pursued a proteomics approach to identify antigens which are associated with protection.

4.1.2 Generation of immune sera

Three different Immune sera from mice protected against parasite challenge following increasing parasite exposure were generated as shown in Figure 4A. The ‘Chloroquine treated’ sera were generated by mice exposed to the parasites for a short duration followed by rapid clearance following injection of chloroquine shortly after parasite challenge [139]. The ‘Attenuated’ sera were generated by mice exposed to parasites transfected with a plasmid containing human dihydrofolate reductase (hDHFR) as a drug selection cassette. Mice were treated with pyrimethamine to select for transfected parasites and a low parasitemia (<5%) was detected in mice for about 2 weeks before clearance. The ‘self cured’ and ‘none-protected’ sera were generated by mice exposed to the parasites for the longest time. Mice were injected with low dose of parasites and sera collected from mice which cured the parasites within 6 weeks were named ‘Self cured’ sera. Sera from mice that were unable to cure the parasites after 6 weeks were classified as ‘none-protected’. All mice from the ‘attenuated’, ‘chloroquine treated’ and ‘self cured’ group were challenged 2 – 4 times with 10^6 *P. yoelii* 17X 1.1 to ensure that mice were sterilely protected. These sera are termed “protective sera”.

The different sera generated were tested for their ability to protect mice from subsequent challenge via passive transfer into 5 naïve mice each. Passive transfer of each immune sera into naïve mice showed that the sera protected the mice with different efficacies (Figure 4B). Out of the 5 mice injected with the protective sera, 80% of the Attenuated sera injected mice survived; 60% of the chloroquine treated sera injected mice survived; 40% of the self cured sera injected mice

survived. 20% of the non-protected sera injected mice survived, suggesting that the protective sera contained antibodies directed against protective antigens.

Figure 5: Generation of sera for identification of antigens associated with protection in mouse malaria model. A) Schematic diagram for the generation of the 4 different immune sera differing in the time the immune system is exposed to the parasites. B) Mice survival after passive transfer of different sera (Total of 5 mice per group). Data showed that the different sera protected the mice with different efficacies with attenuated sera providing the best protection.

4.1.3 Characterization of antigens recognized by protective sera

Further analysis was done to characterize the antigens recognized by the different sera and this is summarized in Figure 5. First, transmembrane protein-enriched fractions extracted under native condition from trophozoites/schizont parasites

were immunoprecipitated with the 4 sera covalently linked to protein A/G agarose beads. The resulting immunoprecipitated proteins were characterized using LC-MS/MS.

Second, to further characterize the protective antigens, a 2D proteomics approach was used to identify protective antigens under denatured condition. 5 aliquots of the same parasite transmembrane protein-enriched extract were fractionated in parallel by 2D gel electrophoresis and probed with sera from the different preparation. The corresponding area from the silver stained gel containing the antigenic spots were excised and LC-MS/MS was performed to identify the different proteins.

Lastly, proteins associated with protection are most likely present on the surface of iRBC and in direct contact with the immune system. To detect potential surface antigens, intact parasite infected erythrocytes were labelled with biotin and native transmembrane protein fractions were extracted. These were passed through a Neutravidin column and bound proteins were eluted and evaluated by LC-MS/MS.

Analysis of all the mass spectrometry results was done by Dr Anthony Siau in the laboratory.

Figure 6: Proteomics approaches to identify antigens associated with protection. A) Transmembrane protein enriched fractions extracted under native conditions during trophozoite/schizont stages were immunoprecipitated with immune sera covalently linked to protein A/G agarose beads. B) Transmembrane protein enriched fractions were subjected to 2D gel electrophoresis and probed with sera from the different groups of protected mice. C) Purified trophozoite culture were labeled with biotin and passed through a neutravidin column to retain labeled proteins. Proteins collected from all 3 approaches were sent for LC MS/MS analysis.

The antigens immunoprecipitated by the protective sera were ranked for their likelihood of being associated with protection. Five criteria were considered and we score the proteins from 1 – 10 with 1 being the least likely and 10 being the most likely to be associated with protection (Table 7). This ranking classified the antigens into 4 groups with increasing likelihood to be associated with protection.

		Found in IP and denatured immunoblot	Enriched relative to abundant parasite proteins	Found on infected erythrocyte surface	Signal peptide/ Transmembrane domain
Dominantly attenuated	+4	+2	+1	+2	+1
Dominantly chloroquine treated	+3				
Dominantly self-protected	+2				
Dominantly non protected	+1				
Group		1	2	3	4
Score		≥3	4 - 6	7 and 8	9 and 10

Table 8: Criteria used for the ranking of the antigens found by immunoprecipitation and the score awarded for each criterion. Antigens were divided into 4 groups based on the overall score.

4.2 Three novel proteins of unknown function were shortlisted

A list of 202 proteins that are potentially associated with protection was generated by a proteomics screen in our laboratory [135]. Using the scoring matrix defined in table 7, the 202 proteins potentially associated with protection were ranked, with proteins having a score of ≥ 8 being considered for subsequent analysis.

The resulting list of 24 proteins (Table 8) consist of known antigens which have been shown to be localized to the merozoites such as Ron2, PyRhopH1A, rhoptry associated protein 1, merozoite surface protein-9 precursor, putative, merozoite surface protein precursor, Py140/RON4 and merozoite surface protein 1 precursor (colored blue in the table). In addition, a number of internal proteins or proteins with cellular functions that is likely not to be related to immunity (colored orange in table) were identified in this group. These proteins might be contaminants due to the relative abundance in the iRBCs. We omitted the results from proteins which do not have an orthologs in *P. falciparum* (colored red in table) to ensure that any proteins found in this study is also relevant to human malaria

ID	Name	Final score	Ortholog	possible function of hypothetical proteins	TM domain	Signal Peptide
PY06813	hypothetical protein	10	Yes	ortholog in PF is RON2	1	Yes
PY05748	merozoite surface protein 1 precursor	10	Yes		1	Yes
PY05892	Exp-2	10	Yes		0	Yes
PY06415	merozoite surface protein precursor	10	Yes		1	No
PY00293	Papain family cysteine protease, putative	10	Yes		0	Yes
PY01014	relinitis pigmentosa GTPase regulator-like protein	10	No		0	Yes
PY02883	merozoite surface protein-9 precursor, putative	9	Yes		0	Yes
PY05420	vacuolar proton translocating ATPase 116 kDa subunit a isoform 1	8	Yes		6	No
PY03783	hypothetical protein	8	No		1	Yes
PY03777	Streptococcus pyogenes AMV256, putative / orthologue to Plasmodium translocon component P TEX88	8	Yes		0	Yes
PY06306	hypothetical protein	8	Yes	unknown plasmodium membrane protein in PF with unknown function	3	No
PY03953	Drosophila melanogaster CG11212 gene product/ orthologue to Plasmodium lipid/sterolH symporte	8	Yes		8	No
PY03970	ATPase, P-type, HAD superfamily, subfamily IC, putative	8	Yes		5	No
PY04372	hypothetical protein	8	Yes		0	Yes
PY05364	clpB protein	8	Yes		1	Yes
PY03882	adenine nucleotide translocase	8	Yes		3	No
PY01995	hypothetical protein / orthologue to Plasmodium protein phosphatase 2C	8	Yes		0	No
PY06203	blood-stage membrane protein Ag-1	8	No		1	No
PY02932	PyRhopH1A-related	8	Yes	PIESP15	0	Yes
PY04147	hypothetical protein	8	Yes	ortholog in PF is EXP1 (circumsporozoite related antigen)	2	Yes
PY04421	U43539 hepatocyte erythrocyte protein 17 kDa	8	Yes		2	Yes
PY00622	rhoptly associated protein 1	8	Yes		0	Yes
PY02159	Py140/RON4 Narum DL et al 2008	8	Yes		0	Yes
PY03372	hypothetical protein	8	Yes		0	Yes

Table 9: List of 24 proteins ranked 8 or higher in the proteomics screen. The list of proteins is organized in terms of the score given with the highest scoring protein at the top and the lowest at the bottom. Color bar represents different localization/functions of the proteins. Blue: merozoite proteins, Orange: parasite internal proteins/ cellular functions. Red: proteins not selected due to not having orthologs or no transmembrane region. Green: the final selected proteins for this work. PY03372 is omitted from downstream work due to the lack of transmembrane domain.

The final 4 remaining hypothetical proteins were checked for the presence of a signal peptide and/ or transmembrane domains might indicate transport into the secretory pathway (colored green in table) [140, 141]. PY03372 was omitted from further analysis due to the lack of a transmembrane domain. Finally, selection of novel proteins not previously characterized would avoid repetition of previous work. With these criteria, three hypothetical proteins (PlasmoDB ID PY04147, PY04372 and PY06306) and the merozoite surface protein 1 (MSP1) coded by gene PY05748 were selected. In particular MSP1 coded by gene PY05748 has previously been shown to induce a protective response in mice [142, 143]. The 19 kDa C-terminal fragment of MSP1 termed MSP1-19 will act as a positive control for the immunization assay.

