

ANALYSIS OF GENETIC DIVERSITY IN A NOVEL VACCINE CANDIDATE, THE  
GLUTAMIC ACID-RICH PROTEIN OF *PLASMODIUM FALCIPARUM* (PfGARP)



Miss Rattanaporn Rojrung

A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science in Medical Sciences

FACULTY OF MEDICINE

Chulalongkorn University

Academic Year 2022

Copyright of Chulalongkorn University

การวิเคราะห์ความหลากหลายทางพันธุกรรมของยีนสำหรับโปรตีนที่มีกรดกลูตามิกสูงของเชื้อพ  
ลาสโมเดียมฟัลซิพารัม (PfGARP) ที่มีศักยภาพเป็นองค์ประกอบของวัคซีนชนิดใหม่



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาวิทยาศาสตร์การแพทย์ไม่สังกัดภาควิชา/เทียบเท่า  
คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ปีการศึกษา 2565  
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title ANALYSIS OF GENETIC DIVERSITY IN A NOVEL  
VACCINE CANDIDATE, THE GLUTAMIC ACID-RICH  
PROTEIN OF *PLASMODIUM FALCIPARUM* (PfGARP)

By Miss Rattanaporn Rojrung

Field of Study Medical Sciences

Thesis Advisor Professor Doctor Somchai Jongwutiwes

Thesis Co Advisor Professor Doctor CHATURONG PUTAPORNTIP

---

Accepted by the FACULTY OF MEDICINE, Chulalongkorn University in Partial  
Fulfillment of the Requirement for the Master of Science

..... Dean of the FACULTY OF MEDICINE  
(Associate Professor CHANCHAI SITTIPUNT, Ph.D.)

THESIS COMMITTEE

..... Chairman  
(Associate Professor Doctor WILAI ANOMASIRI)

..... Thesis Advisor  
(Professor Doctor Somchai Jongwutiwes)

..... Thesis Co-Advisor  
(Professor Doctor CHATURONG PUTAPORNTIP)

..... Examiner  
(Associate Professor Dortor WACHAREE  
LIMPANASITHIKUL)

..... External Examiner  
(Associate Professor Kriengsak Limkittikul)



# # 6370045630 : MAJOR MEDICAL SCIENCES

KEYWORD: Genetic diversity, *Plasmodium falciparum*, pfgarp

Rattanaorn Rojrung : ANALYSIS OF GENETIC DIVERSITY IN A NOVEL VACCINE CANDIDATE, THE GLUTAMIC ACID-RICH PROTEIN OF *PLASMODIUM FALCIPARUM* (PfGARP). Advisor: Prof. Dr. Somchai Jongwutiwes Co-advisor: Prof. Dr. CHATURONG PUTAPORNTIP

Malaria vaccine development confronts high genetic polymorphism in some antigenic components of the vaccine candidates, which could potentially limit the vaccine efficacy. An antigen of *Plasmodium falciparum*, designated glutamic acid-rich protein (PfGARP), has recently been considered as a strong vaccine target against asexual-blood-stages parasite. It has been shown that naturally acquired anti-PfGARP antibodies elicited clinical protection against severe malaria among people living in hyperendemic areas. Furthermore, PfGARP has been involved in binding to human erythrocyte band 3, which induces erythrocyte aggregation and might bring about severe malaria. However, the extent of genetic variation in the *pfgarp* locus among field isolates remains unknown. Herein, the genetic diversity in the PfGARP encoding gene was evaluated using samples collected during 2009 - 2014 from different malaria-endemic areas in Thailand (Chanthaburi ( $n = 20$ ), Tak ( $n = 20$ ), Ubon Ratchathani ( $n = 20$ ), and Yala ( $n = 20$ ) Provinces) by direct sequencing. The analyzed data on 80 sequences revealed high conservation in *pfgarp* among the parasite populations. The gene contains two exons, exon I and exon II where the latter consists of 8 SNPs and 8 repeats. Twenty-six haplotypes distribute among the populations. Selective pressure measured by Tajima's  $D$  was 0.397, not statistically significant ( $p$  value  $> 0.10$ ), suggesting no deviation from selective neutrality at this locus among the parasite populations in Thailand. The lack of high genetic diversity in this gene appears to be a strong component for malaria vaccine.

Field of Study: Medical Sciences

Student's Signature .....

Academic Year: 2022

Advisor's Signature .....

Co-advisor's Signature .....

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my grateful gratitude to my supervisor, Professor. Dr. Somchai Jongwutiwes for his continuous support throughout my thesis. This journey would not be started without the opportunity he gave me to join his research group. I was very impressed by his warmly welcoming, inspiring, enthusiasm, and correcting my writing and understanding, these all encourage me to complete my master's degree. I also deeply appreciated to co-advisor, Associate Professor Dr. Chaturong Puthapornpit. He helps me a lot in the laboratory parts. He also gave me immense knowledge about practical training and solving technical problems.

I would like to express my sincere thank to all my research committees, Associate Professor Dr. Wilai Anomasiri, Associate Professor Dr. Watcharee Limpanasithikul, and Associate Professor Dr. Kriengsak Limkittikul for their helpful recommendations.

My grateful thanks are also extended to all lab members. They cooperate with me to find the sample stock. They are also training me the laboratory skills. I am very appreciative and thank them a lot. I would like to thank all of the patients who own the blood samples that were used in this study.

The accomplishment of this research was financially supported by Ratchadapisek Sompoch Fund, Faculty of Medicine Chulalongkorn University. The Grant number is GA64/18. I would like to thank the DPST scholarship for providing an M.Sc. scholarship. I do grateful for the financial support in many ways.

Thanks to the Medical Science Program and malaria and opportunistic parasite research unit, Department of Parasitology. Faculty of Medicine Chulalongkorn University for providing me the opportunity to apply to the program and facilitate the laboratory equipment.

Friends, persons who are deserved my appreciation. They are one of the best people in life. They are like a power source. They make me feel like everything can be possible.

Last but not the least, million thanks to my family and my beloved parents for always beside and spirit me in whatever ways I choose. I do love you all the most.

Rattanaporn Rojrung



จุฬาลงกรณ์มหาวิทยาลัย  
**CHULALONGKORN UNIVERSITY**

## TABLE OF CONTENTS

	<b>Page</b>
.....	iii
ABSTRACT (THAI).....	iii
.....	iv
ABSTRACT (ENGLISH).....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vii
CHAPTER 1.....	1
INTRODUCTION.....	1
1.1 Background of The Study.....	1
1.2 Objectives.....	4
1.3 Scope.....	4
CHAPTER 2.....	5
LITERATURE REVIEW.....	5
2.1 Biology and Distribution of Malaria Parasites.....	5
2.2 Malaria Mosquito Vectors.....	8
2.3 Malaria Preventive Measures and Malaria Vaccine Development.....	9
2.3.1 Pre-erythrocytic Vaccine Candidates.....	10
2.3.2 Asexual Erythrocytic Vaccine Candidates.....	11
2.3.3 Transmission Blocking Vaccine Candidates.....	12
2.4 Glutamic Acid-rich Protein of <i>Plasmodium falciparum</i> (PfGARP).....	13



2.5 Polymorphism in <i>Plasmodium falciparum</i> Vaccine Candidates and Potential Impacts on Vaccine Efficacy .....	16
2.6 Malaria Diagnostic Methods .....	17
2.6.1 Microscopy-based Detection.....	17
2.6.2 Polymerase Chain Reaction-based Diagnosis.....	19
2.7 Genes for <i>Plasmodium</i> species Identification .....	23
2.7.1 Small subunit ribosomal RNA (SSU rRNA or 18S rRNA) gene .....	23
2.8 Sanger-based DNA Sequencing.....	24
CHAPTER 3 .....	26
MATERIALS AND METHODS.....	26
3.1 Target Population .....	26
3.2 Sample Size .....	26
3.3 Sources of <i>Plasmodium falciparum</i> Isolates .....	26
3.4 Ethic Consideration .....	26
3.5 Slide Preparation and Parasite Density Estimation .....	27
3.5.1 Thin-Blood Film Preparation.....	27
3.5.2 Thick-blood Film Preparation.....	28
3.5.3 Parasite Density Estimation from the Thin Blood Film.....	28
3.5.4 Parasite Density Estimation from the Thick Blood Film.....	29
3.6 DNA Extraction.....	29
3.7 <i>Plasmodium</i> species Identification by PCR .....	30
3.8 Determination of Clonal Mixture in <i>P. falciparum</i> Samples .....	33
3.8.1 Genotyping of Block 2 of <i>pfmsp1</i> .....	33
3.9 Amplification of <i>pfgap</i> .....	38

3.10 PCR Product Purification.....	38
3.11 DNA Sequencing.....	39
3.12 Data Analysis .....	41
CHAPTER 4 .....	43
RESULTS .....	43
4.1 Study Areas and Parasite Populations.....	43
4.2 Demographic Data of Study Populations.....	45
4.3 Comparison of the Parasite Density by Study Area, Sex, Ethnicity, and <i>msp1</i> -genotype...46	
4.4 Amplification of PfGARP-encoding Gene .....	49
4.5 Structure of PfGARP-encoding Gene .....	49
4.6 Genetic Diversity of the PfGARP-encoding Gene.....	53
4.6.1 Nucleotide Diversity of the PfGARP-encoding Gene .....	53
4.6.2 Haplotype Diversity .....	54
4.7 Structural Organization of PfGARP-encoding Gene .....	57
4.7.1 Exon I and Intron .....	57
4.7.2 Non-repeat Regions in Exon II .....	57
4.7.3 Repetitive Regions .....	60
4.7.4 Homopolymeric Glutamic Acid Repeats .....	61
4.8 Recombination .....	61
4.9 Selective Pressure on PfGARP .....	70
4.10 Phylogenetic Analysis .....	72
4.11 Genetic Differentiation of <i>P. falciparum</i> based on <i>pfgarp</i> .....	73
4.12 Association between The Number of Amino Acid and Parasite Density .....	74
4.13 Antigenic Epitope Prediction .....	76

4.13.1 B-cell Epitope Prediction .....	76
4.13.2 T-Cell Epitope Prediction .....	79
CHAPTER 5 .....	84
DISCUSSION AND CONCLUSIONS .....	84
5.1 Discussion and Conclusions .....	84
REFERENCES .....	89
VITA .....	101



# CHAPTER 1

## INTRODUCTION

### 1.1 Background of The Study

Malaria is an enormous public health burden caused by *Plasmodium* infection, affecting more than 241 million people worldwide and responsible for approximately 627,000 deaths (WHO, 2021). The life cycle of malaria parasite involves sexual development in anopheline mosquitoes and asexual reproduction in humans. Natural transmission of malaria occurs when female mosquito bites and inoculates infective malarial sporozoites into human blood stream. *Plasmodium* species that naturally cause malaria in humans are *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, *Plasmodium ovale wallikeri*, *Plasmodium ovale curtisi*, *Plasmodium knowlesi*, *Plasmodium cynomolgi* and *Plasmodium fieldi*, *Plasmodium inui* (Perkins, 2014; Putaporntip *et al.*, 2021). Infections caused by the first four species are limited to humans while the latter four species are considered to be zoonosis with long-tailed (*Macaca fascicularis*) and pig-tailed (*M. nemestrina*) macaques serve as natural reservoir hosts (Fong *et al.*, 1971). Among these species, *P. falciparum* causes the most severe and fatal illnesses (Guerra *et al.*, 2008).

Although antimalarial drugs, including chloroquine, sulfadoxine-pyrimethamine, mefloquine and artemisinin, have been extensively and globally used to treat malaria, the emergence and the spread of drug resistant strains have been a major hurdle for both therapeutic efficacy and disease control, especially the widespread occurrence of multidrug-resistant *P. falciparum* (N. J. White, 1996; Putaporntip *et al.*, 2016; Buppan *et al.*, 2018; Haldar *et al.*, 2018). Meanwhile, anopheline mosquito vectors in various malaria endemic areas develop resistant to several insecticides (WHO, 2012; Moyes *et al.*, 2020). As an alternative, malaria vaccine development is one of the potential strategies to prevent and control malaria. Despite no currently available effective vaccine against malaria, many vaccine candidates have been evaluated and tested beyond pre-clinical trials (Chatterjee & Cockburn, 2021). Currently, there are 3 major approaches for the development of malaria vaccines that correspond to three stages of parasite's

life cycle, i.e. pre-erythrocytic stages, asexual erythrocytic stages and sexual stages in mosquito vector (Hill, 2011; Duffy & Patrick Gorres, 2020). The current leading pre-erythrocytic vaccine candidate is RTS, S vaccine derived from the major surface coat of sporozoite, known as circumsporozoite protein (CSP). The aim of this vaccine is to interrupt the initial infection by targeting the sporozoites introduced by mosquitoes; thereby, no hepatic stage development ensues. However, this vaccine showed varying protective efficacy in African children and the results seem to depend on various factors such as types of adjuvant, age groups of vaccinated volunteers, host genetic backgrounds, geography and duration of follow-up (Nei & Gojobori, 1986; RTS, 2015; Targett, 2015; Nielsen *et al.*, 2018). On the other hand, asexual erythrocytic vaccine aims to block red blood cell entry by malarial merozoites. A number of proteins expressed during late schizogony that appear on the surface of merozoite stage have been identified as potential vaccine targets. Of these, the most attractive candidates belong to merozoite surface proteins (MSP) which include MSP1, MSP2 and MSP3 (Ouattara & Laurens, 2015). Although, MSP1 is the major protein component of vaccine for blood-stage malaria, the strong potential is limited due to extensive polymorphism of antigenic region (Miller *et al.*, 1993). The third vaccine approach is to block the transmission of malaria parasites from mosquitoes to humans by interruption of sexual stage development in the vectors. However, vaccinated individuals with transmission blocking vaccines are not protected from symptomatic malaria while further malaria spreading is halted by the efficacy of vaccine (Nunes *et al.*, 2014).

A high degree of genetic diversity of *P. falciparum* vaccine-candidate antigens encoding loci may affect efficacy of malaria vaccines. Several malaria antigens elicit allele-specific antibody responses to their variant epitopes that may compromise protective efficacy among vaccinated individuals. It is noteworthy that during the initial stage of malaria vaccine development, the candidate immunogens are generally designed based on proteins derived from certain clones or strains of malaria parasites that may not be the best representative of parasites circulating in each endemic area. For example, the circumsporozoite protein component in the RTS, S vaccine has been derived from the culture-adapted and chloroquine-sensitive 3D7 clone of *P. falciparum* while in most malaria endemic areas parasites carrying different variants of this protein have been found in both clinical samples and field isolates (Cohen *et al.*, 2010). Likewise, in human trials, it has been found that antibodies induced by vaccine containing 3D7-type MSP2

of *P. falciparum* could not protect vaccinated subjects from heterologous challenge with parasites carrying MSP2-FC27 variant (Genton *et al.*, 2002). Importantly, sequence analysis of *P. falciparum* MSP2 among clinical isolates in Thailand has revealed that both FC27 and 3D7 types co-circulated while a remarkable sequence variation has been detected among parasite populations from different endemic areas (Putaporntip *et al.*, 2008). These findings have suggested that effective malaria vaccines require incorporation of possible or major variants of vaccine candidate proteins.

Besides merozoite surface antigens as asexual erythrocytic vaccine components, a novel *P. falciparum* protein with an approximate molecular weight of 80 kDa, designated glutamic acid-rich protein (PfGARP), has been considered a promising vaccine target. PfGARP is predominantly expressed during young and growing trophozoites and exported to exofacial surface of infected erythrocytes (Triglia *et al.*, 1988; Almukadi *et al.*, 2019; Raj *et al.*, 2020). Interestingly, PfGARP seems to play a crucial role in mediating the adhesion properties of infected erythrocytes by serving as a ligand for erythrocyte band 3, which may lead to the pathogenesis of severe falciparum malaria (Almukadi *et al.*, 2019). Furthermore, increased expression of PfGARP was observed among *P. falciparum* isolates from symptomatic Tanzanian children, suggesting the potential role of this protein in contributing to the pathogenesis of childhood malaria (Vignali *et al.*, 2011). Proteomic analysis has revealed that PfGARP was predominantly recognized by antibodies from individuals in endemic areas who had lower parasitemia or benign malarial symptoms than those harboring higher parasite density or severe malaria (Raj *et al.*, 2020). More importantly, antibodies raised against recombinant PfGARP elicited intraerythrocytic parasite killing *in vitro* through disruption of food vacuole integrity, alteration of mitochondrial membrane potential, activation of caspase-like proteases and fragmentation of parasite DNA, all of which indicate the process of programmed cell death (Raj *et al.*, 2020). Meanwhile, anti-PfGARP antibodies from either immunized mice or purified from plasma of adults living in Tanzania, a malaria hyperendemic area, elicited 94-99% inhibition of parasite growth *in vitro*. Moreover, vaccination trials in owl (*Aotus*) monkeys using either PfGARP mRNA vaccine or recombinant PfGARP protein have conferred partial protection upon challenge infection with heterologous *P. falciparum* strain. These immunized monkeys had significantly lower levels of parasitemia than those in the control group (Raj *et al.*, 2020).

Therefore, PfGARP has been considered to be a component of multi-stage malaria vaccine against *P. falciparum*.

Although the immunogenicity and protective efficacy of antibodies against PfGARP have been investigated, the extent of genetic polymorphism in the *pfgarp* locus among field isolates remains unknown. Therefore, analysis of the *pfgarp* sequences among clinical isolates is crucial for the basis of vaccine development against *P. falciparum* infection.

## 1.2 Objectives

1. To analyze the extent of sequence diversity in *pfgarp* of *P. falciparum* from Thai isolates.
2. To determine genetic mechanisms contributing to sequence variation in *pfgarp*.
3. To investigate whether *pfgarp* of *P. falciparum* isolates from diverse endemic areas of Thailand displays population structure.
4. To determine the association between *pfgarp* haplotypes and parasite density among *falciparum* malaria patients.

## 1.3 Scope

Observation of the genetic diversity in *pfgarp* in the different areas in Thailand, including Chanthaburi ( $n = 20$ ), Tak ( $n = 20$ ), Ubon Ratchathani ( $n = 20$ ), and Yala ( $n = 20$ ) Provinces. Blood samples used in this study were collected during 2009-2014 from *falciparum*-malaria-infected patients. DNA samples from those bloods were extracted. Microscopy examination was employed for the percent parasitemia and the parasite density estimations. PCR technique was used for the malaria parasite species identification. Clonal detection was performed by PCR targeting *msp1* and *msp2*. PCR for the *pfgarp* amplification was applied. The genetic diversity in *pfgarp* sequences was evaluated.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Biology and Distribution of Malaria Parasites

Malaria is an infectious disease caused by unicellular apicomplexan protozoa in the genus *Plasmodium*. There are around 200 species in the genus *Plasmodium* that can infect diverse classes of vertebrate hosts, including reptiles, birds, ungulates, nonhuman primates and humans. *Plasmodium* species that incriminate in natural infections in humans are *P. falciparum*, *P. vivax*, *P. ovale wallikeri*, *P. ovale curtisi*, *P. malariae*, *P. knowlesi* and *P. cynomolgi*, *P. inui* and *P. fieldi* in which the latter four species have been circulating in Southeast Asian macaques as natural reservoir hosts (Perkins, 2014; Putaporntip *et al.*, 2021). Of these, *P. falciparum* and *P. vivax* are the main causative agents of human malaria that inflict over 200 millions of the world population, particularly those living in tropical region. Both species are sympatric in most malaria endemic areas except Sub-Saharan Africa where *P. vivax* occurs at very low prevalence. Globally, *P. falciparum* is the most common species responsible for morbidity and mortality whereas *P. vivax* can cause chronic relapsing illnesses and has wider geographic distribution (Mendis *et al.*, 2001; Guerra *et al.*, 2008). On the other hand, the epidemiology of malaria caused by *P. malariae*, *P. ovale wallikeri* and *P. ovale curtisi* displays relatively low prevalence worldwide while indigenous ovale malaria cases have not been documented in Central and South America. On the other hand, both simian malaria caused by *P. knowlesi*, *P. cynomolgi*, *P. inui* and *P. fieldi* are confined to Southeast Asia including Thailand (Jongwutiwes *et al.*, 2011; Perkins, 2014; Putaporntip *et al.*, 2021; Putaporntip *et al.*, 2022).

The life cycle of malaria parasites involves both sexual development in mosquitoes and asexual reproduction in vertebrate hosts (Garnham, 1966) (Figure 1). After a female anopheline mosquito takes blood meals from an infected individual, the intraerythrocytic gametocytes are transferred from human circulation into the mosquito's midgut. Subsequently, the erythrocyte membrane covering gametocytes is ruptured, releasing free gametocytes into the mosquito's midgut lumen where subsequent gametogenesis takes place through the process of exflagellation



for male gametocytes and transformation with maturation for female gametocytes, generating male and female gametes, respectively (Bennink *et al.*, 2016). Syngamy or fertilization occurs between male and female gametes and zygote, resulting in the formation of zygotes in the midgut. During this stage, the zygotes pave their ways by transforming to motile ookinetes, and traverse peritrophic matrix and midgut basal lamina. At the outer surface of midgut, the ookinetes develop into sporoblasts which are covered with midgut basal lamina and newly synthesized protective capsule, known as oocysts. Inside the oocyst, sporogonic development ensues by progressive multiplication through the process reminiscent to budding of a large number of spindle shaped sporozoites from the core of sporoblast. Upon maturation, the oocysts, each containing thousands of sporozoites, are ruptured, releasing numerous sporozoites into thoracic cavity and hemolymph of the mosquito vector. However, the majority of sporozoites migrate to and traverse mosquito's salivary glands, ready to be transmitted to their vertebrate hosts (Vaughan, 2007; Smith & Jacobs-Lorena, 2010). Salivary gland sporozoites are fully mature and infective, measured 0.4-2.7 micrometers in width and 9-16.5 micrometers in length, depending on malaria species. At the anterior portion of the infective sporozoite contains apical organelles which are essential for host cell invasion.

Upon the infective bites of female anopheline mosquitoes, most salivary gland sporozoites are inoculated into the blood stream of humans. Approximately within 30 minutes, sporozoites enter hepatocytes by a complex receptor-ligand mediated invasion process. Intrahepatic sporozoites migrate by traversing several liver cells until they reach the desired hepatocyte. A small subset of inoculated sporozoites are delivered outside capillary lumen and deposited in dermal tissue at the mosquito's biting site where asexual development proceeds locally but imperfectly. Some of these extracappillary sporozoites find their ways to the draining lymph nodes and are suggested to be implicated in transient host immune suppression (Amino (Amino *et al.*, 2006). In the liver cells, sporozoites multiply by schizogony, producing several thousands of merozoites (Prudencio *et al.*, 2006; Sturm *et al.*, 2006; Garnham, 1966). No obvious clinical symptom of malaria occurs during intrahepatic developmental period which takes approximately 9 to 14 days. Upon rupture of infected hepatocytes, merozoites are released into blood stream, then invade erythrocytes within a few minutes. Intraerythrocytic development of malaria parasites comprises (i) asexual multiplication and (ii) gametocytogenesis in which the

former is responsible for clinical symptoms of malaria. Erythrocyte entry by malarial merozoites is a complex process involving interaction between several erythrocyte receptors and various parasite ligands, most of which are proteins expressed on the surface or transported proteins to the surface of merozoites. For *P. vivax*, the interaction between Duffy-binding proteins of malaria parasites and Duffy antigen receptor for chemokines (DARC) on erythrocyte membrane as well as reticulocyte-binding proteins and reticulocyte receptors are required during red cell invasion process (Chitnis & Sharma, 2008). For *P. falciparum*, erythrocyte glycoporphins and erythrocyte-binding proteins are crucial for merozoite invasion into erythrocytes (Jaskiewicz *et al.*, 2019) while specific merozoite surface proteins form a protein complex that binds to erythrocyte cytoskeleton such as band 3 and spectrin (Goel *et al.*, 2003). Once the merozoite gains access into the red blood cell, it resides in a vacuole surrounded by host cell-derived parasitophorous vacuole membrane (Aikawa *et al.*, 1978). Inside the erythrocyte, the merozoite develops into actively growing trophozoite stage. Young trophozoites of all *Plasmodium* species take similar morphology, characterized by a small nucleus or chromatin and a large spherical or ovoidal vacuole displacing the nucleus and cytoplasmic mass to the periphery, resulting in a signet ring shape appearance (Garnham, 1966).

Intraerythrocytic parasites mainly feed on host hemoglobin in which globin is digested into amino acids and iron is salvaged from heme. However, breakdown of heme by malarial parasites generates a toxic intermediate, ferriprotoporphyrin IX, and thus requires further detoxification into nontoxic hemozoin, also known as malaria pigments. The amount of hemozoin increases with the growth of parasites. Growing or mature trophozoites undergo multiple nuclear divisions within non-dividing cytoplasmic mass and become schizont stages. Maturation of schizont occurs when cytoplasmic division takes place, producing several merozoites. The numbers of merozoites in mature schizonts are variable between species (Garnham, 1966). Shortly after schizont maturation, merozoites are released from infected erythrocytes along with free hemozoin and glycosylphosphatidyl inositol anchor moiety from certain merozoite surface proteins which are responsible for pyretic symptom of malaria. When free merozoites invade new erythrocytes, another asexual reproduction resumes repeatedly. The intervals of complete asexual erythrocytic development are rather different among malaria species, i.e. 24 hours for *P. knowlesi*, 36-48 hours for *P. falciparum*, 48 hours for *P. vivax*, *P. ovale wallikeri*, *P. ovale curtisi*

and *P. cynomolgi* and *P. fieldi* and 72 hours for *P. malariae* and *P. inui*. These cyclical periods coincide with febrile paroxysm observed in malaria caused by different species. Meanwhile, after certain cycles of asexual reproduction, some merozoites enter new erythrocytes without nuclear division and transform into gametocytes, an infective stage to mosquito vectors (Sinden, 1982).

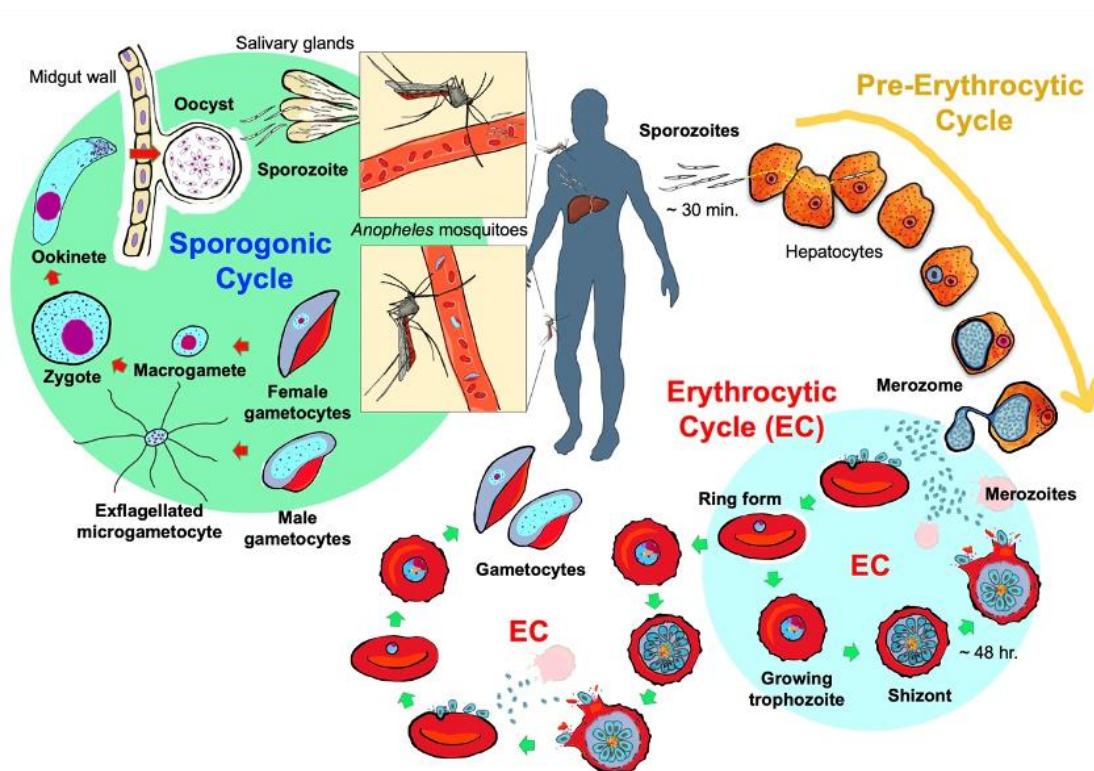


Figure 1 Life cycle of malaria parasite (Rattanaporn Rojrung).