4.2.1 In-silico analysis of the 3 selected proteins

PY04147 is denoted as a putative parasite infected erythrocyte surface protein 15 (PIESP15) and the *P. falciparum* ortholog was first identified as a potential PIESP by an in-silico prediction approach by Florens et al in 2004 [129]. In their screen, they identified 36 potential PIESPs and given a number to each of them respectively. However, no validations were done except for PIESP1 and PIESP2 confirming their surface localization in that paper. The protein has a high amino acid sequence homology to its other orthologs from amino acid 60 to 280 which contains a concanavalin A like lectins motif, and a conserved C terminus which is predicted to be a coil-coiled domain (Figure 7).

```

PF3D7_0103900    MRKTGRLVFYLC-LSINCLFVNICKNSDTKNNSYNDYNTYLSKIHSEKAPFNVENLEN 59
PVX_081565      ---MARLLALLVGLLCWAFANQTCANGSSTNSGS---GGMYVSPAHSFNSPFTVNDLNN 53
PY04147         MNKYARFIHVSII-LIFVLFINCKENE---SPNK---NENFISPIHSFKSPFQVEGLNN 53
PCHAS_0207600   MNKYARFIHILIV-LIFVLFINCKEGE---SKNK---NENVISPIHSFKSPIQIEDLNN 53
               .*: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    GHWVDYSSVSTNKHVVLPISYNNRGVLYNKNPINSDFIIEFNFMERKYYKEIDEKY 119
PVX_081565      GHWINYSTASTNKYLVLPINVYNNRGFLYHSPLLSDTLNIEFSFHIKKQLYREHDDSY 113
PY04147         GHWLDYSAASTNKYLVLPINVYNNRGFLYNSNPISDVFDFINFSFNIYKKYKESIDSNY 113
PCHAS_0207600   GHWIDYSAAPTNNYLVLPINVYNNRGILYNNPTPKSDVDFISFSFNIYKKYKESIDGSY 113
               ** : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    YKDEKNDVHEIVYKTNSDNLTSTNKSNNKEQKIKRNGFALNWLKDPFVKEAGSNV-N 178
PVX_081565      YKTNQV--KETVYKTATE-----EKNKINGFAFWLLPNEFSLQWNAEGE- 156
PY04147         YKNEEG--KHITYNKKID-----EKDKTNGFAFWILENQLSVKSPGVSEEQ 157
PCHAS_0207600   YKNEPG--KNITYNKNID-----EKDKTNGFAFWILENQMSVQSPPTTSNEP 157
               ** : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    LVLQEEYNIIFGYKNRFNGIGILFQKKSTNKSNNKSNNNNNYNNYNDILNVSIIIN 238
PVX_081565      KQLDEEQFGLYGFKKDFNGIGVGFQLRAND-----LVLSGLIN 194
PY04147         IIDEEDVILYGYKIFNGICIFYQLRND-----LSVYALIN 195
PCHAS_0207600   IVLNEEDVILYGYKTFSGICVYFQLREND-----LAVYALIN 195
               : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    NGDKSINSDDFIFKNIPIDIMINNLHTKIICNKKETIAFYYPKSNNYINTLIMKN 298
PVX_081565      NGKKNLNVKEN-VTKSYNLSLSDMLVYVKITTQNEHRYVLFNASHNSYHSLTVKKQ 253
PY04147         NGKNSISLNNN-SAKNYNLSLQNNGLISVKIVAQNDRIRYLFNLKTSTYIHSITIKKN 254
PCHAS_0207600   NGKNSININNN-SAKNYNLGLQNNGLISVKVTAQNDRIRYLFNAKTSTYIHSITIKKI 254
               ** : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    ISTENYIAFTSFNYKEDDNKIDYNDVNNKYENQNGYLPFTVGIKQVLTFLNNNNNN 358
PVX_081565      IPRENYIGFSAFNYKENELLPD-----THAEKYVPTFVGITHVSISSDFPDGQ 303
PY04147         LPKENYIGFSSFNFNEDTSISI-----TNVKNYLPFTVGITDFHVTYNNYVKEKE 304
PCHAS_0207600   LPKENYIGFSSYFNKEDKSISL-----THVKNYLPFTVGITDFHVTYNNYVKEPE 304
               : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    NNININEANINNNNNNNNNNNNNNNKINYDQNNKDISANDINASFNDVLNNINPEKNTIN 418
PVX_081565      QGEYSKKE-ADS--MD--ESEEHLG-NDP-SDVLSGQDS-----SHPNYHGYH 345
PY04147         QKIIEYENDPENLVK---TKNDEEINYNELNGNINLM-NEI-----LSSQNINLS 350
PCHAS_0207600   QEIIAENDNPEPNVKK---ANN-EEANNNDVNVNTLM-NDI-----SSSQNLGQS 349
               : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    QTIDVLSNITKLLQKFMHQINSEKLLQINILNEKLQFVQTELKVIKKNLTNQTEEPK 478
PVX_081565      NQAEALNSLNATLQQFMSYQISSEKILQSLNTLQESIQYLHAEKQVKKMNVTKSEDPK 405
PY04147         SQEEMLKTKLIIQDFISYQVNNDKKVLQNTSFLKENIHNNQDEIKQIRKNITN-KSDPK 409
PCHAS_0207600   SQEETLKSHLKIIDFVTYQVNDKHSQMAVLKENIQNNQEEIKQIRKNINN-KSDPK 408
               . : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    NLQKMFSTELTGLKNLFHTHTQHHKKNIEDITNKLTKKIERNNELKILTEKAQKLEQIIN 538
PVX_081565      HFQKIFTTELSGLKSLFHSHAQHSKKNMEDITDRLTRINENQELKLLAEKAQKLETIIS 465
PY04147         NLQKIFTSELSGLKNLFHSHAQHHKKNIEDITNRLTSKIDNNQELKLLAQKLEHIIN 469
PCHAS_0207600   NIQKIFTSELSGLKNLFHSHAQHHKKNIEDITNRLTSKLDNNYDLKLLAQKLEHIIN 468
               : : : : : : : : : : : : : : : : : : : : : : : : : : : :
PF3D7_0103900    KGNYYTYVFCFAFALLIILTLMLIYKKIRDVEKKHIL 575
PVX_081565      KGNSTSYHFSFAFALLIILTILILIYKKIRDVDKKHIL 502
PY04147         KGNNTSYFFSFAFALIVITLIMIYKKIRD----- 499
PCHAS_0207600   KGNNTSYFFSFAFALIVITLIMIYKKIRDVEKKHIL 505
               ** * : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Figure 7: Alignment of protein sequence of PY04147 with orthologs in *P. falciparum* 3D7, *P. vivax* Sal-1 and *P. chabaudii*.

PY04372 has orthologs in all Plasmodium species and is designated as a parasitophorous vacuole protein 2 (putative). There is no conserved motif within the protein. The protein is more conserved among the rodent malaria parasites than with primate malaria parasites. The ortholog of this protein in *P. falciparum* was identified in a BioID screen as parasitophorous vacuole protein 2 in 2017

domains and generated the recombinant protein based on the longest protein sequence between the transmembrane regions for immunization.

PY06306	MGVDWNLISIKKIEISKQLNQLNKIQNEFGSYGGRKNRSINLE----	IHNFEFK	54
PCHAS_0207400	MGDIDVLSIKKIEATKSLHTKFNWQSEFGYGCGIKNRNINLE----	INSNEDFQ	54
PF307_0104100	MVDNDLVLEKKIEKDFSLTDLNRNISHIINNEFGYQENKINFSNNISSYNDIDFSK	60	
PVX_081555	MDSYENLSIDKKITELLKHLKAIKAMHNEFGYFCSGGVIVQD-----	HPNEFSK	53
	* * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	SIFKRLYSSNRMEDLNFNFGSVIKIMERNPYVIFFFHFMINIFIVYLSIFLYTKNFKK	114	
PCHAS_0207400	TIPIKRLYSSNRMEDLNFNFMVSMIRIMERNPYVIFLLIFVLLIFLYTSFPLYTKCFRR	114	
PF307_0104100	STLNNLYSKLEDFSHFSSDLSIDTLKRNQSYMIFVLLIFTSFYLTLLLYTKCIKT	120	
PVX_081555	TIWISLYTSNRMEDTRNLKSVLNLKRNPNYMGCVLIFTFTCYVLLTFLYTKCFRR	113	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	LLKKFSNSHK-----NNKDEEDWKKNKAYNSNSHTGTINKDNYNQELDHNSENEN	171	
PCHAS_0207400	LLKKLNSPKNNNNNKDEENAKTNKAYNRSTSTNGNSVKENYNNELEDNNSNGNEN	171	
PF307_0104100	TLKKWFCRYCSENINE-----	INSNHNQ	144
PVX_081555	LLKAIRKTKCRKKK-----	RQEESEN	137
	* * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	SNVIVIKKRAYNLVINLVCFSFLICLIFLGWITFIFADTQKGINMNICGLSKTVEQFL	231	
PCHAS_0207400	SNVITIVKKRAYNLVINLVCFSFLCLICLIGWIFIFIDAQKGINMNICGLSKTIEKFL	234	
PF307_0104100	RTVLQWINKSCYFITYSSIIICLLFLLSGITYHYMKITKKGISHINCIYTRDLCF	204	
PVX_081555	KDLIENKGRFNFIMTVFLSLLCVLGLGIVNLSYFFKTRKGIYMWVCASSTIENFL	197	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	IDKCPD-TKNWNPQCYSLHEVINDAVSWNQYQLTKFVNKNTLNNKGLPIVLKYQTG	290	
PCHAS_0207400	IDKCPD-TKNWNSCYSLHEVNDITVTVNQYQLTKFVKIKTNNHNEGFPILLKYQTG	293	
PF307_0104100	LNNKLD-PKVDTSCYSAEHLDTLSNLEYKVKVQKQDOLLDETNPFLPLYRYITT	263	
PVX_081555	TDPCSVQGGVESSCYSLEHVSDAVSIIEQYQIKQLKADLLVEDORAVPLLTGFLTV	257	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	FNNLAKLRDNDIKNNVKLENGLYHTYVPLTKLRFLEIVSKGENLLNQAESIIDSKEE	350	
PCHAS_0207400	FNNLAKLRSNIEKNKKLEDELYHTYVPLTKLKSADNIWITGDDLQAESINTSKNE	353	
PF307_0104100	FNNLNVLNKNNNNTTLENGLYHTYPAKIGVLSLTTIIESENGFNARNVKEVST	323	
PVX_081555	FNNLKLQWVARNNHLEEQYHTYVPLTRLGRALDAVQEGEANLQOATGTLEAKQA	317	
	* * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	IGKIFNNVDNAIANTNNVPSLSKISGLGIYIKQDENLKIRYILNKFTVTMIILSTV	410	
PCHAS_0207400	IGNLNNVDNTIINSVDTNLSNLTKIDNLGLIKAYDKIKFIRSLNKFATAMTLNIM	413	
PF307_0104100	IKYFSVFTDTRNVLGFSVPKLTGLTQAKSGXGINNYKIKFIRPKYNTIILTNII	383	
PVX_081555	VKGAFEDIEIRVFTNFKENYNDKTIITFNKSNIRITHYKIKQNLKQISLIVLKV	377	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	IILFSLVLIGLMSYMYFLIRGHSI-NEKFFSKLLGFSGTGFLAIILITGALLS	469	
PCHAS_0207400	IILISLVILVMSNSMFFVIGNM-DQNFSLGAIIRFSGTGFLAIILITGTVLILS	472	
PF307_0104100	LLLPFLILGLIIGTIFFTLNGYQKNNFFDLGFSAYFGLTIVILISFGLILSTS	443	
PVX_081555	LLPPLPILIGLVLFFIFFLVGDIGNSSHFGLGFSAYFGLTIVILIGLILSTS	437	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	VLGGTSCISDRILKNEFTDFLSENKIGYQLQNPDESINKNIVKKYANTLDSLNTDI	529	
PCHAS_0207400	VLGGTSCISDRILKNEFTDFLSENKIGYQLNPEALDLSLTITMYMESLSDNTDI	532	
PF307_0104100	VLGGTSLSERLKLNFIDLNTLIDYKNNESPAIDDDITTFVAKINSPTDGIH	586	
PVX_081555	ILGGTTCIADRVKLNELFVNLDTLIDYCKNNEESPLSEDITGLVGNMKSLODTEM	497	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	YNSVEGYSYFDIKDEYQKSHKIINENMIIP-TDNNYKVMVSDIKKSLTGCLT	588	
PCHAS_0207400	YNSIAKYNSYFDIKIDGYKQSKQKIDYINWIP-TDNNIENPADEALVKSLQGTCLT	593	
PF307_0104100	DHNIENYEHTFIDKISFFHSLKMFYDINIVKIMERNFTNLIRTEQVKSLITIGI	561	
PVX_081555	ERRVMEYDSYFDMKRTFRENTNFVNMVWITKPNNLVYDRILNLTLLKSLATSIT	557	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	KESAQFEYHLMGTDAMKYINKFGL-LNNYEMCFEDPSCENDRKYNNYNSKVTPDKY	647	
PCHAS_0207400	RENQFEYQYLMGNTIMYNYLNKFGL-SDNYEMCFEDESCENDPNYNYKSVTDKAY	650	
PF307_0104100	RENKIKYENHAEIGTSYLLTTLNKIPFENNGICFENIDICEENNTYNTENSKTDDKY	623	
PVX_081555	NNIKIFGKFNKLEVDYFENLNPHYRGQTALCFENECDQFENKFNIRSSINDPKY	617	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	LDVKRNRVLYQSDDFNLVFLVSKKINNDDKFNISDLETKENITWREYTPKNG--	705	
PCHAS_0207400	ITVNRNKRLYSDRDVDVLEIMALKAKINDKIDFVDRLEDEKTEKLSQEYTPQNE--	708	
PF307_0104100	RNIRDGM-DEHLNDLDAVQLFVYKIRKENIFINDLSDNEKNGISEYTPRNI--	680	
PVX_081555	QMRNHLRNLDRELDNDFVYKISVRTEKIFSVDLSDSMTEKISEYTPRINKR	677	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		
PY06306	-AGENKKSIVQTYFEKAEYMKFENWLTLLKEVMHNSKFNWIEKANSLVDNTNCSRF	764	
PCHAS_0207400	-PGSSNISIIQTYFEKAEYNEFIILTFREVSNIHLKFNWIEKESLENTKCNRL	767	
PF307_0104100	-NGTQKSIINTFLVMEINSEINSEINFDDQNFKNWLDLKLDTLTNTCKNRL	739	
PVX_081555	EGGKQTFLLRYLVEDIENFNKDFVSFFEKIKQFNTRLDDITIKVQMLVKNTNCSRL	737	
	: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *		

Figure 9: Alignment of protein sequence of PY06306 with orthologs in *P. falciparum* 3D7, *P. vivax* Sal-1 and *P. chabaudii*.

4.3 Generation of Halotag recombinant proteins for immunization

In order to generate recombinant proteins for use in immunization of mice, The appropriate online software, such as PlasmoDB, Tmpred and SPLIT, was used to identify the soluble and hydrophobic regions such as signal peptide and transmembrane regions in these proteins. The regions selected for generating recombinant proteins are highlighted in Figure 10B. Using the online motif finding software Pfam, PY04147 is the only protein predicted to have a L-type lectin fold which is similar to the ERGIC system in eukaryotes.

Figure 10: Schematic diagram of the shortlisting process and region of peptide generation. (A) The 202 proteins were ranked according to their scoring in the scoring matrix and given a score of 1 – 10. The 24 proteins with a score of 8 or higher were selected and screened against another set of criteria. This resulted in the final 3 shortlisted protein candidates. (B) Regions selected for making recombinant protein, omitting transmembrane region and signal peptide.

Many *Plasmodium* proteins are insoluble when expressed as recombinant proteins using the *Escherichia coli* (*E. coli*) expression system [146]. Soluble proteins are easier to work with for immunization and purification, insoluble proteins in inclusion bodies can also be used for immunization but most of the time other proteins are also present in these inclusion bodies. In order to increase the solubility of these recombinant proteins for purification, I used the Halotag system, which include the addition of a soluble tag in the fusion protein thereby aiding the solubility of the protein. In addition, the covalent binding of the Halotag to the Halotag resin allows for more stringent washes and thus increasing the purity of the protein yield. To generate recombinant proteins, 1000bp gene segment corresponding to the region outside of the transmembrane domain of PY04147, PY04372 and PY06306 were PCR amplified using genomic *P. yoelii* genomic DNA. The resulting inserts were ligated into the expression vector pH6HTN His6HaloTag T7 in between the SacI and NotI sites and expressed as recombinant proteins (Appendix 2). The recombinant proteins were induced by adding 0.02% rhamnose, 0.005% glucose at 25°C for 18 hours. Successful induction was determined by western blot probed with Halotag TMR Ligand stain. The detection of protein bands around 60 – 70kDa indicated the presence of the induced recombinant proteins (Figure 11). Purification of Halotag proteins in native conformation is required for the Halotag to bind to the resin. For PY04147A, the majority of the expressed proteins were found in the soluble fraction after cell lysis. However, PY04372A and PY06306A were not expressed in soluble fraction after lysis, suggesting that these two proteins are extremely

insoluble such that the addition of a soluble tag also fail to make it soluble. Therefore, only PY04147A was subjected to Halotag purification. The other two proteins were reconstructed into another expression vector pET24a and generated as Histag proteins.

Figure 11: Western blot images of Halotag induction. Protein induction was done using 0.02% rhamnose, 0.005% glucose and incubating cells at 25°C for 18 hours. Lysed bacteria pellets were subjected to SDS-PAGE and probed with Halotag- TMR Ligand. PY04147A was found to be in the soluble fraction after lysis in purification buffer. PY04372A and PY06306A were present but not soluble in the purification buffer. Arrows point to the size of the induced recombinant protein. Molecular sizes are indicated on the left (in kDa).

4.4 Purification of Halotag proteins

Purification of PY04147A was performed using the Halotag purification kit with modified protocol. The cell pellets were first lysed in B-per solution containing 1X protease inhibitors. The supernatant was collected and incubated with Halotag resin for 1 hours with rotation. Bound proteins and resin were collected by centrifugation at 10000 rpm for 15 min. Washes were carried out 3 times using the different wash buffers (Appendix 1) . After washing, the resin was incubated with TEV proteases to release the protein from the resin. The resulting

supernatant containing TEV and recombinant protein was harvested and incubated with Halolink resin to remove TEV.

A modified 2-round TEV removal was used as TEV proteases used in the cleavage of the Halotag were found to be present after the first TEV removal step as shown by the presence of a 50kDa band in the E2A lane (Figure 12). The remaining TEV proteases was removed after a second round of TEV removal, leaving the purified proteins as indicated as the E2B band in the figure.

Figure 12: Coomassie Blue stained gel of different fractions obtained after Halotag purification. Abbreviations: FT, Flowthrough collected after proteins were bound to beads; R, resin alone, to check for incomplete protein elution; TEV, TEV protease used for cleaving the Halotag off the bound protein; E1, eluted sample with TEV protease; E2A, 1st elution after TEV removal; E2B, 2nd elution after TEV removal. The arrow points to the size of the purified tag-less PY04147A proteins. Molecular sizes are indicated on the left (in kDa).

4.5 Generation of Histag recombinant proteins

Proteins PY04372A and PY06306A were not soluble when fused to Halotag and therefore cloned into pET24a and generated as recombinant proteins with a C

terminal Histag instead. MSP1-19 has been previously purified as a Histag fusion protein in our laboratory and therefore new batch of MSP1-19 was generated using the same approach for the current work. Gene segments from PY04372, PY06306 and MSP1-19 were PCR amplified from genomic DNA of *P. yoelii*. The amplified insert was ligated into a pET24a vector in between the XhoI and NdeI sites and expressed as Histag recombinant proteins (Appendix 3).

Protein induction was carried out using 1mM IPTG at 37°C for 4 hours. The expression of the recombinant protein is confirmed by western blot and probed with anti-His antibodies (Figure 13). The expected size of the purified recombinant proteins was detected with a single band of 19 kDa in MSP1-19 and of 30 kDa in PY04372A, respectively. A major portion of purified protein with a correct size of 34 kDa was detected in PY06306A. While there was a smaller fragment also detected by anti-His antibodies in PY06306A, suggesting that this protein is partially degraded during the expression process. The purity of the protein required for polyclonal antibody production is >70% purity. Since majority of PY06306A protein is still the full-length protein of interest, was used in the following immunization studies.

Figure 13: Recombinant protein expression in *E. coli* cells after induction. Induction was carried out by using 1mM IPTG and incubating cells at 37°C for 4 hours. Western blot analysis with anti-His antibodies confirm the expression of the recombinant proteins. Proteins were probed with mouse anti-His primary antibodies at 1:5000 dilution and goat α mouse IgG peroxidase conjugated secondary antibodies at 1:5000 dilution. Expected size of the recombinant protein was detected at 19 kDa for MSP1-19 and 30 kDa for PY04372. A smaller fragment was detected by the anti-His antibodies in PY06306A, suggesting that partial degradation of the protein had occurred during the expression process. Molecular sizes are indicated on the left (in kDa).

4.6 Buffer optimization of Histag proteins

To produce soluble proteins, buffer optimization was carried out to determine the buffer conditions for purification. Depending on the isoelectric point (pI) of the proteins, different buffers were tested (Table 8). The pI of the proteins was predicted based on the online software ExPASy. For PY04372A and PY06306A, Tris-HCl based buffers with effective pH range between 7.5 to 9.0 was used. For MSP1-19 which has a relatively lower pI, buffers with effective range of 6.5 – 7.4 were used.

Protein	Isoelectric point	Buffers used
PY04372A	9.34	B1: 50mM Tris HCL, 200mM NaCl, 0.5mM DTT, pH 8

		B2: 50mM Tris HCL, 200mM NaCl, 0.5mM DTT, pH 8.5
		B3: 50mM Tris HCL, 500mM NaCl, pH8.5
PY06306A	6.16	B1: 50mM Tris HCL, 200mM NaCl, 0.5mM DTT, pH 8
		B2: 50mM Tris HCL, 200mM NaCl, 0.5mM DTT, pH 8.5
		B3: 50mM Tris HCL, 500mM NaCl, pH8.5
MSP1-19	5.19	B4: 50mM Na ₂ HPO ₄ , 150mM NaCl, pH 7.4
		B5: 50mM HEPES, 150mM NaCl, pH 7.0
		B6: 50mM MOPS-KOH, 250mM NaCl, pH6.5

Table 10: List of proteins, their corresponding pI and the type of buffers used for the buffer optimization.

Equal amount of bacteria pellets were resuspended in the different purification buffers and subjected to sonication for 20 min with amplitude 35%, pulser 6 sec on & 6 sec off. The resulting solution was centrifuged at 13000g at 4°C to separate the pellet from the supernatant. SDS loading dye was added to the pellet and the supernatant and equal amount of each fraction were used to run the SDS-PAGE. Presence of protein in each fraction was determined by western blot probed with anti-His antibodies (Figure 14).

The western blot images showed that MSP1-19 was soluble in all the buffers used while PY04372A and PY06306A were not. For MSP1-19 the preferred buffer used in the purification is Buffer 5 under native condition while PY04147A and PY06306A Buffer 3 was used under the denaturing condition.

Figure 14: Western blot image of buffer optimization of proteins probed with anti-His antibodies. Bacteria were lysed in different purification buffers and the pellet (P) and supernatant (S) were loaded onto SDS-PAGE. MSP1-19 is soluble in all buffers tested while the PY04147A and PY06306A were not. Molecular sizes are indicated on the left (in kDa).

4.7 Purification of Histag proteins

Large scale purification of the proteins was done by inoculating cells at 1:100 dilution in 2 liter of LB media. When OD reaches 0.6 – 0.8, 1mM IPTG was added to induce protein at 37°C for 4 hours. After incubation, the cells were centrifuged at 5000rpm to pellet the cells and approximately 12g of cell pellet can be obtained per 1 liter culture. Cells were sonicated for 20 minutes with amplitude 35%, pulser 6 sec on & 6 sec off in the appropriate buffer as determined in the buffer optimization. The cell lysate was centrifuged at 13000g, 4°C for 20 minutes to separate the supernatant from the pellet. Depending on