## 2.2 Malaria Mosquito Vectors

*Anopheles* mosquitoes are the main transmission vector for human malaria. Mosquitoes have been known to be vectors of many pathogens (Tolle, 2009). There are more than 3,500 species of mosquitoes that currently distributed worldwide (Ferraguti *et al.*, 2016). About 430 *Anopheles* species were recognized. There are only 40 *Anopheles* species responsible for malaria transmission. Of these, there are 3 species of anopheline mosquitoes that have been known to be primary vector of human malaria in Thailand, which are *Anopheles minimus*, *An. maculatus*, *An. dirus*. The first species is the most abundant species in Thailand (Wamaket *et al.*, 2021). The secondary vectors are *An. acronitus*, *An. sundaicus*, and *An. pseudowillmori*.

Mosquitoes are generally developed in water. There are four developmental stages, including egg, larva, pupa, and adult. The first three stages are aquatic stages that take roughly 5-14 days to complete. The period of aquatic development depends on the species and water temperature. Lifespan of adult is approximately 1-2 weeks. Generally, female mosquitoes required blood meal for serving their eggs while male mosquitoes take nectar, plant sap or honeydew for food. The initial cycle of malaria parasite transmission begins when an anopheline mosquito takes a blood meal from a *Plasmodium*-infected human host.

### 2.3 Malaria Preventive Measures and Malaria Vaccine Development

Avoiding the biting of the mosquitoes is the best preventive way from infection. However, it might be unavoidable from mosquito bites, especially in the areas with that high density of mosquitoes. Currently, several preventive measures have been employed, including using insecticide-treated nets (ITNs), indoor residual spraying (IRS), prophylactic drugs (PD), and untreated nets (UN). While these interventions may not be universally applied, malaria vaccine employment is an important alternative strategy for prevention.

Although malaria can be treated and prevented by anti-malarial drugs and insecticides, the emergence and widespread of drug-resistant parasite strains and insecticide-resistant anopheline mosquitoes have hindered effective control measures by means of currently available compounds. Therefore, malaria vaccine development, especially against the most pernicious species *P. falciparum*, is a promising alternative control measure. Initially, vaccine developmental strategies aim to interrupt or destroy parasite stages that are free from host cells, i.e. sporozoite, merozoite and gametocyte (the host cell is ruptured during gametogenesis), mainly by humoral immune responses. However, it has been shown that intracellular parasites can also be attacked by cell-mediated immune system, especially cytotoxic T cell responses. Therefore, targets for malaria vaccines comprise (i) pre-erythrocytic vaccine against sporozoites and hepatic stages, (ii) asexual erythrocytic stages, including merozoites and intraerythrocytic stages and (iii) sexual stage or transmission blocking vaccine targeting gametocytes and other sexual stages in the mosquito vectors. To date, a number of malaria vaccine candidates have been identified.

However, a few have strong potential to be incorporated into subunit vaccines. Herein, only some of these prime vaccine candidates are reviewed.

### 2.3.1 Pre-erythrocytic Vaccine Candidates

The circumsporozoite protein or CSP is the most abundant protein on the surface coat of infective sporozoites. Anti-CSP antibodies can cause precipitation reaction around the sporozoite stage, resulting in loss of infectivity to hepatocytes. The gene encoding CSP of *P. falciparum* is a single copy gene, located on chromosome 3 (Dame *et al.*, 1984). Structure of CSP comprises relatively conserved N- and C-terminal domains flanking the central immunodominant repetitive region. CSP is immunogenic upon natural infections and uniformly appears in all species of *Plasmodium* at the sporozoite stage. The central repeats of CSP differ in sequence and number of repeat units across malaria species. Immunizations of monkeys with gamma or X-irradiated sporozoites of a simian malaria parasite, *P. knowlesi*, have conferred protection against challenge infections with intact sporozoites. This finding has been replicated in human malaria, *P. falciparum* which showed similar results (Clyde *et al.*, 1973). Vaccination trials in humans using synthetic or recombinant peptides derived from the central repeats of CSP of *P. falciparum* elicited partial protection against challenge infection in vaccinated volunteers while complete or almost complete protection was observed among those immunized with live attenuated sporozoites (Rieckmann, 1990; Hoffman *et al.*, 2002). These findings suggest that immunoprotective epitopes in CSP could include non-repeat regions of the protein while other sporozoite proteins, e.g. thrombospondin-related adhesive protein and sporozoite threonine asparagine-rich protein, could also induce additive protective immunity among volunteers immunized with attenuated sporozoites. Anti-CSP antibodies could disrupt the infectivity of live sporozoites during hepatocyte invasion by parasite ligand competition (Kumar *et al.*, 2006; Zou *et al.*, 2013). The most advanced malaria vaccine development, RTS, S/AS01, is based on peptides derived from the central repeat regions of CSP from *P. falciparum* 3D7 clone (Garçon *et al.*, 2003). Vaccination trials in different cohorts have shown variable percentages of protection, ranging from low to moderate levels (up to 46% protective efficacy) among African children despite using multiple booster doses (Targett, 2015). To-date, RTS, S or Mosquirix is the first malaria vaccine endorsed by WHO in 2021. The recommendation based on the benefitting result in young children living in sub-Saharan Africa (WHO, 2021). This vaccine was proved to provide up-to 86% protective

effect by fractional dose boosting (Regules *et al.*, 2016). The efficacy of this vaccine could also be limited by the presence of CSP variants among circulating parasite isolates that differ from the immunogen derived from the 3D7 clone, particularly in the C-terminal region which contains helper T cell epitopes, i.e. Th2R and Th3R. It is therefore likely that a vaccine derived from a single allele may have lower efficacy against parasites carrying distantly related variant antigens (Chaudhury *et al.*, 2021). Interestingly, the most effective malaria vaccine, R21/Matrix-M was reported with 77% effective in initial trials (Roxby, 2021). Although, it showed higher efficacy than RTS, S vaccine, it undergoes trials and approval processes.

### 2.3.2 Asexual Erythrocytic Vaccine Candidates

Merozoite is a blood stage of malaria parasite that is free in circulation prior to erythrocyte invasion. During this brief period, the merozoites are vulnerable to host immune-mediated attack, especially by innate and humoral immunity. Merozoite surface proteins (MSPs) are heterogeneous groups of blood-stage antigens that are expressed on or associated with the merozoite surface. Some of these proteins have been shown to play a crucial role in erythrocyte entry that mainly appear on the surface of merozoites or are accumulated within rhoptries and micronemes of the merozoite. Merozoite surface proteins can be (i) glycosylphosphatidylinositol (GPI)-anchored proteins, e.g. merozoite surface protein 1 (MSP1), merozoite surface protein 4 (MSP4) and merozoite surface protein 5 (MSP5), (ii) integral membrane proteins, e.g. erythrocyte binding protein 175 (EBA175), reticulocyte binding protein family (RH) and apical membrane antigen 1 (AMA1) and (iii) peripherally-associated proteins, e.g. merozoite surface protein 3 family (MSP3) and merozoite surface protein 7 family (MSP7). Of these, MSP1 is the most abundant GPI-anchored merozoite surface protein that forms complex with other merozoite proteins involving in binding to erythrocyte band 3 (Goel *et al.*, 2003; Kariuki *et al.*, 2005). Several lines of evidence suggest that MSP1 is a promising vaccine candidate because (i) anti-MSP1 antibodies inhibited merozoite invasion into erythrocytes *in vitro* (Guevara Patino *et al.*, 1997), (ii) immunizations of mice with MSP1 of murine malaria parasites and owl monkeys with *P. falciparum* MSP1 have conferred protection against challenge infections with the respective parasite species (Hirunpetcharat *et al.*, 1997; Singh *et al.*, 2006), and (iii) individuals living in malaria endemic areas who developed clinical immunity against malaria had significantly higher levels of anti-MSP1 antibodies than symptomatic malaria patients while the levels of parasitemia

are inversely correlated with the levels of antibodies (al-Yaman *et al.*, 1996). However, MSP1 of *P. falciparum* exhibits extensive sequence diversity among isolates, an issue that could compromise vaccine development. On the other hand, polymorphism in merozoite surface proteins has been deployed as genetic markers to determine clonal mixture within parasite isolates such as MSP1 and MSP2 of *P. falciparum*.

Meanwhile, apical membrane antigen 1 (AMA1) is also an integral membrane protein considered to be a leading vaccine candidate. AMA1 plays an important role in initial formation of the moving junction upon erythrocyte entry of the merozoite by interacting with the rhoptry neck protein (RON) complex that are parasite-derived receptors secreted from the rhoptry to erythrocyte surface (Lamarque *et al.*, 2011). Like MSP1, the potential immunoprotective epitopes of AMA1 have been mapped in the polymorphic region of the protein albeit strain-transcending immunity has been suggested to occur among people living in malaria endemic areas (Takala *et al.*, 2009).

Besides merozoite surface proteins, malarial proteins synthesized during early or mature asexual blood stages and translocated to the surface of infected erythrocyte have also been targets by host immune responses. These proteins have not been extensively investigated while some of them display extensive sequence variation among isolates precluding vaccine development such erythrocyte membrane protein 1 of *P. falciparum* (PfEMP1) that mediates cellular adhesion between infected erythrocytes with mature stages parasites and endothelial cells in various internal organs (Ortolan *et al.*, 2022). However, recent studies have shown that a *P. falciparum* protein, designated glutamic acid-rich protein (PfGARP), is considered to be a potential vaccine target (Raj *et al.*, 2020).

### 2.3.3 Transmission Blocking Vaccine Candidates

Immunity against sexual stages does not confer benefit to those who receive vaccination. However, it is anticipated that a large coverage of individuals immunized with sexual stage vaccine can reduce or terminate malaria transmission in a given community. Although sexual stage vaccine may not as attractive as other vaccine targets, several proteins derived from gametocytes, gametes and ookinetes have been identified and some have been considered to be promising candidates such as Pfs230, Pfs48/45, Pfs25 and Pfs28 against sexual stages of *P. falciparum* (Kaslow, 2002).

## 2.4 Glutamic Acid-rich Protein of *Plasmodium falciparum* (PfGARP)

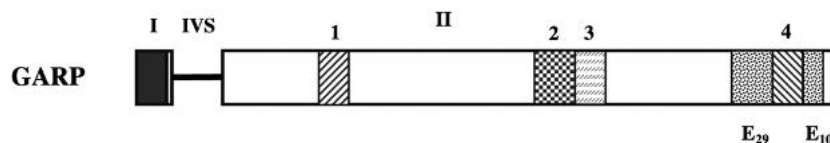
Modification of infected red blood cell membrane (iRBC) by *Plasmodium* proteins has been associated with pathogenesis of malaria or progression to severe malaria. During growth and development in erythrocyte, a number of malarial proteins are transported via complex trafficking mechanism to express on iRBC membrane (Maier *et al.*, 2008). These proteins play an essential role in enhancing iRBC adhesiveness which can be categorized into (i) binding of iRBC to endothelial cells (cytoadherence), (ii) adhesion between iRBCs to uninfected RBCs (uRBCs) (erythrocyte rosetting), (iii) binding of iRBCs to others iRBCs (autoagglutination) mediated by CD36 on platelets, (iv) clumping of iRBCs to platelets (platelet-mediated clumping), and (v) aggregation of uRBCs to each other (RBCs aggregation) (Treutiger *et al.*, 1992; Dondorp *et al.*, 2000; Balaji & Trivedi, 2013). One of the malarial proteins exported to the iRBC membrane that serves as a potential ligand for human erythrocyte band 3 and has been suggested to be associated with cytoadherence during *P. falciparum* infection is glutamic acid-rich protein or PfGARP (Almukadi *et al.*, 2019).

The gene encoding PfGARP has been identified by screening complementary DNA (cDNA) libraries of *P. falciparum* FC27 strain with Papua New Guinean immune sera (Triglia *et al.*, 1988). The *pfgarp* gene locates on chromosome 1 and contains 673-678 codons. PfGARP is transcribed and translated from 2 exons spanning 25 and 648-653 codons with 77% and 71% of AT-content, respectively. The first 29 amino acids at the N-terminus of PfGARP contain a signal peptide while amino acid residues 48-52 are RLLNE which is a characteristic motif of *Plasmodium* export element (PEXEL) (Almukadi *et al.*, 2019). Like most other malarial antigen genes, 4 repeat domains have been identified in the coding regions of *pfgarp* (Figure 2). The first region (repeat domain I) is near the 5' portion of the second exon, consisting of 15 repetitive units of KXX-encoding motif where X is E, D, H or K. Repeat domains II and III are found 3' to the center and the remaining repeat region (repeat domain IV) is located near the 3' end of the gene. Repeats in domains II-IV are characterized by 9, 5, and 7 repetitive units of EGHKE/GGHKE, KGKKX where X is D, K, E or H, and EDDXEED/DEEXDDD where X is A, D or E, respectively (Triglia *et al.*, 1988). The predicted 3D structure of PfGARP was showed in Figure 3. *In silico* modeling of the protein produced by predicted Local Distance Difference Test

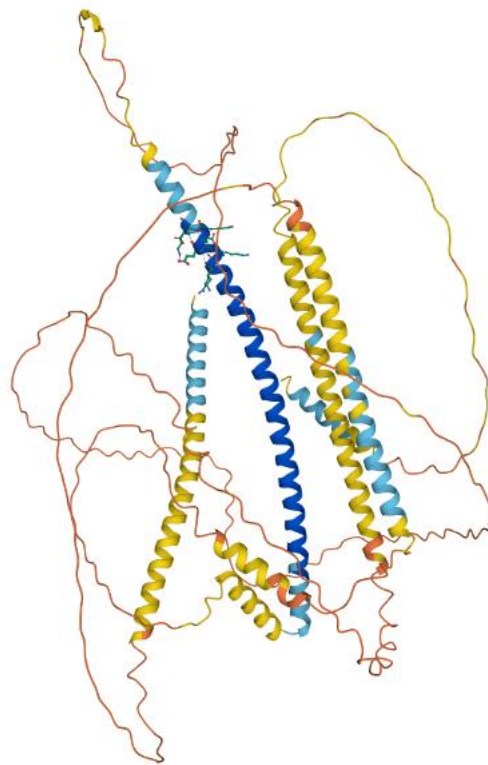


(pLDDT) has revealed eight coil-coil motif chains. The regions with high pLDDT scores are shown in dark blue color (pLDDT > 90) while those having lower scores are in light blue (90 > pLDDT > 70), yellow (70 > pLDDT > 50) and orange (pLDDT < 50), respectively.

It is noteworthy that the majority of amino acids in PfGARP are glutamic acid and lysine; together they constitute 44.7% of the protein and most of them are in repeat domains (Triglia *et al.*, 1988). Meanwhile, a number of malarial proteins containing *Plasmodium* export element or host-targeting (HT) motif are believed to be exported by the parasite to the periphery of infected erythrocyte membrane. These repeat domains are considered to be low complexity region in which the lysine-rich tandem repeats in PfGARP have been shown to play an important role in binding to erythrocyte cytoskeleton (Davies *et al.*, 2016). Fine mapping of the interaction region between PfGARP and erythrocyte cytoskeleton by using phage display cDNA libraries has revealed that repeat domains II and III in PfGARP serve as a secreted binding-ligand to an ectodomain of human erythrocyte anion-exchanger or band 3. Importantly, synthetic peptides spanning repeat domains II and III could induce erythrocyte aggregation akin to erythrocyte rosetting, a clumping phenomenon of uninfected erythrocytes around infected erythrocyte containing mature asexual parasite stages, suggesting that PfGARP may enhance the adhesion of infected and uninfected erythrocytes (Almukadi *et al.*, 2019). Therefore, PfGARP is an attractive target for immunological and therapeutic intervention for severe malaria and vaccine development while this protein could serve as a biomarker for disease progression.



**Figure 2** Schematic diagram of the *garp* gene of *Plasmodium falciparum* FC27 strain depicting exons I and II, the four blocks of repeats (1-4) and the two polyglutamic acid (E<sub>29</sub> and E<sub>10</sub>) strings. IVS represents intervening sequence (intron) (Modified from Triglia *et al.*, 1988).



**Figure 3** Picture showing 3D model of PfGARP structured by a per-residue confidence score (pLDDT) (UniProt Database; <https://www.uniprot.org/uniprot/Q9U0N1>).

The role of PfGARP as a potential malaria vaccine candidate has been supported from *in vitro*, *in vivo* and field studies (Raj *et al.*, 2020). Mice antibodies raised against recombinant protein derived from the C-terminal half of PfGARP have conferred efficient inhibition (94%-99%) of *P. falciparum* growth *in vitro*. Likewise, human antibodies from clinically immune individuals in Tanzania exhibited similar extent of parasite growth inhibition. Although a repertoire of mouse monoclonal antibodies against PfGARP have been obtained upon immunization with the recombinant protein, only monoclonal antibody direct against VKNVIEDEDKDGVEIIN located after repeat domain II could elicit parasite killing, suggesting that this region could be essential for parasite survival. Importantly, parasites treated with anti-PfGARP antibodies showed ‘crisis forms’ in Giemsa-stained blood smears, characterized by dysmorphic and pyknotic appearance. Furthermore, a remarkable loss of mitochondrial membrane potential has been demonstrated by flow cytometric analysis while activation of parasite caspase-like enzymes has been observed using specific fluorescent dye. Therefore, it is likely that anti-

PfGARP antibodies mediate parasite killing by activation of programmed cell death of parasites; which in turn, potentially leading to reduced severity of malaria (Raj *et al.*, 2020). Meanwhile, comparative analysis of Tanzanian cohorts has shown that individuals having high levels of anti-PfGARP antibodies were associated with a significantly decreased risk of severe malaria while a similar study in Kenyan cohorts has shown a significant inverse correlation between the levels of anti-PfGARP antibodies and parasite density. Furthermore, vaccination trials in owl monkeys using immunogens from recombinant PfGARP spanning the N-terminal part of repeat domain II to the C-terminal end of the protein has conferred partial protection against heterologous parasite challenge (Raj *et al.*, 2020). Taken together, PfGARP is considered to be a promising therapeutic and vaccine targets against falciparum malaria.

## **2.5 Polymorphism in *Plasmodium falciparum* Vaccine Candidates and Potential Impacts on Vaccine Efficacy**

The presence of sexual cycle in mosquito vectors has enabled genetic recombination between malaria parasites carrying distinct genomes, resulting in the generation of genetically different progenies (Walliker *et al.*, 1987). Recombination involves the pairing of homologous chromosomes from male and female gametes during syngamy. The recombination events outside the protein-coding genes may generate re-assortment of genetic loci while intragenic recombination can give rise to different haplotypes within loci. If the rates of intragenic recombination at loci carrying several mutations are exceptionally high, an enormous number of variants can be found in the population. Besides recombination, point mutations in protein-coding genes can lead to the generation of different variants and usually occur under the stochastic process known as random genetic drift caused by errors during DNA replication and repairing efficiency. Non-silent point mutations may affect protein functions when they occur at critical residues. Insertion or deletion of one or two nucleotides causes frame-shift mutation of the downstream coding sequence that can be lethal unless an appropriate compensatory insertion or deletion preserves the remaining portion of the coding region (Futuyma, 2006). Meanwhile, repetitive DNA sequences are abundant in the genomes of living organisms including *Plasmodium* species. These repetitive sequences may undergo expansion or retraction during

cellular multiplication due to the process of slipped-strand mispairing. However, epigenetic factors, especially selection pressures, are crucial to shape the gene pools of infectious pathogens, such as host immune responses, drug pressure, competition among variants for growth advantage and geographic isolation (Hughes, 2000). All of these mechanisms have been documented in malarial genome. Therefore, it is not unexpected to observe genetic diversity within a given species of *Plasmodium*. Genetic diversity in vaccine candidate-encoding loci of malaria parasites may compromise vaccine development if the magnitude is extensive and most variants induce strain-specific immune responses. For example, variation in T cell epitopes in the C-terminal part of PfCSP could alter immune recognition (Plebanski *et al.*, 1999) while antibodies against the central repeats of PfMSP2 were family-specific and non-cross-reactive (Taylor *et al.*, 1995). These issues are crucial for the design of subunit vaccines against malaria.

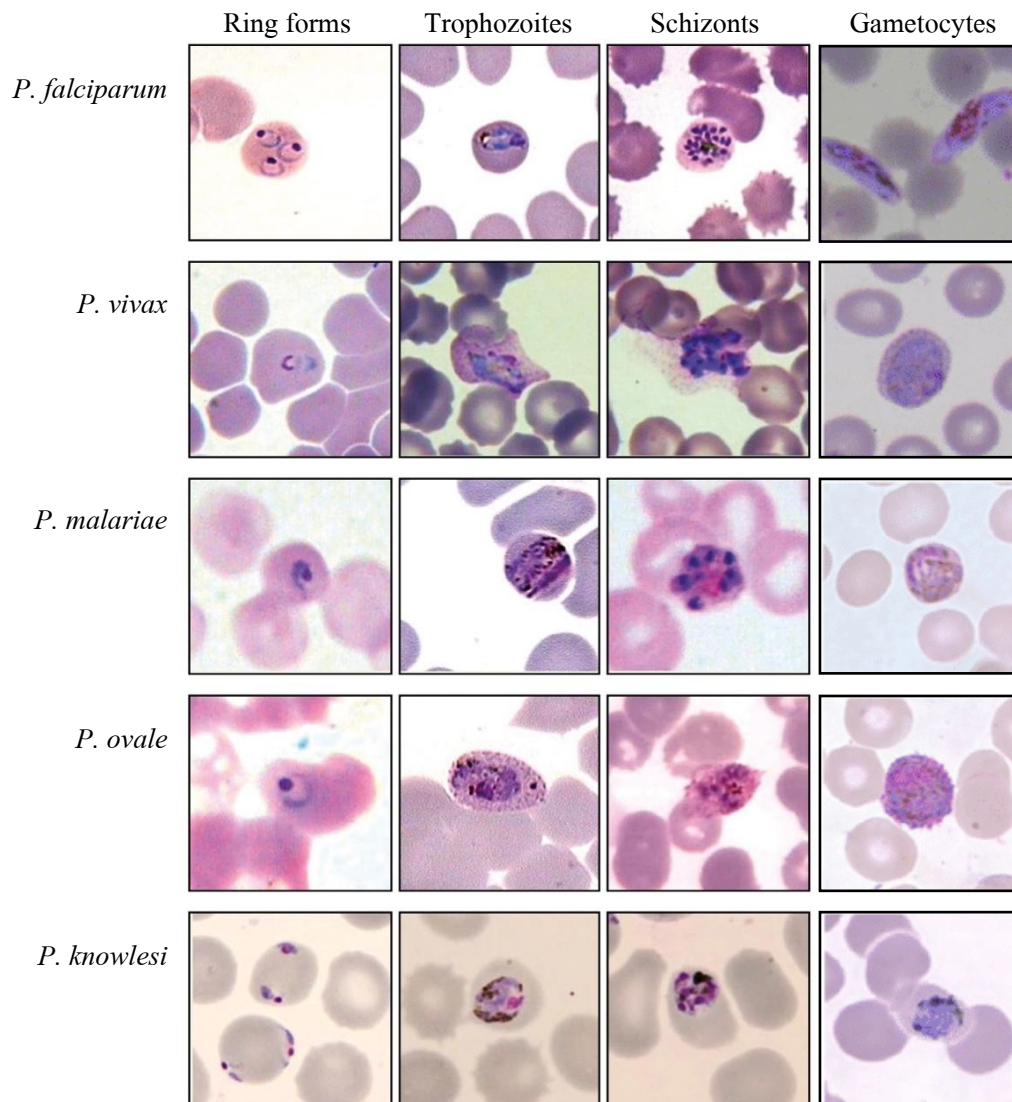
To date, little is known about the extent of sequence diversity in *pfgarp*, the newly defined malaria vaccine candidate protein. Although an attempt has been reported by retrieving available whole genome sequencing data, such information is partial and does not thoroughly delineate the patterns of diversity in repeat domains of PfGARP (Raj *et al.*, 2020). Therefore, it remains to be of prime importance to investigate and analyze the extent of sequence diversity among field isolates, especially in Thailand where multi-drug resistant *P. falciparum* are abundant. The results from this study will be the basis for vaccine design targeting PfGARP.

## 2.6 Malaria Diagnostic Methods

### 2.6.1 Microscopy-based Detection

Malaria parasites could be identified and quantified under microscope by examination of thin and thick blood films. These techniques are considered as gold standard for malaria identification in field trials. Additionally, malaria microscopy intends to be used for screening the infection before confirming with the molecular technique. Thin blood films are generally employed for qualitative observation while thick blood films are for quantitative propose. Parasite density can be estimated from both thin and thick blood films which are reported as the percentage of infected erythrocytes per total erythrocytes examined and the total number of parasites in a microliter of blood, respectively.

Basically, for the best quality slide, the glass slide should be cleaned initially before blood smearing. After that smeared slide is stained with Giemsa solution and left for drying. The infection is generally identified by the distinction between malaria parasites and other blood cells, which are mainly considered by the different colors of the stained components, and the characteristics of parasites. The parasite species also could be determined by this protocol. Characterization of the malaria parasite is based on various features, including, erythrocyte morphology (normocytic or ovalocytic or macrocytic erythrocytes), specific stages of the parasites in peripheral blood, shape and size of the parasites, etc. For *P. falciparum*, the salient characteristics are (I) iRBC is normocytic whereas both smooth and rough-erythrocyte membrane could be found, (II) almost all asexual stages in peripheral blood are early trophozoites, also known as ring stages, and (III) multiple infections of ring stages is more commonly observed than in other species (CDC, 2020). Typical stages in chronologic order of common human malaria parasites are shown in Figure 4. The sensitivity of microscopy detection of malaria parasites depends on many factors, including the expertise and experience of the microscopists, qualities of blood smears and the number of microscopic fields examined. The information obtained by this technique might be poorly accurate when the level of the parasites is very low. Meanwhile, mixed infections of different parasite species can compromise correct diagnosis. Thus, the PCR-based method might offer better diagnostic performance than microscopy. More details on the microscopic examination procedure can be accessed in the manual report from World Health Organization (WHO).



**Figure 4** Differential characteristic stages of human malaria parasites on thin blood film (Poostchi *et al.*, 2018).

### 2.6.2 Polymerase Chain Reaction-based Diagnosis

Polymerase Chain Reaction (PCR) is a molecular technique that has been widely used in many laboratory settings. The PCR technique was invented by an American chemist, Kary Mullis in 1985 (Saiki, 1989). The principle of this technique is that it can amplify a large quantity of DNA of a target region in the genome within a short period of time. Initially, PCR was used for diagnostic test of sickle cell anemia by amplification of the beta-globin DNA using Klenow fragment of *E. coli* DNA

polymerase I for polymerization process (Saiki *et al.*, 1985). The drawback of this enzyme is that it is thermosensitive. Consequently, it must be re-added in every cycle. This obstacle has been gone after thermostable DNA polymerase (*Taq* polymerase) was introduced (Saiki, [chapter 1], 1989, p. 1).

To date, PCR has been increasingly used for both medical and research purposes, for instance pathogen detection, species identification, etc. Detection of *Plasmodium* parasite by PCR was initially applied in 1990 (Jaureguiberry *et al.*, 1990). The application of this technique has conferred more sensitive and specific detection of *Plasmodium* species than microscopy examination, especially when species-specific PCR primers are used in the specimens with low parasite level and mixed infections of malaria parasites (Snounou *et al.*, 1993). Sensitivity and specificity are increased by adopting nested PCR (Haqqi *et al.*, 1988; Nolte, 2011). It typically involves the use of two sets of primers. The outer primers are designed to cover the interested gene region. It is deployed in the first round of amplification. Then, the products of primary amplification are served as the DNA template for secondary round of PCR using the inner set of primers. This assay not only reduces non-specific amplification but also substantially increase the amount of the products. Low concentration of DNA template is one of the problems for PCR amplification although theoretically a single DNA target can be amplified to reach enough amounts for detection, e.g., blood sample from a malaria patient who carries low parasitemia. Therefore, nested PCR is one of the useful approaches used to handle certain limitations. Since the PCR is highly sensitive, the contamination should be careful. To get only the target PCR product, the specificity of primers should be evaluated in which they do not match with other genomic regions of malaria parasites or other organisms. Even through, there are many factors that should be considered to reduce or minimize the errors, the PCR technique is widely used due to its convenience, fast, and ease to use (Garibyan & Avashia, 2013).

Each PCR protocol mainly requires 7 precursors which are thermostable DNA polymerase, primers, deoxyribonucleotide triphosphates (dNTPs), buffer, deionized water, magnesium chloride ( $MgCl_2$ ), and DNA template. The nucleotides comprise four bases that are adenosine (A), cytosine (C), guanine (G), and thymine (T). Each component plays the different function as described in Table 1. DNA polymerase is a key enzyme to develop the DNA strand by adding nucleotides complementary to the single stranded DNA templates. These components are

mixed in microcentrifuge tube in appropriate portion and placed on thermocycler, allowing the PCR amplification to occur (Garibyan & Avashia, 2013; Rahman, 2013).

The principle of the PCR technique is the cycle of temperature which is controlled for suitability with the function of precursors used for DNA amplification. There are three steps of the PCR assay. The first is denaturation phase, in which the DNA template molecule is denaturalized from its natural structure to single stranded DNA. The temperature required in this step ranges from 90-96°C. The next step is annealing phase. For this period, the temperature is decreased to 50-68°C which allows forward and reverse primers to bind the complementary DNA template regions specifically. These binding regions defined as an extension point of DNA polymerase in an extension phase. The temperature used in annealing phase depends on the melting temperature ( $T_m$ ) of forward and reverse primers in which both primers should have similar levels of  $T_m$ . The third step is extension phase. The temperature is raised again to 90-96°C, allowing DNA polymerase to extend the primers by adding nucleotides that are matched with the template strands while the new DNA strand is developed. After a cycle of PCR from one DNA strand, two DNA molecules are generated which are hemi-conservative strands, one of which is the prototype. The other strand is the newly created. The target DNA fragments are increased by  $2^n$  (n is the number of PCR's cycle). The target amplicons will be generated in the third cycle of the PCR (Powledge, 2004; Joshi & Deshpande, 2010).



**Table 1** The PCR precursors and their functions.

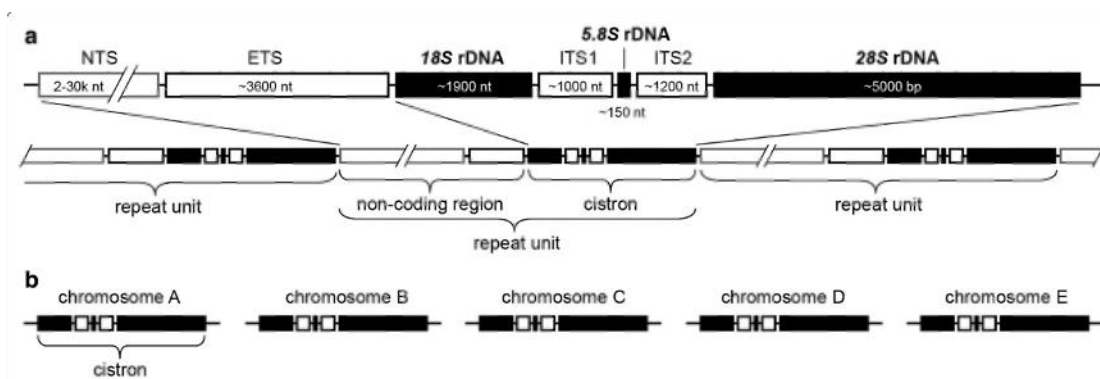
<b>Reagents</b>	<b>Functions</b>
1. DNA template	DNA that contains the target gene region to be amplified.
2. Primers	Single short DNA strands that are designed complementary with the upstream and downstream of the target gene region. These primers define the polymerized point on DNA template.
3. dNTPs	There are four types of nucleotides, dATP, dTTP, dCTP, and dGTP. These nucleotides will be polymerized with the nucleotides on the DNA template by DNA polymerase.
4. DNA polymerase	Nucleotide polymerization
5. Buffer	The PCR condition regulator to appropriate with enzyme activity.
6. MgCl <sub>2</sub>	Co-factor of DNA polymerase

## 2.7 Genes for *Plasmodium* species Identification

### 2.7.1 Small subunit ribosomal RNA (SSU rRNA or 18S rRNA) gene

Eukaryotic ribosomes consist of two inequivalent protein subunits and ribosomal RNA (rRNA). Those two subunits are typically called large (60S) and small (40S) subunits in which the larger one is composed of over 50 proteins and three rRNA elements which are 28S, 5.8S, and 5S rRNA while the smaller one contains 33 proteins and 18S rRNA string. Those rRNAs (5.8S, 18S, and 28S) are transcribed by RNA polymerase I from strongly conserved tandem repeat gene whereas 5S rRNA is generally transcribed by RNA polymerase III (R. J. White, 2013; Martinez-Calvillo *et al.*, 2019; Clark, 2012), *rRNA* gene encodes both conserved and hypervariable regions. The conserved regions are the target areas for primer design to detect the genus of a given organism. On the other hand, there are two hypervariable regions with a high level of sequence difference among strains or species of organisms called internal transcribed spacer 1 (ITS1) and internal transcribed spacer 2 (ITS2) which are located between 18S and 5.8S and 5.8S and 28S, respectively (Figure 5). These two areas are remarkably variable than other regions. Therefore, they are usually used as target markers for species- or subspecies-level classifications in fungi and lower-class eukaryotes. Although 5.8S, 18S, and 28SrRNA contain relatively conserved sequences, there are some regions exhibit extremely variable as well (Hadziavdic *et al.*, 2014; Mitreva, 2017; Ghosh, 2018).

The feature of 18S rDNA of *Plasmodium* species consists of four to eight units per haploid genome, which locate on different chromosomes. The ribosomal units of *Plasmodium* spp. are not arranged in cluster and are not contain external-transcribed spacer (EST) and non-transcribe spacer (NTS) like in eukaryotic rDNA (Figure 5) (Harl *et al.*, 2019). In this study, the 18S rDNA was the target gene for *Plasmodium* species classification.



**Figure 5** Arrangements of rDNA of Eukaryotes (a) and *Plasmodium* species (b) (Harl *et al.*, 2019). [NTS: Non-transcribed spacer, ETS: External transcribed spacer, ITS: Internal transcribed spacers 1 and 2]

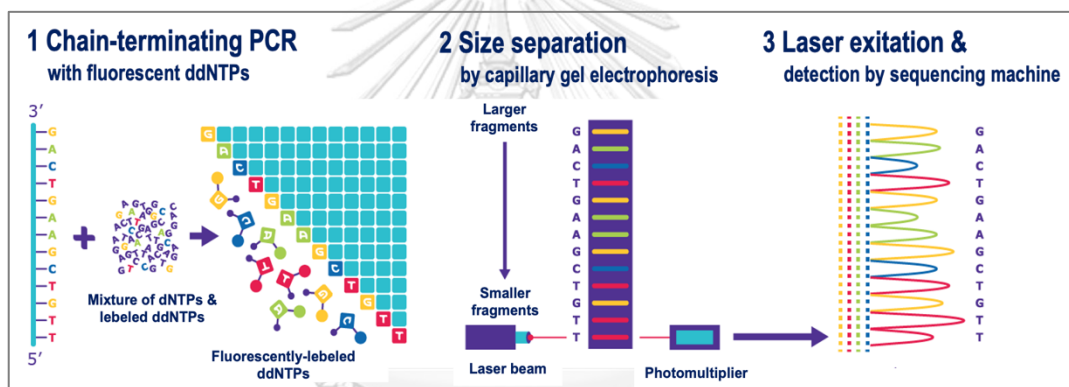
## 2.8 Sanger-based DNA Sequencing

Chain termination method DNA sequencing also known as Sanger sequencing has been widely used in genetic study since this technology provides the genetic sequence data for nucleotide acid analysis. Sequencing analysis is essential in a laboratory for genetic diversity study for the identification of the evolutionary changes in the genome or existing phenotypes of pathogens.

Sanger sequencing plays a role in the verification of PCR results. This technology can be performed manually or, more commonly, in an automated fashion via a sequencing machine. There are mainly 3 basic steps for obtaining the results, which are chain-terminating PCR, size separation, and detection by gel electrophoresis (Figure 6). 1) Chain-terminating PCR: it is DNA amplification based on standard PCR which composes nucleotide primers for annealing with DNA template, DNA polymerase enzyme for extending by incorporation of either mixture of 4 deoxynucleoside triphosphate or dNTPs (dATP, dGTP, dCTP, and dTTP) and low ration of dideoxynucleoside triphosphate or ddNTPs (ddATP, ddGTP, ddCTP, ddTTP). Incorporation of ddNTPs in elongation phase resulting in stop the reaction according to lack of the 3'-OH group required for phosphodiester bond formation. Then, variable DNA fragment lengths are obtained and distinguished in the further step. In manual protocol, four PCR reaction are set up with each ddNTP, while in automated approach, all ddNTPs are mixed in single reaction, and each of them is labeled with the different unique fluorescent. 2) Size separation: the PCR products were separated by gel electrophoresis which run separately in different lanes with each reaction of

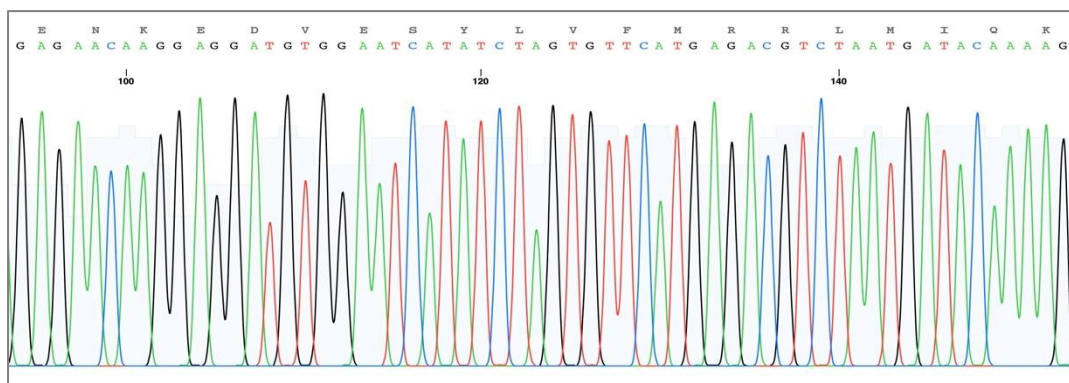
manual protocol, while all oligonucleotide products from automated approach are performed in a single capillary gel electrophoresis within the sequencing machine. 3) Determination of DNA sequence: In manual protocol, using four lanes in the gel to read the terminal ddNTPs at once. The sequence is reading from bottom to top. In automated approach, each band is read in order by computer. The fluorescence is used to identify the terminal ddNTPs. The output is called a chromatogram, which shows the fluorescent peak of each nucleotide along the length of the template DNA (please rearrange the sentence) (Figure 7).

In this study, we employed Sanger-based genetic sequencing on an ABI3100 Genetic Analyzer using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). for obtaining the genetic sequence data from the PCR results.



**Figure 6** Workflow of three basic steps of automated Sanger-sequencing.

(Modified from this source: <https://www.sigmaaldrich.com/TH/en/technical-documents/protocol/genomics/sequencing/sanger-sequencing>).



**Figure 7** Chromatogram from automate Sanger-sequencing demonstrates fluorescence signal of nucleotides.

## CHAPTER 3

### MATERIALS AND METHODS

#### 3.1 Target Population

*P. falciparum*-infected individuals from Chanthaburi, Tak, Ubon Ratchathani, and Yala Provinces were included in this study.

#### 3.2 Sample Size

This study involves molecular evolutionary analysis. Therefore, appropriate sample size calculation does not follow general statistical purposes which requires data on disease prevalence. Although more taxa included in the analysis will lead to more robust analytical results, it has been shown that a minimum of 20 taxa from each sampled population is essential for a meaningful analysis of the relationship between allelic diversity and molecular population genetics (Pruett, 2008). Therefore, this study recruited a minimum of 20 *P. falciparum* isolates from each selected province; a total minimum of 80 single-clone isolates was analyzed. The samples were selected based on different haplotypes of *msp1* and *msp2* loci.

#### 3.3 Sources of *Plasmodium falciparum* Isolates

Blood samples that have been collected during 2009-2014 were used in this study. Approximately 1 mL of venous blood samples were collected from febrile patients who attended malaria clinics or local district hospitals in Tak, Chantaburi, Ubon Ratchathani and Yala Provinces. Those are pre-treatment participants with antimalarial drugs.

#### 3.4 Ethic Consideration

The study protocol was approved by ethics committees of the institutional review board of the Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand (IRB. No. 193/64).

Since blood samples used in this study were collected during 2009-2014 that have been approved, there is not new formal written consent obtained from those patients.

### 3.5 Slide Preparation and Parasite Density Estimation

Both thin- and thick-blood films were prepared from each blood sample and stained with working Giemsa solution. Examination of Giemsa-stained blood films were performed under light microscopy using 100X objective with at least 200 microscopic fields for thin blood films, and at least 200 leukocytes for thick blood films. The parasite density was estimated by assuming a leukocyte count of 8,000/ $\mu$ l. The procedures were modified from method manual by World Health Organization (WHO, 2015). Thick and thin slide preparations were depicted in Figure 8. For parasitic identification, the slides could be used for distinguishing either the parasite within infected red blood cells (iRBCs) or other blood cells. Normal RBCs and iRBCs seem to have different appearances. An iRBC appeared rough membrane compared with the uninfected one that seem smoother. However, smooth membrane of iRBCs is also observed. The size of iRBCs of *P. falciparum* infection is commonly normalcytic, Meanwhile, iRBC of *P. vivax* infection commonly larger than uRBCs. Blood films of *P. falciparum* infection commonly find two parasite stages which are ring-form and banana-shaped gametocytes. Giemsa's dye stained the nucleus and cytoplasm of parasite characterized in dark purple and pink, respectively.

#### 3.5.1 Thin-Blood Film Preparation มหาวิทยาลัย

(1) Slides were labeled with patient's ID and cleaned with 90% ethanol (EtOH), then leave on rack for air drying.

(2) A small amount of blood, roughly 2-3  $\mu$ L was dropped on the cleaned slide. Held another cleaned slide at 45-degree angle on the specimen slide. Wait until the blood spreads along entire width of the spreader slide. Then push it toward smoothly and rapidly.

(3) The slides were leaved on rack for air drying. After completed drying, the slides were fixed with absolute methanol for 1 minute and leaved them dried again before staining.

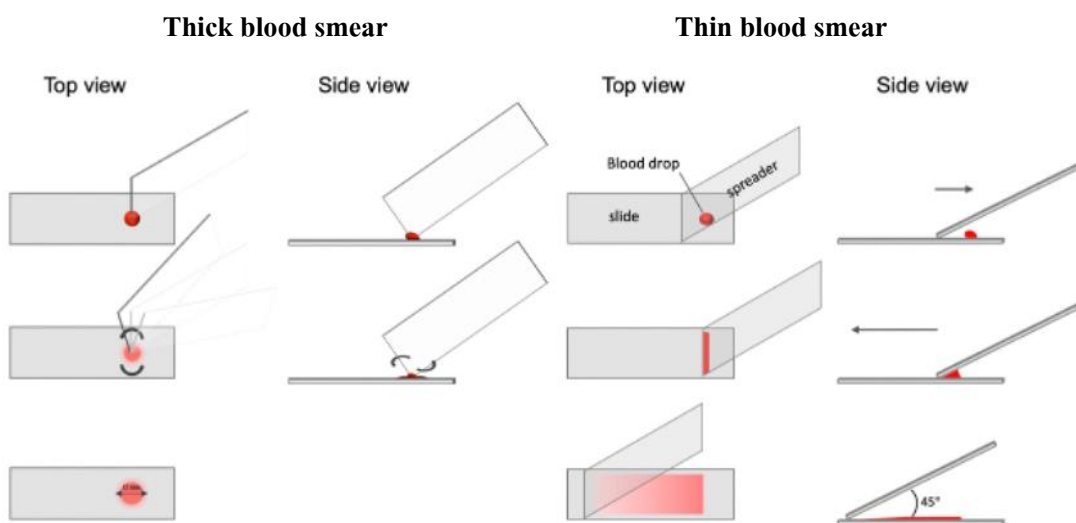
(4) The fixed slides were stained with 3% v/v Giemsa working solution for 45 minutes (WHO, 2015). The excess dye was washed with gentle running water. Allow the stained slides to dry completely. The slides were kept in boxes.

### 3.5.2 Thick-blood Film Preparation

Thick blood smear procedure was almost alias as thin blood film preparation except the steps as followed.

(1) The drop of blood (6  $\mu\text{L}$ ) was spread by the corner of a spreader slide in circular pattern, roughly 1-2 cm diameter in size.

(2) The thick film should not be fixed. It allowed to stain after air drying.



**Figure 8** Diagram of thick and thin blood smear.

### 3.5.3 Parasite Density Estimation from the Thin Blood Film

The parasite density was estimated by observing the slide under simple light microscope at 1000X total magnification. Screening direction was demonstrated herein (Figure 9).

(1) A small amount of immersion oil was drop on the tip of smear where blood cells are monolayer.

(2) Around 200 fields were scanned, and malaria parasites were counted accordingly on thin film at 1000X total magnification. Roughly each microscopic field contained 200 RBCs.

(3) Roughly 200 RBCs were counted. The number of infected RBCs and total counted RBCs were filled in Excel data sheet for further calculation.

(4) The parasite density was calculated using the following formula:

$$\text{The percentage of parasitemia} = \frac{\text{Number of infected RBCs counted} \times 100}{\text{Total number of RBCs counted}}$$

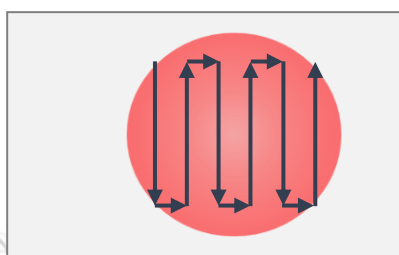
The criteria of parasite density were divided into 2 groups, including high parasitemia ( $\geq 2\%$ ) and low parasitemia ( $<2\%$ ). In this study, 200 microscopic field (100X objective) were counted and reported as the percentage of infected RBCs per 100 RBCs counted.

### 3.5.4 Parasite Density Estimation from the Thick Blood Film

After staining and leaving it dried, the smear was observed under simple light microscope at 1000X total magnification. The parasite density was calculated using the following formula:

$$\text{The parasite per } \mu\text{L} = \frac{\text{Number of parasites counted} \times 8000 \text{ WBC}/\mu\text{L}}{\text{Number of WBCs counted}}$$

In this study, WBCs per  $\mu\text{L}$  was assumed to be 8,000 cells (WHO, 2015). WBCs were counted 200 cells and reported as number of the parasite per  $\mu\text{L}$  of blood.



**Figure 9** Diagram shows the direction of microscopic observation.

### 3.6 DNA Extraction

DNA of each isolate was extracted from 0.3 ml of blood samples using QIAGEN DNA mini kit. DNA was extracted following the manufacturer's recommendation as followed.

(1) Blood sample was pipetted into 0.6 microcentrifuge tube and followed by adding 180  $\mu\text{L}$  ATL for tissue lysis. The mixture was mixed briefly before adding Proteinase K (20 mg/mL) 20  $\mu\text{L}$ . After that the mixture was mixed well by vertexing for 15 second. The cap was wrapped with paraffin preventing evaporation. Then, the mixture was incubated in incubator shaker at  $56^{\circ}\text{C}$  over nigh (12-18 hrs.) for more complete enzymatic activity.

(2) The tube was centrifuged to remove drops from the lid. AL buffer 200  $\mu\text{L}$  was added, mix thoroughly by vertexing for 15 s. DNA was precipitated by adding absolute ethanol 200  $\mu\text{L}$ , gently mixed by tipping the tube upside down. The tube was briefly centrifuged again.



(3) The mixture was transferred into the column tube. Centrifugation was performed. The filtrated solution and the collection tube were discarded while the column was transferred into the new collection tube. Buffer AW1 500  $\mu$ L was added followed by centrifugation. All centrifugation steps were performed at 13,000 rpm for 3 min. After centrifugation, the filtrated solution was discarded.

(4) Buffer AW2 500  $\mu$ L was added into the column. Centrifugation was performed at 18,000 rpm for 3 min. The filtrated solution and collection tube were discarded while the column was transferred to a 1.5 mL tube. After that centrifugation was applied at full speed for 1 min to eliminate the remaining buffer AW2.

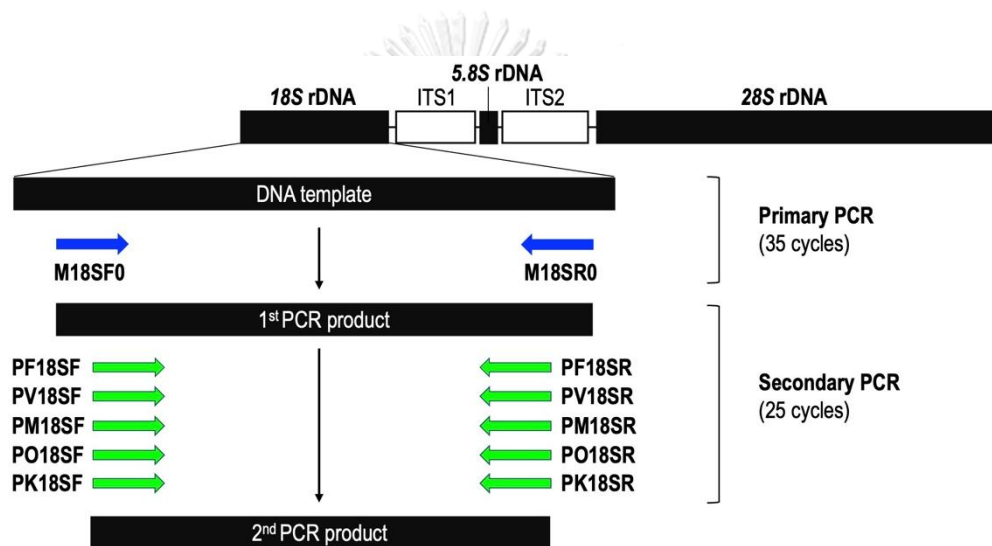
(5) The column tube was transferred to a new 1.5 mL microcentrifuge tube. DNA eluting buffer AE was added and incubated at room temperature for 5 min. Centrifugation was performed at 13,000 rpm for 2 min. The column tube was discarded while the flow-through solution was pipetted into a 0.6 microcentrifuge tube. Then, the tube was stored at  $-20^{\circ}\text{C}$  until used.

(6) Step 5 was repeated to increase DNA yield with a further eluting buffer.

### 3.7 *Plasmodium* species Identification by PCR

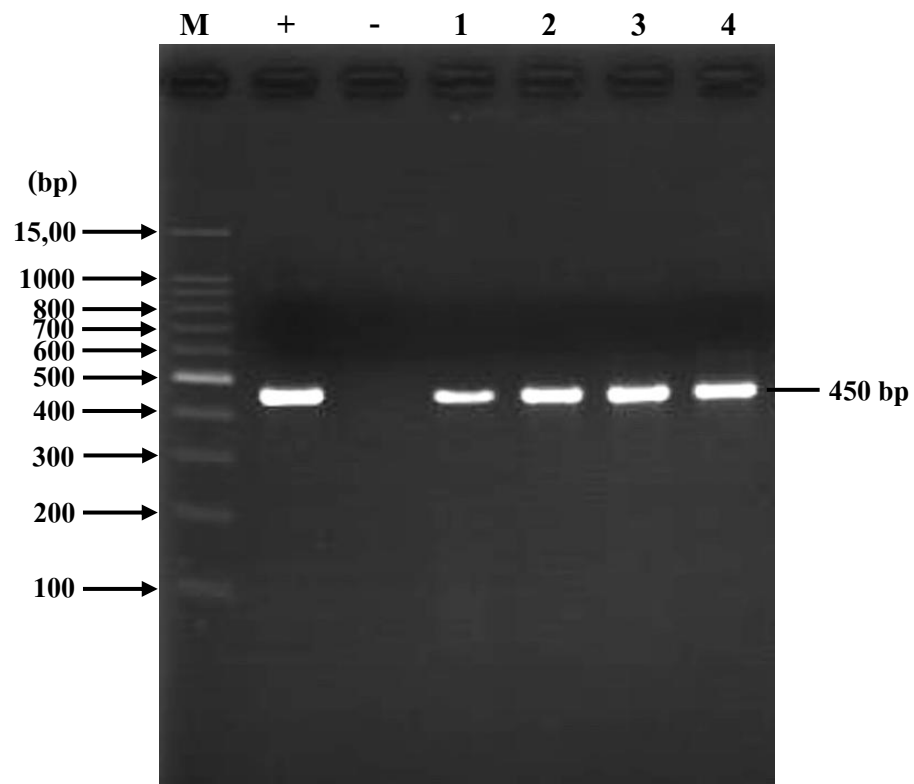
The samples that were positive for *P. falciparum* by microscopic observation were confirmed by PCR. Detection of *Plasmodium* species was done by nested PCR amplification targeting 18S rDNA (Figure 10) implemented protocol as previously described (Putaporntip *et al.*, 2009). Briefly, DNA amplification was carried out in a total volume of 20  $\mu$ l containing template DNA, 0.5  $\mu$ l of 10 mM each deoxynucleoside triphosphate, 2  $\mu$ l of 10X PCR buffer, 0.13  $\mu$ l of 0.3  $\mu$ M each primer, and 0.4 units of rTaq DNA polymerase (Takara, Seta, Japan). Thirty-five cycles ( $94^{\circ}\text{C}$  for 40 s,  $60^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 60 s) were performed for primary PCR, and 25 cycles for secondary PCR. The amplification products were analyzed by agarose gel electrophoresis. Primers for primary PCR target small subunit ribosomal RNA gene (18S rRNA), M18SF0: 5'-CCATTAATCAAGAACGAAAGTTAAGG-3' and M18SR0: 5'-CAAGGAAGTTTAAGGCAACAACA-3'. Secondary PCR was performed in separate reaction tubes for each pair of species-specific primers; *P. falciparum*, PF18SF: 5'-CATCTTTCGAGGTGACTTTTAG-3' and PF18SR: 5'-GTTTTTACTCTATTTCTCTCTTC-

3'; *P. vivax*, PV18SF: 5'-GAATTTTCTCTTCGGAGTTTATTC-3' and PV18SR: 5'-TAACAGTTTCCCTTTCCCTTTTCTAC-3'; *P. malariae*, PM18SF:5'-GAGACATTCATATATATGAGTG-3' and PM18SR:5'-GTTTTTTTTAATAAAAACGTTCTTTTCCC-3'; *P. ovale*, PO18SF: 5'-GAAAAFFCCTTTTGGAAATTTCTTAG-3' and PO18SR: 5'-GATACATTATAGTGTCTTTTCCC-3' and *P. knowlesi*, PK18SF: 5'-GAGTTTTTCTTTTCTCTCCGGAG-3' and PK18SR: 5'-ACGTTAAATGTGATTCCTTTCCC-3' (Putaporntip *et al.*, 2009).