protein solubility in the different buffers, the purification steps differ for each of the proteins.

4.7.1 Purification of MSP1-19

For MSP1-19 which is soluble in Buffer 5, the supernatant was collected for Ni-NTA purification. Ni-NTA beads were added to the supernatant and the mixture was incubated at 37°C for 20 minutes to allow the Histag-proteins to bind onto the NI-NTA beads. After binding, the Ni-NTA mixture was loaded onto a column and allowed to drain. The flow through was collected and kept for SDS-PAGE (FT1). After the column is drained, it was washed extensively with buffers as stated below. All the washing flow through were kept.

1. 20ml buffer 5 and collect the flow through (FT2)
2. 20ml buffer 5 + 10mM imidazole and collect the flow through (FT3)

Elution of the proteins using imidazole gradient was done as follows:

1. 4ml buffer 5 x 40mM imidazole (S1)
2. 4ml buffer 5 x 60mM imidazole (S2)
3. 4ml buffer 5 x 75mM imidazole (S3)
4. 4ml buffer 5 x 100mM imidazole (S4)
5. 4ml buffer 5 x 200mM imidazole (S5)

The protein purity was verified by SDS-PAGE. Each fraction was loaded onto the 12% SDS gel and followed by the Coomassie Blue staining (Figure 15). Wash-through fractions FT1, FT2 and FT3 contains many proteins which did not bind the beads (Figure 15A, left panel). An expected size of 19 kDa band was

shown in the Elution fractions of S1 to S5 (Figure 15A, right panel). However, there were many contaminant proteins in fractions S1 and S2 and therefore these fractions were not pooled together for further procedures. Fractions S3, S4 and S5 were pooled together as they were relatively pure and was further passed through 30kDa-cut off concentrator to remove any contaminant proteins bigger than 30kDa. The protein was then passed through a 10kDa-cut off concentrator to concentrate the proteins. The final purified protein was subjected to SDS-PAGE to check for purity (Figure 15B). After Coomassie Blue staining, a single band of about 19kDa was shown, indicating that the MSP1-19 protein is pure and suitable for subsequent vaccination.

Figure 15: Coomassie Blue stained gels of MSP1-19 purification. (A) Fractions from each of the purification steps. FT1 – 3 shows the flow through during the initial washing steps. S1 – S5 shows the elution profile of MSP1-19 during elution with an imidazole gradient. (B) The purified MSP1-19 protein after concentration. Molecular sizes are indicated on the left (in kDa).

4.7.2 Purification of PY04147A and PY06306A

For the insoluble PY04147A and PY06306A proteins, after sonication in buffer 3, the cell lysate was centrifuged at 13000g, 4°C for 20 minutes. The pellet was

collected and incubated with 8M urea in Buffer 3 for 1.5 hours at 37 °C to solubilize the pellet. The solution was centrifuged at 13000g, 4°C for 20 minutes and the supernatant was collected and incubated with Ni-NTA at 37 °C for 20 minutes to allow the Histag-proteins to bind onto the NI-NTA beads. After binding, the Ni-NTA mixture was loaded onto a column and allowed to drain. The flow through was collected and kept for subsequent SDS-PAGE analysis (labeled as FT1). After the column is drained, it was washed extensively with buffers as stated below.

1. 60ml buffer 3 + 10mM imidazole and collect the flow through (FT2)
2. 60ml buffer 3 + 20mM imidazole and collect the flow through (FT3)

Elution of the proteins using imidazole gradient was done as follows:

1. 4ml buffer 3 x 40mM imidazole (S1)
2. 4ml buffer 3 x 60mM imidazole (S2)
3. 4ml buffer 3 x 75mM imidazole (S3)
4. 4ml buffer 3 x 100mM imidazole (S4)
5. 4ml buffer 3 x 200mM imidazole (S5)
6. 4ml buffer 3 x 500mM imidazole (S6)
7. 4ml buffer 3 x 1M imidazole (S7)

Each fraction was loaded onto SDS-PAGE for checking of protein purification (Figure 16A and 16C). For PY04372A, majority of the contaminant proteins were washed away in fractions FT1 and FT2. Fraction FT3 is relatively clear and indicates that most unbound contaminant proteins were removed by the washes.

Elution Fractions S1 – S5 were pooled together as the elute is relatively free of contaminant proteins.

For PY06306A, majority of the contaminant were washed away in FT1 and FT2. In fraction FT3, with 20mM imidazole we managed to remove more unbound contaminant proteins with some of the bound PY06306 protein eluting out as well. Elution with increasing concentration of imidazole resulted in the elution of pure PY06306A proteins in fractions S1 – S4 and these were pooled together.

The protein was then passed through a 10 kDa-cut off concentrator to concentrate the proteins. The final purified protein was subjected to SDS-PAGE to test for purity (Figure 16B and 16D). Due to the use of the concentrators, we can remove most of the larger contaminant proteins and the resulting PY04372A and PY06306A proteins are relatively pure and suitable for subsequent vaccination.

Figure 16: Coomassie Blue stained gels of PY04372A and PY06306A purification. Fractions from each of the purification steps. FT1 – 3 shows the flow through during the initial washing steps. S1 – S7 shows the elution profile of the proteins during elution with an imidazole gradient. (A) Purification profile of PY04372A. (B) The purified PY04372A protein after concentration. (C) Purification profile of PY06306A. (D) The purified PY06306A protein after concentration. Molecular sizes are indicated on the left (in kDa).

4.8 Recombinant proteins were recognized by protective immune sera

The recognition of the purified proteins by the protective sera generated previously in our laboratory was validated by western blot (Figure 17). All the purified proteins were recognized by the three protective sera except PY04147A which is recognized by all but the self-cured sera. None of the proteins were

recognized by the none-protected sera, indicating that mice not able to cure the infection did not develop antibodies against these antigens. The differential recognition by the immune sera gave us an indication that these proteins may play a role in protective immunity since they are only recognized in mice protected from the infection. Similar to PY04147A, MSP1-19 was only recognized by Chloro-treated sera and the Attenuated sera.

Figure 17: Recombinant proteins were probed with mice sera generated in previous screen. Both PY04372 and PY06306 were recognized by all 3 protective immune sera, PY04147 is recognized by 2 out of 3 of the sera. None of the proteins were recognized by the non-protected sera. Molecular sizes are indicated on the left (in kDa).

4.9 Immunization regime and antibody titer

Six groups of mice (n = 5) were immunized according to the protocol in Figure 18A. Male BALB/C mice were immunized 5 times with recombinant proteins emulsified in either CFA or InFA (1:1). In control group, mice were injected with PBS emulsified in adjuvants. The amount of antigens injected per mice is 50ug per injection. For PY06306, immunization trials with 25μg and 50μg of protein

did not induce high antibody titers in some mice (Appendix 4) and thus 100µg was used for injection (Figure 18B). All injections were performed by subcutaneous injections to mimic the delivery methods for use in humans.

Figure 18: Immunization regime and amount of antigens injected. (A) Mice were immunized at 2 weeks interval and antibody titer against the protein was measure at week 10. Challenge of immunized mice with YM was done 2 weeks after the last injection. (B) Amount of antigens injected during each injection. CFA = Complete Freund’s adjuvant, InFA = Incomplete Freund’s adjuvant.

4.10 Antibodies generated from mice are specific to the proteins

ELISA was carried out on week 10 to test the antibody titer. The ELISA plate was coated with 200ng of proteins per well. 100ul of blood from the immunized mice were collected via retro-orbital bleeding and centrifuged at 1000g for 5 minutes to separate the serum from the blood. The sera were collected and diluted in PBS and different dilutions were used in the ELISA. The final antibody titer of the ELISA is defined as the final dilution that gives a reading more than 3x the SD of the mean of sera from mice in the PBS control group. All 3 antibodies give

a titer between the range of 1:32000 to 1:128000, although there was one mouse in each of the 6306 immunized group that showed a lower antibody titer of 1:4000 and 1:8000. Titer for MSP1-19 is about 1:4000. ELISA results showed that the antibodies generated from the mice are specific to the proteins injected into the mice (Table 11).

Immunization set 1					
	4147	4372	6306	Mix	MSP1-19
Mouse 1	1:64000	1:128000	1:64000	1:16000	1:4000
Mouse 2	1:64000	1:64000	1:32000	1:8000	1:4000
Mouse 3	1:64000	1:64000	1:64000	1:4000	1:4000
Mouse 4	1:64000	1:64000	1:8000	1:32000	1:4000
Mouse 5	-	1:64000	1:64000	1:8000	1:8000
Immunization set 2					
	4147	4372	6306	Mix	MSP1-19
Mouse 1	1:16000	1:128000	1:4000	1:4000	1:4000
Mouse 2	1:64000	1:128000	1:16000	1:32000	1:4000
Mouse 3	1:64000	1:128000	1:32000	1:4000	1:4000
Mouse 4	1:64000	1:128000	1:32000	1:4000	1:4000
Mouse 5	1:64000	1:128000	1:64000	1:32000	1:4000

Table 11: ELISA results of antibody titer against the individual proteins used during the immunization. Two sets of mice were injected immunized with the proteins and the corresponding antibody titer at week 10 were noted.

4.11 Antibodies recognized full-length proteins in *P. yoelii* schizont lysate

The ELISA result showed that the mouse antibodies specifically recognized the respective proteins. To further confirm the specificity of the polyclonal antibodies, the *P. yoelii* parasite lysates were tested in the western blot by these polyclonal antibodies (Figure 19). A 53 kDa band was detected in PY04147 and PY04372 respectively, while a 91 kDa band was detected in PY06306, which is in line with the predicted full-length protein size obtained from the online tool ExPASy. Taken together, these data indicates that the polyclonal antibodies are specific.

Figure 19: Western blot using polyclonal antibodies probed against *P. yoelii* extract. The antibodies generated from immunization recognized a single band in the lysate which corresponds to the size of the full-length protein. Numbers in brackets represent the dilution factor of the serum used. Arrows point to the full-length protein. Molecular sizes are indicated on the left (in kDa).

4.12 Localization studies of proteins of interest in *P. yoelii*

The specificity of these antibodies was confirmed by ELISA and western blot of *P. yoelii* lysate. The antibodies were further tested in IFA to determine the localization of each protein within the parasite throughout its lifecycle.

4.12.1 Proteins were not located on the surface of infected erythrocytes

To visualize the localization of the three proteins, IFA was performed on different stages of the parasites. The three proteins were observed as a dot pattern within the parasite during the ring stage (Figure 20A). In the trophozoite stage, the localization of the 3 proteins differ from each other (Figure 20B). PY04147 gave a punctate like staining within the parasite while PY06306 showed a diffused staining pattern. For PY04372, the protein mostly locate to the PVM and formed a circle around the parasite. None of these three proteins were seen in the HCC either in the ring or trophozoite stages indicating that these proteins were not exported during early stages of intraerythrocytic development. All three proteins

have diffused staining pattern in the schizont stage (Figure 20C) but are still localized within the parasites as live imaging of intact schizonts probed with antibodies does not yield any surface staining.

B

Trophozoite Stage

Figure 20: Antibodies generated in mice immunized by the different recombinant proteins recognized the native proteins expressed by *P. yoelii*. Pooled sera from immunized mice were diluted to a concentration of 1:400 and probed against the different stages of *P. yoelii* on glass slides. A) Ring stage, B) Trophozoite stage, C) Schizont stage. Bound antibodies were stained with α -mouse IgG Dylight 594 labeled Secondary antibodies, the parasite nuclei were stained with Hoechst.

4.12.2 Proteins were located within individual merozoite

Since proteins were not expressed on the surface of infected erythrocytes, the expression of these proteins in merozoites were investigated. Schizonts were allowed to mature after Histodenz purification. Thin blood smears of the schizonts were made on glass slides and in the process the schizonts will burst releasing the merozoites onto the slides. The slides were fixed in 4% PFA and probed with antibodies against the proteins of interest (labelled Green) as well as MSP1-19 (labelled red). MSP1-19 labels the surface of merozoites and helps to locate merozoite boundaries. As shown in Figure 21A, PY04147 is seen as punctate dots staining within the merozoites. Similar punctate staining within the merozoite was observed in PY04372 suggesting that these two proteins are localized in membrane bound organelles (Figure 21B). Whereas PY06306 showed a diffuse pattern of staining within the merozoites (Figure 21C).

Figure 21: IFA showing co-staining of MSP1-19 and proteins of interest. MSP1-19 proteins were stained with α -MSP1-19 rabbit antibodies and α -rabbit IgG Dylight 594 labeled Secondary antibodies. Pooled sera from immunized mice were diluted to a concentration of 1:400 and probed against the different stages of *P. yoelii* on glass slides. Bound antibodies against the proteins of interest were stained with α -mouse IgG Dylight 488 labeled Secondary antibodies. Parasite nuclei was labeled with Hoechst. All three proteins were expressed in the merozoite stage and were located within the merozoite.

4.12.3 Proteins were not localized in the rhoptries of *P. yoelii* merozoites

As the proteins were found in the merozoites, we are interested to determine if it is localized to any of the known organelles in the merozoites. MAEBL is a protein located in the rhoptries of the merozoites[147] and is used as a marker in IFAs.

Co-staining with anti-MAEBL antibodies showed that the PY04147 protein was not localized together with the distinct two-dot staining pattern typical of rhoptry staining (Figure 22A). Similar staining was also seen for PY04372 in which the punctate staining of the PY04372 protein was not co-localized with MAEBL in the rhoptry (Figure 22B). PY06306 has diffused staining within the merozoite and was not localized to the rhoptry (Figure 22C).

Figure 22: Antibodies specific against the recombinant proteins were co-stained with MAEBL to check for co-localization in the rhoptries. MAEBL proteins were stained with α -MAEBL rabbit antibodies and α -rabbit IgG Dylight 488 labeled Secondary antibodies. Pooled sera from immunized mice were diluted to a concentration of 1:400 and probed against the *P. yoelii* on glass slides. Bound antibodies against the proteins of interest were stained with α -mouse IgG Dylight 594 labeled Secondary antibodies. Parasite nuclei was labeled with Hoechst.

4.13 Proteins did not confer protection against a lethal YM challenge

ELISA analysis showed that antibodies against the recombinant proteins were successfully produced in immunized mice on week 10 (Table 9). To explore the whether these antibodies have any protective effect on the parasite infection, immunized animals were challenged with 10^2 *P. yoelii* YM parasites two weeks after the last immunization (Figure 23). YM parasites have been shown to be lethal [148] and usually kills the mice within 8 dpi. Although the proteins used in this immunization were found using infections with 17X 1.1 mice, we decided to use YM for our challenge studies to subject the immunization to a more stringent test. In addition, it has been demonstrated that mice protected against 17X 1.1 infection were also similarly protected when challenged with YM parasites [149]. This implies that the proteins mediating protection against these two parasite clones are similar. Furthermore with YM parasites infection, the protective phenotype can be easily observed as mice which are not protected will succumb to the infection.

Mice in the control group which were injected with PBS died from YM infection at 7 dpi (Figure 23Ai, Aii) while mice immunized with MSP1-19 had a survival rate of 80% with the majority of the mice showing very low levels of parasitemia which was eventually cleared after 2 weeks (Figure 23Bi, Bii). None of these proteins identified in this proteomic screen confers any protection in mice and all the mice died at the same time as the PBS control group. Both of the parasitemia growth from mice immunized with PY04147A (Figure 23Ci, Cii) and PY04372A (Figure 23Di, Dii) were similar to that of the control, with parasitemia detected at 4 dpi and mice dying at 7dpi. Although 4 out of 5 mice immunized with

PY06306A died at the same time as the control, there is one mouse in the group that showed a delayed in parasitemia (Figure 23Ei, Eii). When the challenge was repeated, this delayed in parasitemia was again observed in one of the mice (Appendix 5). However when we used a cocktail mixture of both PY04147A and PY06306A for immunization, the delayed in parasitemia seen in mice immunized with only PY06306A was abolished and all the mice died by 7 dpi (Figure 23Fi, Fii).

We initially speculated that the antibody titer of the mice which showed delay in parasitemia should be higher than that of mice which do not show the delay. We checked the ELISA results for the antibody titer of mice which had delayed parasitemia and those which didn't and found that mice that showed a delay in parasitemia did not have higher antibody titer than the ones that died at the same time as the control. The mice with a delay in parasitemia had antibody titer of 1:32000. Although 1 of the mice that has no delay in parasitemia has antibody titer of 1:8000, the other 3 had antibody titer of 1:64000. This indicated that higher antibody titer is not the main reason for the delayed parasitemia observed.

Since the antibody titer is not the main reason for the parasitemia delay, the other reason could be the difference in epitopes that the antibodies recognize. The mice that had delayed parasitemia might have generated antibodies targeting a specific epitope that slows down the parasite growth initially.

Figure 23: Growth of *P. yoelii* YM parasites in mice immunized with different proteins and the corresponding survival curve. Mice from all the immunized groups were challenged with 1×10^2 *P. yoelii* YM 2 weeks after the last immunization. Immunization with MSP1-19 protected the mice from lethal YM challenge while all other proteins do not confer any protection.

4.14 Discussion

My work aims to validate the antigen identification process previously conducted by our laboratory. The proteomic screen revealed a set of 202 proteins that were recognized by the different protective sera and therefore might be associated with protection. To increase the relevance of the screen for analysis in human malaria in the future, only proteins with orthologs in *P. falciparum* were considered and selected in my work.