**Figure 10** PCR workflow for *Plasmodium* species identification targeting 18S rDNA.

The products of *Plasmodium falciparum* were analyzed together with positive control (Figure 11). Samples that showed mixed bands of PCR products with other human *Plasmodium* species (*P. vivax*, *P. ovale*, *P. malariae*, and *P. knowlesi*) were excluded.



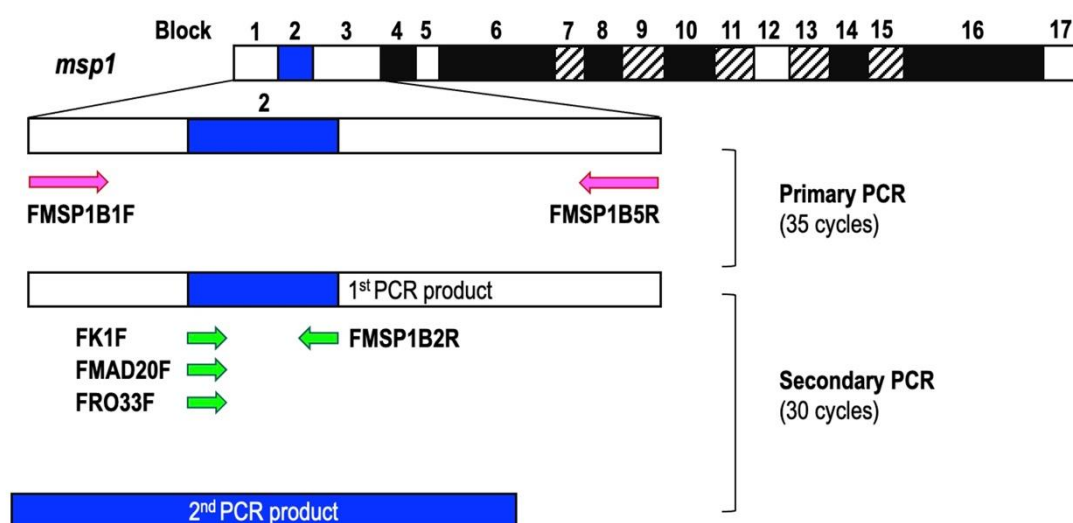
**Figure 11** Species identification of *Plasmodium falciparum* by nested-PCR targeting 18S rRNA. Symbols M, + and - represent the 100-bp DNA ladder, positive control, and negative control, respectively. The positive control and tested samples (Lanes 1-4) are identical bands.

### 3.8 Determination of Clonal Mixture in *P. falciparum* Samples

Because direct sequencing requires homogenous DNA templates, only *P. falciparum* isolates containing single clones were included for analysis. Clonal mixtures were determined from genotyping of two polymorphic loci, i.e. the polymorphic block 2 of *pfmsp1* (Saichanapun (Saichanapun, 2013) and the central repeats of *pfmsp2* (Putaporntip *et al.*, 2008).

#### 3.8.1 Genotyping of Block 2 of *pfmsp1*

The highly polymorphic *pfmsp-1* block 2 locus is used for genotyping and determination of *P. falciparum* clones using PCR technique. Blocks 2 of *pfmsp-1* contains 3 major allelic types (families) represented by K1, MAD20 and RO33 alleles (Tanabe *et al.*, 1987; Jongwutiwes *et al.*, 1992). Both K1 and MAD20 alleles contain tripeptide repeats with sequence and size variation among isolates while RO33 is relatively less polymorphic and lacks repeats (Tanabe, *et al.* 1987; Jongwutiwes, *et al.* 1992). The primers for primary amplifications (outer PCR), are derived from the flanking conserved sequences and the secondary PCR (nested PCR) primers from allelic-specific sequences corresponding to the K1, RO33 and MAD20 families of the *pfmsp-1* gene (Figure 12). All primary and nested PCR reactions were performed in a total volume of 20  $\mu$ L using primers listed in Table 2. The amplification profiles were essentially the same containing a preamplification denatiation at 94°C for 1 minute, followed by 96°C for 30 seconds, 55 °C for 30 seconds, and 61°C for 1.30 minutes in 35 and 25 cycles for primary and nested PCR, respectively. Nested PCR products were visualized by electrophoresis on 2% agarose gel for various lengths of time depending of the predicted size of the PCR products and visualized under UV trans-illumination. Size polymorphism of the PCR products was measured by comparing with the 50 or 100 bp DNA ladder marker. Samples that contained more than one allelic type or the presence of more than one band of a single allele were considered to have multiple clone infections.

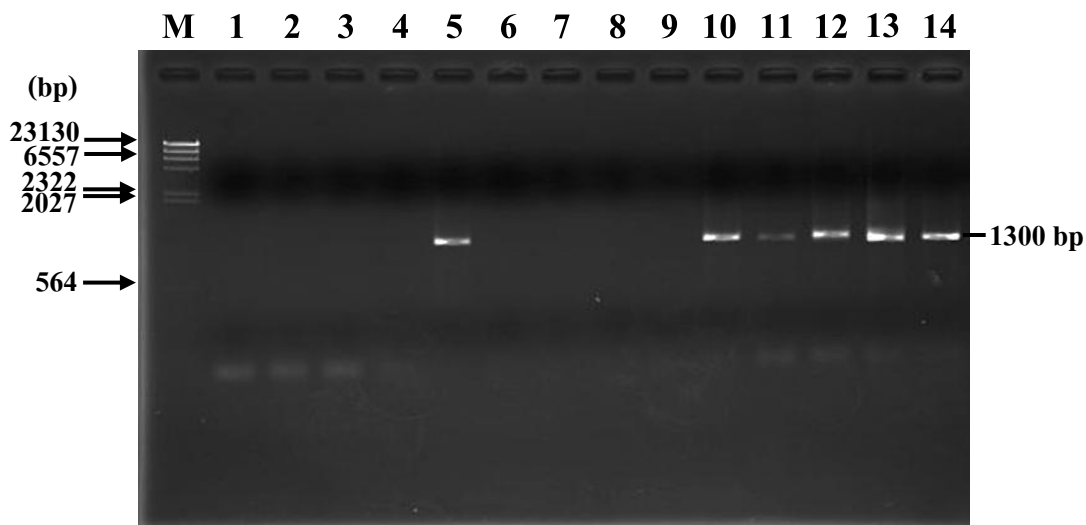


**Figure 12** PCR workflow of clonal mixture determination targeting *msp1* gene.

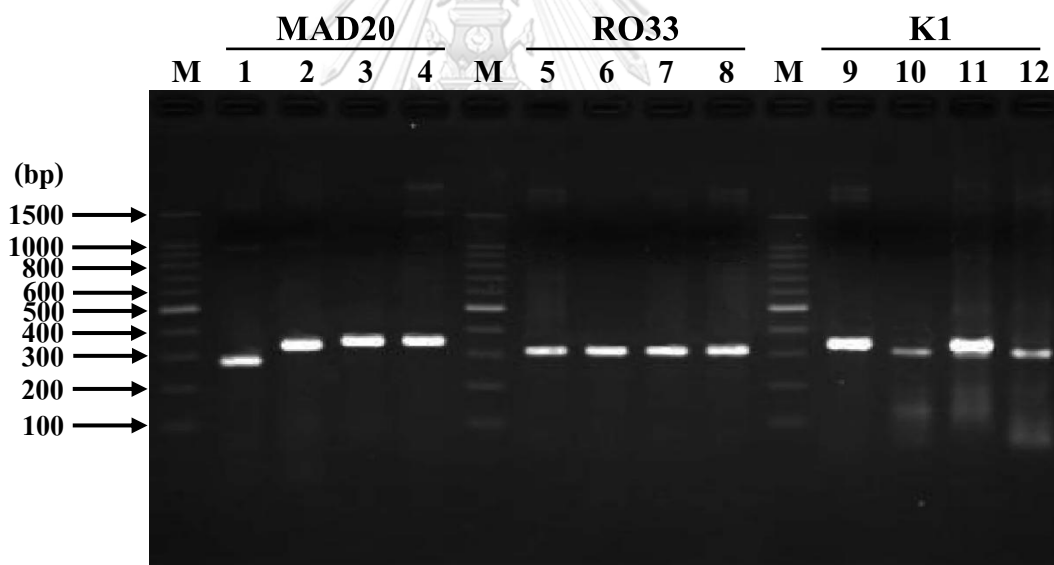
**Table 2** PCR primers and product sizes for genotyping of the *pfmsp-1* block 2.

Allele	Primer	Sequence 5' → 3'	Amplicon Size (bp)
Conserved	FMSP1B1F	CACAATGTGTAACACATGAAAG	1,300
	FMSP1B5R	CAAGTGGATCAGTAAATAAACTATC	
K1	FK1F	GAAATTACTACAAAAGGTGCAAGTG	300 - 350
	FMSP1B2R	CCATCAATTAAATATTTGAAACC	
MAD20	FMAD20F	GAACAGCTGTTACAACACTAGTACAC	250 - 360
	FMSP1B2R	CCATCAATTAAATATTTGAAACC	
RO33	FRO33F	GGAGCAAATACTCAAGTTGTTGC	310
	FMSP1B2R	CCATCAATTAAATATTTGAAACC	

The primary amplification revealed the PCR product size as roughly 1,300 base pairs (Figure 13). The secondary PCR was carried out with allelic-specific primers corresponding to the K1, RO33 and MAD20 allelic families of the *pfmsp-1* gene. The PCR products were visualized under UV trans-illuminator, showing various sizes which were 300 - 350, 310, and 250 - 360 bp for K1, RO33, and MAD20 alleles, respectively (Figure 14). Isolates containing mixed-*pfmsp1* clones were excluded while single-clone isolates were further affirmed by analysis of the *pfmsp2* locus.



**Figure 13** Detection of the primary amplification of *pfmsp-1*. Lambda-DNA marker digested with *Hind* III is loaded in lane 1.

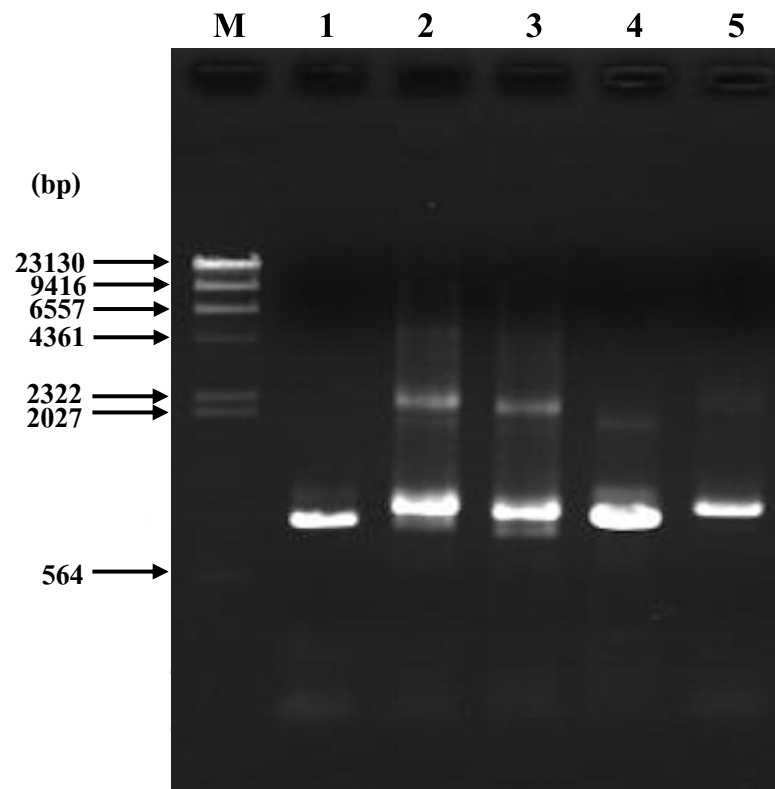


**Figure 14** Genotyping of *Plasmodium falciparum* clones by nested-PCR amplification on twelve representative samples with allele-specific primers, MAD20 (Lanes 1-4), RO33 (Lanes 5-8), and K1 (Lanes 9-12) primer pairs. M lanes are 100-bp DNA ladder.

### 3.8.2 Genotyping of the Central Region of *pfmsp2*

In addition to *Pfmsp1* analysis, the samples were subject to further analysis using PCR targeting the *pfmsp-2* locus. The *pfmsp-2* gene contains central repeats belonging to 2 distinct families, FC27 and 3D7, each of which exhibits sequence and length polymorphism among isolates (Putaporntip *et al.*, 2008). Genotyping of *pfmsp-2* is determined by considering the size polymorphism using the published method (Putaporntip, *et al.*, 2008). Single clones appeared as the single bands of PCR product while mixed clones presented two bands with different size (Figure 15). The primary set of PCR primers were FMSP2F0: 5'-AAAGAATTGTATTTATTAATTCTT AAC-3', and FMSP2R0: 5'-CTCTTCATTTTAAAACATTGAC-3'. PCR products of the primary round were screened on agarose gel. Then, one microlitre of the first-round products was used as DNA templates for secondary PCR. The second primer set included FMSP2F1: 5'-GTATTTATTAATTCTTAACATATTATATTAG -3' and FMSP2R0 as mentioned above. Amplification conditions of both procedures are as described (Putaporntip, *et al.*, 2008). The semi-nested PCR products were analyzed on 2% agarose gel. DNA samples containing a single clone were included in this study while those containing mixed clones were excluded.





**Figure 15** Determination of *Plasmodium falciparum* clones by *pfmsp2*-based PCR. Single clone represented in a single band (Lane 5) while mixture clones appeared in double bands in different size (Lanes 1-4). Lambda DNA maker digested with *Hind* III is run in Lane M.



### 3.9 Amplification of *pfgarp*

The entire coding region of *pfgarp* was amplified by PCR using respective forward and reverse primers, PFGARPF0; 5'-ATAAATAAAGATTAGTATATTTAAAACG-3' and PFGARPR0; 5'-AAATAGCTTTGATTTAACACATTAC-3'. The sequences of these primers were derived from the 5' untranslated regions and the 3' untranslated regions of the complete *pfgarp* sequence of the FC27 strain (GenBank accession number J03998) (Triglia *et al.*, 1988). DNA amplification was carried out in a total volume of 20  $\mu$ L of a single reaction mixture containing DNA template 2  $\mu$ L, 200  $\mu$ M each dNTP, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.0 mM MgCl<sub>2</sub>, 2.5 units Ex-Taq DNA polymerase (Takara, Shiga, Japan) and 1.0  $\mu$ M of each primer. PCR profile contained preamplification denaturation at 94°C for 1 min, followed by 35 cycles of denaturation phase at 94°C for 40 sec, annealing phase at 50°C for 30 sec, extension phase at 62°C for 3 min, and final extension cycle at 72°C for 10 min. The amplification process was performed in PCR thermal cycler (Applied Biosystems, Foster City, CA). Amplicons were analyzed by 1% agarose gel electrophoresis, stained with ethidium bromide and visualized under UV transilluminator.

### 3.10 PCR Product Purification

The solution of PCR product was purified using QIAquick PCR purification kits. The product was purified following the manufacturer's recommendation as followed.

(1) Buffer PB 90  $\mu$ L was added to 18  $\mu$ L of the PCR solution (5 volumes of PB/1 volume of PCR solution). After mixing, if the mixture appeared in orange color, 10  $\mu$ L of 3 M sodium acetate would be added. After that the mixture turned yellow.

(2) The mixture was pipetted into a column tube which was placed in a collection tube. Centrifugation was performed at 13,000 rpm for 1 min. The DNA bound on the filter. The flow-through solution was discarded. After that the column tube was placed back on the same collection tube.

(3) Buffer PE was applied to wash other components out of the DNA by pipetting 750  $\mu$ L into the column tube. Centrifugation was performed. The flow-through solution and collection tube were discarded while the column tube was transferred into the new 1.5 mL tube.

Centrifugation was performed for removing the residual of washing buffer. All centrifugation steps were performed at 13,000 rpm for 1 min. The collection tube was discarded.

(4) The column was placed in a new 1.5 mL microcentrifuge tube. DNA was eluted by adding 40  $\mu$ L buffer EB (10 mM Tris-Cl, pH 8.5) to the center of filter membrane and incubated at room temperature for 15 min. The tube was centrifuged at 18,000 rpm for 2 min.

(5) Step 4 was repeated for increasing the DNA concentration by adding 30  $\mu$ L elution buffer.

(6) The purified DNA was analyzed on 1% gel by electrophoresis technique. The DNA was mixed with loading dye and loaded in a well. Lambda-*Hind* III digested DNA was loaded together with the tested DNA for the target size affirmation.

### 3.11 DNA Sequencing

After purification of amplicons using QIAquick PCR purification kit (QIAGEN, Germany) by following the manufacturer's protocol, DNA sequences were determined bidirectionally from nested PCR-purified templates. Sequencing analysis was performed on an ABI3100 Genetic Analyzer using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Overlapping sequences were obtained by appropriate sequencing primers. The core sequencing primer sequences (PFGARPF1, PFGARPF2, PFGARPF01, PFGARPR2, and PFGARPR3) and sequencing primer sequences for nucleotide affirmation (PFGARPR1, PFGARPR4, and PFGARPR5) were shown in Table 3.

**Table 3** Nucleotide sequences of sequencing primer.

Primer's name	Sequence (5' → 3')
PFGARPF1	CGGGTCAACATAAACCAAAAAACGC
PFGARPF2	ACACGTAGTTAAAAATGTTATAGAAGA
PFGARPF01	AATGAATGTGCTATTTCTTTTCGTAT
PFGARPR2	TACATTGTGGTTGGCTTAGTGGTCTAC
PFGARPR3	ACTAGATATGATTCCACATCCTCC
PFGARPR1	AAATATATTA AAAAAGGGATGG
PFGARPR4	ACAGAGTTTTTCATTATCTTTATC
PFGARPR5	CAAGTGTTTTCAACATTCTTC



### 3.12 Data Analysis

Preliminary sequence alignment and sequence comparison were performed by using CLUSTAL X program (Thompson *et al.*, 1997) and UGENE software version 38.1 (Okonechnikov *et al.*, 2012). Alignment platform was set as a default (Trimming quality threshold is 30, mapping minimal similarity is 80%). Comparative analysis of *pfgarp* sequences was performed in total 99 sequences in which 80 reads from Thai isolates, 18 reads from non-Thai isolates (Accession number [Strain, country of origins]; AF251290 [FCC-1HN, China], AL844501 [3D7, Netherlands from West Africa], J03998 [FC27, Papua New Guinea], LR131290 [Dd2, Indochina], LR131306 [KH2, Cambodia], LR131322 [It, Brazil], LR131338 [HB3, Honduras], LR131354 [KE01, Kenya], LR131386 [GA01, Gambia], LR131402 [GB4, Ghana], LR131418 [KH1, Cambodia], LR131434 [SN01, Senegal], LR131450 [TG01, Togo], LR131466 [SD01, Sudan], LR131481 [ML01, Mali], LR129686 [CD01, Congo], GG665681 [IGH-CR14, India], and KE124372 [UGT5.1, Vietnam]), and one clinical isolate (MDCU32) from Guinea. FC27 was used as reference sequence. Haplotype diversity ( $Hd$ ), nucleotide diversity ( $\pi$ ), and the minimum number of recombination events (RM) were analyzed by using the DnaSP version 5 (Librado & Rozas, 2009) and Genetic Algorithm for Recombination Detection (GARD) (Kosakovskiy *et al.*, 2006). Repeat-containing regions were explored by using the Tandem Repeat Finder (Benson, 1999). Rates of nucleotide substitution in pairwise comparisons between alleles were computed by the numbers of synonymous substitutions per synonymous site ( $ds$ ) and non-synonymous substitutions per non-synonymous site ( $dn$ ) (Nei & Gojobori, 1986) using the MEGA version 5 (Tamura *et al.*, 2011). Neighbor-joining phylogenetic tree based on nucleotide sequences was constructed by using maximum composite likelihood parameter and Tamura-Nei model for building a maximum likelihood tree. Population genetic structure was examined by analysis of molecular variance (AMOVA) to obtain interpopulation variance indices ( $F_{ST}$ ) using the Arlequin software version 3.5.2.2 (Excoffier, 2005). The neutrality test using Tajima's  $D$  statistic was also evaluated by Arlequin software version 3.5.2.2 (Excoffier & Lischer, 2010). The significant level was set at  $p$  value  $< 0.05$  by a permutation test. Prediction of linear B cell epitopes in *PfGARP* was performed by using a sequence similarity to know experimentally verified epitopes from the Immune Epitope DataBase (IEDB) implemented in the BepiBlast Web Server (Ras-Carmona *et al.*, 2022). Furthermore, linear B cell epitopes were also predicted based

on protein language models implemented in BepiPred-3.0 (Clifford *et al.*, 2022). Potential T cell epitopes were selected based on presenting of percentile rank cutoff of 10% and  $IC_{50}$  less than 1000. Binding affinity to T cell was predicted by NetMHCIIpan  $IC_{50}$ , lower number indicates higher affinity. Peptides with  $IC_{50} < 50$  nM are considered high affinity (strong-binder),  $< 1,000$  nM intermediated affinity (weak-binder), and  $> 1,000$  nM low affinity (non-binder) (Zhao & Sher, 2018). The analysis mainly concerned the common HLA class II haplotypes among Thai populations with allele frequency  $> 0.1$  (Satapornpong *et al.*, 2020). The Association between the number of amino acid of repeats and the parasite density was analyzed by Kruskal-Wallis test using IBM SPSS Statistics (Version 28.0.0) with  $p$  value  $< 0.5$  as a significant level.

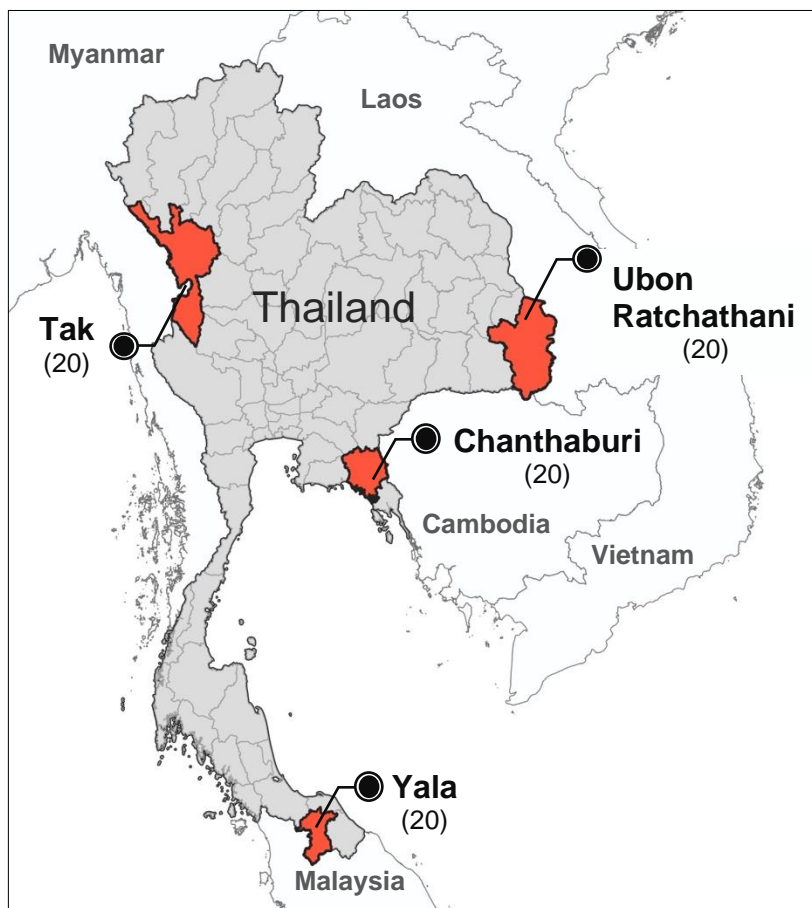


## CHAPTER 4

### RESULTS

#### 4.1 Study Areas and Parasite Populations

In total, 80 isolates of *P. falciparum* blood samples were obtained in this analysis in which 20 each was randomly selected from different target regions, including Tak, Ubon Ratchathani, Chanthaburi, and Yala Provinces (Figure 16). In total, 210 isolates were screened to obtain the appropriate samples in this analysis. For clonal detection by PCR targeting *msp1*, 3 strains of the parasite were identified, including the K1, MAD20, and RO33 types. In total, MAD20 ( $n = 40$ ) was the most abundant, followed by RO33 ( $n = 26$ ), and K1 ( $n = 14$ ), respectively. MAD20 was dominant in Tak and Ubon Ratchathani while RO33 and K1 seemed to predominate in Yala and Chanthaburi, respectively.



**Figure 16** Map of Thailand shows sample collection sites. Number of samples are in the parentheses.

## 4.2 Demographic Data of Study Populations

The baseline characteristics of the patients included in this study, were shown in Table 4. The patients who gave their blood samples for this study were from both genders with male predominance. The ages range from 6 - 78 years old. Among these populations, they were Cambodian, Karen, Laotian, and Thai. Thai contributed to the major ethnic group across the study populations. Karen was distributed in Tak Province while Laotian and Cambodian were confined to Ubon Ratchathani and Chanthaburi Provinces. The distributions of gender, age and ethnicity by province are shown in Table 4.

**Table 4** Demographics of populations from different study areas.

Characteristics	Province (%)				Total n = 80
	Tak n = 20	Ubon Ratchathani n = 20	Chanthaburi n = 20	Yala n = 20	
Gender					
Male	14 (70)	20 (100)	16 (80)	12 (60)	62 (77.50)
Female	6 (30)	0 (0.00)	4 (20)	8 (40)	18 (22.50)
Age (year)					
< 20	12 (60)	4 (20)	8 (40)	9 (45)	33 (41.25)
20 – 40	6 (30)	14 (70)	10 (50)	10 (50)	40 (50.00)
> 40	2 (10)	2 (10)	2 (10)	1 (5)	7 (8.75)
Range	6 - 50	17 - 45	10 - 78	6 - 64	6 - 78
Mean $\pm$ S.D.	19.85 $\pm$ 10.52	26.60 $\pm$ 6.73	28.50 $\pm$ 11.38	24.10 $\pm$ 9.56	24.76 $\pm$ 12.92
Ethnicity					
Thai	2 (10)	19 (95)	10 (50)	20 (100)	51 (63.75)
Cambodian*	0	0	10 (50)	0	10 (12.50)
Karen*	18 (90)	0	0	0	18 (22.50)
Laotian*	0	1 (5)	0	0	1 (1.25)

\* Local settlements in Thailand and transmigration.



### 4.3 Comparison of the Parasite Density by Study Area, Sex, Ethnicity, and *msp1*-genotype

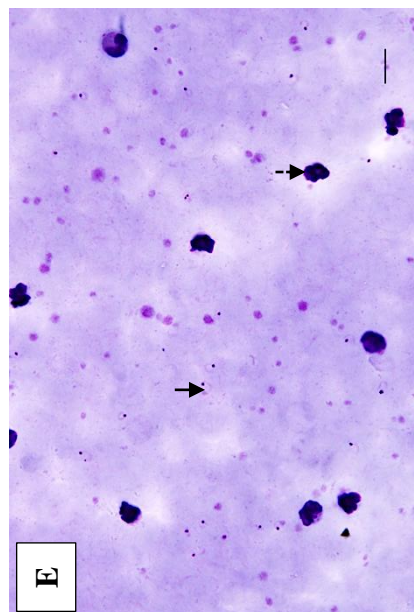
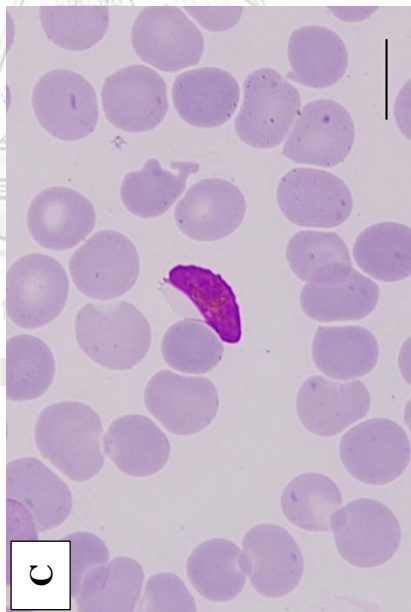
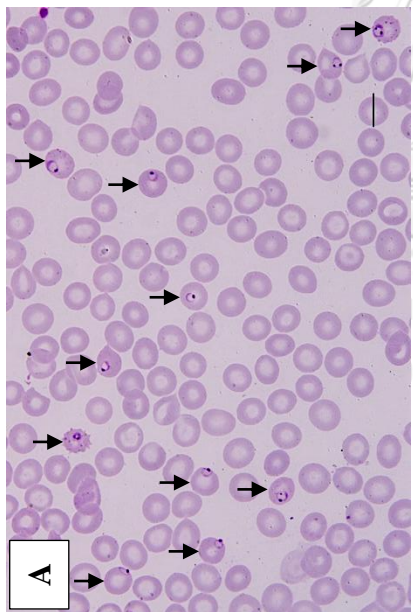
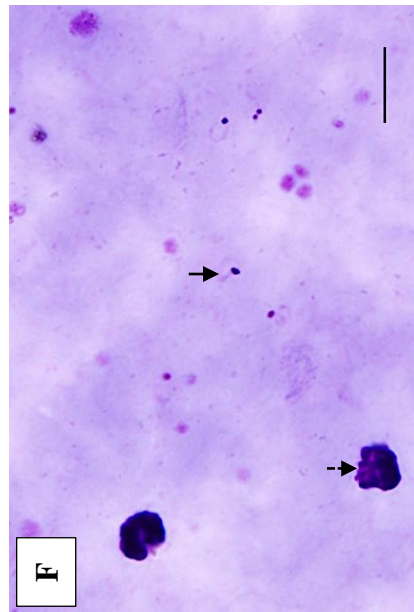
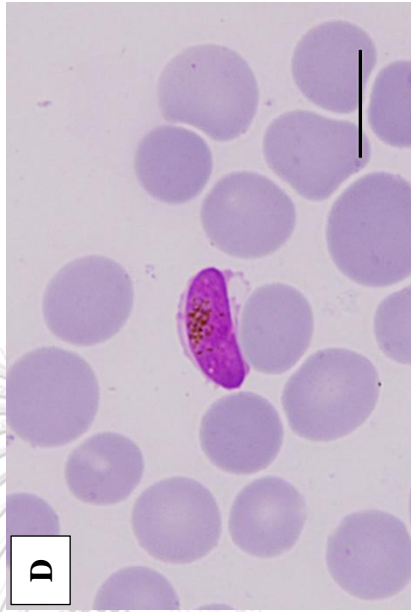
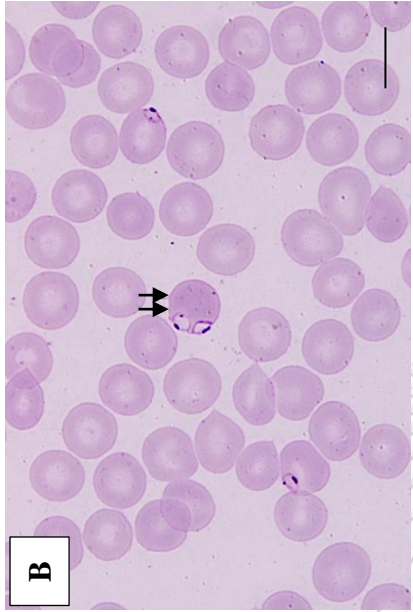
Parasitemia and parasite density per  $\mu\text{L}$  of blood were determined by microscopic observation of thin- and thick- blood films, respectively. According to microscopic observation of all the slides, ring forms (Figure 17A, B, E, and F) of the parasite were more observable than gametocytes (Figure 17C and D) in both thick- and thin-blood smears. Overall, the parasite density ranged from 199 to 864,000 parasite per microliter of blood. The parasite density among parasite populations by study areas was seemingly not different from central tendency (Table 5). Likewise, there were no differences in the parasite density from the central tendency among the parasite populations by sex, ethnicity, and the *msp1*-genotypes (Table 5).

**Table 5** The parasite density of *Plasmodium falciparum* in the study areas

Factors	Parasite density, Median (Q1, Q3)	<i>p</i> value*
Study areas		0.072
Tak	3936 (1753, 26209)	
Ubon Ratchathani	17849 (9307, 42981)	
Chanthaburi	21028 (7910, 101169)	
Yala	3713 (1922, 11747)	
Sex		0.783
Male	11111 (3649, 24387)	
Female	6950 (1835, 28359)	
Ethnicity#		0.259
Thai	10272 (3669, 17897)	
Cambodian	31267 (14133, 159886)	
Karen	3936 (1221, 28359)	
<i>msp1</i> -genotypes		0.590
MAD20	17775 (3744, 32942)	
K1	12167 (3784, 158186)	
RO33	7446 (2373, 61151)	

# A parasite density of Laotian has been omitted regarding statistically unanalyzable.

\* Test of the hypothesis that the medians of the parasite density from which populations are drawn are identical by Median test, significant *p* value < 0.05.

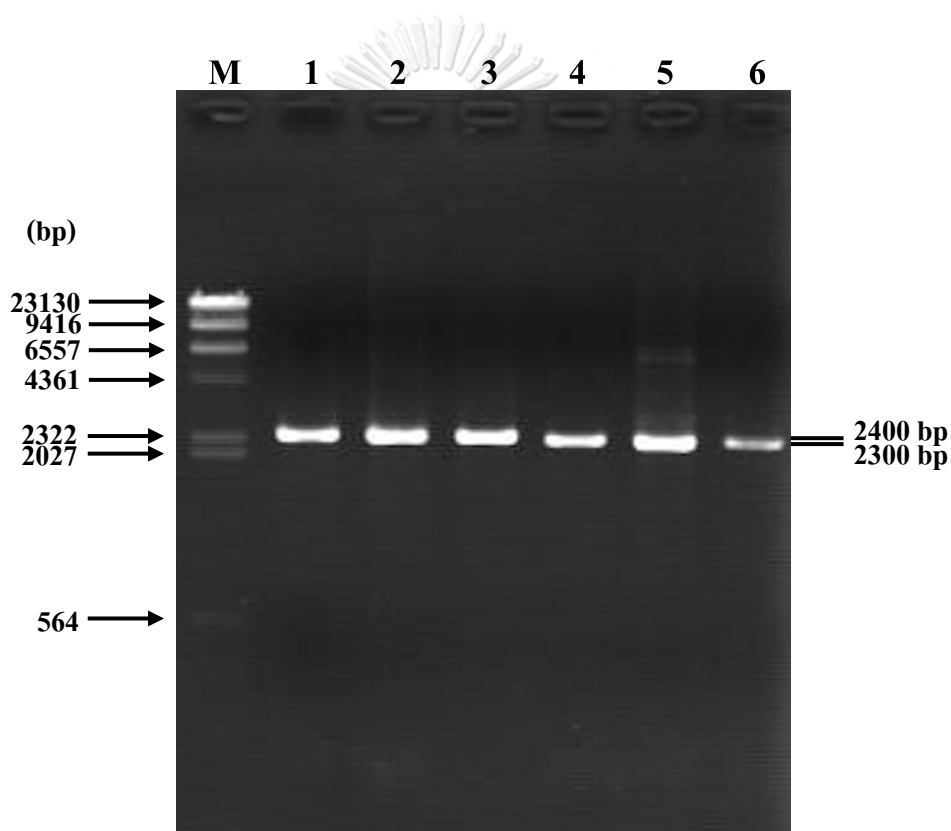


**Figure 17** Giemsa's stained (A-D) thin- and (E-F) thick-blood films show the characteristics of *Plasmodium falciparum* under 100X objective: (A) ring-form of single (single arrow) and (B) double infections (double pointed arrows), (C) male and (D) female gametocytes, (E-F) ring stages (pointed arrow), and leucocytes (dash arrow) on thick blood films. Scale bar is 10  $\mu$ m at 100X objective.



#### 4.4 Amplification of PfGARP-encoding Gene

DNA samples that were identified as single clones of *P. falciparum* based on *msp1* and *msp2* genotypes were amplified by PCR targeting *pfgarp*. All amplicons were separated by 1% agarose gel electrophoresis which showed identical size, measured 2,300 bp. (Figure 18). Isolates that have low parasitemia yielded low concentration of the *Plasmodium* DNA. Therefore, a small amount of DNA template for the *pfgarp* PCR, has resulted in thinner band appearance on agarose gel than those generated from higher amount of DNA in the samples.



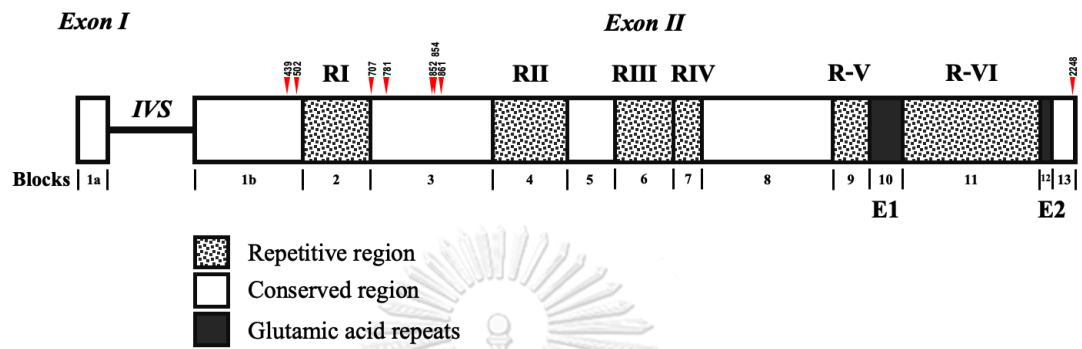
**Figure 18** Agarose gel electrophoresis reveals the slight difference in size of the *pfgarp*-PCR products between lanes 1-3 and 4-6. Lambda DNA digested with *Hind* III is run in lane M.

#### 4.5 Structure of PfGARP-encoding Gene

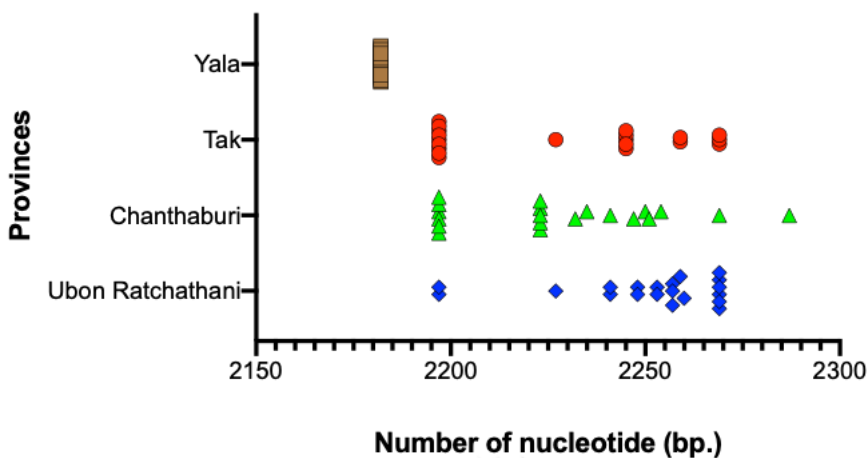
Primarily, all reads of *pfgarp* from 80 Thai, and 19 non-Thai isolates were aligned and mapped with the FC27 reference sequence (GenBank accession no. J03998), before manual

rearrangement. Analyzed data revealed that the sequences of PfGARP encoding gene contain 2 exons (exons I and II) and an intron (Figure 19). Among Thai isolates, the total length of sequences ranged from 2182 to 2287 bp, which could be assigned to 26 alleles. Besides, size and sequence variations in this locus also were observed in previous report which possessed size variation from 2212 – 2269 bp. (Table 6). The exon I was identical in size containing 75 bp, with AT content 77%, while the intron and the exon II ranged from 214 – 216 bp, with 88% AT, and 1893 – 1956 bp, with 73% AT. Variations in sequence length mainly depend on the repetitive unit variation in the exon II. The longest length of nucleotide sequence was noticed in Chanthaburi while the shortest one was limited only in Yala (Figure 20 and Table 6). The exon I of *pfgap* has been predicted to encode 25 amino acids, followed by the intron and the exon II encoded 630-665 amino acids. Based on 99 available sequences, the gene could be divided into 13 blocks (Figure 19). There were five non-repeat and eight repeat regions spanning along the sequences. The latter comprised six heteropolymeric repeats (RI to RVI) and two strings of homopolymeric glutamic acids (E1 and E2) (Figure 19). The glutamic acid was estimated as 24% that was the most abundant amino acid in this gene. Meanwhile, in non-repeat regions, only 8 SNPs were identified (Table 10). Among 26 haplotypes distributed in Thai isolates, the number of haplotypes and haplotype diversity of isolates from Yala Province was remarkably less than other study areas (Table 6). Likewise, the nucleotide diversity of the parasite population in Yala Province was considerably lower than those of the other populations (Table 6). Since the first eight codons of the exon I were location of the PCR primer region, sequences in this study were nearly completed (Figure 21). Therefore, reverse primer was used to obtain the missing 5' codon. Nucleotide sequences and amino acid sequences of Thai isolates were presented herein (Supplementary data 1 and 2). The nucleotide positions were after based on the reference FC27 sequence (J03998)

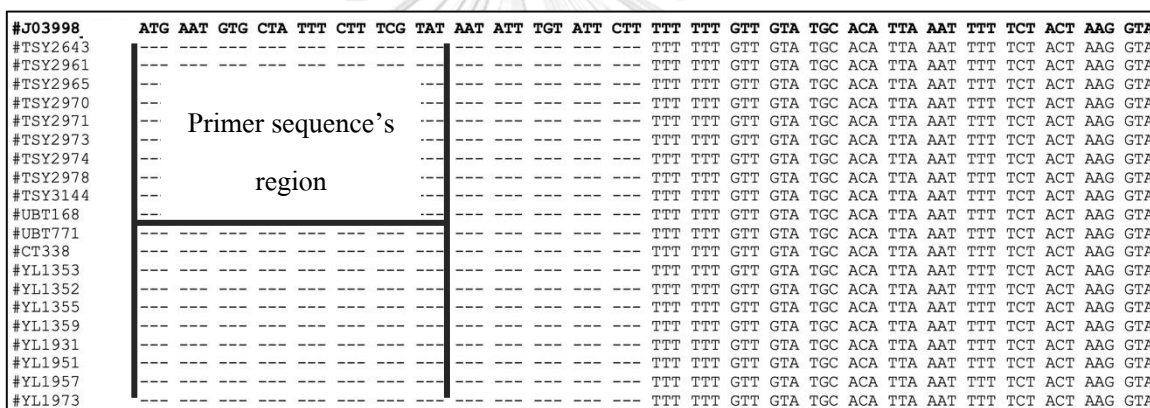
available from the *National Center for Biotechnology Information* (NCBI) database. The first nucleotide of the start codon (ATG) of the reference sequence was set as the first nucleotide position.



**Figure 19** Schematic diagram of *pfgarp* of *P. falciparum* depicts exons I and II, the six blocks of repeats (RI to RVI) and the two polyglutamic acid (E1 and E2) strings. IVS represents intervening sequence (intron). The upside-down triangles point single nucleotide polymorphisms (SNPs) locations. The nucleotide positions for blocks 1a, 1b, and 2-13 are 1 – 75, 290 – 570, 571 – 706, 707 - 997, 998 – 1215, 1216 – 1327, 1328 – 1462, 1463–1537, 1538–1855, 1856–1939, 1940–2026, 2027–2173, respectively.



**Figure 20** Length of the *pfgarp* sequences of *Plasmodium falciparum* from different study areas. Individual shape and color denote the samples in each region.



**Figure 21** Representative sequence alignment illustrates the first 26 codons of *pfgarp*. The primer sequence region is depicted as the first eight codons of the exon I plus base A on 5'UTR. The representative sequence names are shown on the left column.

## 4.6 Genetic Diversity of the PfGARP-encoding Gene

### 4.6.1 Nucleotide Diversity of the PfGARP-encoding Gene

Analysis of genetic variation in the near complete sequence of *pfgarp* were demonstrated herein (Table 6). The shortest length of the sequence was noticed in Yala while the longest one was spotted in Chanthaburi. Nucleotide diversity ( $\pi$ ) of the sequences was relatively low especially in Yala Province which was  $0.00005 \pm 0.00005$ , significantly lower than those in the other areas in Thailand (Table 6). Overall nucleotide diversity in both Thai and non-Thai isolates was  $0.42400 \pm 0.00098$ .

**Table 6** Nucleotide diversity in *pfgarp* among different *Plasmodium falciparum* populations

Populations	<i>n</i>	Number of nucleotides	$\pi \pm$ S.E.
Thai	80	2182 - 2287	$0.00366 \pm 0.00100$
Chanthaburi	20	2197 - 2287	$0.00333 \pm 0.00081$
Tak	20	2197 - 2269	$0.00349 \pm 0.00091$
Ubon Ratchathani	20	2197 - 2269	$0.00365 \pm 0.00082$
Yala	20	2182	$0.00005 \pm 0.00005\#$
Non-Thai	19	2212 - 2269	$0.00576 \pm 0.00083$
Global	99*	2182 - 2287	$0.00424 \pm 0.00098$

\* Complete and near complete sequence analysis, *n* = number of sequences,  $\pi$  = nucleotide diversity, S.E. = standard error. Test of the hypothesis that  $\pi$  for one province equals  $\pi$  for another province in Thailand: #  $p < 0.0005$ .



#### 4.6.2 Haplotype Diversity

To evaluate the genetic diversity of *pfgarp*, all *pfgarp* sequences from different malaria-endemic areas in Thailand and worldwide were analyzed. The analyzed data demonstrated 44 haplotypes (Table 7). Among these haplotypes, 26 haplotypes distributed among different parasite populations in Thailand, including Chanthaburi, Tak, Ubon Ratchathani, and Yala Provinces (Figure 22) while the rest occurred in non-Thai isolates (Table 8). Among Thai isolates, more haplotypes were observed in Chanthaburi Province ( $n = 11$ , 42%) than other endemic areas. Haplotype 10 ( $n = 19$ ), 6 ( $n = 17$ ), and 5 ( $n = 9$ ) were the most three dominant haplotypes that accounted for 44% of all *P. falciparum* samples (Figure 22). There were only two haplotypes (haplotypes 5 and 6) were shared among parasite populations from Chanthaburi, Tak, and Ubon Ratchathani while the remaining seemed to be unique for each endemic province. It is noteworthy that the parasite population from Yala was limited in only two haplotypes that were haplotypes 10 ( $n = 19$ ) and 11 ( $n = 1$ ). Remarkably, haplotype diversity of the parasite isolates from Yala seemed to be lower than those from of the other provinces (Figure 22 and Table 7).

**Table 7** Haplotype frequency and haplotype diversity among Thai and global *Plasmodium falciparum* populations.

Origin of parasite isolates	<i>n</i>	M	h	<i>Hd</i> ± S.D.	Haplotypes
Thai	80	45	26	0.887 ± 0.022	1 - 26
Chanthaburi	20	35	11	0.868 ± 0.057	5, 6, 12 - 20
Tak	20	22	9	0.795 ± 0.087	1 - 9
Ubon Ratchathani	20	32	8	0.868 ± 0.049	5, 6, 21 - 26
Yala	20	1	2	0.100 ± 0.088	10, 11
Non-Thai	19	52	19	1.000 ± 0.017	5, 27 - 44
Global	99	58	44	0.925 ± 0.016	1 - 44

h = the number of haplotypes, *Hd* = haplotype diversity, M = the number of mutation sites, *n* = the number of isolates, S.D. = Standard deviation

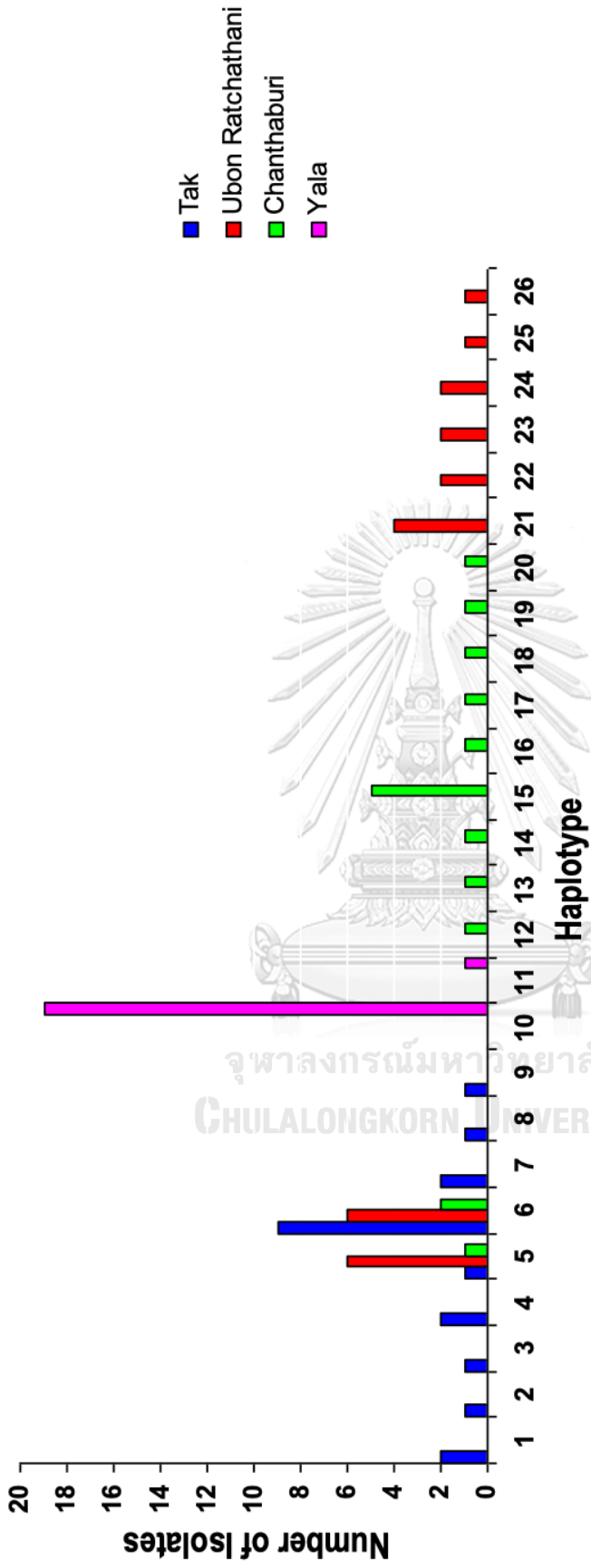


Figure 22 Distribution of *PfGARP* haplotypes inferred from coding regions among 80 Thai isolates.

**Table 8** Haplotypes of *pfgarp* among Thai and worldwide isolates.

H	n	Isolates
1	2	APH80, TSY2638
2	1	TSY3136
3	1	TSY2639
4	2	AP24, TSY2647
5	9	CT223, TSY2947, UBT1004, UBT126, UBT146, UBT280, UBT454, UBT176, KH2
6	17	TSY2961, TSY2965, TSY2970, TSY2971, TSY2973, TSY2974, TSY2978, TSY3144, UBT168, UBT771, CT338, AP2460, CT1459, CT1565, CT18, CT435, CT489
7	2	TSY2977, TSY2976
8	1	TSY2643
9	1	TSY3168
10	19	YL1352, YL1355, YL1359, YL1931, YL1951, YL1957, YL1973, YL1974, YL1977, YL1987, YL1988, YL1999, YL2000, YL2003, YL2004, YL2005, YL2007, YL2009, YL2013
11	1	YL1353
12	1	CT150
13	1	CT154
14	1	CT1598
15	5	CT339, CT344, CT350, CT447, CT450
16	1	CT346
17	1	CT258
18	1	CT1597
19	1	CT1334
20	1	CT446
21	4	UBN112, UBN93, UBT118, UBT340
22	2	UBT113, UBT209
23	2	UBN259, UBT312
24	2	UBT139, UBT784
25	1	UBT303
26	1	UBT144
27	1	FC27
28	1	3D7
29	1	SD01
30	1	Dd2
31	1	KH1
32	1	FCC1HN
33	1	ML01
34	1	TG01
35	1	GA01
36	1	SN01
37	1	GB4
38	1	HB3
49	1	It
40	1	KE01
41	1	CD01
42	1	IGHCR14
43	1	UGT5
44	1	MDCU32

H = haplotypes, n = The number of haplotypes, Thai isolates with initials TSY and AP are from Tak, UBN and UBT from Ubon Ratchathani, YL from Yala and CT from Chanthaburi Provinces. The numbers following these initials are used to label individual isolates.

## 4.7 Structural Organization of PfGARP-encoding Gene

### 4.7.1 Exon I and Intron

The 75-bp coding region in exon I exhibited perfect sequence identity among Thai and worldwide isolates. The adjacent intron region displayed two variants due to short insertion/deletion of TA residues. Four palindromic sequences flanking indel region were identified based on the reference sequence (J03998) (Table 9).

**Table 9** Palindromic sequences of *pfgarp*

Palindromes	Start position	Sequence 5'-->3'	Start position	Sequence 5'-->3'
1	81	AAAAATATAT	191	TTTTTATATA
2	141	TATAATAAAT	230	ATATTATTTA
3	145	ATAAATATGTA	265	TATTTATACAT
4	243	TACATATATATT	277	ATGTATATATAA

Nucleotide positions based on J03998 sequence.

### 4.7.2 Non-repeat Regions in Exon II

The nonrepetitive sequences in exon II were highly conserved. Among Thai and non-Thai isolates, eight SNPs were identified (designated SNP1 [c. 439A>G (E75)], SNP2 [c. 502A>T (I96)], SNP3 [c. 707A>G (K165E)], SNP4 [c.781G>T (D193Y)], SNP5 [c. 852C>T (P213L)], SNP6 [c. 854T>G (Y214D)], SNP7 [c. 861A>G (Y216C)], and SNP8 [c. 2248C>T (I678)]) (Tables 10 and 11). SNP1 and SNP2 are located in block 1b while SNP3 – SNP7 and SNP8 occurred in conserved block 3 and block 13, respectively (Figure 19). Among these, SNP3 - SNP7 were identified as non-synonymous SNPs which have amino acid alteration, while the remainders (SNP1, SNP2, and SNP8) were synonymous SNPs which did not change an amino acid. Three singleton SNPs were noticed, SNP1, SNP2, and SNP5, which all were non-Thai sequences obtained from GenBank. Observation on 80 sequences of Thai isolates revealed that there were five out of eight SNPs that all sequences limited in only one genotype, GAA, ATT, TAT, CCA, and TAT for the respective SNP1, SNP2, SNP4, SNP5, and SNP6 (Table 10). Noted that only GAA genotype of SNP3 was detected only in Yala, while the other genotype (AAA) of

this SNP occurred predominantly in Chanthaburi, Tak, and Ubon Ratchathani. Both genotypes of SNP7 and SNP8 distributed in all parasite populations but was not found in Yala where the genotypes limited in TAT and ATC for SNP7 and SNP8, respectively (Table 10). The distribution of SNPs among isolates was demonstrated in Supplementary data 3. The remaining non-repeat regions including blocks 5 and 8 were perfectly conserved among Thai and worldwide isolates.