In this study, three proteins (PY04147, PY04372 and PY06306) which are highly associated with protection were identified and generated as recombinant proteins. All 3 proteins were uncharacterized previously and online domain prediction tools did not manage to identify any functional domains in these proteins except for PY04147. The PY04147 protein is approximately 59 kDa in size and contains a predicted Concanavalin A-like domain near the N terminal of the protein. Proteins belonging to the Concanavalin A-like lectin super family are carbohydrate binding proteins which are involved in many different types of cellular function in both prokaryotes and eukaryotes. One of the function is for adhesion to carbohydrates found on cell surface and associated with binding of sialic acid in rotavirus invasion[150].

The Concanavalin A-like lectin domain within PY04147 was highly homologous (>40% similarity) to other *Plasmodium* species that can infect humans such as *P. falciparum*, *P. vivax*, *P. cynomolgi* and *P. knowlesi*. We generated and expressed this domain in PY04147 as a recombinant protein named PY04147A. For the

other two proteins, regions excluding the transmembrane domains were generated as recombinant proteins.

Once the proteins were successfully expressed and purified, they were validated by with the protective immune sera generated previously. All the 3 proteins were recognized by at least 2 out of 3 of the immune sera, indicating that the proteins might be associated with protection.

Immunization of mice was carried out at two weeks interval with the recombinant proteins emulsified in Freuds adjuvant. ELISA was performed to test the antibody titers from each mouse at one week after the last immunization. The ELISA results showed that these proteins are immunogenic, and the antibodies were successfully generated from the immunized mice. The specificity of these antibodies was validated by western blot and IFAs. Pooled sera from the immunized mice with the same protein were used. All the sera specifically recognized the corresponding full-length protein in the *P. yoelii* lysate by western blot.

We next proceeded to use the antibodies to check for the protein localization in the asexual blood stage. IFA was conducted on acetone fixed parasites and all three proteins were detected throughout the asexual lifecycle (Figure 20). Moreover, all the 3 proteins were found inside the merozoite when co-stained with anti MSP1 antibodies (Figure 21). The diffused staining of PY06306 within the merozoite indicated that it might be within the parasite cytoplasm. The clustered staining of PY04147 and PY04372 proteins indicates possible localization to a specific organelle in the merozoite. Co-staining with anti

MAEBL (a rhoptry protein) antibodies showed that all three proteins are not localized in the rhoptries.

Our previous data showed that these 3 recombinant proteins are immunogenic and antibodies against these 3 proteins are specific. Therefore, we would like to address whether these antibodies have any protective effect on the parasite infection. After mice immunized with the proteins were challenged with 10^2 YM parasites, none of the proteins protected the mice from the infection and most of the mice died at the same time as the control group. However, for mice immunized with PY06306A, one of the mice has a delayed parasitemia and this result is reproducible. When compared with the ELISA antibody titer results, the antibody which has a higher titer had no protection and/or effect on the delay in parasitemia and all the mice died. This suggests that the epitope specificity of the antibodies generated might play a crucial role in slowing down the initial growth of the parasites.

The PY06306 protein has been recently patented as an antigen for a cocktail subunit vaccine targeting the pre-erythrocytic stage parasites [145]. According to the patent, the cocktail was able to protect 100% of the CD1 mice from *P. yoelii* 17X 1.1 sporozoite challenge, of which PY06306 alone was able to prevent malaria in 71% of the mice. In mice that still develop blood stage malaria in their sporozoite challenge study, immunization with PY06306 was able to delay the onset of parasitemia. When immunized mice were challenged with 10,000 blood stage parasites, they were able to show that PY06306 alone is able to achieve sterile protection. However, when I did the infection studies, immunization with

PY06306A was only able to delay parasitemia in some mice and no sterile protection was observed. The difference in the protective effect seen might be due to the fact that I used a different method to generate antibodies to the PY06306 protein.

In the study by Aguiar et al., instead of using recombinant PY06306 protein for immunization, they utilized Adenovirus as a vector for PY06306 expression. With this approach they were able to get 100% sterile protection in immunized mice challenged with 10^4 PY17X 1.1 infected erythrocytes. This is in contrast with our results in which I did not see any sterile protection although this can be due to the use of the lethal YM clone instead of non-lethal PY17X 1.1. Another difference between their approach and mine that might affect the protection was that they used the entire protein for immunization while we only use a portion of the protein. As the full-length PY06306 protein is large (~91 kDa), I only generated a 30kDa fragment and thus I may lose the potential protective epitopes presented on other parts of the protein.

When generating recombinant proteins as subunit vaccines I encountered the problem of the protein insolubility as well as the difficulty in the full-length protein expression in *E. coli*. The ability of Adenovirus to accommodate large genomic sequence insertion (up to 8Kb) is an added advantage that allows for the full-length expression of proteins [151]. In addition, adenovirus can mediate strong antibody and T cell responses in malaria [152, 153] and can be easily scaled up for vaccine production. It might be a better vector than the use of

protein in adjuvant subunit vaccine delivery method that we are utilizing currently.

In summary, the reason for the failure of my current selection of antigens to protect mice is multi-factorial. Firstly, in order make recombinant proteins which are easier to be expressed in *E. coli*, I omitted the transmembrane region and this potentially resulted in the omission of potential protective epitopes within these region. A second problem that I have is that these proteins are expressed as insoluble proteins and thus I lose conformational epitopes as well. If I were to start this work from the beginning, I would utilized other vaccination methods such as the DNA vaccine approach utilized by Aguiar et al. DNA vaccines can produce effective immunization by presenting antigens on both MHCI and MHCII [154].

Although my current work did not show that any of the three proteins are protective, it does gives us an indication that the blood stage antigens which are associated with protection might be merozoite proteins. This is supported by the fact that many of the current leading blood stage vaccine candidates are merozoite proteins such as Apical membrane antigen 1(AMA1) and MSP3 [155-157]. From the list of 24 proteins shortlisted in the previous proteomic screen, almost 50% are merozoite proteins (inclusive of the 3 in this screen). It will be interesting to employ more stringent selection criteria on the proteomic data to better filter out proteins which might be protective. As the original mass spec analysis was done several years ago, reanalyzing the mass spec data by incorporating new and updated information of genes from PlasmoDB might yield more information.

Chapter 5

Endogenous expression of a single *yir* gene in *Plasmodium yoelii*

5.1 Expression of constitutively expressed *yir* gene in *P. yoelii*

The *pir* multi-gene family is the largest multigene family conserved in *Plasmodium* species affecting humans, primates and rodents. The inability to do any in-depth studies on *pirs* is mainly due to the lack of reagents and methods to study individual *pirs* during the course of an infection. A particular challenge is that each parasite expresses a different *pir* and that expression of *pirs* can change rapidly. Moreover, the expression profile of *pirs* from a singly cloned parasite can differ from the initial parental profile as early as 10- 12 cycles into the infection [95], making it hard to maintain a homogenous population of cells expressing a particular *pir*. The high sequence homology among members of this family also makes genetic manipulation problematic.

To study these *pirs* in isolation, it is critical to generate distinct *pir* expressing parasites. We have now developed an experimental approach to overcome some of these critical challenges, thereby opening up the potential platform to initiate a comprehensive and systematic study of *pirs*. In this study, the *yir* homolog in *P. yoelii* was used. The main strategy of this study is to endogenously tag a single *yir* gene and subsequently force it to be constitutively expressed thereby creating a parasite line always actively expressing this single *yir*. Using this experimental strategy, we are now able to generate a range of parasite lines that enable us to dissect the function of different members of *pir* and characterize their biological functions.

5.1.1 Generation of *P. yoelii* constitutively expressing YIR parasite lines.

Generation of transgenic parasites requires the cloning of part of the gene in question as well as a drug selection marker for subsequent selection of transfected parasites. These vectors are usually constructed via introduction of a homologous region followed by an independent promoter driving the drug selection cassette. The main disadvantage of this method is that it can lead to the appearance of resistant parasites without transgene integration due to the drug selection cassette being transcribed by its own promoter. Therefore, there is great value in getting the drug selection cassette activated by a single promoter to minimize the wildtype escape. More importantly, the promoter that controls the transgene should be from the gene that we are targeting to maintain the expression pattern of the gene within the parasite.

Skip peptides are small self-processing peptides that are cleaved post translationally in cells and many viruses use it for processing proteins [158, 159]. Skip peptides have recently been utilized as a tool for dual expression of genes from a single promoter in *P. falciparum* [160]. Here, we incorporated the skip peptide approach into the creation of a tagging vector for single crossover integration. The resulting plasmid will have a skip peptide linking a GFP tag to the TgDHFR selection marker (Figure 24A). Multiple cloning sites were inserted just before the GFP tag allowing the addition of any gene sequences that we are interested in tagging. As the entire plasmid does not have any promoter, for the cells to survive pyrimethamine selection, the plasmid must be successfully integrated into the correct locus in the genome for the entire construct to be transcribed.

With this tool, we now intended to use the skip peptide to tag members of the *pir* multigene family in the rodent malaria model *P. yoelii*. The main advantage of using the skip peptide plasmid to tag *yir* genes is that only parasites which have integrations in frame will express the drug selection cassette and survive under the drug selection, thereby reducing the appearance of wildtype escape. Being the largest multigene family in *Plasmodium*, the genetic sequence of *yirs* are highly conserved. Due to the high sequence homology among members of the *yir* multigene family, there is a risk that integration might occur in a closely related *yir* with similar sequence. It is therefore important that the parasites obtained are cloned and the integration is validated by PCR and sequencing.

To construct the tagging plasmid, the last 500bp from the 3' end of the gene of interest (GOI) was cloned into HPXY T2A-GFP plasmid. The resulting plasmid was linearized with HindIII. 12µg of the cut plasmid was used for transfection (Figure 24A). Upon transfection, the entire plasmid will be integrated into the 3' end of the gene by homologous recombination (Figure 24B). The successfully integrated parasites will have a modified locus containing the GOI fused to a GFP tag. The T2A peptide links the GFP tagged GOI with the drug selection cassette and the entire gene will be transcribed and translated as a single protein. Upon translation of the gene, the protein will be cleaved at the T2A site to yield the GFP tagged GOI and the drug cassette. Drug selection was done with pyrimethamine injection via i.p for 3 days post transfection and subsequently on water supplemented with pyrimethamine. The time taken for transfected parasites to appear is approximately ten days post transfection.

We based the selection of these *yir* genes on the phylogenetic tree generated by Otto et al. in which *pir* genes were classified into 8 ‘short’ and 4 ‘long’ form *pirs* [93]. Six *yir* genes (PYYM_0017400, PYYM_0020000, PYYM_0035300, PYYM_0300800, PYYM_0400700 and PYYM_0401200) were selected from PlasmoDB for use in this study. The 6 selected *yirs* represented 2 from the ‘long’ *yir* and 4 from the ‘short’ *yir*. Transfection of the 6 constructs gave us viable parasites for 3 constructs after drug selection and these 3 genes were used in this study. All the 3 *yir* genes successfully tagged belong to the short form *pirs* in which PYYM_0017400 belongs to the clade S4 while both PYYM_0035300 and PYYM_0400700 belong to the clade S8.

Figure 24: Plasmid map of skip peptide and schematic diagram of recombination. Ampicillin resistance, drug-selectable markers and GFP tags are indicated. (A) Tagging plasmid is digested at the HindIII site as indicated before transfection. (B) Schematic diagram of the crossing over integration of the transgene. After single crossover integration, the whole plasmid is integrated at the 3' end of the endogenous gene of interest. Diagnostic PCR primer binding sites are indicated by FP and RP. After translation, the protein is cleaved to yield the tagged GOI and the DHFR resistance cassette. FP = forward primer binding site, RP = reverse primer binding site. GOI = gene of interest, eGFP = enhanced green fluorescence protein.

5.1.2 GFP tagged YIRs are expressed and exported

After selection of the parasites, I am able to get close to 100% of the parasites with in frame integration, which stained positive for Dapi and expressed the GFP fusion YIR (Figure 25). The use of the HPXY plasmid for integration greatly reduces the occurrence of wildtype escape parasites, leading to a large population of parasites that express the GFP fusion protein.

Figure 25: Live cell imaging of transfected PYYM_0017400 parasites at day 7 post transfection. Majority of the cells that survived the selection process expressed the GFP fusion YIR protein.

All three genes were successfully transfected into the parasites with GFP expression seen in all parasites using this approach (Figure 26). In live cell imaging, GFP fusion proteins of PYYM_0017400 was seen to be exported into the host cell cytoplasm (HCC), which is in line with published work showing that YIRs are exported into the host cell [108]. Moreover, PYYM_0017400 also showed a punctate staining within the HCC, suggesting that these proteins might

be contained in some vesicles or membranous organelles. For PYYM_0035300 and PYYM_0400700, the GFP was mainly located within the parasite, in line with observations that not all YIRs are exported and YIRs have different localization as reported previously using episomal tagging [108].

Figure 26: Live cell imaging of HPXY T2A-GFP transfected parasites. Transfected parasites were visualized under fluorescence microscope. GFP signal shows exported tagged YIR proteins as punctate dots within the HCC.

5.1.3 Single cell cloning and validation of clones

Due to the high DNA sequence homology between members of the yir multigene family, it is possible that the integration of the transgene might occur at the wrong locus. Therefore, we performed single cell cloning on all the 3 parasite lines to isolate individual parasites. Genomic DNA from cloned transfectants was recovered and used for diagnostic PCR to check for correct integration. The forward diagnostic PCR primer is gene specific and binds within the gene of interest while the reverse primer is a universal primer binding within the GFP

region (Figure 27A). Only clones with the targeting construct integrated into the correct locus will yield a band in this diagnostic PCR (Figure 27B). Clone 1 and 3 from PYYM_0017400 transfected parasites had the correct band size of 900bp (Figure 27B, left panel). All 3 clones from PYYM_0035300 and PYYM_0400700 had the correct band size of 650bp (Figure 27B, middle and right panel). All PCR positive products were sent for sequencing and confirmed that all clones had the correct integration. One clone from each transfectant were used for further analysis and these transfected parasites will be referred to as 0017400CA, 0035300CA and 0400700CA.

Figure 27: Diagnostic PCR for checking of correct integrant. (A) Schematic diagram showing integrated gene and binding sites of diagnostic primers. Numbers at the bottom of the gel picture indicate clone number. (B) 2 clones from PYYM_0017400 has the correct bands of around 900bp after PCR. All clones of PYYM_0035300 and PYYM_0400700 have correct bands around 650bp after PCR. All PCR products from positive clones were sent for sequencing and found to be integrated into the correct locus.

5.1.4 YIR-GFP is expressed in transfected parasites

To examine the presence of the YIR-GFP fusion proteins in transfected parasites, trophozoites were collected and lysed with saponin (0.05%). After centrifugation

at 1800g, two fractions were obtained; the supernatant which contains proteins in the HCC and the PV; and the pellet which contains proteins from the host cell membrane, PVM and within the PPM. The supernatant (S) and pellet (P) were subjected to immunoblotting and probed with anti GFP antibodies (Red bands) and anti EXP2 antibodies (Green bands) as control.

A band of approximately 63 kDa was detected in both fractions which corresponds to the predicted size of the full-length PYYM_0017400, PYYM_0035300 and PYYM_0400700 proteins (Figure 28Ai, Aii and Aiii). The control band for the EXP2 shows up only in the pellet fraction for all 3 samples. More of the protein was detected in the S fraction in 0017400CA than that in the other 2 transfectants 0035300CA and 0400700CA, indicating that more of the PYYM_0017400 fusion protein was exported into the HCC. The western blot results showed that the proteins were successfully expressed in these transfected parasite lines.

When live trophozoites were visualized, the PYYM_0017400 can be seen clearly localized to the PVM/PPM forming a ring within the erythrocyte and also exported as small fast-moving vesicles within the HCC (Figure 28B). Similar observations of YIR localizing in fast moving vesicles have been observed by Siau et al. [84]. In contrast, the majority of the PYYM_0035300 fusion proteins was localized to the PVM/PPM and within the parasite, with little GFP seen in the HCC (Figure 28C). A third localization pattern was seen in the PYYM_0400700 fusion proteins where the majority of the protein is found within the parasite interior and was not localized to the PVM/PPM (Figure 28D).

In late schizonts, PYYM_0017400 and PYYM_0035300 had similar localization pattern of enveloping individual merozoites (Figure 28B and C). However, expression of PYYM_0400700 is low in schizonts and it has a diffused staining pattern within the parasites (Figure 28D).