**Table 10** The genotype frequency for three SNPs of the PfGARP-encoding gene.

SNPs	Nucleotide position (Codon)	Nucleotide substitution (amino acid)	Number of isolates					Number of worldwide isolates
			TK	UBN	CT	YL	Total	
1	437 – 439 (75)	GAA (E)	20	20	20	20	80	18
		. . G (E)	0	0	0	0	0	1
2	500 – 502 (96)	ATA (I)	0	0	0	0	0	1
		. . T (I)	20	20	20	20	80	18
3	707 – 709 (165)	AAA (K)	20	20	20	1	61	9
		G . . (E)	0	0	0	19	19	9
4	781 – 783 (193)	GAT (D)	0	0	0	0	0	5
		T . . (Y)	20	20	20	20	80	14
5	851 – 853 (213)	CCA (P)	20	20	20	20	80	18
		. T . (L)	0	0	0	0	0	1
6	854 – 856 (214)	TAT (Y)	20	20	20	20	80	16
		G . . (D)	0	0	0	0	0	3
7	860 – 862 (216)	TAT (Y)	10	18	14	20	62	19
		. G . (C)	10	2	6	0	18	0
8	2246 – 2248 (678)	ATC (I)	10	2	7	20	39	3
		. . T (I)	10	18	13	0	41	16

CT = Chanthaburi, TK = Tak, UBN = Ubon Ratchathani, YL = Yala

**Table 11** Genotypes of eight SNPs in the *pgarp* sequence collecting from different resources

Origin of isolates	SNPs
	c. 439A>G c. 502A>T c. 707A>G c. 781G>T c. 852C>T c. 854T>G c. 861A>G c. 2248C>T (E75) (I96) (K165E) * (D193Y) * (P213L) * (Y214D) * (Y216C) * (I678)
J03998 ( <i>n</i> = 1)	A A A G C C T T A A C
Non-Thai ( <i>n</i> = 17)	A/G T A/G G/T C/T T/G A A C/T
MDCU32 ( <i>n</i> = 1)	A T T G T C G T A A T
Chanthaburi ( <i>n</i> = 20)	A T T A T T C C T T A/G C/T
Tak ( <i>n</i> = 20)	A T T A T T C C T T A/G C/T
Ubon Ratchathani ( <i>n</i> = 20)	A T T A T T C C T T A/G C/T
Yala ( <i>n</i> = 20)	A T T A/G T T C C T A C

\* Non-synonymous SNPs. *n* = the total sample size.

### 4.7.3 Repetitive Regions

Repeat-containing regions of all sequences collected from different areas in Thailand and those retrieved from GenBank database were explored by using the Tandem Repeat Finder. In this study, six repetitive (designed as RI - RVI) domains and two strings of glutamic acids (Es) were identified, all of them were distributed in the exon II. Previously reported repeat sequence motifs in blocks 1-4 are corresponding to domains RI, RIII, RIV and RVI, respectively (Trigilia *et al.*, 1988). The first domain of repeats (**Repeat I: RI, block 2**) commenced at base 571 and consists of 12 – 19 repetitive units of KXX-encoding motif, where X is D, E, H or K. Based on RI, 13 alleles were assigned (Table 12). Designation of alleles in all repeat-containing regions was referred here according to the number of amino acid residues. When different sequences contained identical number of amino acids, alleles were further subdivided by adding an alphabet following the number to indicate variants; thereby, new alleles could be included in alphabetical order. Of these, eight alleles found in Thai isolates in which RI-45A allele was the most prevalent variant (Table 12). The second repeat domain (**Repeat II: RII, block 4**) was newly identified in this present study. It contained 2 complex units commencing at base 998. Based on 33 codon-encoding KKERKQKEKEMKE(or K)QE(or K)KIEKK(or E)K(or R)KKQ(or K)EEKEKKKQ(or K)E(or K), a short region encoding KERKKQE intervened between the units. Deletion of the last two lysine residues of the repeat were found in all Thai and most worldwide isolates (Table 13). The third repeat domain (**Repeat III: RIII, block 6**) commenced at base 1228. This repeat region ranged from 6 to 9 pentapeptides of E(or G/K)EH(or D)K(or E/K)E(or K/S). The repeats comprised EEHKE, GEHKE, GEDKE, GEHKK, EEHKK, GEHEE, EEHKS, GEHKS and KEHKE. There were 19 alleles in total (Table 14). Of these, 11 alleles were found in Thailand. Allele RIII-30 was most common, followed by alleles RIII-45A and RIII-35A (Table 14). After RIII, it was instantly followed by the fourth repeats (**Repeat IV: RIV, block 7**). This repeat region comprised 5 pentapeptide units of the KG(or A/V)KKX, where X is D, E, H or K. All repeat units comprised KGKKD, KGKHK, KAKKE, KVKKH. This repeat was perfectly identical among all analyzed sequences. The fifth imperfect repeats (**Repeat V: RV, block 9**) commenced at base 1856. It was newly identified in this study. RV comprised 2 to 4 units of XEE(or D)S(or E)K(or E)EV where X is E, K or Q. The repeats were flanked by EEDK and EED or QEDED at the N- and C- termini, respectively. Four alleles have been found in Thai isolates (Table 15). The sixth

repeats (**Repeat VI: RVI, block 11**) based on 5 – 9 heptapeptide of E(or D)E(or D)E (or D)XE(or D)E(or D)D where X is A, D, E or V, following by (E)<sub>n</sub>(D)<sub>m</sub> residues where n and m varied from 1-5 and 1-3, respectively. RVI was the most polymorphic region in PfGARP. Based on this repeat domain, 27 alleles were assigned and 13 of them circulated among Thai isolates (Table 16).

#### 4.7.4 Homopolymeric Glutamic Acid Repeats

There is a string of 16 - 29 Es (E1) commencing at base 1940 running after the RV. This region contained repeat-motif of GAX where X is A or G (Table 17). The other string of glutamic acids (E2) was followed instantly by RVI. It consisted of 5-11 amino acid residues all of which encoded by GAA (Table 17). The spanning of repeats along the gene sequence was illustrated as Figure 19. A comparison of the number of repetitive units in each repeat among Thai and Non-Thai isolates were shown in Table 18.

#### 4.8 Recombination

The minimum number of recombination events (RM) of the PfGARP-encoding gene in all 80 sequences was explored by two algorithms which were the DnaSP version 5.10.01 and the Genetic Algorithm for Recombination Detection (GARD) (Kosakovsky Pond *et al.*, 2006) on Data Monkey Web Server. No evidence of RM has been detected.





**Table 13** Diversity and distribution of repeat alleles in RII of PfGARP.

Allele	Amino acid sequence	Thai isolates, n					Non-Thai isolates/strains*	Total
		Tak	Ubon Ratchathani	Chanthaburi	Yala	Total		
RII-73	KKERKQKEKEMK <b>E</b> QEKIEKKK <b>K</b> QEEKEKKK <b>Q</b> E KERKKQ <b>E</b> KKERKQKEKEMK <b>K</b> QKKIEKER <b>R</b> KKKEKEKKK <b>K</b>	0	0	0	0	0	FC27, 3D7, SD01 and CD01	4
RII-71	KKERKQKEKEMK <b>E</b> QEKIEKKK <b>K</b> QEEKEKKK <b>Q</b> E KERKKQ <b>E</b> KKERKQKEKEMK <b>K</b> QKKIEKER <b>R</b> KKKEKEKKK <b>K</b> --	20	20	20	20	80	Dd2, FCC1/HN, GA01, GB4, HB3, IGH-CR14, IT, KE01, KH1, KH2, MDCU32, ML01, SN01, TG01 and UGT5.1	15
Total		20	20	20	20	80		19

MDCU32 is from a Guinean patient.

\* GenBank accession numbers are listed in data analysis section.

**Table 14** Diversity and distribution of repeat alleles in RIII of PfGARP.

Allele	Amino acid sequence	Thai isolates, n					Non-Thai isolates/strains*	Total
		Tak	Ubon Ratchathani	Chanthaburi	Yala	Total		
RIII-45A	EEHKEGEHKEEEHKEGEDKGEHKKKEEHHKKEEHHKSKEHKS	3	10	4		17	KH1, KH2	2
RIII-45B	EEHKEGEHKEEEHKEEEHKKKEEHHKKEEHHKKEEHHKSKEHKS	3	3			6		
RIII-45C	EEHKEGEHKEEEHKEGEHKEGEHKEGEHKEGEHKEEHHKKEEHHKSKEHKS	3				3	IT	1
RIII-45D	EEHKEGEHKEEEHKEGEHKEGEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS		2	1		3	FCC1/HN	1
RIII-45E	EEHKEGEHKEEEHKEEEHKEGEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS			1		1		
RIII-45F	EEHKEGEHKEEEHKEGEHKEGEHKEEHHKKEEHHKKEEHHKKEEHHKSKEHKS						3D7, FC27, SD01	3
RIII-45G	EEHKEGEHKEEEHKEEEHKEGEHKEGEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS						KE01	1
RIII-40A	EEHKEGEHKEEEHKEGEHKEGEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS			6		6		
RIII-40B	EEHKEGEHKEEEHKEEEHKEEHHKKEEHHKKEEHHKKEEHHKSKEHKS	2		1		3	Dd2	1
RIII-40C	EEHKEGEHKEEEHKEEEHKEGEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS						GA01, GB4, HB3, SN01	4
RIII-40D	EEHKEGEHKEEEHKEGEHKEEHHKKEEHHKKEEHHKKEEHHKSKEHKS						CD01, TG01	2
RIII-40E	EEHKEGEHKEEEHKEEEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS						IGH-CR14	1
RIII-40F	EEHKEGEHKEEEHKEEEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS						ML01	1
RIII-40G	EEHKEGEHKEEEHKEEEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS						MDCU32	1
RIII-35A	EEHKEGEHKEEEHKEGEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS	9	2	6		17		
RIII-35B	EEHKEGEHKEEEHKEEEHKEEHHKKEEHHKKEEHHKSKEHKS		3			3		
RIII-35C	EEHKEGEHKEEEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS			1		1		
RIII-35D	EEHKEGEHKEEEHKEGEHKEEHHKKEEHHKKEEHHKSKEHKS						UGT5.1	1
RIII-30	EEHKEGEHKEEEHKEGEHKSKEHKSKEHKS				20	20		
Total		20	20	20	20	80		19

MDCU32 is from a Guinean patient. \* GenBank accession numbers are listed in data analysis section.

**Table 15** Diversity and distribution of repeat alleles in RV of PfGARP.

Allele	Amino acid sequence	Thai isolates, n					Non-Thai isolates/strains*	Total
		Tak	Ubon Ratchathani	Chanthaburi	Yala	Total		
RV-35	<b>EEDKKEESKEVEEESKEVQEEESKEVQDEDEEEVFEED</b>						FC27, ML01	2
RV-28A	<b>EEDKKEESKEVEEESKEVQDEDEEEVFEED</b>	10	18	5		33	3D7, FCC1/HN, GB4, HB3, IGH-CR14, IT, KE01, KH1, KH2, MDCU32, TG01, UGT5.1	12
RV-28B	<b>EEDKKEESKEVQEEESKEVQDEDEEEVFEED</b>			2		2	Dd2, GA01, SD01, SN01	4
RV-23	<b>EEDKKEESKEVEEESKEVQDEDEEEVFEED</b>			7		7	CD01	1
RV-21	<b>EEDKKEESKEVQDEDEEEVFEED</b>	10	2	6	20	38		
Total		20	20	20	20	80		19

MDCU32 is from a Guinean patient.

\* GenBank accession numbers are listed in data analysis section.



**Table 16** Diversity and distribution of repeat alleles in RVI of PfGARP (Cont.).

Allele	Amino acid sequence	Thai isolates, n					Non-Thai isolates/strains*	Total
		Ubun Ratchathani						
		Tak	Chanthaburi	Yala	Total	Total		
RVI-54A	EDEVEEDDAEEDEDDAAEEDEDDAAEEDDDDAEEDDDDEDEDEDEDEDEDEEED		1				1	
RVI-54B	VEEDEDAAEEDEDDAAEEDEDDAAEEDDDDEDEDEDEDEDEDEDEDEDEEED					SN01	1	
RVI-53	EDEVEEDDAEEDEDDAAEEDEDDAAEEDDDDAEEDDDDAEEDDDDEDEDEDEEED					CD01	1	
RVI-51	EDEVEEDDAEEDEDDAAEEDEDDAAEEDDDDAEEDDDDEDEDEDEDEDEDEEED	10	2	6	20	38		
RVI-49	EDEVEEDDAEEDEDDAAEEDEDDAAEEDDDDEDEDEDEDEDEDEDEDEDEEED	20	20	20	20	80	1	
Total							19	

MDCU32 is from a Guinean patient.

\* GenBank accession numbers are listed in data analysis section.



**Table 17** Sequence variation in polyglutamic acid repeats in E1 and E2 domains of *PjGARP*.

Domain/ Allele	Sequence#	Thai isolates					Non-Thai isolates/strains*	Total	
		Tak	Ubon Ratchathani	Chanthaburi	Yala	Total			
E1	11111121112111211121112111	10	2	6	20	38	FC27, CD01	2	
	11111121112111211121112111	4	2	0	0	6	-	0	
	111111211121112111211121	0	0	0	0	0	IGH-CR14	1	
	111111211121112111211121	0	0	0	0	0	GA01, SN01	2	
	11111211111111112111	0	0	0	0	0	SD01	1	
	111112111111112111	0	0	0	0	0	3D7	1	
	111112111111112111	0	0	11	0	11	UGT5.1, Dd2, KH1, ML01, TG01	5	
	111112111121112111	0	2	0	0	2	-	0	
	111111111121112111	0	0	0	0	0	KE01	1	
	1111121111211121	6	14	3	0	23	FCC1/HN, KH2, GB4, HB3, IT, MDCU32	6	
	Total	20	20	20	20	80		19	
	E2	111111111111	0	0	0	0	0	KE01	1
		1111111111	0	0	1	0	1	FC27, Dd2	2
		1111111111	10	2	7	20	39	ML01, CD01	2
		1111111111	0	0	9	0	9	KH1, MDCU32	2
		1111111111	6	18	3	0	27	FCC1HN, KH2, IT	3
1111111111		0	0	0	0	0	3D7, IGH-CR14, SD01, GA01, SN01	5	
1111111111		4	0	0	0	4	UGT5.1, TG01, GB4, HB3	4	
Total		20	20	20	20	80		19	

# 1 and 2 represent GAA and GAG, respectively. Alleles are assigned based on the number of codons and their variants in alphabetical orders.

MDCU32 is from a Guinean patient.

Dash denotes none among non-Thai isolates/strains.

\* GenBank accession numbers are listed in data analysis section.

**Table 18** The repeats of the *plgarp* sequences.

Reference Seq (J03998)	Thai isolates (80 seq) + MDCU32	Non-Thai isolates (17 seq)	Worldwide (99 seq)
<b>Repeat I</b> KKX X; E, D, H, K 15 units	<b>RI</b> KKX X; D, E, H, K 15-19 units	<b>KKX</b> X; D, E, H, K 12-17 units	<b>KKX</b> X; D, E, H, K 12-19 units
	<b>RII</b> KKERKQKEKEMKE(or K)QE(or K)KIEKK(or E)K(or R)KKQ(or K)EEKEKKQQ(or K)E(or K) KERKKQE-motif inserted between the units 2 units		
<b>Repeat II</b> E/G/KEHKE Variants 2S, 1K in 5 <sup>th</sup> position 9 units	<b>RIII</b> E/G/KEHKE Variants 2S, 1-3K in the 5 <sup>th</sup> position 2D in the 3 <sup>rd</sup> position 6-9 units	<b>E/G/KEHKE</b> Variants 2S, 1-2K in the 5 <sup>th</sup> position 2D in the 3 <sup>rd</sup> position 8-9 units	<b>E/G/KEHKE</b> Variants 2-3S, 1-3K in the 5 <sup>th</sup> position 2D in the 3 <sup>rd</sup> position 6-9 units
<b>Repeat III</b> KG(or A/V)KXX X; D, E, H, K 5 units	<b>RIV</b> KG(or A/V)KXX X; D, E, H, K 5 units		
	<b>RV</b> XEE(or D)S(or E)K(or E)EV X; E, K, Q Franked by EEDK and EEED or QEED at the N- and C- termini, respectively. 2-4 units		
<b>E1 string</b>	16-29 residues (16, 21, 27, 29)		
<b>Repeat IV</b> E(or D)E(or D)E(or D)XE(or D)E(or D)D(E)n(D)m X; A, D, E n: 1-5, m: 1-3	<b>RVI</b> E(or D)E(or D)E (or D)XE(or D)E(or D)D(E)n(D)m X; A, E, V n: 1-5, m: 1-3	E(or D)E(or D)E(or D)XE(or D)E(or D)D(E)n(D)m X; A, E, V n: 1-5, m: 1-3	E(or D)E(or D)E(or D)XE(or D)E(or D)D(E)n(D)m X; A, D, E, V n: 1-5, m: 1-3
<b>E2 string</b>	5-11 residues (5, 7, 8, 9, 10, 11)		



#### 4.9 Selective Pressure on PfGARP

To measure the selective pressure on PfGARP-encoding gene among *P. falciparum* populations in Thailand, the neutrality test was applied for *pfgarp* coding sequences. The Tajima's *D* test based on the segregating sites in non-repeat regions was computed using Arlequin software (Table 19). The analyzed data revealed that the overall Tajima's *D* value was positive (0.397), statistically not significant, *p* value > 0.05. These finding indicated no deviation from selective neutrality at this locus among the parasite populations in Thailand base on *pfgarp*. Meanwhile, the rate of nonsynonymous substitutions per nonsynonymous site ( $d_N \pm \text{S.E.} = 0.00860 \pm 0.00380$ ) significantly exceeded that of synonymous substitutions per synonymous site ( $d_S \pm \text{S.E.} = 0.00000 \pm 0.00000$ ) in block 3 ( $p = 0.024$ ) whereas no significant difference between these parameters occurred in other blocks based on the FC27 sequence (Z-test) (Table 20).

**Table 19** Tajima's *D* statistics in different regions of *pfgarp* by province.

Resources	<i>n</i>	M	S	Tajima's <i>D</i> values	<i>p</i> values*
Chanthaburi	20	35	10	1.48611	0.93600
Tak	20	22	15	2.03336	0.99000
Ubon Ratchathani	20	32	12	-0.76857	0.24700
Yala	20	1	1	-1.16439	0.14400
Overall	80	45	18	0.39663	0.57925

*n* = Number of sequences, M = Number of mutation sites, S = Number of polymorphic (segregating) sites, \* Test of hypothesis that there is no deviation from neutrality by Tajima's *D* test, significant *p* value < 0.05.

**Table 20** Neutrality test in PfGRAP.

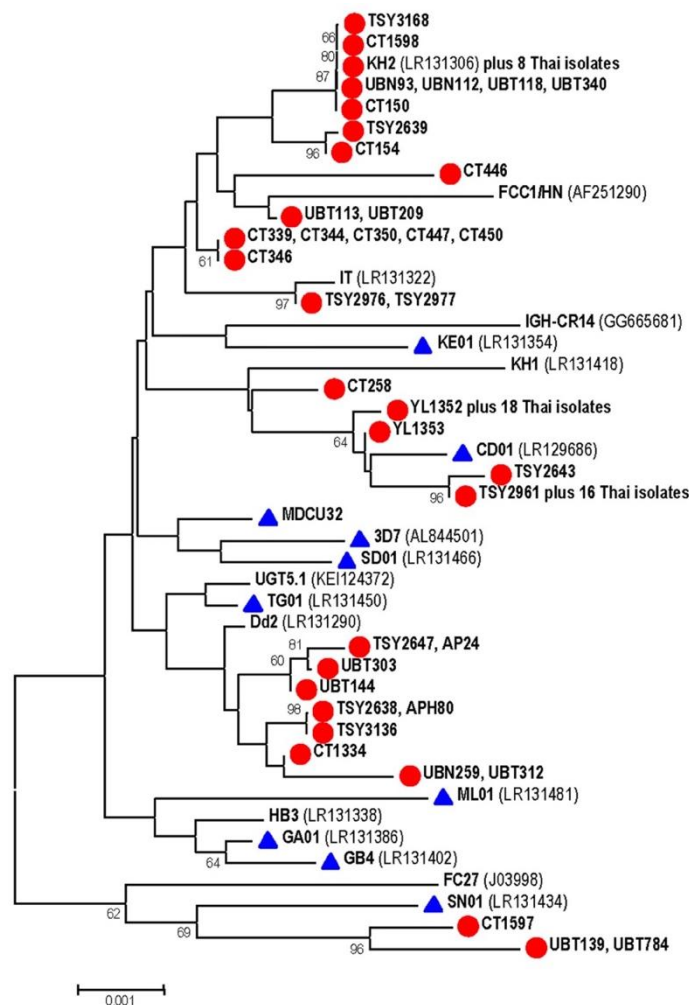
<b>Region#</b>	<b><i>n</i></b>	<b>dN ± S.E.</b>	<b>dS ± S.E.</b>	<b><i>p</i> value*</b>
All non-repeat blocks	99	0.00106 ± 0.00067	0.00208 ± 0.00202	0.492
Block 3	99	0.00860 ± 0.00380	0.00000 ± 0.00000	0.024

# Repeat regions were excluded from analysis. *n* = number of sequences, dN = the number of nonsynonymous substitutions per nonsynonymous site, dS = the number of synonymous substitutions per synonymous site, S.E. = standard error, \* *p* value < 0.05 is considered significance for Z-tests of the hypothesis that mean dS equals to mean dN.



#### 4.10 Phylogenetic Analysis

Phylogenetic tree inferred from *pfgarp* revealed closely related lineages of Thai isolates while taxa belonging to African isolates seemed to be widely distributed in the phylogenetic tree (Figure 23). Due to high sequence conservation in non-repeat regions, the topology of the tree was mainly affected by sequence variation in repeat domains I-III, V, and VI.



**Figure 23** Phylogenetic tree of *pfgarp* sequences from Thai and worldwide isolates constructed by Neighbor-joining model. Thai isolates with initials TSY and AP are from Tak, UB from Ubon Ratchathani, YL from Yala and CT from Chanthaburi Provinces. The numbers following these initials are used to label individual isolates. Bootstrap values greater than 60% are shown along the branches. Thai and African isolates are marked with circles and triangles, respectively. Scale denotes nucleotide substitutions per site.

#### 4.11 Genetic Differentiation of *P. falciparum* based on *pfgarp*

The geographic differentiation of *P. falciparum* population based on the PfGARP-encoding gene among four malaria-endemic areas in Thailand, including Chanthaburi, Tak, Ubon Ratchathani, and Yala Provinces was evaluated by using Arlequin software version 3.5.2.2 (Excoffier *et al.*, 2005). Analysis of molecular variance (AMOVA) to obtain variance indices among populations for calculation of fixation index ( $F_{ST}$ ). Interpopulation analysis by pairwise difference method showed that almost all  $F_{ST}$  values between pairs of parasite populations were significantly exceeded zero which indicated genetic differentiations between a pair of the parasite populations ( $p$  value  $< 0.05$ ). Meanwhile, the variance of the parasite populations between Tak and Chanthaburi were not statistically meaningful ( $p$  value = 0.009) (Table 21). It suggested that the genetic structure between the parasite populations of these two provinces seemed to occurred.

**Table 21** Pairwise differences in the  $F_{ST}$  values between populations.

	Tak	Ubon Ratchathani	Chanthaburi	Yala
Tak		0.0090*	0.0991	$<10^{-5}$ *
Ubon Ratchathani	0.1153		$<10^{-5}$ *	$<10^{-5}$ *
Chanthaburi	0.0359	0.0907		$<10^{-5}$ *
Yala	0.5526	0.5158	0.5158	

$F_{ST}$  indices and their respective  $p$  values are in lower and upper diagonals, respectively.

\* Test of hypothesis that the pairwise difference among the populations was not different from zero, significant level at  $p$  value  $< 0.05$ .

## 4.12 Association between The Number of Amino Acid and Parasite

### Density

It has been observed that specific number of repeat units in PfGARP has affected parasite's trafficking function of the protein (Davies *et al.*, 2016). Herein, association between the variations in the number of the length of RIII, RV, and E1-RVI-E2 of *pfgarp* and the parasite density were determined. Analysis was performed by Kruskal-Wallis H test implementing IBM SPSS Statistics (Version 28.0.0). The total number of isolates which was performed in this analysis was 76, including 19, 19, 18, and 20 isolates from Tak, Ubon Ratchathani, Chanthaburi and Yala, respectively. The parasitemia of those isolates ranged from 200 to 864,000 parasites per  $\mu\text{L}$ . The statistical analysis revealed the length of RIII and E1-RVI-E2 appear to be associated with the parasite density with  $p$  values = 0.011 and 0.028, respectively (Table 22). There was enough statistical evidence to support an association between RV and parasite density ( $p$  value = 0.098) (Table 18). This analysis omitted a categorical data in which the number of isolates was limited. The remaining repeats that have no variations in size also precluded from the analysis.

**Table 22** Statistical analysis of association between the parasite density and length polymorphism in repeat domains of *PfGARP*

Repeat domain	No. codons	No. isolates	Parasite density (parasites/ $\mu$ L)		Kruskall-Wallis $H$	$p$ value
			Range	Geometric mean		
RIII	30	20	1,081-17,225	4,348	11.080	0.011
	35	21	1,143-354,857	14,689		
	40	7	199-200,000	9,387		
	45	28	1,476-864,000	16,119		
RV*	21	42	1,081-200,000	7,164	4.641	0.098
	23	5	199-32,743	5,082		
	28	28	1,476-864,000	16,987		
E1-RVI-E2#	85-87	9	1,553-354,857	14,004	9.093	0.0281
	88	39	1,081-119,771	7,047		
	89	21	199-459,000	13,018		
	91-96	7	17,849-864,000	38,418		

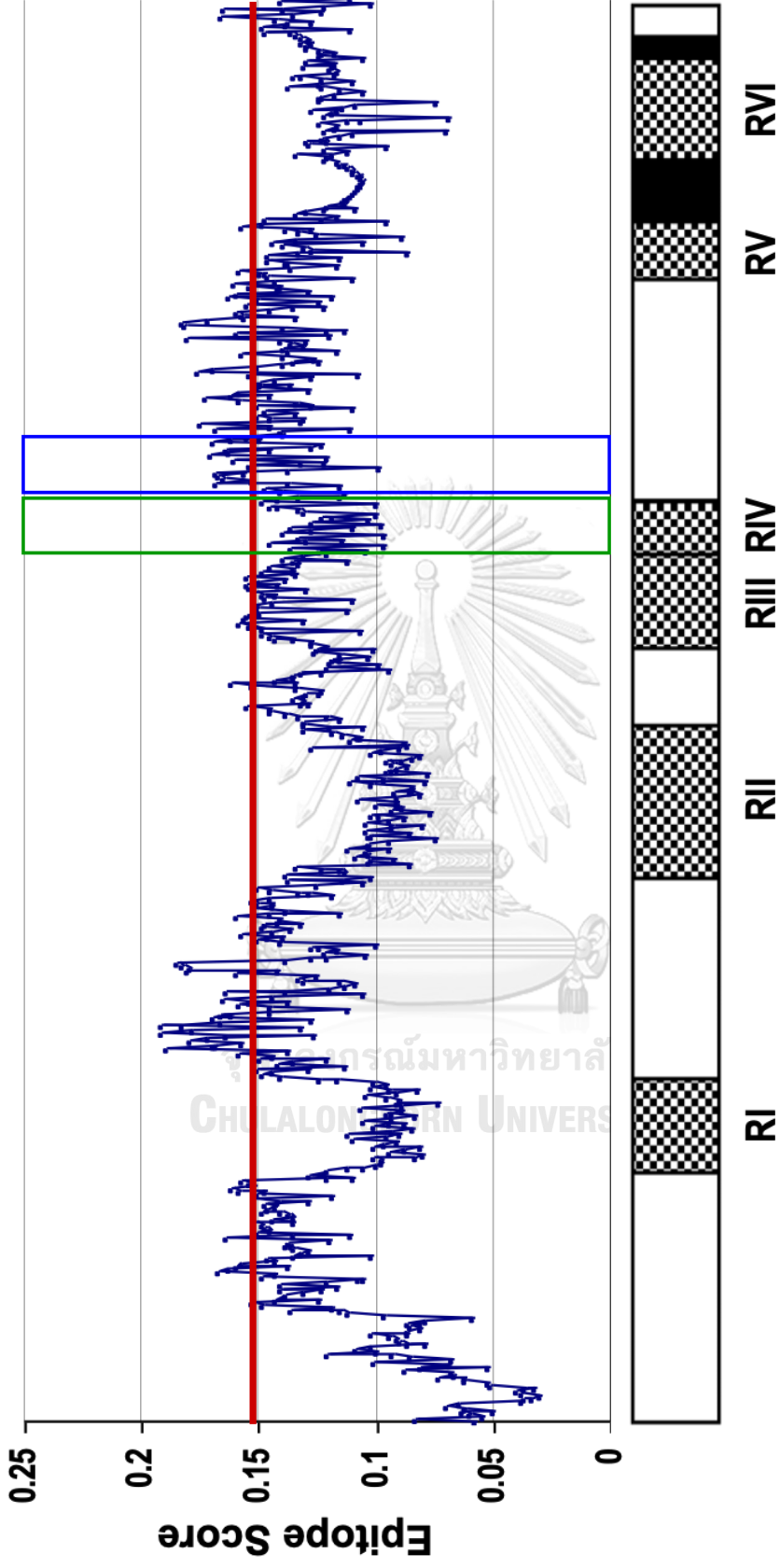
\* One isolate containing 35 codons was omitted from analysis. # Repeats with number of isolates < 5 were combined due to extensive length polymorphism in repeat domain VI. Domains RI, RII and RIV were excluded due to insufficient categorical data.

## 4.13 Antigenic Epitope Prediction

Although experimental evidence is required to document both B- and T-cell epitopes, *in silico* prediction could be helpful for further identification of actual epitopes in given antigen.

### 4.13.1 B-cell Epitope Prediction

All 44 distinct haplotype sequences of *pfgarp* were translated in amino acid sequences which further predicted linear B-cell epitopes. Linear B cell epitopes in PfGARP were predicted based on similarity of known epitope sequences implemented in BepiBlast web server (Ras-Carmona *et al.*, 2022) and protein language models implemented in BepiPred-3.0 (Clifford *et al.*, 2022). In total, nine B cell epitopes were predicted by the BepiBlast method, most of which spanned repeat domains. Three of these predicted epitopes, i.e., NDKENISE, KQKKIEKE and KKQEEKEK, were perfectly conserved across isolates whose sequences were similar to known epitopes in ankyrin repeat-containing protein of *Ehrlichia chaffeensis*, spike glycoprotein of severe acute respiratory syndrome coronavirus 1 and glutathione S-transferase isozyme of *Schistosoma mansoni*, respectively (Table 23). Repeat domain III contained four predicted epitopes that possessed sequence similarity either with genome polyprotein of dengue virus or M protein of *Streptococcus pyogenes*. Furthermore, two predicted epitopes were identified in repeat domain VI (Table 23). Meanwhile, prediction based on BepiPred-3.0 has identified linear B cell epitopes mostly in conserved blocks 3 and 8. All repeat-containing domains received epitope scores below the cut-off threshold by this method (Figure 24). The epitope score for monoclonal antibody mAb7899 capable of killing *P. falciparum* *in vitro* (Raj *et al.*, 2020) was remarkably above the epitope threshold (the N-terminal part of block 8) albeit the sequence did not share similarity to known epitopes based on the BepiBlast method (Table 23 and Figure 24).



**Figure 24** Linear B cell epitope scores across PfGARP predicted by BepiPred version 3.0. Blocks are after Fig. 19. Red line represents epitope threshold. Green and blue boxes indicate ligand for erythrocyte band 3 and epitope for mAb7899, respectively.



**Table 23** Predicted linear B cell epitopes spanning 8 amino acids in PfGARP and their distribution among variant alleles

No.	Epitope	Block*	Known epitope (IEDB ID)#	Similarity of known epitope	Prevalence among Thai isolates, n = 80 (%)	Prevalence among non-Thai isolates, n = 19 (%)
1	NDKENISE	3	EPDLEEIVSILKNDKEGISE (119567)	Ankyrin repeat-containing protein of <i>Ehrlichia chaffeensis</i> #	100	100
2	KQKKIEKE	4 (RII)	ESKQKKIENEIA (1429913)	Spike glycoprotein chain A of severe acute respiratory syndrome coronavirus 1#	100	100
3	KKQEEKEK	4 (RII)	KPQEEKEKITKEILNGK (32844)	Glutathione S-transferase class-mu 28 Kda isozyme of <i>Schistosoma mansoni</i> #	100	100
4	HKEGEHKK	6 (RIII)	VTNHMEGEHKLAEA (1642389) and NEEMVTNHMEGEHKK (1640001)	Genome polyprotein of dengue virus#	12.5	5.26
5	EGEHKEGE	6 (RIII)	LEGEWKEGEEVQVLA (1639118) and GGWKLEGEWKEGEEV (1637720)	Genome polyprotein of dengue virus#	15	63.16
6	EGEDKEGE	6 (RIII)	LEGEWKEGEEVQVLA (1639118) and GGWKLEGEWKEGEEV (1637720)	Genome polyprotein of dengue virus#	21.25	10.53
7	EEEHKKEE	6 (RIII)	LFEKLDKVEEHKKVE (1465890)	M Protein of <i>Streptococcus pyogenes</i> serotype 2.1	15	36.84
8	DEEDEDDA	11 (RVI)	MYCSFYPPDEEEEDDA (1680661)	Orf1Ab polyprotein (Pp1Ab) of severe acute respiratory syndrome-related coronavirus Tor2#	-	5.26
9	AEEDEDDD	11 (RVI)	AEEEEDDDDMGFGLFD (876)	Ribosomal protein P-J15 of <i>Trypanosoma cruzi</i> #	2.5	-

\* After Fig. 19. # Immune epitope database and analysis resource identity document.

#### 4.13.2 T-Cell Epitope Prediction

Antigenic peptide prediction for CD4+ T-cell receptor (TCR) was evaluated by IEDB web server. The platform for prediction was adjusted following the default option. The reference sequence of PfGARP (GenBank accession number J03998) was used for preliminary analysis. Due to a wide repertoire of HLA class II alleles, analysis has been configured to 9 common HLA class II alleles (allele frequency >10%) among Thai population, including HLA-DRB1\*12:02, \*15:02, -DQA1\*01:01, \*01:02, \*03:02, \*06:01, -DQB1\*03:03, \*05:01, \*05:02 (Satapornpong *et al.*, 2020). The results showed three regions of PfGARP that were recognized by specific alleles. The first region (commencing at amino acid 1) was predicted to be recognized by four different HLA supertypes, including HLA-DQA1\*01:01, \*01:02, -DQB1\*05:01, and -DRB1\*15:02, (ii) HLA-DQA1\*01:01, \*01:02, \*03:02, and \*06:01 (commencing at amino acid 199) and (iii) HLA-DQA1\*01:02, and -DQB1\*03:03 (from amino acid 351) (Figure 25). The first and the third predicted areas located in conserved region while the second spanned in an area where amino acid substitutions have been observed, containing four peptide variants: SEINNNAQGGLLLSSPYQY, SEINNNAQGGLLLSSLYQY, SEINNNAQGGLLLSSPDQY, and SEINNNAQGGLLLSSPYQC. Amino acid substitutions occurred at positions 213, 214, and 216, all of which were dimorphic. Amino acid sequences, J03998, LR131322, LR131481, and TSY2643 were chosen as representative peptide variants 1, 2, 3, and 4, respectively for analysis with those nine allele supertypes mentioned above. The overall results displayed five HLA-DQ supertypes that could recognize four different haplotypes (Table 24). Among four different variants, nine predicted epitopes (designed as 1a, 1b, 2a, 2b, 2c, 3a, 3b, 3c, and 4a) were recognized by T-cell receptors in different HLA haplotypes (Table 25). All of them had immediate to low binding affinity (Table 25). The number in each epitope's name indicate the corresponding haplotypes, epitopes were further subdivided by adding an alphabet following the number to indicate variants. The predicted results have shown that only HLA-DQA1\*03:02 and \*06:01 could potentially bind all variants. Interestingly, not only amino acid replacement contributed to different HLA-D allele binding but also the epitope size difference. For instance, epitope 1b was recognized by three HLA-D supertypes, more amino acid residue (epitope 1a) have led to different supertype recognition. Length spectrum of the epitopes of the variants 2 and 3 was also given the similar pattern. Remarkably, 4a epitope found only in Thai isolates.





**Table 24** Overall HLA-D supertypes recognize different variants in the middle-predicted region of PfGARP

No.	HLA-D supertypes	Variant 1	Variant 2	Variant 3	Variant 4
1	DQA1* 01:01	*	-	-	-
2	DQA1* 01:02	*	*	*	-
3	DQA1* 03:02	*	*	*	*
4	DQA1* 06:01	*	*	*	*
5	DQB1* 03:03	-	*	-	-
6	DQB1* 05:01	-	-	-	-
7	DQB1* 05:02	-	-	-	-
8	DRB1* 12:02	-	-	-	-
9	DRB1* 15:02	-	-	-	-

Symbols "\*" and "-" represent recognized and non-recognized regions by HLA-D supertype, respectively.

Table 25 Predicted HLA-D allele binding peptides in PfGARP

Epitopes	Predicted Peptides	HLA allele pair		PR	NetMHCIIp an_IC <sub>50</sub>	Isolates (n = 99)
		DQA1*	DQB1*			
1b	SEINNAQGGLLSSPYQY	0101	X   X ∈ {0301, 0309, 0310, 0313, 0316, 0319, 0321, 0322, 0324, 0327, 0328, 0329, 0335, 0336, 0601, 0635, 0643}	3.7-9.2	321-895	GenBank (15), Thai (62)
		0102	X ∪ {03.12, 06.02, 06.11, 06.16, 06.19, 06.24, 06.33}	4.9-9.8	169-613	
		0601	X	7.4-9.9	162-724	
1a	SEINNAQGGLLSSPYQYR	0302	X ∪ {03.04, 03.12, 03.14}	6.1-10	122-679	
2a	SEINNAQGGLLSSLYQYR	0102	X ∪ {0303, 0304, 0305, 0308, 0312, 0314, 0315, 0320, 0323, 0326, 0330, 0331, 0333, 0334, 0403, 06.02, 0611, 0616, 0619, 0623, 0624, 0629, 0633}	3-10	148-932	GenBank (1)
		0302	X ∪ {0304, 0312, 0314}	4.6-10	112-677	
2b	SEINNAQGGLLSSLYQY	0601	X ∪ {0304, 0312, 0314}	6.1-9.9	136-621	
2c	INNAQGGLLSSLYQ	0102, 0106, 0108, 0109, 0501, 0503, 0504, 0505, 0506, 0507, 0508, 0509, 0510, 0511	0303	9-10	253-932	
3a	SEINNAQGGLLSSPDQY	0102	X ∪ {06.02, 06.16, 06.19}	6.3-9.9	187-544	GenBank (2), MDCU32
3b	SEINNAQGGLLSSPDQ	0302	X	7.7-9.9	139-617	
3c	SEINNAQGGLLSSPD	0601	0313	8.8-9.9	189-215	
4a	SEINNAQGGLLSSPYQC	0302	X ∪ {0304, 0312, 0314}	6.1-10	122-679	Thai (18)
		0601	X	7.4-10	162-698	

HLA = human leukocyte antigen, NetMHCIIpan\_IC<sub>50</sub> = the 50% maximal binding concentration calculated from NetMHCIIpan prediction method, PR = percentile rank, X = HLA-DQB1 allele

## CHAPTER 5

### DISCUSSION AND CONCLUSIONS

#### 5.1 Discussion and Conclusions

PfGARP was firstly identified by [Triglia \*et al.\* \(1998\)](#). Studies demonstrated that this protein expresses on cell membrane of infected erythrocyte in early to late trophozoite stages of *P. falciparum* ([Davies \*et al.\*, 2016](#); [Raj \*et al.\*, 2020](#)). It was recognized by human erythrocyte band 3 protein receptor on iRBC membrane, which might lead to rosetting-like formation ([Maier \*et al.\*, 2008](#); [Davies \*et al.\*, 2016](#)). A recent study by Raj and colleagues (2020) suggested that anti-PfGARP antibody has partial protective effect against falciparum malaria trialed in non-human primates. Additionally, it may have the potential to be synergized with other protein subunits of *P. falciparum* for blood-stage malaria vaccines that target the parasite invasion. Confrontation with high genetic variations of many parasitic antigens mainly affects the effectiveness of the malaria vaccines. Thus, genetic diversity in the PfGARP-encoding gene should be primarily considered before a vaccine design.

The present study was designed to evaluate the genetic diversity in *pfgarp* of *P. falciparum* clinical isolates from different malaria-endemic areas in Thailand (Chanthaburi, Tak, Ubon Ratchathani, and Yala Provinces) by direct sequencing. Sequence analysis *pfgarp* revealed high conservation among Thai isolates. Although this gene was highly conserved, some variations were spotted, and those variation were limited to exon II of the *pfgarp* sequences. The size variations of the *pfgarp* were difficultly observed by gel electrophoresis, while the sequence data obtained from direct sequencing clearly revealed polymorphism in this locus. Previous study of [Fandeur and colleagues \(1996\)](#) considered *pfgarp* as conserved locus by size. They genotyped the gene of 15 parasite strains and then visualized the size by migration of PCR products on agarose gels ([Fandeur \*et al.\*, 1996](#)). This outcome might be resulted from priming a small region covered only exon I and intron where length variation was limited. This study indicated that direct sequencing gave more sensitive results than agarose gel electrophoresis.

Comparison of the nucleotide diversity and the number of haplotypes of *pfgarp* across malaria endemic areas in Thailand revealed the lowest values of these two indicators observing in Yala. Previous studies of the genetic diversity in *P. vivax* genes, apical membrane protein 1 (AMA1), merozoite surface protein 1 (MSP1), MSP4, and MSP5 indicated a lower number of haplotype and nucleotide diversity in Yala than that of Tak (Jongwutiwes *et al.*, 2010). However, population structure of the parasite populations in Thailand based on *pfgarp* was significantly distinct except the parasite populations between Tak and Chanthaburi Provinces. The results were evidenced by  $F_{ST}$  values in which the average pairwise difference of this pair was not significantly different from zero ( $p$  value = 0.099). This finding might be explained from a massive migration of malaria patients who routinely traveled between these two endemic areas for gem trading during the past 1980 to 1990 (Thimasarn *et al.*, 1995). Suggesting that gene flow between these two parasite populations may generate existing gene variants across provinces.

Haplotype diversity seems to be mainly contributed by variations in repetitive regions rather than those from the SNPs in non-repeat regions. Genetic mechanisms drove *pfgarp* to generate variants in several repeats and some of them vary in size among isolates. Two additional repeats were newly identified (RII and RV). More sequences analyzed in this study, the number of repetitive units of RI, RIII, and RVI was found wider spectra than those previously observed (Trigilia *et al.*, 1998). This finding could be potentially explained by slippage mispairing in DNA replication process of mitotic division, resulting in expanding or contracting repetitive sequences. This event is commonly a dynamic process in DNA replication episode of organisms and also occur in repeat sequences of several malaria parasite genes (Jongwutiwes *et al.*, 1994; Jongwutiwes *et al.*, 2008; Putaporntip *et al.*, 2008; Putaporntip *et al.*, 2013; Putaporntip *et al.*, 2022). Interestingly, four pairs of palindromic sequence were noticed flanking 241-242delAT in an intron. This finding may be the biological process evidence of the deletion or insertion of AT in this region. No evidence of a recombination event was found, then this finding might not be explained by unequal crossing over of chromosomes in meiosis. Phylogenetic tree analysis based on this gene showed a related relationship among most Thai isolates while a few isolates were spotted outside of most Thai lineage. Co-lineage between Thai isolates and non-Thai isolates seems to be resulted from accidentally generating some similar alleles across the parasite populations. Meanwhile, concerted evolution seems to be a potential biological process that



drives the similarity of the repeat sequences within parasite populations. Repetitive DNA sequences are likely to be disturbed by many complex processes (i.e., DNA replication, repair, transcription) which prone to break the sequence stability, especially in homopolymeric repeats. The length of homopolymeric glutamic acid repeats in domain E1 containing stretches of GAA interrupted by GAG was approximately three times longer than the perfect GAA repeats in domain E2. It seemed likely that long perfect triplet repeats encoding the same amino acids could be affected by structural instability at the DNA level unless they were interrupted by another triplet encoding the same amino acid as previously described (Dorsman *et al.*, 2002; Brazda *et al.*, 2020; Brown & Freudenreich, 2021).

Several factors contribute to the virulence of the *Plasmodium* parasite including genetic diversity in the antigenic proteins. A recent study unveiled a variable variant of *var2csa*, a virulent protein of *P. falciparum*, inversely associated with placental parasitemia in falciparum-infected pregnant women (Talundzic *et al.*, 2022). Noteworthy, the present study revealed that the length of RIII and RVI appear to be associated with the parasite density (Table 22). However, to reaffirm this finding, more samples may be required to improve statistical power. This finding raises the possibility that if the repetitive size of those repeats is modified, the protein might be proceeded and behaved in a similar way in the experiment of Davies and colleagues (2016), truncation of lysine-rich repeat of PfGARP affect the protein localization on iRBC membrane. This in turn way affect the parasite carrying a truncated repeat of this protein by alteration of the parasite aggregation and parasite load. Importantly, rosette formation has been suggested to be responsible for increasing parasitemia in *P. falciparum* infections (Rowe *et al.*, 2002). Moreover, this protein might link to the virulence of the parasite as like have been found in *Entamoeba sp.* that its lysine and glutamic acid repeat-containing protein served as a virulent factor (Weedall, 2020). Expanding or contracting of the repetitive regions of *pfgarp* might affect the protein's functionality or pathogenesis process. However, the relationship between the parasite density and the repeat length of *pfgarp* might be driven by other unidentified factors. Hence, suggesting further study might focus on the experimental design for functional clarification of this protein

and its influence on the parasite growth. However, analysis of the parasite density among Thai, Cambodians, and Karens revealed no statistical difference. Similar results were also found when analysis concerned *msp1* genotypes and gender of patients. These findings likely suggested that the ethnicity, parasite strain, and gender might not contribute to the parasite density. However, further research for the clarification may be needed.

Appropriate selection of epitopes and variants is considerably important for malaria vaccine development. Some studies revealed that restricted antigenic diversity in blood-stage proteins of *P. falciparum* supports an effective malaria vaccine development. A specific conserved region of EBA-175 elicits an immune system that potentially inhibits the parasite invasion across several parasite strains (Healer *et al.*, 2013). Furthermore, another study revealed that even though high polymorphism was investigated, antigenic diversity of a leading vaccine candidate antigen, AMA1, was restricted suggesting only a few alleles are feasible to produce a vaccine against the naturally circulating parasite (Terheggen *et al.*, 2014). It was not until recently that the significance of PfGARP has been unveiled as an important target for host antibody responses capable of protecting African children with falciparum malaria from high parasitemia and severe symptoms (Raj *et al.*, 2020). It has been shown that anti-PfGARP antibodies conferred parasite killing through the induction of programmed cell death as evidenced by the activation of caspase-like proteases and the fragmentation of parasite DNA of late trophozoites and schizonts. Importantly, the epitope for mAB7899 antibody conferring parasite killing *in vitro* has been mapped to a perfectly conserved block of the protein in which high scores for linear B cell epitopes were predicted in this region by the BepiPred 3 algorithm (Figure 24). Although additional B cell epitopes await further investigations, little evidence in naturally acquired immunity against this molecule was found to be recognized by human and mouse antibodies (Almukadi *et al.*, 2019). Prediction of linear B cell epitopes by sequence similarity with known epitopes implemented in the BepiBlast web server has identified nine potential linear B cell epitopes spanning eight amino acids, eight of these predicted epitopes were found in repeat domains and most of which exhibited sequence variation among isolates (epitopes nos. 4-9 in Table 23). Intriguingly, sequence variation in repeat domains RIII and RVI could probably be influenced by host immune pressure.

T-cell predicted epitopes were also identified in block 3 where four peptide variants were observed (Table 25). Alteration of amino acid residue and size of the epitopes contributed to the recognition of different HLA-D alleles. These results may indicate that substitution in this region might be profitable for the parasite to avoid the host's defensive immune system. Among those four peptide variants of T cell epitope, only two peptide variants were potentially recognized by common HLA class II in Thai population. This finding suggested that vaccines based on a few heterogenous variants of PfGARP might be sufficient for coverage of major circulated strains of the parasite in Thailand. Although the parasite-infected RBC has not been clearly clarified to induce a cell-mediated immune system since RBCs not express HLA class II, depletion of CD8<sup>+</sup> and CD4<sup>+</sup> T-cells impeded the parasitic clearance in *P. chabaudi* infected mice, suggesting that those T-cells may have potential role for blood-stage malaria (Podoba & Stevenson, 1991). Besides, identification of T-cell epitopes of other blood-stage antigens, e.g., AMA1, EBA-175, Rh5, etc. have been evaluated immune accessibility beyond *in silico* experiment (Kotraiah *et al.*, 2021). Meanwhile, a recent immunoinformatic and structural approach have suggested that a vaccine construct derived from PfGARP was predicted to induce both humoral and cellular immune responses (Atapour *et al.*, 2022). However, its functionality is underway to evaluate *in vitro* and *in vivo*. Taken together, these findings might be important for future work on the development of antibody-based therapies and vaccine design against blood stage of falciparum malaria. Above all else, prediction of the functional epitope in PfGARP is required for further research.

In conclusion, this thesis has provided the genetic information of the PfGARP-encoding gene of *P. falciparum* from different malaria-endemic areas in Thailand. Albeit global malaria cases have been reduced by concerted efforts, polymorphism in malaria vaccine candidate proteins has been raising concern about the impact the vaccine efficacy. With limited number of isolates analyzed, it seemed that expansion or reduction of lysine-rich and glutamic acid-rich repeat regions seemed to influence parasite density of malaria patients. Based on limited diversity in *pfgarp* observed in this study and predicted immunogenic epitope regions, it is plausible that PfGARP-derived vaccines from conserved and less polymorphic domains could be sufficient to affect outbred parasite populations.

## REFERENCES

- Aikawa, M., Miller, L. H., Johnson, J., & Rabbege, J. (1978). Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J Cell Biol*, 77(1), 72-82.
- al-Yaman, F., Genton, B., Kramer, K. J., Chang, S. P., Hui, G. S., Baisor, M., & Alpers, M. P. (1996). Assessment of the role of naturally acquired antibody levels to *Plasmodium falciparum* merozoite surface protein-1 in protecting Papua New Guinean children from malaria morbidity. *Am J Trop Med Hyg*, 54(5), 443-448.
- Almukadi, H., Schwake, C., Kaiser, M. M., Mayer, D. C. G., Schiemer, J., Baldwin, M. R., Hegde, S., Lu, Y., Hanada, T., & Chishti, A. H. (2019). Human erythrocyte band 3 is a host receptor for *Plasmodium falciparum* glutamic acid-rich protein. *Blood*, 133(5), 470-480
- Amino, R., Thiberge, S., Martin, B., Celli, S., Shorte, S., Frischknecht, F., & Menard, R. (2006). Quantitative imaging of *Plasmodium* transmission from mosquito to mammal. *Nat Med*, 12(2), 220-224.
- Atapour, A., Vosough, P., Jafari, S., & Sarab, G. A. (2022). A multi-epitope vaccine designed against blood-stage of malaria: an immunoinformatic and structural approach. *Sci Rep*, 12(1), 11683.
- Balaji, S. N., & Trivedi, V. (2013). Extracellular methemoglobin primes red blood cell aggregation in malaria: an in vitro mechanistic study. *FEBS Lett*, 587(4), 350-357.
- Bennink, S., Kiesow, M. J., & Pradel, G. (2016). The development of malaria parasites in the mosquito midgut. *Cell Microbiol*, 18(7), 905-918.
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*, 27(2), 573-580.
- Brazda, V., Fojta, M., & Bowater, R. P. (2020). Structures and stability of simple DNA repeats from bacteria. *Biochem J*, 477(2), 325-339.
- Brown, R. E., & Freudenreich, C. H. (2021). Structure-forming repeats and their impact on genome stability. *Curr Opin Genet Dev*, 67, 41-51.
- Buppan, P., Seethamchai, S., Kuamsab, N., Harnyuttanakorn, P., Putaporntip, C., & Jongwutiwes, S. (2018). Multiple novel mutations in *Plasmodium falciparum* chloroquine resistance transporter gene during implementation of artemisinin combination therapy in Thailand. *Am*

*J Trop Med Hyg*, 99(4), 987-994.

CDC. (2020). Malaria. Retrieved from <https://www.cdc.gov/dpdx/malaria/>

Chatterjee, D., & Cockburn, I. A. (2021). The challenges of a circumsporozoite protein-based malaria vaccine. *Expert Rev Vaccines*, 20(2), 113-125.

Chaudhury, S., MacGill, R. S., Early, A. M., Bolton, J. S., King, C. R., Locke, E., Pierson, T., Wirth, D. F., Neafsey, D. E., & Bergmann-Leitner, E. S. (2021). Breadth of humoral immune responses to the C-terminus of the circumsporozoite protein is associated with protective efficacy induced by the RTS,S malaria vaccine. *Vaccine*, 39(6), 968-975.

Chitnis, C. E., & Sharma, A. (2008). Targeting the *Plasmodium vivax* Duffy-binding protein. *Trends Parasitol*, 24(1), 29-34.

Clark, D. P., & Pazdernik, N. J. (2012). *Molecular biology*: Elsevier.

Clifford, J. N., Hoie, M. H., Deleuran, S., Peters, B., Nielsen, M., & Marcatili, P. (2022). BepiPred-3.0: Improved B-cell epitope prediction using protein language models. *Protein Sci*, 31(12), e4497.

Clyde, D. F., Most, H., McCarthy, V. C., & Vanderberg, J. P. (1973). Immunization of man against sporozoite-induced falciparum malaria. *Am J Med Sci*, 266(3), 169-177.

Cohen, J., Nussenzweig, V., Nussenzweig, R., Vekemans, J., & Leach, A. (2010). From the circumsporozoite protein to the RTS, S/AS candidate vaccine. *Hum Vaccin*, 6(1), 90-96.

Dame, J. B., Williams, J. L., McCutchan, T. F., Weber, J. L., Wirtz, R. A., Hockmeyer, W. T., Maloy, W. L., Haynes, J. D., Schneider, I., Roberts, D., & et al. (1984). Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. *Science*, 225(4662), 593-599.

Davies, H. M., Thalassinou, K., & Osborne, A. R. (2016). Expansion of lysine-rich repeats in *Plasmodium* proteins generates novel localization sequences that target the periphery of the host erythrocyte. *J Biol Chem*, 291(50), 26188-26207.

Dondorp, A. M., Kager, P. A., Vreeken, J., & White, N. J. (2000). Abnormal blood flow and red blood cell deformability in severe malaria. *Parasitol Today*, 16(6), 228-232.

Dorsman, J. C., Bremmer-Bout, M., Pepers, B., van Ommen, G. J., & Den Dunnen, J. T. (2002). Interruption of perfect CAG repeats by CAA triplets improves the stability of glutamine-encoding repeat sequences. *Biotechniques*, 33(5), 976-978.