Figure 28: GFP fusion protein is expressed in transfected parasites. (A) Western blot using anti-GFP antibodies (Red band) and anti EXP2 antibodies (Green band) on *P. yoelii* trophozoite lysed with saponin (P: pellet, S: supernatant). Band of size 63kDa which correspond to the size of the full-length YIR-GFP protein is detected. (B) Live trophozoites and schizonts observed under florescence microscopy for 0017400CA (C) Live trophozoites and schizonts observed under florescence microscopy for 0035300CA. (D) Live trophozoites and schizonts observed under florescence microscopy for 0400700CA.

5.1.6 *In vivo* growth phenotype

To determine whether constitutive expression of a single YIR variant has any impact on the growth of the parasites, the *in vivo* growth of the 0017400CA, 0035300CA and 0400700CA parasites were compared to that of wildtype *P. yoelii* YM. Since the transfected parasites must be maintained under pyrimethamine selection, *P. yoelii* YM-mCherry (YMmC) was used as the control as it is resistant to pyrimethamine. *P. yoelii* YMmC was generated in our laboratory by integration of a pL0016 plasmid into the parasite genome at the small ribosomal RNA site [161]. Previous studies in *P. berghei* have shown that integration into this locus does not affect parasite development [162] and is therefore suitable as a control in our studies.

From the experiments, *P. yoelii* YMmC kills the mice within 9 days post infection(pi) (Figure 29Ai, Bi, black lines) while the majority of the 0017400CA infected mice died one day later (Figure 29Ai, Bi, orange lines). The corresponding mice survival curve are shown in Figure 29Aii and Bii. We observed that the initial parasitemia for 0017400CA was detectable only after day 5pi while the parasitemia for the wildtype was detected on day 4. The difference in parasitemia between 0017400CA and wildtype YM parasites can be seen very clearly on day 6pi and day 7pi (Figure 29Aiii, Biii).

From two independent infections, we consistently observed that 20% of the mice infected with 0017400CA survived, while all the wildtype YM infected mice died by day 9. Mice that survive the infection had parasitemia plateau at about 30% on day 8-9 before the parasitemia started declining.

All mice that survived past day 9pi were able to control parasitemia but continued to suffer from anemia due to excessive RBC lysis. To minimize any disturbance caused to the mice during the recovery period, blood smears were only done every 3 days instead of daily. In both of the 2 independent infection, 1 mouse from each group died on day 11pi even though the parasitemia was dropping and this could be due to the inability to recover from anemia. The remaining mouse within the group eventually cleared the infection about 28 days post infection.

Figure 29: *In vivo* growth assay and mice survival after infection with wildtype YmMc or 0017400CA parasites. A) Growth assay run 1 B) Growth assay run 2. Ai, Bi) All mice injected with YM died within 9 days while mice infected with 0017400CA has a survival rate of about 20%. Aii, Bii) Mice survival chart showing a survival rate of 20% for 0017400CA infected mice. Aiii, Biii) Parasitemia difference between 0017400CA and YM. Data are mean $\pm$ SEM results of 2 independent experiments with 5 mice in each group. Statistical analysis was done using Student's t-test. * indicates p-value <0.05 .

In contrast, infection with 0035300CA resulted in the detection of parasites in the blood of 3 mice as early as 5 dpi (Figure 30A, orange lines) which is earlier than

the rise of parasitemia seen in YMmC parasites (Figure 30A, black lines). The remaining 2 mice had parasitemia curve similar to that of the YMmC parasites. *P. yoelii* YMmC kills the mice within 9 days post infection(pi) (Figure 30A, black line) while all the 0035300CA infected mice died either one day earlier or at the same time as the control (Figure 30B, orange line). Initially we speculated that this might be due to the increased virulence of the transfectants, however t-test analysis on the difference in the parasitemia on day 5 and day 6 of these 2 parasite lines showed no significant difference ($p > 0.05$) (Figure 30C) Due to the small number of mice used in our current experiment we cannot conclusively rule out the possibility that 0035300CA is more virulent than the YMmC parasites. A larger sample size is required for the confirmation of this result.

Figure 30: *In vivo* growth assay and mice survival after infection with wildtype YMmC or 0035300CA parasites. A) Growth assay of 0035300CA Vs YMmC. B) All mice injected with 0035300CA and YMmC died within 9 days. C) Parasitemia difference between 0035300CA and YMmC. Data are mean $\pm$ SEM results of experiment with 5 mice in each group. Statistical analysis was done using Student's t-test. T-test showed that the difference in the growth of the 2 strains were not statistically significant.

Finally, the parasitemia curve of mice injected with either 0400700CA or YMmC was similar with the detectable parasites appearing on day 5 (Figure 31A). All the mice in both experimental groups died on day 9 post infection (Figure 31B).

Figure 31: *In vivo* growth assay and mice survival after infection with wildtype YMmC or 0400700CA parasites. A) Growth assay of 0400700CA Vs YMmC. B) All mice injected with 0400700CA and YMmC died within 9 days.

5.1.7 Growth defect in 0017400CA is not due to loss of fitness of transfected cells

The data from the *in vivo* growth assay showed that overexpression of a specific *yir* can result in differences in the development of parasitemia and the rate of survival. It is clear that the blood parasitemia of 0017400CA was reduced in the first few days of infection. We hypothesized that this could be due to either the transfection causing a fitness defect, or the constitutive expression of YIR can be targeted by the mice innate immune system more efficiently. To determine if there is a fitness defect in 0017400CA parasites, we monitored the *ex vivo* maturation of intra-erythrocytic stage of 0017400CA and YMmC over a duration of 24 hours. Although *P. yoelii* cannot be grown *in vitro*, *ex vivo* infection assay using *P. yoelii* has been carried out successfully in the past for drug screening[134, 163] and invasion inhibition assays[164].

The single-step enrichment of *P. yoelii* ring stage was adapted from Chang et al. [134]. After enrichment for ring stage using a 65% Percoll gradient, the majority of the late stage parasites are eliminated from the enriched fraction. No schizonts

was seen in the culture. However, there is still some contamination with trophozoites in the enriched ring fraction (approximately 5 – 10% of total parasitemia at 0hr). The proportion of rings, trophozoites and schizonts were counted and determined at each time point and was represented in a bar chart (Figure 32). Comparison of the two *ex vitro* parasite cultures showed no significant differences in intra-erythrocytic development. From 0 – 8 hours, a decrease in the proportion of rings and an increase in trophozoites of both parasite lines was observed, confirming that the parasites were growing in the culture. Schizonts from both parasite lines started appearing at about 12 hours post-incubation and burst at about 18 hours. This result is consistent with what was published [134] and thus showed that there is no difference in the growth of these two parasite lines. This data provides us with critical information that constitutive expression of PYYM_0017400 did not affect the intrinsic growth of the parasites and gave us the confidence to speculate that the delay in parasitemia seen *in vivo* is probably due to host-parasite interactions.

Figure 32: Ex vivo growth assay. Growth of *P. yoelii* rings were observed over a period of 24 hours in *ex vitro* cultures. Cells were collected every three hours for FACS analysis to determine the proportion of cells at each stage.

5.2 Discussion

Although *pirs* have been discovered for more than a decade, little is known about the function of these genes. The difficulty in studying *pirs* is mainly due to the lack of suitable reagents and powerful tools. A particular challenge is the fact that *pirs* represent the largest multigene family in *Plasmodium* making genetic manipulation of a specific *pir* gene unworkable because of the highly conserved gene sequence among members of this family. Moreover, *pir* expression in daughter parasites has been shown to differ from the cloned parental parasite as early as 10 generations *in vivo* [95] suggesting a high switching rate. Therefore, without an effective way to maintain the expression of a specific *pir* gene, functional studies become nearly impossible.

In this study, we have developed a new approach which utilized the T2A skip peptide to link the GFP fusion GOI to the TgDHFR selection marker. This new method that enables us to activate a unique *yir* from its endogenous promoter. With this novel tool, we were able to generate 3 transfected parasite lines with population close to 100% of the parasites expressing a distinct tagged *yir*.

This is in contrast with other integration approaches in which vectors containing their own promoter have a high rate of wildtype escape leading to survival of parasites that only express the resistance cassette but not the transgene. The efficiency of this new approach also speeds up the downstream process of single cell cloning as we do not need to perform other selection steps such as FACS before cloning.

Current works on the selection for unique members of the multigene relies on the prior knowledge of functional roles of the protein in order to design experiments to study them. For example, PfEMP1 are known to bind endothelial cells [165] and thus the selection for a particular member of *vars* [166, 167] rely on the use of laborious panning techniques, which includes multiple rounds of binding of the iRBC onto the substrates. Panning selects the proteins which has a high binding affinity to the substrate. However, study on multigene families using this panning method will omit many proteins of unknown biological functions. The major advantage of our method is that it allows the possibility to select any member of the multigene family without any prior knowledge or characterization of the protein of interest. This powerful approach thereby opens up an avenue for the in-depth study of multigene families across the *Plasmodium* species.

These 3 GFP-tagged YIRs were characterized by western blot and IFA. Different YIRs have different localization pattern within the iRBC, which is in line with the observations reported previously in which episomally tagged CIRS have different localization within the RBC [108]. In addition to the difference in the localization of the 3 YIRS, our result also showed that the localization of specific YIR changes during parasite development within the RBC. In parasite lines PYYM_0017400 and PYYM_0035300, the protein was initially localized in the PVM and in punctate dots in the HCC during the trophozoite stage. Interestingly, the protein finally surrounds the individual merozoite in mature schizonts. The punctate staining seen in the HCC may belong to membranous structures [168] which are described in *P. berghei* and could perform similar functions as the Maurer's cleft in *P. falciparum*. However, PYYM_0400700 was only detected in

the intracellular stages and not in merozoites, suggesting that distinct members of this large multigene family have differential developmental expression patterns.

Although proteins from both PYYM_0017400 and PYYM_0035300 parasites are enveloping individual merozoites in late schizonts, their localization in the trophozoite stage is different. The formation of the merozoite membrane in blood stage is currently unknown and no data shows whether it is formed from the existing PVM/PPM or synthesized *de novo*. However, there is evidence showing that in the liver stages the PM invaginate around individual nuclei and eventually forms the merozoite membrane [169]. If the results from the liver stage merozoite can be extrapolated to the blood stage, I speculate that the insertion of the 2 YIR proteins into the merozoite might be different. PYYM_0017400 proteins are exported in the trophozoite stage and could form an envelope covering the merozoite from outside the PVM/PPM. On the other hand, PYYM_0035300 which is not exported in the trophozoite stage associates with merozoites from within the parasite.

The different insertion of YIRs into the merozoite membrane can explain why the 2 proteins have similar staining pattern in the merozoites but have different *in vivo* growth phenotypes. The 0017400YIR-GFP proteins would be likely displayed on the surface of the merozoites and exposed to the immune system. Due to the inability to switch expression of the surface YIR, the 0017400CA parasites cannot evade the immune system as effectively. As the delay in parasitemia was observed very early in infection (4dpi – 5dpi), the innate

immunity or potential increase in splenic trapping of the parasites presumably contributed to the delay. Innate lymphocytes such as $\gamma\delta$ T cells and Natural killer cells have been shown to be able to limit parasite growth by producing pro-inflammatory cytokines such as IFN γ [170-172] or killing merozoites directly with cytotoxic granulysin [173]. I speculate that the 0017400CA parasites which cannot vary the surface expressed YIR might be easier targets for these immune cells thereby resulting in lower parasitemia.

Several important features of *pirs* suggest that different members of this multi-gene family play different roles during the blood stage infection. The presence of distinct phylogenetic clades between members of *pirs* highly suggest that functional diversity exist within the family [93, 96]. In addition, different cellular localization and timing of expression further indicates the presence of functional diversity [108, 174].

In our work, we have clearly shown that the endogenously expressed YIRs have different localization during blood stage development and our results support the hypothesis of functional diversification in the *pir* multigene family. Similar findings were reported in *P. vivax* where *virs* are expressed in different subcellular location and surface expressed VIRs can bind to ICAM-1 to mediate cytoadherence [110]. In *P. falciparum*, STEVOR proteins were exported into Maurer's Cleft in late trophozoite and schizont stages [116, 175] and subsequently found on surface [176] as well as the apical end of merozoites [177, 178]. STEVOR binds to glycophorin C on the surface of RBC and antibodies against STEVOR can block merozoite invasion [111]. In our study, the presence

of YIRs on the merozoites is compatible with potential roles in invasion or immune evasion. As the localization profile of 0017400CA and 0035300CA in our study are similar to what was observed for STEVOR, we hypothesize that the some YIRs might perform similar functions as STEVOR and could play a role in erythrocyte invasion. Our hypothesis is supported by similar observations published by Yam et al. in which they reported the presence of subsets of CIRs which are located on the edges of merozoites and can bind to RBC via protein-protein interactions [108]. This interaction can be abolished with the addition of anti-CIR antibodies which reduces the binding by ~40%. To further investigate this finding, antibodies can be generated against these YIR proteins and used in invasion inhibition assay to determine whether these antibodies could block invasion of these transfected parasite lines.

Overall, our result demonstrated that constitutive expression of some *yirs* has an impact on the growth of the parasites *in vivo* and this happens early in the infection. However, it is important to note that this delay in parasitemia was only observed in the 0017400CA parasite line and not the other 2 transfectants. The main difference in the localization between the 0017400CA parasite line and the other 2 transfectants is that majority of the proteins in the other 2 transfectants are predominantly localized within the parasite in the early stages of parasite development. There are evidences showing that surface expressed RIFINs are immunogenic in natural infection and high levels of antibodies against RIFINs can be detected in adults living in endemic areas. [179, 180]. However, currently we are not able make any conclusive statement whether the export of YIRs has

any impact on the delay in parasitemia and more experiments needs to be conducted to confirm this hypothesis.

Chapter 6

Evidence of co- regulation in *yir* expression

The malaria parasites export many proteins into the host cell cytoplasm to modify it for its own survival. Many of these exported proteins are displayed on the surface of iRBC and are termed VSAs. The most well-studied VSA is the *var* gene family which encodes for PfEMP1. The *var* genes are under monoallelic regulation in which a single parasite expresses only 1 single variant of the PfEMP1 protein [166]. Unlike the *var* genes which are present only in *P. falciparum*, the *pir* multigene family is present in all malaria species affecting human, non-human primates and rodents. Despite being discovered for more than a decade, the transcriptional regulation of *pils* are currently unknown. In this study, we are able to use our single *yir* forced expression cell line to show for the first time that the activation of a single *yir* has an impact on other *yir* expression within the parasites.

6.1 RNA preparation for RNA sequencing

The different transfected parasite clones were named according to the gene ID of the activated YIR. To maintain consistency in the naming of the wildtype and transfected parasites, for the rest of this section I will refer to the wildtype *P. yoelii* YM parasites as PYYM and the transfected parasites as PYYM_0017400CA.

In *P. chabaudii*, the level and timing of transcription of specific *cirs* had been previously reported by Lawton et al to be different during the lifecycle [96], thereby showing the complex expression profile of *pir* gene expression. In my current work, I focused on the trophozoite stage of the parasite as it is the most metabolically active stage and many surface expressed antigens are exported at

this stage [181, 182]. To obtain synchronous parasites, donor mice infected with each parasite lines were sacrificed when parasitemia reached 40%. Whole blood was then passed through a Histodenz gradient to isolate the ring stage parasites. Ring stage parasites of each parasite lines were injected into 3 mice via i.v and the blood was collected 30 – 32 hours after inoculation. A total of 9 RNA samples from trophozoites were purified (triplicates of PYYM_0017400CA SC1, PYYM_0017400CA SC3 and parental PYYM obtained from independent infections). The 9 RNA samples were sent for RNA sequencing on the Illumina RNA-seq next generation sequencer. The sequencing generated a total of ~1.7 million to 4 million paired end reads per sample. For RNA-seq analysis, we will be utilizing Kallisto which is a light weight alignment program that uses pseudoalignment to estimate transcript abundance [183]. Before the alignment process, the *P. yoelii* 17x 1.1 CDS was converted into an index file which is required for Kallisto to map the RNA-seq data. The fastq files which contain the raw reads were then mapped to this reference index with Kallisto v0.44.0. To avoid missing any *yir* genes in our downstream analysis, we mapped the RNA seq raw data to the reference genome of *P. yoelii* 17X 1.1 due to the larger number of *yir* gene annotation in PlasmoDB. Kallisto generates the abundance file in the form of Transcript per million (TPM), which is a normalization method to normalize the difference in read counts resulting from different length of the genes. The formula for TPM is given as:

$$TPM_i = \frac{X_i}{l_i} \cdot \left(\frac{1}{\sum_j \frac{X_j}{l_j}} \right) \cdot 10^6$$

$\frac{X_i}{l_i}$ gives the read per base reading of the gene of interest where X_i is the read count of the gene, l_i is the length of the gene in base pair. Divide this number by the summation of all the read per base reading in the sample $\sum_j \frac{X_j}{l_j}$ and multiple by 10^6 to get the TPM count.

The resulting TPM counts were imported using the R program *tximport* to Deseq2 for differential gene expression analysis [184]. Since the annotations used were from 17X 1.1 (The 17x 1.1 annotation for PYYM_0017400 is PY17X_1100077), I will refer all the genes name according to the 17X 1.1 annotation for the rest of this chapter.

6.2 Endogenous expression of a single *yir* gene results in the global downregulation of *yir* expression.

Distribution of the average counts for the *yir* gene family in *P. yoelii* YM is different from the transfected cell line SC1 and SC3. The counts from each cell line were sorted into bins of width 0.5 and represented as a histogram (Figure 33A). The blue bars represent PYYM, pink and green bars represent the transfected cell line SC1 and SC3 respectively. Counts towards the left of the x-axis indicates low TPM which corresponds to low level of transcripts being detected in the sample. As can be seen from Figure 33A, the majority of the TPM counts for all 3 cell lines were <60 indicating that the majority of the *yirs* in the population were transcribed at low levels. Only 3 *yirs* in the parental YM, 5 *yirs* in SC1 and 11 *yirs* in SC3 are highly transcribed and have TPM >60 and this includes the constitutive expressed *yir* PY17X_1100077 (Figure 33B). As the majority of the genes in all 3 cell lines fall below 60 TPM, the histogram was

truncated at 60 TPM. The full histogram is included in Appendix 6. In the YM cell line, the plot was skewed to the left, but we can also see a wider spread of the data towards the right. In contrast, the plots of SC1 and SC3 are skewed more to the left, with many *yir* genes falling into bins <1 , indicating the narrowing of transcription of *yir* genes in the transfected cell lines (799 and 786 *yir* genes have TPM <1 in SC1 And SC3 respectively as compared to 509 in the parental YM).

To better illustrate the differences seen in the *yir* gene expression between the parental and transfected cell line, the TPM of each cell line were represented as a pie chart. We categorized genes with TPM <1 as no expression, 1 – 10 as low expression, 11 – 59 as medium expression and >60 as high expression. The distribution for *yir* expression in PYYM showed that a higher proportion of these genes (~50%) were expressed between 1 – 60 TPM counts (Figure 33C). In contrast, SC1 and SC3 displayed a tight global repression of *yir* expression as only 23 – 25% of the total *yirs* had a read count between 10 – 60 TPM. This difference represents a drop in ~50% of the *yir* transcription in the transfected cell lines as compared to the parental YM. This strongly suggest that constitutive expression of a single *yir* causes the narrowing of transcription of *yir* genes, thereby resulting in fewer *yir* genes being transcriptionally active in the population.

Figure 33: Global *yir* expression is downregulated in PYYM_0017400CA cell lines. A) Histogram depicting the distribution of global *yir* expression in the wildtype YM and transfected cell lines. Blue bars represent PYYM, pink and green bars represent SC1 and SC 3 respectively. Histogram was created based on bin range of 0.5. B) List of *yirs* in the 3 cells lines with TPM >60. C) Pie chart showing the percentage of *yir* genes in each cell line sorted according to the TPM value.

6.3 Co-regulation of *yir* gene expression

The transcript level abundance estimate generated as TPM using Kallisto was input into Deseq2 using the tximport pipeline [133] to perform differential gene

expression analysis. Genes that were considered to be statistically differentially expressed between PYYM and SC1 or SC3 are set at a cutoff of $\text{padj} < 0.05$ and more than 2-fold difference. To increase the robustness of the screening, only genes with counts of >20 were selected for in our differential gene analysis.

Using these set of criteria, we identified 59 yir genes which can be further divided into 3 groups (upregulation, downregulation and no change) based on their expression profile when compared to the parental YM control (Figure 34A). Within SC1 and SC3, 4 yir genes were upregulated, 35 yir genes were downregulated and 20 yir genes showed no change in expression (Figure 34A).

This suggest that forcing expression of a single yir gene can lead to the upregulation of four transcripts (Figure 34A), namely the tagged gene PY17X_1100077 and 3 other yir genes, PY17X_0401000, PY17X_0114700, PY17X_1100081 (TPM values shown in Appendix 7). Unlike the var gene families which are under monoallelic exclusion [166, 185], control of yir gene expression seems to affect only a subset of the yirs in the population; 39 yirs either up or down regulated, while another subset of 20 yirs were not affected.

To ensure the robustness of the RNA sequencing analysis, we selected seven genes with the highest expression from the 3 groups for validation using qRT-PCR (Figure 34B). PCR primers were designed with a melting temperature of $\sim 60^{\circ}\text{C}$. Prior to running the qRT-PCR, primer efficiency was established by plotting a dilution curve using the cDNA of the samples to be tested. The dilution used for testing were 1:10, 1:100 and 1:1000. The primer efficiency was calculated from the slope using the following formula ($10^{(-1/\text{the slope value})}$)-

1)*100. The efficiency of each primers can be found in Appendix 8. Expression of the selected yir genes was normalized to the HSP70 gene which acts as an internal reference control [186]. The advantage of having an internal control is that it compensates for any differences that would arise from the PCR preparation steps because the control is exposed to the same treatment as the other genes.

The qPT-PCR analysis is consistent with the profile and trend obtained from the RNA sequencing analysis (Figure 34B) indicating that the RNA sequencing results and analysis is robust and reproducible.

Figure 34: Expression of *yir* genes are co-regulated. Heatmap of *yir* expression compared to that of the parental YM. A) RNA sequencing showed that 59 *yir* genes are expressed (count >20) and are divided into 3 groups based on the *yir* expression relative to that of the wildtype YM parasites. The blue arrow points to the gene that was tagged in this transfected cell line. Orange arrows point to the genes selected for qRT-PCR validation. B) Expression of *yirs* in each group were validated using qRT-PCR and results are in line with RNA sequencing results. Error bars are $\pm$ SEM results of qRT-PCR triplicates of each sample.

6.4 Presence of motif in *yirs* under regulatory control

The presence of differential regulation of *yir* expression indicates that there might be subgroups of *yirs* present in the parasite. Certain transcription factors such as

the *P. falciparum* (Pf)NF-YB transcription factor can bind to specific motifs upstream of the start codon to mediate transcription [187]. In order to identify potential motifs in these *yirs* which might give us a clue on how they are regulated, we aligned the promoter region 1kb before each *yir* gene, but we fail to see any pattern in the *yir* promoter alignment.

As an alternative approach, we used the published phylogenetic databases to determine whether *yirs* which are co-regulated are evolutionary related. Recently Otto et al. published a comprehensive phylogenetic analysis of *pirs* in RMP [93]. In their analysis, *pirs* can be classified into 4 long form clades and 8 short form clades depending on the size of the protein.

Using the results of their analysis, we mapped the 59 genes that were identified in our RNA sequencing analysis into their phylogenetic tree (Table 12). Of all the 39 genes that were co-regulated, we managed to map 35 of the genes to Otto's phylogenetic tree and all the co-regulated *yirs* fall into the 'short' *yir* clade. For the *yir* genes with no change in expression, 12 mapped to the phylogenetic data and these genes fall into both the 'short' and 'long' *yir* clades. From this data, it seems that transcriptional control of 'long' and 'short' form *yirs* are distinct from each other as none of the co-regulated *yirs* are from the 'long' form *yirs*. This finding suggests that 'short' and 'long' form *yirs* are regulated independently of each other and might perform different functional roles in the parasites.

Co-regulated YIR ID	Gene ID in Otto screen (YM ID)	Clade in Otto screen	Not co-regulated YIR ID	Gene ID in Otto screen (YM ID)	Clade in Otto screen
PY17X_0100049.1	PYYM_00436	s1	PY17X_0102100	N.D	
PY17X_0114700.1	PYYM_01141	S2	PY17X_0117100	PYYM_02006	S6
PY17X_0202000.1	PYYM_02036	s1	PY17X_0202800	N.D	
PY17X_0223101.1	PYYM_00123	s5	PY17X_0217100	PYYM_00202	L2
PY17X_0319701.1	PYYM_00483	s4	PY17X_0222501	N.D	
PY17X_0320701.1	PYYM_04001	s4	PY17X_0322601	PYYM_00052	S1
PY17X_0320801.1	N.D		PY17X_0500031	PYYM_00020	S1
PY17X_0320901.1	PYYM_00412	S1	PY17X_0600081	N.D	
PY17X_0400700.1	PYYM_04009	S1	PY17X_0600261	PYYM_00025	S1
PY17X_0401000.1	PYYM_14012	S1	PY17X_0700200	PYYM_01162	L2
PY17X_0527901.1	PYYM_00395	S4	PY17X_0800500	PYYM_08004	S1
PY17X_0627300.1	PYYM_06263	S4	PY17X_0802500	PYYM_08024	L1
PY17X_0627400.1	N.D		PY17X_0840700	PYYM_08404	L1
PY17X_0627501.1	PYYM_00301	S8	PY17X_0846101	PYYM_00254	S1
PY17X_0700003.1	PYYM_00290	S5	PY17X_0900027	N.D	
PY17X_0801500.1	PYYM_08014	S4	PY17X_0951401	N.D	
PY17X_0840500.1	PYYM_08402	S5	PY17X_1050901	N.D	
PY17X_0950001.1	PYYM_00167	S1	PY17X_1300077	N.D	
PY17X_0950401.1	PYYM_00213	S4	PY17X_1400019	N.D	
PY17X_0952201.1	PYYM_11010	S1	PY17X_1400400	PYYM_14024	S1
PY17X_0952301.1	PYYM_00236	S5			
PY17X_1000900.1	N.D				
PY17X_1043400.1	PYYM_10430	S4			
PY17X_1046000.1	PYYM_00039	S5			
PY17X_1046201.1	PYYM_00041	S4			
PY17X_1046301.1	PYYM_00042	S1			
PY17X_1047201.1	PYYM_00097	S4			
PY17X_1047601.1	PYYM_00310	S1			
PY17X_1100015.1	PYYM_02187	S1			
PY17X_1100043.1	N.D				
PY17X_1100077.1	PYYM_00174	s4			
PY17X_1100081.1	PYYM_00333	S1			
PY17X_1101400.1	PYYM_11025	S5			
PY17X_1149100.1	PYYM_00074	S2			
PY17X_1200121.1	PYYM_00021	S2			
PY17X_1200151.1	PYYM_00170	S4			
PY17X_1200900.1	PYYM_00367	S1			
PY17X_1400100.1	PYYM_14021	S4			
PY17X_1400700.1	PYYM_14027	S1			

Table 12. List of *yir* genes found in the RNAseq and the corresponding ID and clade in Otto's screen. Table on the left are list of *yir* genes that are either up or down regulated in our RNAseq while table to the right are list of *yir* genes which have no change in expression. *Yir* in bold red font was the *yir* tagged in this study. N.D stands for No data found in Otto's screen. S stands for 'short' form *pirs* and L stands for 'long' form *pirs*.

Next, we proceeded to look at potential motifs at the protein level which might provide us with additional information about protein features of *yirs* which are transcriptionally regulated. The tagged PY17X_1100077 gene encodes for a YIR protein which contains a *Plasmodium* Lipophilic and Secondary structure Mediated Export Domain (PLASMED) motif near the C-terminus of the protein. The PLASMED motif consists of about 26 amino acids that spans the transmembrane domain of PLASMED proteins and can mediate protein export of PNEPs in RMP [84].

To determine if any of the co-regulated YIRs contains the PLASMED motif, an alignment was done using Clustal Omega using the last 100 amino acid of each YIR which contains the transmembrane domain (Figure 35). The *yir* genes which

are either up or down regulated in our RNAseq are boxed in the blue box. From the Clustal Omega results, YIRs which are co-regulated have a preference for polar amino acids (green letters) before the transmembrane region (red letters). The series of serine residues before the transmembrane domain of these differentially regulated YIRs were not found in other YIRs which have no change in expression level.

The presence of these polar residues before the transmembrane domain is indicative of the PLASMED motif which has been previously reported by Siau et al. [84]. Examination of the 59 YIR protein sequence, showed that the 37 out of 39 (94.8%) of the co-regulated *yir* genes contain the PLASMED, as compared to those that showed no change in transcriptional expression. Our results indicate that PLASMED containing YIRs are transcriptionally co-regulated suggesting that PLASMED containing YIRs form a subgroup that are co-regulated independently from other *yirs*.

Figure 35: ClustalW alignment of the last 100 amino acids of the 59 proteins detected in the RNA sequencing analysis. Blue box indicates *yir* genes which have the PLASMED motif. 95% of the *yirs* which are differentially regulated are highlighted in the blue box indicating presence of PLASMED in these proteins.

6.5 Discussion

The constitutive expression of a *yir* gene resulted in the global down regulation of other *yir* genes in the population. RNA sequencing results indicated that there are subsets of *yirs* which are transcriptionally regulated. Unlike the monoallelic regulation of *var* genes in which only 1 variant exist within a parasite, our results showed that multiple *yirs* can be maintained in the population even when all the parasites in the population have one specific *yir* gene constitutively activated. However, considering the large number of *yir* genes in the genome, the control of *yir* expression seems to be tightly regulated.

Of the 4 *yir* genes upregulated in our transfected cell line, one of the gene PY17X_1100081 is located just upstream of the *yir* gene which was constitutively expressed. Although the exact mechanism of how *yir* genes are regulated is unknown, we postulated that the upregulation of this particular *yir* might be due to the open chromatin structure resulting in the creation of transcriptionally active site in that region. Others have also speculated on the existence of transcriptionally active sites in at least two chromosomal locations in *P. vivax* in which *yirs* located adjacent to each other are highly expressed [188].

However, the other two upregulated *yir* genes are located on other chromosomes (Chromosome 4 for PY17X_0401000 and chromosome 1 for PY17X_0114700) and their upregulation might be due to other transcriptional control mechanism. Research on other organisms have shown that the three-dimensional architecture of chromatin plays an important role through the formation active and repressive

subcompartments [189, 190]. By moving through the different subcompartments, activation or repression of specific chromosomal domains can be achieved. This has also been shown for *var* genes in which the active *var* occupy a single, perinuclear *var* expression site, away from the other *vars* which are silent [191]. The activation of *yir* genes which are located on other chromosomes might be due to such three-dimensional transcriptional control.

Fonager et al. utilized *in silico* analysis to identify a motif upstream of *yir*, *bir* and *cir* genes which could potentially be regulatory elements for rodent *pvr* transcription [94]. However, since all the rodent *pvr*s contains this motif upstream of the start codon, it is highly unlikely to play a differential role in transcription. In a recent publication on the transcriptional control of *rifins*, it was reported that the sequence upstream of the *rif* gene cannot predict promoter activation or silencing [192]. Similar to what others had previously described, the 5' UTR of the 60 *yir* genes identified in our RNAseq are very conserved and we did not manage to identify any motif that might suggest how these genes are differentially regulated. These results together seem to suggest that upstream regulatory elements might not be the reason for differential expression of *yirs*.

Therefore, we speculate that differential regulation of these *yir* genes might not be due to upstream regulatory elements, instead other forms of control such as alternative splicing, non-coding RNA (ncRNA) or epigenetic controls might be involved. Alternative splicing has been shown in the case of *var* and *sicavar* to produce truncated transcript which may exert regulatory control on transcription [193, 194]. ncRNA has been shown to be able to regulate transcription at multiple

loci [195], suggesting that ncRNA could play a role in the transcriptional activation of multiple *yirs*. Early studies on the mutually exclusive expression of *var* genes showed that differential expressions are marked with differential histone modification [166, 191]. Presence of these histone modifications might play a role in the differential gene regulation of the *yir* multigene family.

In other multigene families such as PfEMP1 and STEVOR which are exported and expressed on the erythrocyte surface, tight regulation of transcriptional control of these genes are important for immune evasion. Since PLASMED containing YIRs are exported and might potentially be expressed on the surface of iRBC, they could be regulated in tandem to coordinate their expression in the iRBC to promote immune evasion. Furthermore, we have shown in Chapter 5 that the tagged YIR in 0017400CA parasites are expressed on the merozoite periphery. Since merozoite surface proteins are major targets of immunity [124, 196], transcriptional regulation to ensure antigenic variation is important and this has been demonstrated in STEVOR in which only one 1 STEVOR variant is present on individual merozoite [176]. The need to coordinate expression of YIRS exposed to the immune system to maintain antigenic variation might be the reason why PLASMED containing YIRs are transcriptionally regulated.

However, the current finding is still preliminary, and we need to conduct more experiments to determine if the regulation of *yirs* is indeed dependent on its localization. We have currently forced the expression of 2 other *yir* genes which do not contain a PLASMED motif. Results from the RNA sequencing results of

these cell lines will enable us to shed some light on possible regulatory control of PLASMED negative *yirs*.

7 Conclusion and future perspective

Our work here has demonstrated the effective use of skip peptide in forcing the expression of a single *yir* within the parasite population. Using this method, we were able to get close to 100% transfection efficiency and this facilitates downstream single cell cloning of the parasites. Through forcing the endogenous expression of a single *yir* variant, we were able to show that subgroups of *yirs* are under co-regulation. In particular, *yirs* with the PLASMED motif seems to be co-regulated. This work to our knowledge is the first documented evidence of co-regulation in *yir* gene expression.

Our skip peptide system provides a platform for the further study of the *pir* multigene family of which the progress of research in this field has been kind of stagnant due to difficulties in getting the parasites to express a single *pir* variant. With this toolset available, researchers can start to look at questions such as the role that specific *pir* play in immune evasion.

Although we did not manage to identify any YIRs which are expressed in the surface of iRBC in our IFA, we cannot rule out the possibility that only a small amount of YIR proteins are inserted into the iRBC membrane and cannot be detected by IFA. A novel approach using HiBit which can be fused to the ends of proteins has been used in mammalian system to detect surface proteins. If the protein tagged is displayed on the cell surface, HiBit will react with the reagent

to give off a luminescent reading which can be quantified. Using this approach, we can see whether the YIRs which are exported in our study are also expressed on the iRBC surface.

A high throughput method to screen for all potential surface expression of YIRs is to utilize random Hibit tagging by initiating homologous recombination with targeting sequence which is highly conserved in the multigene family. When transfected into the parasites, the Hibit construct will randomly integrate into any *yir* gene thus creating a pool of parasites each expressing a differently tagged *yir*. All the parasites that survived the drug selection will be serially diluted and cloned to isolate single parasites. These parasites will then be screened for surface expression of the tagged *yir* gene. Parasites with surface expressing YIRs will be sequenced using forward primers which can bind to majority of the *yir* genes and reverse primers binding to the Hibit sequence to identify the *yir* gene that the Hibit had integrated into. With this method I can narrow down to focus on all potential surface expressed YIRs and subsequently validate their expression *in vivo*.

In this study, I have showed that there are 2 YIRs (PYYM_0017400 and PYYM_0035300) which surround individual merozoites. As these YIRs were potentially found on the surface of merozoites, it allows us to address the question of whether these YIRs play a part in invasion of RBCs. Antibodies against these *yir* variants will be generated and use in *ex vivo* invasion assay to determine whether these antibodies have any invasion inhibitory effects on these transfected cell lines.

Although the aspect of *pirs* highlighted in this thesis is a small contribution to the understanding of this large multigene family, it is the first documented evidence that subsets of *pirs* are under transcriptional exclusion. The discovery that certain groups of *yirs* are regulated differently suggested that there is functional diversity in *yirs*. Our finding that PLASMED containing *yirs* are co-regulated suggests the possibility that *yirs* which are exported might be regulated in tandem to generate antigenic variation. Further analysis of *yirs* without PLASMED is required to confirm this finding.

8 References

1. Ta, T.H., et al., *First case of a naturally acquired human infection with Plasmodium cynomolgi*. Malar J, 2014. **13**: p. 68.
2. Singh, B., et al., *A large focus of naturally acquired Plasmodium knowlesi infections in human beings*. Lancet, 2004. **363**(9414): p. 1017-24.
3. Contacos, P.G., et al., *QUARTAN-TYPE MALARIA PARASITE OF NEW WORLD MONKEYS TRANSMISSIBLE TO MAN*. Science, 1963. **142**(3593): p. 676.
4. Contacos, P.G., et al., *Transmission of Plasmodium schwetzi from the chimpanzee to man by mosquito bite*. Am J Trop Med Hyg, 1970. **19**(2): p. 190-5.
5. Deane, L.M., M.P. Deane, and J. Ferreira Neto, *Studies on transmission of simian malaria and on a natural infection of man with Plasmodium simium in Brazil*. Bull World Health Organ, 1966. **35**(5): p. 805-8.
6. Bartoloni, A. and L. Zammarchi, *Clinical aspects of uncomplicated and severe malaria*. Mediterr J Hematol Infect Dis, 2012. **4**(1): p. e2012026.
7. W.H.O, *Malaria Fact Sheets*. Geneva: World Health Organization. 2016.
8. Facer, C.A. and D. Rouse, *Spontaneous splenic rupture due to Plasmodium ovale malaria*. Lancet, 1991. **338**(8771): p. 896.
9. Moreno-Perez, D.A., J.A. Ruiz, and M.A. Patarroyo, *Reticulocytes: Plasmodium vivax target cells*. Biol Cell, 2013. **105**(6): p. 251-60.
10. Krotoski, W.A., et al., *Demonstration of hypnozoites in sporozoite-transmitted Plasmodium vivax infection*. Am J Trop Med Hyg, 1982. **31**(6): p. 1291-3.
11. Cogswell, F.B., *The hypnozoite and relapse in primate malaria*. Clin Microbiol Rev, 1992. **5**(1): p. 26-35.
12. Richter, J., et al., *What is the evidence for the existence of Plasmodium ovale hypnozoites?* Parasitol Res, 2010. **107**(6): p. 1285-90.
13. Shute, P.G. and M. Maryon, *Studies in the transmission of Plasmodium malariae by Anopheles mosquitoes*. Parasitology, 1951. **41**(3-4): p. 292-300.
14. Young, M.D. and R.W. Burgess, *The transmission of Plasmodium malariae by Anopheles maculipennis freeborni*. Am J Trop Med Hyg, 1947. **27**(1): p. 39.
15. Vinetz, J.M., et al., *Plasmodium malariae infection in an asymptomatic 74-year-old Greek woman with splenomegaly*. N Engl J Med, 1998. **338**(6): p. 367-71.
16. Gilles, H.M. and R.G. Hendrickse, *Nephrosis in Nigerian children. Role of Plasmodium malariae, and effect of antimalarial treatment*. Br Med J, 1963. **2**(5348): p. 27-31.
17. Walker, K.C., M.B. Goette, and G.S. Batchelor, *Pesticide Residues in Foods, Dichlorodiphenyltrichloroethane and Dichlorodiphenyldichloroethylene Content in Prepared Foods*. Journal of Agricultural and Food Chemistry, 1954. **2**(20): p. 1034-1037.
18. Wolff, M.S., et al., *Risk of breast cancer and organochlorine exposure*. Cancer Epidemiol Biomarkers Prev, 2000. **9**(3): p. 271-7.
19. Wurster, C.F., Jr., D.H. Wurster, and W.N. Strickland, *BIRD MORTALITY AFTER SPRAYING FOR DUTCH ELM DISEASE WITH DDT*. Science, 1965. **148**(3666): p. 90-1.
20. Fry, D.M., *Reproductive effects in birds exposed to pesticides and industrial chemicals*. Environ Health Perspect, 1995. **103 Suppl 7**: p. 165-71.
21. Roberts, D.R. and R.G. Andre, *Insecticide resistance issues in vector-borne disease control*. Am J Trop Med Hyg, 1994. **50**(6 Suppl): p. 21-34.

22. Rehwagen, C., *WHO recommends DDT to control malaria*. BMJ : British Medical Journal, 2006. **333**(7569): p. 622-622.
23. Bhatt, S., et al., *The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015*. Nature, 2015. **526**(7572): p. 207-211.
24. Steinhardt, L.C., et al., *Effectiveness of insecticide-treated bednets in malaria prevention in Haiti: a case-control study*. Lancet Glob Health, 2017. **5**(1): p. e96-e103.
25. Hemingway, J., et al., *Averting a malaria disaster: will insecticide resistance derail malaria control?* Lancet, 2016. **387**(10029): p. 1785-8.
26. Dong, Y., et al., *CRISPR/Cas9 -mediated gene knockout of Anopheles gambiae FREP1 suppresses malaria parasite infection*. PLoS Pathog, 2018. **14**(3): p. e1006898.
27. Wang, S., et al., *Driving mosquito refractoriness to Plasmodium falciparum with engineered symbiotic bacteria*. Science, 2017. **357**(6358): p. 1399-1402.
28. Achan, J., et al., *Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria*. Malar J, 2011. **10**: p. 144.
29. White, N.J., *The treatment of malaria*. N Engl J Med, 1996. **335**(11): p. 800-6.
30. Yayon, A., Z.I. Cabantchik, and H. Ginsburg, *Susceptibility of human malaria parasites to chloroquine is pH dependent*. Proc Natl Acad Sci U S A, 1985. **82**(9): p. 2784-8.
31. Sullivan, D.J., Jr., et al., *A common mechanism for blockade of heme polymerization by antimalarial quinolines*. J Biol Chem, 1998. **273**(47): p. 31103-7.
32. Pagola, S., et al., *The structure of malaria pigment beta-haematin*. Nature, 2000. **404**(6775): p. 307-10.
33. Doherty, T., *Anti-malarial drugs for Africa*. The Lancet, 2000. **356**(9226): p. 339-340.
34. Lakshmanan, V., et al., *A critical role for PfCRT K76T in Plasmodium falciparum verapamil-reversible chloroquine resistance*. Embo j, 2005. **24**(13): p. 2294-305.
35. Bray, P.G., et al., *Defining the role of PfCRT in Plasmodium falciparum chloroquine resistance*. Mol Microbiol, 2005. **56**(2): p. 323-33.
36. Payne, D., *Spread of chloroquine resistance in Plasmodium falciparum*. Parasitol Today, 1987. **3**(8): p. 241-6.
37. Adam, I., Y. Ibrahim, and G.I. Gasim, *Efficacy and safety of artemisinin-based combination therapy for uncomplicated Plasmodium falciparum malaria in Sudan: a systematic review and meta-analysis*. Malar J, 2018. **17**(1): p. 110.
38. Dondorp, A.M., et al., *Artemisinin Resistance in Plasmodium falciparum Malaria*. New England Journal of Medicine, 2009. **361**(5): p. 455-467.
39. Phyo, A.P., et al., *Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study*. Lancet, 2012. **379**(9830): p. 1960-6.
40. Ariey, F., et al., *A molecular marker of artemisinin-resistant Plasmodium falciparum malaria*. Nature, 2014. **505**(7481): p. 50-5.
41. Spring, M.D., et al., *Dihydroartemisinin-piperaquine failure associated with a triple mutant including kelch13 C580Y in Cambodia: an observational cohort study*. Lancet Infect Dis, 2015. **15**(6): p. 683-91.
42. Clyde, D.F., et al., *Immunization of man against sporozite-induced falciparum malaria*. Am J Med Sci, 1973. **266**(3): p. 169-77.

43. Clyde, D.F., et al., *Specificity of protection of man immunized against sporozoite-induced falciparum malaria*. Am J Med Sci, 1973. **266**(6): p. 398-403.
44. Hoffman, S.L., et al., *Development of a metabolically active, non-replicating sporozoite vaccine to prevent Plasmodium falciparum malaria*. Hum Vaccin, 2010. **6**(1): p. 97-106.
45. Hoffman, S.L., et al., *Protection of humans against malaria by immunization with radiation-attenuated Plasmodium falciparum sporozoites*. J Infect Dis, 2002. **185**(8): p. 1155-64.
46. Seder, R.A., et al., *Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine*. Science, 2013. **341**(6152): p. 1359-65.
47. Lyke, K.E., et al., *Attenuated PfSPZ Vaccine induces strain-transcending T cells and durable protection against heterologous controlled human malaria infection*. Proc Natl Acad Sci U S A, 2017. **114**(10): p. 2711-2716.
48. Labaied, M., et al., *Plasmodium yoelii sporozoites with simultaneous deletion of P52 and P36 are completely attenuated and confer sterile immunity against infection*. Infect Immun, 2007. **75**(8): p. 3758-68.
49. Spring, M., et al., *First-in-human evaluation of genetically attenuated Plasmodium falciparum sporozoites administered by bite of Anopheles mosquitoes to adult volunteers*. Vaccine, 2013. **31**(43): p. 4975-83.
50. Belnoue, E., et al., *Protective T cell immunity against malaria liver stage after vaccination with live sporozoites under chloroquine treatment*. J Immunol, 2004. **172**(4): p. 2487-95.
51. Schats, R., et al., *Heterologous Protection against Malaria after Immunization with Plasmodium falciparum Sporozoites*. PLoS ONE, 2015. **10**(5): p. e0124243.
52. Bijker, E.M., et al., *Sporozoite immunization of human volunteers under mefloquine prophylaxis is safe, immunogenic and protective: a double-blind randomized controlled clinical trial*. PLoS One, 2014. **9**(11): p. e112910.
53. Nussenzweig, V. and R.S. Nussenzweig, *Development of a sporozoite malaria vaccine*. Am J Trop Med Hyg, 1986. **35**(4): p. 678-88.
54. Casares, S., T.D. Brumeanu, and T.L. Richie, *The RTS,S malaria vaccine*. Vaccine, 2010. **28**(31): p. 4880-94.
55. Didierlaurent, A.M., et al., *Enhancement of adaptive immunity by the human vaccine adjuvant AS01 depends on activated dendritic cells*. J Immunol, 2014. **193**(4): p. 1920-30.
56. Crompton, P.D., S.K. Pierce, and L.H. Miller, *Advances and challenges in malaria vaccine development*. J Clin Invest, 2010. **120**(12): p. 4168-78.
57. Kester, K.E., et al., *Randomized, double-blind, phase 2a trial of falciparum malaria vaccines RTS,S/AS01B and RTS,S/AS02A in malaria-naive adults: safety, efficacy, and immunologic associates of protection*. J Infect Dis, 2009. **200**(3): p. 337-46.
58. Agnandji, S.T., et al., *A phase 3 trial of RTS,S/AS01 malaria vaccine in African infants*. N Engl J Med, 2012. **367**(24): p. 2284-95.
59. Baer, K., et al., *Release of hepatic Plasmodium yoelii merozoites into the pulmonary microvasculature*. PLoS Pathog, 2007. **3**(11): p. e171.

60. Sauerwein, R.W., M. Roestenberg, and V.S. Moorthy, *Experimental human challenge infections can accelerate clinical malaria vaccine development*. Nat Rev Immunol, 2011. **11**(1): p. 57-64.
61. Desai, S.A., S.M. Bezrukov, and J. Zimmerberg, *A voltage-dependent channel involved in nutrient uptake by red blood cells infected with the malaria parasite*. Nature, 2000. **406**(6799): p. 1001-5.
62. Alkhalil, A., et al., *Plasmodium falciparum likely encodes the principal anion channel on infected human erythrocytes*. Blood, 2004. **104**(13): p. 4279-86.
63. Chotivanich, K., et al., *Central role of the spleen in malaria parasite clearance*. J Infect Dis, 2002. **185**(10): p. 1538-41.
64. Khunrae, P., et al., *Full-length recombinant Plasmodium falciparum VAR2CSA binds specifically to CSPG and induces potent parasite adhesion-blocking antibodies*. J Mol Biol, 2010. **397**(3): p. 826-34.
65. Lavstsen, T., et al., *Plasmodium falciparum erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children*. Proc Natl Acad Sci U S A, 2012. **109**(26): p. E1791-800.
66. Maier, A.G., et al., *Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes*. Cell, 2008. **134**(1): p. 48-61.
67. Marti, M., et al., *Targeting malaria virulence and remodeling proteins to the host erythrocyte*. Science, 2004. **306**(5703): p. 1930-3.
68. !!! INVALID CITATION !!! {}.
69. de Koning-Ward, T.F., et al., *A newly discovered protein export machine in malaria parasites*. Nature, 2009. **459**(7249): p. 945-9.
70. Matz, J.M., K. Matuschewski, and T.W. Kooij, *Two putative protein export regulators promote Plasmodium blood stage development in vivo*. Mol Biochem Parasitol, 2013. **191**(1): p. 44-52.
71. Matthews, K., et al., *The Plasmodium translocon of exported proteins (PTEX) component thioredoxin-2 is important for maintaining normal blood-stage growth*. Mol Microbiol, 2013. **89**(6): p. 1167-86.
72. Bhattacharjee, S., et al., *Endoplasmic reticulum PI(3)P lipid binding targets malaria proteins to the host cell*. Cell, 2012. **148**(1-2): p. 201-12.
73. Boddey, J.A., et al., *An aspartyl protease directs malaria effector proteins to the host cell*. Nature, 2010. **463**(7281): p. 627-31.
74. Russo, I., et al., *Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte*. Nature, 2010. **463**(7281): p. 632-6.
75. Blisnick, T., et al., *Pfsbp1, a Maurer's cleft Plasmodium falciparum protein, is associated with the erythrocyte skeleton*. Mol Biochem Parasitol, 2000. **111**(1): p. 107-21.
76. Spycher, C., et al., *MAHRP-1, a novel Plasmodium falciparum histidine-rich protein, binds ferriprotoporphyrin IX and localizes to the Maurer's clefts*. J Biol Chem, 2003. **278**(37): p. 35373-83.
77. Pachlatko, E., et al., *MAHRP2, an exported protein of Plasmodium falciparum, is an essential component of Maurer's cleft tethers*. Mol Microbiol, 2010. **77**(5): p. 1136-52.
78. Kraemer, S.M. and J.D. Smith, *A family affair: var genes, PfEMP1 binding, and malaria disease*. Curr Opin Microbiol, 2006. **9**(4): p. 374-80.
79. Beck, J.R., et al., *PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes*. Nature, 2014. **511**(7511): p. 592-5.

80. Elsworth, B., et al., *PTEX is an essential nexus for protein export in malaria parasites*. Nature, 2014. **511**(7511): p. 587-91.
81. Spielmann, T. and T.W. Gilberger, *Protein export in malaria parasites: do multiple export motifs add up to multiple export pathways?* Trends Parasitol, 2010. **26**(1): p. 6-10.
82. Sargeant, T.J., et al., *Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites*. Genome Biol, 2006. **7**(2): p. R12.
83. Hiller, N.L., et al., *A host-targeting signal in virulence proteins reveals a secretome in malarial infection*. Science, 2004. **306**(5703): p. 1934-7.
84. Siau, A., et al., *Identification of a new export signal in Plasmodium yoelii: identification of a new exportome*. Cell Microbiol, 2014. **16**(5): p. 673-86.
85. de Bruin, D., M. Lanzer, and J.V. Ravetch, *The polymorphic subtelomeric regions of Plasmodium falciparum chromosomes contain arrays of repetitive sequence elements*. Proc Natl Acad Sci U S A, 1994. **91**(2): p. 619-23.
86. Carlton, J.M., et al., *Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii*. Nature, 2002. **419**(6906): p. 512-9.
87. Carlton, J.M., et al., *Comparative genomics of the neglected human malaria parasite Plasmodium vivax*. Nature, 2008. **455**(7214): p. 757-63.
88. Janssen, C.S., et al., *Plasmodium interspersed repeats: the major multigene superfamily of malaria parasites*. Nucleic Acids Res, 2004. **32**(19): p. 5712-20.
89. Pain, A., et al., *The genome of the simian and human malaria parasite Plasmodium knowlesi*. Nature, 2008. **455**(7214): p. 799-803.
90. Cunningham, D., et al., *The pir multigene family of Plasmodium: antigenic variation and beyond*. Mol Biochem Parasitol, 2010. **170**(2): p. 65-73.
91. Tachibana, S., et al., *Plasmodium cynomolgi genome sequences provide insight into Plasmodium vivax and the monkey malaria clade*. Nat Genet, 2012. **44**(9): p. 1051-5.
92. del Portillo, H.A., et al., *A superfamily of variant genes encoded in the subtelomeric region of Plasmodium vivax*. Nature, 2001. **410**(6830): p. 839-42.
93. Otto, T.D., et al., *A comprehensive evaluation of rodent malaria parasite genomes and gene expression*. BMC Biol, 2014. **12**: p. 86.
94. Fonager, J., et al., *Transcription and alternative splicing in the yir multigene family of the malaria parasite Plasmodium y. yoelii: identification of motifs suggesting epigenetic and post-transcriptional control of RNA expression*. Mol Biochem Parasitol, 2007. **156**(1): p. 1-11.
95. Cunningham, D., et al., *Rapid changes in transcription profiles of the Plasmodium yoelii yir multigene family in clonal populations: lack of epigenetic memory?* PLoS One, 2009. **4**(1): p. e4285.
96. Lawton, J., et al., *Characterization and gene expression analysis of the cir multi-gene family of Plasmodium chabaudi chabaudi (AS)*. BMC Genomics, 2012. **13**: p. 125.
97. Bozdech, Z., et al., *The transcriptome of Plasmodium vivax reveals divergence and diversity of transcriptional regulation in malaria parasites*. Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16290-5.
98. Fernandez-Becerra, C., et al., *Variant proteins of Plasmodium vivax are not clonally expressed in natural infections*. Mol Microbiol, 2005. **58**(3): p. 648-58.
99. Fischer, K., et al., *Ten families of variant genes encoded in subtelomeric regions of multiple chromosomes of Plasmodium chabaudi, a malaria species*

- that undergoes antigenic variation in the laboratory mouse. *Mol Microbiol*, 2003. **48**(5): p. 1209-23.
100. Dearsly, A.L., R.E. Sinden, and I.A. Self, *Sexual development in malarial parasites: gametocyte production, fertility and infectivity to the mosquito vector*. *Parasitology*, 1990. **100 Pt 3**: p. 359-68.
 101. Mackinnon, M.J. and A.F. Read, *Selection for high and low virulence in the malaria parasite Plasmodium chabaudi*. *Proc Biol Sci*, 1999. **266**(1420): p. 741-8.
 102. Knowles, G. and D. Walliker, *Variable expression of virulence in the rodent malaria parasite Plasmodium yoelii yoelii*. *Parasitology*, 1980. **81**(1): p. 211-9.
 103. Mackinnon, M.J., A. Bell, and A.F. Read, *The effects of mosquito transmission and population bottlenecks on virulence, multiplication rate and rosetting in rodent malaria*. *Int J Parasitol*, 2005. **35**(2): p. 145-53.
 104. Spence, P.J., et al., *Vector transmission regulates immune control of Plasmodium virulence*. *Nature*, 2013. **498**(7453): p. 228-31.
 105. Cunningham, D.A., et al., *Host immunity modulates transcriptional changes in a multigene family (yir) of rodent malaria*. *Mol Microbiol*, 2005. **58**(3): p. 636-47.
 106. Hall, N., et al., *A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses*. *Science*, 2005. **307**(5706): p. 82-6.
 107. Ebbinghaus, P. and J. Krucken, *Characterization and tissue-specific expression patterns of the Plasmodium chabaudi cir multigene family*. *Malar J*, 2011. **10**: p. 272.
 108. Yam, X.Y., et al., *Characterization of the Plasmodium Interspersed Repeats (PIR) proteins of Plasmodium chabaudi indicates functional diversity*. *Sci Rep*, 2016. **6**: p. 23449.
 109. Fonager, J., et al., *Reduced CD36-dependent tissue sequestration of Plasmodium-infected erythrocytes is detrimental to malaria parasite growth in vivo*. *J Exp Med*, 2012. **209**(1): p. 93-107.
 110. Bernabeu, M., et al., *Functional analysis of Plasmodium vivax VIR proteins reveals different subcellular localizations and cytoadherence to the ICAM-1 endothelial receptor*. *Cell Microbiol*, 2012. **14**(3): p. 386-400.
 111. Niang, M., et al., *STEVAR is a Plasmodium falciparum erythrocyte binding protein that mediates merozoite invasion and rosetting*. *Cell Host Microbe*, 2014. **16**(1): p. 81-93.
 112. Bachmann, A., et al., *A comparative study of the localization and membrane topology of members of the RIFIN, STEVAR and PfMC-2TM protein families in Plasmodium falciparum-infected erythrocytes*. *Malar J*, 2015. **14**: p. 274.
 113. Fernandez, V., et al., *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*. *J Exp Med*, 1999. **190**(10): p. 1393-404.
 114. Petter, M., et al., *Variant proteins of the Plasmodium falciparum RIFIN family show distinct subcellular localization and developmental expression patterns*. *Mol Biochem Parasitol*, 2007. **156**(1): p. 51-61.
 115. Saito, F., et al., *Immune evasion of Plasmodium falciparum by RIFIN via inhibitory receptors*. *Nature*, 2017. **552**(7683): p. 101-105.

116. Kaviratne, M., et al., *Small variant STEVOR antigen is uniquely located within Maurer's clefts in Plasmodium falciparum-infected red blood cells*. Eukaryot Cell, 2002. **1**(6): p. 926-35.
117. Araujo, R.B.D., et al., *Independent regulation of Plasmodium falciparum rif gene promoters*.
118. Fernandez-Becerra, C., et al., *Plasmodium vivax and the importance of the subtelomeric multigene vir superfamily*. Trends Parasitol, 2009. **25**(1): p. 44-51.
119. Pasini, E.M., et al., *Proteomic and genetic analyses demonstrate that Plasmodium berghei blood stages export a large and diverse repertoire of proteins*. Mol Cell Proteomics, 2013. **12**(2): p. 426-48.
120. Siau, A., et al., *Proteome mapping of Plasmodium: identification of the P. yoelii remodelome*. Sci Rep, 2016. **6**: p. 31055.
121. Cohen, S., G.I. Mc, and S. Carrington, *Gamma-globulin and acquired immunity to human malaria*. Nature, 1961. **192**: p. 733-7.
122. Webster, H.K., et al., *Characterization of antibodies to sporozoites in Plasmodium falciparum malaria and correlation with protection*. Journal of Clinical Microbiology, 1988. **26**(5): p. 923-927.
123. Diggs, C.L. and A.G. Osler, *Humoral immunity in rodent malaria. II. Inhibition of parasitemia by serum antibody*. J Immunol, 1969. **102**(2): p. 298-305.
124. Daly, T.M. and C.A. Long, *Humoral response to a carboxyl-terminal region of the merozoite surface protein-1 plays a predominant role in controlling blood-stage infection in rodent malaria*. J Immunol, 1995. **155**(1): p. 236-43.
125. Osier, F.H., et al., *Opsonic phagocytosis of Plasmodium falciparum merozoites: mechanism in human immunity and a correlate of protection against malaria*. BMC Med, 2014. **12**: p. 108.
126. Boyle, M.J., et al., *Human antibodies fix complement to inhibit Plasmodium falciparum invasion of erythrocytes and are associated with protection against malaria*. Immunity, 2015. **42**(3): p. 580-90.
127. Butcher, G.A., G.H. Mitchell, and S. Cohen, *Antibody mediated mechanisms of immunity to malaria induced by vaccination with Plasmodium knowlesi merozoites*. Immunology, 1978. **34**(1): p. 77-86.
128. ten Hagen, T.L., et al., *Role of adjuvants in the modulation of antibody isotype, specificity, and induction of protection by whole blood-stage Plasmodium yoelii vaccines*. J Immunol, 1993. **151**(12): p. 7077-85.
129. Florens, L., et al., *Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes*. Mol Biochem Parasitol, 2004. **135**(1): p. 1-11.
130. Nilsson Bark, S.K., et al., *Quantitative Proteomic Profiling Reveals Novel Plasmodium falciparum Surface Antigens and Possible Vaccine Candidates*. Mol Cell Proteomics, 2018. **17**(1): p. 43-60.
131. Bozdech, Z., et al., *Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray*. Genome Biol, 2003. **4**(2): p. R9.
132. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data*. Bioinformatics, 2014. **30**(15): p. 2114-20.
133. Soneson, C., M.I. Love, and M.D. Robinson, *Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences*. F1000Res, 2015. **4**: p. 1521.

134. Chang, Z.W., et al., *Ex Vivo Maturation Assay for Testing Antimalarial Sensitivity of Rodent Malaria Parasites*. Antimicrob Agents Chemother, 2016. **60**(11): p. 6859-6866.
135. Siau, A., et al., *Immunomic identification of malaria antigens associated with protection in mice*. Molecular & Cellular Proteomics, 2019: p. mcp.RA118.000997.
136. Chan, J.A., F.J. Fowkes, and J.G. Beeson, *Surface antigens of Plasmodium falciparum-infected erythrocytes as immune targets and malaria vaccine candidates*. Cell Mol Life Sci, 2014. **71**(19): p. 3633-57.
137. Pasternak, N.D. and R. Dzikowski, *PfEMP1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite Plasmodium falciparum*. Int J Biochem Cell Biol, 2009. **41**(7): p. 1463-6.
138. Cheng, Q., et al., *stevor and rif are Plasmodium falciparum multicopy gene families which potentially encode variant antigens*. Mol Biochem Parasitol, 1998. **97**(1-2): p. 161-76.
139. Belnoue, E., et al., *Vaccination with live Plasmodium yoelii blood stage parasites under chloroquine cover induces cross-stage immunity against malaria liver stage*. J Immunol, 2008. **181**(12): p. 8552-8.
140. Burghaus, P.A. and K. Lingelbach, *Luciferase, when fused to an N-terminal signal peptide, is secreted from transfected Plasmodium falciparum and transported to the cytosol of infected erythrocytes*. J Biol Chem, 2001. **276**(29): p. 26838-45.
141. Waller, R.F., et al., *Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway*. Embo j, 2000. **19**(8): p. 1794-802.
142. Renia, L., et al., *Immunization with a recombinant C-terminal fragment of Plasmodium yoelii merozoite surface protein 1 protects mice against homologous but not heterologous P. yoelii sporozoite challenge*. Infect Immun, 1997. **65**(11): p. 4419-23.
143. Daly, T.M. and C.A. Long, *A recombinant 15-kilodalton carboxyl-terminal fragment of Plasmodium yoelii yoelii 17XL merozoite surface protein 1 induces a protective immune response in mice*. Infect Immun, 1993. **61**(6): p. 2462-7.
144. Khosh-Naucke, M., et al., *Identification of novel parasitophorous vacuole proteins in P. falciparum parasites using BioID*. Int J Med Microbiol, 2018. **308**(1): p. 13-24.
145. Aguiar, J.C.P., MD, US), *NOVEL ANTIGEN FOR USE IN MALARIA*. 2017, CAMRIS International, Inc. (Bethesda, MD, US): United States.
146. Flick, K., et al., *Optimized expression of Plasmodium falciparum erythrocyte membrane protein 1 domains in Escherichia coli*. Malar J, 2004. **3**: p. 50.
147. Noe, A.R. and J.H. Adams, *Plasmodium yoelii YM MAEBL protein is coexpressed and colocalizes with rhoptry proteins*. Mol Biochem Parasitol, 1998. **96**(1-2): p. 27-35.
148. Yoeli, M., et al., *Sudden increase in virulence in a strain of Plasmodium berghei yoelii*. Ann Trop Med Parasitol, 1975. **69**(2): p. 173-8.
149. Raja, A.I., et al., *Chemically Attenuated Blood-Stage Plasmodium yoelii Parasites Induce Long-Lived and Strain-Transcending Protection*. Infect Immun, 2016. **84**(8): p. 2274-2288.
150. Etzler ME, S.A., Cummings RD. , *L-type Lectins*. In: Varki A, Cummings RD, Esko JD, et al., editors. *Essentials of Glycobiology*. 2nd edition. Cold Spring Harbor

- (NY): Cold Spring Harbor Laboratory Press; 2009, 2009. **Chapter 29 Available from: <http://www.ncbi.nlm.nih.gov/books/NBK1935/>.**
151. Vannucci, L., et al., *Viral vectors: a look back and ahead on gene transfer technology*. New Microbiol, 2013. **36**(1): p. 1-22.
 152. Ophorst, O.J., et al., *Immunogenicity and protection of a recombinant human adenovirus serotype 35-based malaria vaccine against Plasmodium yoelii in mice*. Infect Immun, 2006. **74**(1): p. 313-20.
 153. Draper, S.J., et al., *Effective induction of high-titer antibodies by viral vector vaccines*. Nat Med, 2008. **14**(8): p. 819-21.
 154. Flingai, S., et al., *Synthetic DNA vaccines: improved vaccine potency by electroporation and co-delivered genetic adjuvants*. Front Immunol, 2013. **4**: p. 354.
 155. Thera, M.A., et al., *A field trial to assess a blood-stage malaria vaccine*. N Engl J Med, 2011. **365**(11): p. 1004-13.
 156. Spring, M.D., et al., *Phase 1/2a study of the malaria vaccine candidate apical membrane antigen-1 (AMA-1) administered in adjuvant system AS01B or AS02A*. PLoS One, 2009. **4**(4): p. e5254.
 157. Sirima, S.B., S. Cousens, and P. Druilhe, *Protection against malaria by MSP3 candidate vaccine*. N Engl J Med, 2011. **365**(11): p. 1062-4.
 158. Palmenberg, A.C., et al., *Proteolytic processing of the cardioviral P2 region: primary 2A/2B cleavage in clone-derived precursors*. Virology, 1992. **190**(2): p. 754-62.
 159. Ryan, M.D., A.M. King, and G.P. Thomas, *Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence*. J Gen Virol, 1991. **72 (Pt 11)**: p. 2727-32.
 160. Straimer, J., et al., *Site-specific genome editing in Plasmodium falciparum using engineered zinc-finger nucleases*. Nat Methods, 2012. **9**(10): p. 993-8.
 161. Huang, X., et al., *Differential Spleen Remodeling Associated with Different Levels of Parasite Virulence Controls Disease Outcome in Malaria Parasite Infections*. mSphere, 2016. **1**(1).
 162. Franke-Fayard, B., et al., *A Plasmodium berghei reference line that constitutively expresses GFP at a high level throughout the complete life cycle*. Mol Biochem Parasitol, 2004. **137**(1): p. 23-33.
 163. Claser, C., et al., *Adaptive immunity is essential in preventing recrudescence of Plasmodium yoelii malaria parasites after artesunate treatment*. Cell Microbiol, 2017. **19**(11).
 164. Leite, J.A., et al., *Immunization with the MAEBL M2 Domain Protects against Lethal Plasmodium yoelii Infection*. Infect Immun, 2015. **83**(10): p. 3781-92.
 165. Vogt, A.M., et al., *Heparan sulfate on endothelial cells mediates the binding of Plasmodium falciparum-infected erythrocytes via the DBL1alpha domain of PfEMP1*. Blood, 2003. **101**(6): p. 2405-11.
 166. Scherf, A., et al., *Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum*. Embo j, 1998. **17**(18): p. 5418-26.
 167. Subramani, R., et al., *Plasmodium falciparum-infected erythrocyte knob density is linked to the PfEMP1 variant expressed*. MBio, 2015. **6**(5): p. e01456-15.

168. Ingmundson, A., et al., *The exported Plasmodium berghei protein IBIS1 delineates membranous structures in infected red blood cells*. Mol Microbiol, 2012. **83**(6): p. 1229-43.
169. Graewe, S., et al., *Hostile takeover by Plasmodium: reorganization of parasite and host cell membranes during liver stage egress*. PLoS Pathog, 2011. **7**(9): p. e1002224.
170. Inoue, S., et al., *Roles of IFN-gamma and gammadelta T Cells in Protective Immunity Against Blood-Stage Malaria*. Front Immunol, 2013. **4**: p. 258.
171. De Souza, J.B., et al., *Early gamma interferon responses in lethal and nonlethal murine blood-stage malaria*. Infect Immun, 1997. **65**(5): p. 1593-8.
172. Choudhury, H.R., et al., *Early nonspecific immune responses and immunity to blood-stage nonlethal Plasmodium yoelii malaria*. Infect Immun, 2000. **68**(11): p. 6127-32.
173. Costa, G., et al., *Control of Plasmodium falciparum erythrocytic cycle: gammadelta T cells target the red blood cell-invasive merozoites*. Blood, 2011. **118**(26): p. 6952-62.
174. Fougere, A., et al., *Variant Exported Blood-Stage Proteins Encoded by Plasmodium Multigene Families Are Expressed in Liver Stages Where They Are Exported into the Parasitophorous Vacuole*. PLoS Pathog, 2016. **12**(11): p. e1005917.
175. Przyborski, J.M., et al., *Trafficking of STEVOR to the Maurer's clefts in Plasmodium falciparum-infected erythrocytes*. Embo j, 2005. **24**(13): p. 2306-17.
176. Khattab, A. and S. Meri, *Exposure of the Plasmodium falciparum clonally variant STEVOR proteins on the merozoite surface*. Malar J, 2011. **10**: p. 58.
177. Khattab, A., et al., *Plasmodium falciparum variant STEVOR antigens are expressed in merozoites and possibly associated with erythrocyte invasion*. Malar J, 2008. **7**: p. 137.
178. Blythe, J.E., et al., *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*. Infect Immun, 2008. **76**(7): p. 3329-36.
179. Abdel-Latif, M.S., et al., *Recognition of variant Rifin antigens by human antibodies induced during natural Plasmodium falciparum infections*. Infect Immun, 2002. **70**(12): p. 7013-21.
180. Abdel-Latif, M.S., et al., *Antibodies to rifin: a component of naturally acquired responses to Plasmodium falciparum variant surface antigens on infected erythrocytes*. Am J Trop Med Hyg, 2004. **71**(2): p. 179-86.
181. Josling, G.A. and M. Llinas, *Sexual development in Plasmodium parasites: knowing when it's time to commit*. Nat Rev Microbiol, 2015. **13**(9): p. 573-87.
182. Magowan, C., et al., *Cytoadherence by Plasmodium falciparum-infected erythrocytes is correlated with the expression of a family of variable proteins on infected erythrocytes*. J Exp Med, 1988. **168**(4): p. 1307-20.
183. Bray, N.L., et al., *Near-optimal probabilistic RNA-seq quantification*. Nat Biotechnol, 2016. **34**(5): p. 525-7.
184. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.
185. Voss, T.S., et al., *A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria*. Nature, 2006. **439**(7079): p. 1004-8.

186. Williams, C.T. and A.F. Azad, *Transcriptional analysis of the pre-erythrocytic stages of the rodent malaria parasite, Plasmodium yoelii*. PLoS One, 2010. **5**(4): p. e10267.
187. Lima, W.R., et al., *Genome-wide analysis of the human malaria parasite Plasmodium falciparum transcription factor PfNF-YB shows interaction with a CCAAT motif*. Oncotarget, 2017. **8**(69): p. 113987-114001.
188. Zhu, L., et al., *New insights into the Plasmodium vivax transcriptome using RNA-Seq*. Sci Rep, 2016. **6**: p. 20498.
189. Brown, K.E., et al., *Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division*. Mol Cell, 1999. **3**(2): p. 207-17.
190. Gasser, S.M., *Positions of potential: nuclear organization and gene expression*. Cell, 2001. **104**(5): p. 639-42.
191. Duraisingh, M.T., et al., *Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum*. Cell, 2005. **121**(1): p. 13-24.
192. Araujo, R.B.D., et al., *Independent regulation of Plasmodium falciparum rif gene promoters*. Sci Rep, 2018. **8**(1): p. 9332.
193. Taylor, H.M., et al., *A study of var gene transcription in vitro using universal var gene primers*. Mol Biochem Parasitol, 2000. **105**(1): p. 13-23.
194. al-Khedery, B., J.W. Barnwell, and M.R. Galinski, *Antigenic variation in malaria: a 3' genomic alteration associated with the expression of a P. knowlesi variant antigen*. Mol Cell, 1999. **3**(2): p. 131-41.
195. Guizetti, J., A. Barcons-Simon, and A. Scherf, *Trans-acting GC-rich non-coding RNA at var expression site modulates gene counting in malaria parasite*. Nucleic Acids Res, 2016. **44**(20): p. 9710-9718.
196. Jaschke, A., et al., *Merozoite Surface Protein 1 from Plasmodium falciparum Is a Major Target of Opsonizing Antibodies in Individuals with Acquired Immunity against Malaria*. Clin Vaccine Immunol, 2017. **24**(11).

9 Appendix

Appendix 1

Buffer list for HaloTag purification

1. Equilibration step (HaloTag buffer + 0.05% IPEGAL)
2. Binding step (HaloTag buffer + 1mM DTT + 2% glycerol)
3. 1st wash (HaloTag buffer + 0.1% TWEEN20 + 0.05% IPEGAL)
4. 2nd wash (HaloTag buffer + 500mM NaCl)
5. 3rd wash (HaloTag buffer + 0.05% Triton X114 + 0.05% IPEGAL)
6. Cleavage (TEV + 1mM DTT)

Appendix 2

Figure 1. pH6HTN His₆HaloTag® T7 Vector circle map and sequence reference points.

Plasmid map of Halotag vector.

Appendix 3

Plasmid map of pET-24a

Appendix 4

	PY06306 (25ug)	PY06306 (50ug)
Mouse 1	< 1000	20000
Mouse 2	< 1000	20000
Mouse 3	40000	1000
Mouse 4	< 1000	5000
Mouse 5	5000	1000

ELISA result of trial immunization with PY06306A

Appendix 5

YM challenge run 2

Appendix 6

Appendix 7

	TPM of yir genes			
	PY17X_1100077	PY17X_0401000	PY17X_0114700	PY17X_1100081
Parental YM	5.95	3.46	1.05	1.04
SC1	513.34	210.90	221.52	38.16
SC2	557.94	547.61	418.62	38.52

TPM counts of upregulated yir genes in RNAseq.

Appendix 8

Gene specific qPCR Primer	Efficiency
PY17X_1100077	98.13%
PY17X_0401000	111.72%
PY17X_0114700	109.94%
PY17X_1046201	110.51%
PY17X_1046301	103.59%
PY17X_0102100	111.72%
PY17X_1400400	111.45%

Primer efficiency for qPCR

Primer list

For generation of Halotag construct			
4147A-SacI-F	5-	TCAGAGCTCAAAATAAATGTGAGAATGAATCACCAAATAA	-3
4147A-NotI-R	5-	AAGCGGCCGCTATTCTTCTTGAGAAGATAAATTTATATTTTGA GAT	-3
For generation of Histag construct			
4372A-NheI-F	5-	TTGCTAGCAATGCCGAAACACAAAAAGATGCAAATAATGG	-3
4372A-XhoI-R	5-	TTCTCGAGTAAATTTGAAGAAGTAAGTTTATATGGGT	-3
6306A-NheI-F	5-	TTGCTAGCGGAACAAGTTGTATTATATCTGATCGAATATT	-3
6306A-XhoI-R	5-	TTCTCGAGTGAATTTTTAAATGAATTTATATGATTAT	-3
For generation of HPXY construct			
T2A F for XY ApaI	5-	ATAGGGCCCCGGCGCCGTCGACGGAGAAGGAAGAGG	-3
T2A R for XY NotI	5-	AGTGTATGCGGCCGCGGATCCAACGGGATCAAAGCT	-3
eGFP-ApaI F	5-	ATAGGGCCCCGTGAGCAAGGGCGAGGAGCTG	-3
eGFP-KasI R	5-	TATGGCGCCCTTGTACAGCTCGTCCATG	-3
For checking integration of T2a into genomic locus			
Int check T2A0017400	5-	ATGAGGAAACATTTACCCGATAAATTAGG	-3
0035300 int check	5-	GATTATCCCTGGTTTTATAATGTTAC	-3
0400700 int check	5-	TATAGGATTGCTAGGTTTCTGTAAGG	-3
GFP p for sequencing T2A R	5-	CCGGTGAACAGCTCCTCGCCCTTGCTCAC	-3
For construction of HPXY plasmid			
T2A F for XY ApaI	5-	ATAGGGCCCCGGCGCCGTCGACGGAGAAGGAAGAGG	-3
T2A R for XY NotI	5-	AGTGTATGCGGCCGCGGATCCAACGGGATCAAAGCT	-3
eGFP-ApaI F	5-	ATAGGGCCCCGTGAGCAAGGGCGAGGAGCTG	-3

eGFP-KasI R	5-	TATGGCGCCCTTGTACAGCTCGTCCATG	-3
PCHAS PY6306 AvrII EcorV HindIII F	5-	ACCAAGCTTGATATCCCTAGGAAAGTTACAGACCC	-3
PCHAS PY6306 ApaI R	5-	TATGGGGCCCTTTGATGGTATATCCAAAAAAAAGG	-3
For qPCR reaction			
PYYM_040 1200 F1	5-	TGATAGTAGAGATCCTAGTAAGTC	-3
PYYM_040 1200 R1	5-	ATTTGGCTTGTTTGCATCAAGTTC	-3
PYYM_010 1500 F1	5-	ACAAACTTAATGAAAATGTACCCG	-3
PYYM_010 1500 R1	5-	ATGCTTTCCATATAATTTATCACC	-3
PYYM_001 7400 F2	5-	GATACTGATTTTCGATAAAAATACGGC	-3
PYYM_001 7400 R2	5-	ATCATTTATTTTAGTGAATTTGTTCCC	-3
PY06158 qPCR F	5-	GCTGATAACCAACCAGGAGTC	-3
PY06158 qPCR R	5-	CAGGTGGGATACCATCTAAATG	-3
PYYM_011 4100 F1	5-	GGTTTTTCAGCTGGTCAAGATG	-3
PYYM_011 4100 R1	5-	ATTTTATGGTTAGAGTACTCCACG	-3
PYYM_140 2400 F1	5-	TAATATCCTCCAGAAAGTAACGAAG	-3
PYYM_140 2400 R1	5-	CATCAAACAAAAATAAACATGCAGC	-3
Py17X_104 6201 F1	5-	TGAGGAAGCATTTCCCTGATG	-3
Py17X_104 6201 R1	5-	ATCCAGTATTGCCTTTAATTGGG	-3
Py17X_104 6301 F1	5-	ATGTATGGGATGATTTCCCCG	-3
Py17X_104 6301 R1	5-	AACCAGAAGGCCCAAAAAATTG	-3