- Duffy, P. E., & Patrick Gorres, J. (2020). Malaria vaccines since 2000: progress, priorities, products. *NPJ Vaccines*, 5(1), 48.
- Excoffier, L. (2005). The detection of regions of our genome under selection has increasingly relied on the use of genome scans. *Hum Genomics*, 2(3), 155-157.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*, 10(3), 564-567.
- Fandeur, T., Mercereau-Puijalon, O., & Bonnemains, B. (1996). *Plasmodium falciparum*: genetic diversity of several strains infectious for the squirrel monkey (*Saimiri sciureus*). *Exp Parasitol*, 84(1), 1-15.
- Ferraguti, M., Martinez-de la Puente, J., Roiz, D., Ruiz, S., Soriguer, R., & Figuerola, J. (2016). Effects of landscape anthropization on mosquito community composition and abundance. *Sci Rep*, 6, 29002.
- Fong, Y. L., Cadigan, F. C., & Coatney, G. R. (1971). A presumptive case of naturally occurring *Plasmodium knowlesi* malaria in man in Malaysia. *Trans R Soc Trop Med Hyg*, 65(6), 839-840.
- Futuyma, D. J. (2006). *Evolution* (third ed.). United Kingdom: Oxford University Press.
- Garibyan, L., & Avashia, N. (2013). Polymerase chain reaction. *J Invest Dermatol*, 133(3), 1-4.
- Garnham, P. C. C. (1966). *Malaria parasites and other Haemosporidia*. *Malaria parasites and other Haemosporidia*. Oxford: Blackwell Scientific Publications.
- Genton, B., Betuela, I., Felger, I., Al-Yaman, F., Anders, R. F., Saul, A., Rare, L., Baisor, M., Lorry, K., Brown, G. V., Pye, D., Irving, D. O., Smith, T. A., Beck, H. P., & Alpers, M. P. (2002). A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea. *J Infect Dis*, 185(6), 820-827.
- Ghosh, A., Mehta, A., & Khan, A. M. (2018). Metagenomic analysis and its applications.
- Goel, V. K., Li, X., Chen, H., Liu, S. C., Chishti, A. H., & Oh, S. S. (2003). Band 3 is a host receptor binding merozoite surface protein 1 during the *Plasmodium falciparum* invasion of erythrocytes. *Proc Natl Acad Sci U S A*, 100(9), 5164-5169.
- Guerra, C. A., Gikandi, P. W., Tatem, A. J., Noor, A. M., Smith, D. L., Hay, S. I., & Snow, R. W. (2008). The limits and intensity of *Plasmodium falciparum* transmission: implications for

- malaria control and elimination worldwide. *PLoS Med*, 5(2), e38.
- Guevara Patino, J. A., Holder, A. A., McBride, J. S., & Blackman, M. J. (1997). Antibodies that inhibit malaria merozoite surface protein-1 processing and erythrocyte invasion are blocked by naturally acquired human antibodies. *J Exp Med*, 186(10), 1689-1699.
- Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., & Troedsson, C. (2014). Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS One*, 9(2), e87624.
- Haldar, K., Bhattacharjee, S., & Safeukui, I. (2018). Drug resistance in *Plasmodium*. *Nat Rev Microbiol*, 16(3), 156-170.
- Haqqi, T. M., Sarkar, G., David, C. S., & Sommer, S. S. (1988). Specific amplification with PCR of a refractory segment of genomic DNA. *Nucleic Acids Res*, 16(24), 11844.
- Harl, J., Himmel, T., Valkiunas, G., & Weissenbock, H. (2019). The nuclear 18S ribosomal DNAs of avian haemosporidian parasites. *Malar J*, 18(1), 305.
- Healer, J., Thompson, J. K., Riglar, D. T., Wilson, D. W., Chiu, Y. H., Miura, K., Chen, L., Hodder, A. N., Long, C. A., Hansen, D. S., Baum, J., & Cowman, A. F. (2013). Vaccination with conserved regions of erythrocyte-binding antigens induces neutralizing antibodies against multiple strains of *Plasmodium falciparum*. *PLoS One*, 8(9), e72504.
- Hill, A. V. (2011). Vaccines against malaria. *Philos Trans R Soc Lond B Biol Sci*, 366(1579), 2806-2814.
- Hirunpetcharat, C., Tian, J. H., Kaslow, D. C., van Rooijen, N., Kumar, S., Berzofsky, J. A., Miller, L. H., & Good, M. F. (1997). Complete protective immunity induced in mice by immunization with the 19-kilodalton carboxyl-terminal fragment of the merozoite surface protein-1 (MSP1[19]) of *Plasmodium yoelii* expressed in *Saccharomyces cerevisiae*: correlation of protection with antigen-specific antibody titer, but not with effector CD4<sup>+</sup> T cells. *J Immunol*, 159(7), 3400-3411.
- Hoffman, S. L., Goh, L. M., Luke, T. C., Schneider, I., Le, T. P., Doolan, D. L., Sacci, J., de la Vega, P., Dowler, M., Paul, C., Gordon, D. M., Stoute, J. A., Church, L. W., Sedegah, M., Heppner, D. G., Ballou, W. R., & Richie, T. L. (2002). Protection of humans against malaria by immunization with radiation-attenuated *Plasmodium falciparum* sporozoites. *J Infect Dis*, 185(8), 1155-1164.

- Hughes, A. L. (2000). *Adaptive evolution of genes and genomes*: Oxford University Press, USA.
- Jaskiewicz, E., Jodłowska, M., Kaczmarek, R., & Zerka, A. (2019). Erythrocyte glycophorins as receptors for *Plasmodium* merozoites. *Parasit Vectors*, *12*(1), 317.
- Jaureguiberry, G., Hatin, I., d'Auriol, L., & Galibert, G. (1990). PCR detection of *Plasmodium falciparum* by oligonucleotide probes. *Mol Cell Probes*, *4*(5), 409-414.
- Jongwutiwes, S., Buppan, P., Kosuvin, R., Seethamchai, S., Pattanawong, U., Sirichaisinthop, J., & Putaporntip, C. (2011). *Plasmodium knowlesi* Malaria in humans and macaques, Thailand. *Emerg Infect Dis*, *17*(10), 1799-1806.
- Jongwutiwes, S., Putaporntip, C., & Hughes, A. L. (2010). Bottleneck effects on vaccine-candidate antigen diversity of malaria parasites in Thailand. *Vaccine*, *28*(18), 3112-3117.
- Jongwutiwes, S., Putaporntip, C., Karnchaisri, K., Seethamchai, S., Hongsriruang, T., & Kanbara, H. (2008). Positive selection on the *Plasmodium falciparum* sporozoite threonine-asparagine-rich protein: analysis of isolates mainly from low endemic areas. *Gene*, *410*(1), 139-146.
- Jongwutiwes, S., Tanabe, K., Hughes, M. K., Kanbara, H., & Hughes, A. L. (1994). Allelic variation in the circumsporozoite protein of *Plasmodium falciparum* from Thai field isolates. *Am J Trop Med Hyg*, *51*(5), 659-668.
- Jongwutiwes, S., Tanabe, K., Nakazawa, S., Yanagi, T., & Kanbara, H. (1992). Sequence variation in the tripeptide repeats and T cell epitopes in P190 (MSA-1) of *Plasmodium falciparum* from field isolates. *Mol Biochem Parasitol*, *51*(1), 81-89.
- Joshi, M., & Deshpande, J. D. (2010). Polymerase chain reaction: methods, principles and application. *Int J Biomed Res*, *1*(2), 81-97.
- Kariuki, M. M., Li, X., Yamodo, I., Chishti, A. H., & Oh, S. S. (2005). Two *Plasmodium falciparum* merozoite proteins binding to erythrocyte band 3 form a direct complex. *Biochem Biophys Res Commun*, *338*(4), 1690-1695.
- Kaslow, D. C. (2002). Transmission-blocking vaccines. *Chem Immunol*, *80*, 287-307.
- Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H., & Frost, S. D. (2006). GARD: a genetic algorithm for recombination detection. *Bioinformatics*, *22*(24), 3096-3098.
- Kotraiah, V., Phares, T. W., Terry, F. E., Hindocha, P., Silk, S. E., Nielsen, C. M., Moise, L., Tucker, K. D., Ashfield, R., Martin, W. D., De Groot, A. S., Draper, S. J., Gutierrez, G. M.,



- & Noe, A. R. (2021). Identification and immune assessment of T cell epitopes in five *Plasmodium falciparum* blood stage antigens to facilitate vaccine candidate selection and optimization. *Front Immunol*, *12*, 690348.
- Kumar, K. A., Sano, G., Boscardin, S., Nussenzweig, R. S., Nussenzweig, M. C., Zavala, F., & Nussenzweig, V. (2006). The circumsporozoite protein is an immunodominant protective antigen in irradiated sporozoites. *Nature*, *444*(7121), 937-940.
- Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., Dubremetz, J. F., Fauquenoy, S., Tomavo, S., Faber, B. W., Kocken, C. H., Thomas, A. W., Boulanger, M. J., Bentley, G. A., & Lebrun, M. (2011). The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathog*, *7*(2), e1001276.
- Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451-1452.
- Maier, A. G., Rug, M., O'Neill, M. T., Brown, M., Chakravorty, S., Szeszak, T., Chesson, J., Wu, Y., Hughes, K., Coppel, R. L., Newbold, C., Beeson, J. G., Craig, A., Crabb, B. S., & Cowman, A. F. (2008). Exported proteins required for virulence and rigidity of *Plasmodium falciparum*-infected human erythrocytes. *Cell*, *134*(1), 48-61.
- Martinez-Calvillo, S., Florencio-Martinez, L. E., & Nepomuceno-Mejia, T. (2019). Nucleolar Structure and Function in Trypanosomatid Protozoa. *Cells*, *8*(5).
- Mendis, K., Sina, B. J., Marchesini, P., & Carter, R. (2001). The neglected burden of *Plasmodium vivax* malaria. *Am J Trop Med Hyg*, *64*(1-2 Suppl), 97-106.
- Miller, L. H., Roberts, T., Shahabuddin, M., & McCutchan, T. F. (1993). Analysis of sequence diversity in the *Plasmodium falciparum* merozoite surface protein-1 (MSP-1). *Mol Biochem Parasitol*, *59*(1), 1-14.
- Mitreva, M. (2017). The Microbiome in Infectious Diseases. In *Infectious Diseases, 2-Volume Set* (pp. 68-74.e62): Elsevier.
- Moyes, C. L., Athinya, D. K., Seethaler, T., Battle, K. E., Sinka, M., Hadi, M. P., Hemingway, J., Coleman, M., & Hancock, P. A. (2020). Evaluating insecticide resistance across African districts to aid malaria control decisions. *Proc Natl Acad Sci U S A*, *117*(36), 22042-22050.
- Nei, M., & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and

- nonsynonymous nucleotide substitutions. *Mol Biol Evol*, 3(5), 418-426.
- Nielsen, C. M., Vekemans, J., Lievens, M., Kester, K. E., Regules, J. A., & Ockenhouse, C. F. (2018). RTS,S malaria vaccine efficacy and immunogenicity during *Plasmodium falciparum* challenge is associated with HLA genotype. *Vaccine*, 36(12), 1637-1642.
- Nolte, F. S., and Hill, C. E. . (2011). Polymerase chain reaction and other nucleic acid amplification technology. *Henry's Clinical Diagnosis and Management by Laboratory Methods*, 21, 1239-1249.
- Nunes, J. K., Woods, C., Carter, T., Raphael, T., Morin, M. J., Diallo, D., Lebouilleux, D., Jain, S., Loucq, C., Kaslow, D. C., & Birkett, A. J. (2014). Development of a transmission-blocking malaria vaccine: progress, challenges, and the path forward. *Vaccine*, 32(43), 5531-5539.
- Okonechnikov, K., Golosova, O., Fursov, M., & team, U. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8), 1166-1167.
- Ortolan, L. S., Avril, M., Xue, J., Seydel, K. B., Zheng, Y., & Smith, J. D. (2022). *Plasmodium falciparum* parasite lines expressing DC8 and group a PfEMP1 bind to brain, intestinal, and kidney endothelial cells. *Front Cell Infect Microbiol*, 12, 813011.
- Ouattara, A., & Laurens, M. B. (2015). Vaccines against malaria. *Clin Infect Dis*, 60(6), 930-936.
- Perkins, S. L. (2014). Malaria's many mates: past, present, and future of the systematics of the order Haemosporida. *J Parasitol*, 100(1), 11-25.
- Plebanski, M., Flanagan, K. L., Lee, E. A., Reece, W. H., Hart, K., Gelder, C., Gillespie, G., Pinder, M., & Hill, A. V. (1999). Interleukin 10-mediated immunosuppression by a variant CD4 T cell epitope of *Plasmodium falciparum*. *Immunity*, 10(6), 651-660.
- Podoba, J. E., & Stevenson, M. M. (1991). CD4+ and CD8+ T lymphocytes both contribute to acquired immunity to blood-stage *Plasmodium chabaudi* AS. *Infect Immun*, 59(1), 51-58.
- Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., & Thoma, G. (2018). Image analysis and machine learning for detecting malaria. *Transl Res*, 194, 36-55.
- Powledge, T. M. (2004). The polymerase chain reaction. *Adv Physiol Educ*, 28(1-4), 44-50.
- Prudencio, M., Rodriguez, A., & Mota, M. M. (2006). The silent path to thousands of merozoites: the *Plasmodium* liver stage. *Nat Rev Microbiol*, 4(11), 849-856.
- Pruett, C. L. W., K. (2008). The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*. *J Avian Biol*, 39(2), 252-256.
- Putaporntip, C., Hongsriruang, T., Seethamchai, S., Kobasa, T., Limkittikul, K., Cui, L., &

- Jongwutiwes, S. (2009). Differential prevalence of *Plasmodium* infections and cryptic *Plasmodium knowlesi* malaria in humans in Thailand. *J Infect Dis*, 199(8), 1143-1150.
- Putaporntip, C., Hughes, A. L., & Jongwutiwes, S. (2013). Low level of sequence diversity at merozoite surface protein-1 locus of *Plasmodium ovale curtisi* and *P. ovale wallikeri* from Thai isolates. *PLoS One*, 8(3), e58962.
- Putaporntip, C., Jongwutiwes, S., & Hughes, A. L. (2008). Differential selective pressures on the merozoite surface protein 2 locus of *Plasmodium falciparum* in a low endemic area. *Gene*, 427(1-2), 51-57.
- Putaporntip, C., Kuamsab, N., Kosuwin, R., Tantiwattanasub, W., Vejakama, P., Sueblinvong, T., Seethamchai, S., Jongwutiwes, S., & Hughes, A. L. (2016). Natural selection of K13 mutants of *Plasmodium falciparum* in response to artemisinin combination therapies in Thailand. *Clin Microbiol Infect*, 22(3), 285 e281-288.
- Putaporntip, C., Kuamsab, N., Pattanawong, U., Yanmanee, S., Seethamchai, S., & Jongwutiwes, S. (2021). *Plasmodium cynomolgi* co-infections among symptomatic malaria patients, Thailand. *Emerg Infect Dis*, 27(2), 590-593.
- Putaporntip, C., Kuamsab, N., Seethamchai, S., Pattanawong, U., Rojrung, R., Yanmanee, S., Weng Cheng, C., & Jongwutiwes, S. (2022). Cryptic *Plasmodium inui* and *Plasmodium fieldi* infections among symptomatic malaria patients in Thailand. *Clin Infect Dis*, 75(5), 805-812.
- Rahman, M. T., Uddin, M. S., Sultana, R., Moue, A., & Setu, M. (2013). Polymerase chain reaction (PCR): a short review. *Anwer Khan Mod Med Coll J*, 4(1), 30-36.
- Raj, D. K., Das Mohapatra, A., Jnawali, A., Zuromski, J., Jha, A., Cham-Kpu, G., Sherman, B., Rudlaff, R. M., Nixon, C. E., Hilton, N., Oleinikov, A. V., Chesnokov, O., Merritt, J., Pond-Tor, S., Burns, L., Jolly, G., Ben Mamoun, C., Kabyemela, E., Muehlenbachs, A., Lambert, L., Orr-Gonzalez, S., Gnadig, N. F., Fidock, D. A., Park, S., Dvorin, J. D., Pardi, N., Weissman, D., Mui, B. L., Tam, Y. K., Friedman, J. F., Fried, M., Duffy, P. E., & Kurtis, J. D. (2020). Anti-PfGARP activates programmed cell death of parasites and reduces severe malaria. *Nature*, 582(7810), 104-108.
- Ras-Carmona, A., Lehmann, A. A., Lehmann, P. V., & Reche, P. A. (2022). Prediction of B cell epitopes in proteins using a novel sequence similarity-based method. *Sci Rep*, 12(1), 13739.
- Regules, J. A., Cicatelli, S. B., Bennett, J. W., Paolino, K. M., Twomey, P. S., Moon, J. E., Kathcart,

- A. K., Hauns, K. D., Komisar, J. L., Qabar, A. N., Davidson, S. A., Dutta, S., Griffith, M. E., Magee, C. D., Wojnarski, M., Livezey, J. R., Kress, A. T., Waterman, P. E., Jongert, E., Wille-Reece, U., Volkmuth, W., Emerling, D., Robinson, W. H., Lievens, M., Morelle, D., Lee, C. K., Yassin-Rajkumar, B., Weltzin, R., Cohen, J., Paris, R. M., Waters, N. C., Birkett, A. J., Kaslow, D. C., Ballou, W. R., Ockenhouse, C. F., & Vekemans, J. (2016). Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: a phase 2a controlled human malaria parasite infection and immunogenicity study. *J Infect Dis*, *214*(5), 762-771.
- Rieckmann, K. H. (1990). Human immunization with attenuated sporozoites. *Bull World Health Organ*, *68 Suppl*, 13-16. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/2094578>
- Rowe, J. A., Obiero, J., Marsh, K., & Raza, A. (2002). Short report: Positive correlation between rosetting and parasitemia in *Plasmodium falciparum* clinical isolates. *Am J Trop Med Hyg*, *66*(5), 458-460.
- Roxby, P. (Producer). (2021, 2022, May 23). Malaria vaccine hailed as potential breakthrough. Retrieved from <https://www.bbc.com/news/health-56858158>
- RTS, S. C. T. P. (2015). Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. *Lancet*, *386*(9988), 31-45.
- Saichanapun, W., Seethamchai, S., Tia, T., & Putaporntip, C. (2013). Genotyping of *Plasmodium falciparum* by allele-specific amplification of the merozoite surface protein-1 locus. *Med J*, *57*(3), 305-320.
- Saiki, R. K. (1989). Technology: Principles and applications for DNA amplification In H. A. Erlich (Ed.), *The design and optimization of the PCR* (pp. 7-16): United States and Canada: Macmillan Publishers Ltd.
- Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., & Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, *230*(4732), 1350-1354.
- Satapornpong, P., Jinda, P., Jantararoungtong, T., Koomdee, N., Chaichan, C., Pratoomwun, J., Na Nakorn, C., Aekplakorn, W., Wilantho, A., Ngamphiw, C., Tongsimas, S., & Sukasem, C. (2020). Genetic diversity of HLA class I and Class II alleles in Thai populations:

- contribution to genotype-guided therapeutics. *Front Pharmacol*, 11, 78.
- Sinden, R. E. (1982). Gametocytogenesis of *Plasmodium falciparum* in vitro: an electron microscopic study. *Parasitology*, 84(1), 1-11.
- Singh, S., Miura, K., Zhou, H., Muratova, O., Keegan, B., Miles, A., Martin, L. B., Saul, A. J., Miller, L. H., & Long, C. A. (2006). Immunity to recombinant *Plasmodium falciparum* merozoite surface protein 1 (MSP1): protection in Aotus nancymai monkeys strongly correlates with anti-MSP1 antibody titer and in vitro parasite-inhibitory activity. *Infect Immun*, 74(8), 4573-4580.
- Smith, R. C., & Jacobs-Lorena, M. (2010). *Plasmodium*-mosquito interactions: a tale of roadblocks and detours. *Adv In Insect Phys*, 39, 119-149.
- Snounou, G., Viriyakosol, S., Jarra, W., Thaithong, S., & Brown, K. N. (1993). Identification of the four human malaria parasite species in field samples by the polymerase chain reaction and detection of a high prevalence of mixed infections. *Mol Biochem Parasitol*, 58(2), 283-292.
- Sturm, A., Amino, R., van de Sand, C., Regen, T., Retzlaff, S., Rennenberg, A., Krueger, A., Pollok, J. M., Menard, R., & Heussler, V. T. (2006). Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science*, 313(5791), 1287-1290.
- Takala, S. L., Coulibaly, D., Thera, M. A., Batchelor, A. H., Cummings, M. P., Escalante, A. A., Ouattara, A., Traore, K., Niangaly, A., Djimde, A. A., Doumbo, O. K., & Plowe, C. V. (2009). Extreme polymorphism in a vaccine antigen and risk of clinical malaria: implications for vaccine development. *Sci Transl Med*, 1(2), 2ra5.
- Talundzic, E., Scott, S., Owino, S. O., Campo, D. S., Lucchi, N. W., Udhayakumar, V., Moore, J. M., & Peterson, D. S. (2022). Polymorphic molecular signatures in variable regions of the *Plasmodium falciparum* var2csa DBL3x domain are associated with virulence in placental malaria. *Pathogens*, 11(5).
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*, 28(10), 2731-2739.
- Tanabe, K., Mackay, M., Goman, M., & Scaife, J. G. (1987). Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. *J Mol Biol*, 195(2), 273-287.
- Targett, G. (2015). Phase 3 trial with the RTS,S/AS01 malaria vaccine shows protection against

- clinical and severe malaria in infants and children in Africa. *Evid Based Med*, 20(1), 9.
- Taylor, R. R., Smith, D. B., Robinson, V. J., McBride, J. S., & Riley, E. M. (1995). Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of the immunoglobulin G3 subclass. *Infect Immun*, 63(11), 4382-4388.
- Terheggen, U., Drew, D. R., Hodder, A. N., Cross, N. J., Mugenyi, C. K., Barry, A. E., Anders, R. F., Dutta, S., Osier, F. H., Elliott, S. R., Senn, N., Stanistic, D. I., Marsh, K., Siba, P. M., Mueller, I., Richards, J. S., & Beeson, J. G. (2014). Limited antigenic diversity of *Plasmodium falciparum* apical membrane antigen 1 supports the development of effective multi-allele vaccines. *BMC Med*, 12, 183.
- Thimasarn, K., Jatapadma, S., Vijaykadga, S., Sirichaisinthop, J., & Wongsrichanalai, C. (1995). Epidemiology of malaria in Thailand. *J Travel Med*, 2(2), 59-65.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 25(24), 4876-4882.
- Tolle, M. A. (2009). Mosquito-borne diseases. *Curr Probl Pediatr Adolesc Health Care*, 39(4), 97-140.
- Treutiger, C. J., Hedlund, I., Helmbj, H., Carlson, J., Jepson, A., Twumasi, P., Kwiatkowski, D., Greenwood, B. M., & Wahlgren, M. (1992). Rosette formation in *Plasmodium falciparum* isolates and anti-rosette activity of sera from Gambians with cerebral or uncomplicated malaria. *Am J Trop Med Hyg*, 46(5), 503-510.
- Triglia, T., Stahl, H. D., Crewther, P. E., Silva, A., Anders, R. F., & Kemp, D. J. (1988). Structure of a *Plasmodium falciparum* gene that encodes a glutamic acid-rich protein (GARP). *Mol Biochem Parasitol*, 31(2), 199-201.
- Vaughan, J. A. (2007). Population dynamics of *Plasmodium* sporogony. *Trends Parasitol*, 23(2), 63-70.
- Vignali, M., Armour, C. D., Chen, J., Morrison, R., Castle, J. C., Biery, M. C., Bouzek, H., Moon, W., Babak, T., Fried, M., Raymond, C. K., & Duffy, P. E. (2011). NSR-seq transcriptional profiling enables identification of a gene signature of *Plasmodium falciparum* parasites infecting children. *J Clin Invest*, 121(3), 1119-1129.
- Walliker, D., Quakyi, I. A., Wellems, T. E., McCutchan, T. F., Szarfman, A., London, W. T.,

- Corcoran, L. M., Burkot, T. R., & Carter, R. (1987). Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science*, 236(4809), 1661-1666.
- Wamaket, N., Khamprapa, O., Chainarin, S., Thamsawet, P., Ninsaeng, U., Thongsalee, S., Suwan, V., Sakolvaree, J., Takhampunya, R., Davidson, S. A., McCardle, P. W., Sa-Angchai, P., Mukaka, M., Kiattibutr, K., Khamsiriwatchara, A., Nguitragool, W., Sattabongkot, J., Sirichaisinthop, J., & Kobylinski, K. C. (2021). Anopheles bionomics in a malaria endemic area of southern Thailand. *Parasit Vectors*, 14(1), 378.
- Weedall, G. D. (2020). The Entamoeba lysine and glutamic acid rich protein (KERP1) virulence factor gene is present in the genomes of *Entamoeba nuttalli*, *Entamoeba dispar* and *Entamoeba moshkovskii*. *Mol Biochem Parasitol*, 238, 111293.
- White, N. J. (1996). The treatment of malaria. *N Engl J Med*, 335(11), 800-806.
- White, R. J. (2013). RNA polymerase III transcription.
- WHO. (2012). Global plan for insecticide resistance management in malaria vectors.
- WHO. (2015). *Microscopy for the detection, identification and quantification of malaria parasites on stained thick and thin blood films in research settings (version 1.0): procedure: methods manual*. Retrieved from
- WHO. (2021). *World malaria report 2021*. Retrieved from <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021>
- Zhao, W., & Sher, X. (2018). Systematically benchmarking peptide-MHC binding predictors: From synthetic to naturally processed epitopes. *PLoS Comput Biol*, 14(11), e1006457.
- Zou, X., House, B. L., Zyzak, M. D., Richie, T. L., & Gerbasi, V. R. (2013). Towards an optimized inhibition of liver stage development assay (ILSDA) for *Plasmodium falciparum*. *Malar J*, 12, 394.

## VITA

**NAME** Rattanaporn Rojrung

**DATE OF BIRTH** 16 January 1997

**PLACE OF BIRTH** Yasothon Hospital

**INSTITUTIONS ATTENDED** Chulalongorn University

**HOME ADDRESS** 17 m.2 Nongsamsee Subdistrict, Senangkhanikhom District, Amnatcharoen Province, Thailand 37290

**PUBLICATION** Phimphan, S., Rojrung, R., Aiumsumang, S., Koomsab, K., and Tanomtong, A. (2020). Karyological Analysis of Lesser Bamboo Rat, *Cannomys badius* (Rodentia, Rhizomyinae) by Classical and Molecular Cytogenetic Techniques. *KKU Sci J*, 48(3), 318-325.

Putaporntip, C., Kuamsab, N., Nuprasert, W., Rojrung, R., Pattanawong, U., Tia, T., Yanmanee, S., and Jongwutiwes, S. (2021). Analysis of *Acanthamoeba* genotypes from public freshwater sources in Thailand reveals a new genotype, *T23 Acanthamoeba bangkokensis* sp. nov. *Sci Rep*, 11(1), 1-13.

Putaporntip, C., Kuamsab, N., Rojrung, R., Seethamchai, S., & Jongwutiwes, S. (2022). Structural organization and sequence diversity of the complete nucleotide sequence encoding the *Plasmodium malariae* merozoite surface protein-1. *Sci Rep*, 12(1), 1-14.

Putaporntip, C., Kuamsab, N., Seethamchai, S., Pattanawong, U.,



Rojrung, R., Yanmanee, S.,  
Cheng, CW., & Jongwutiwes, S. (2021). Cryptic Plasmodium inui  
and P. fieldi infections among symptomatic malaria patients in  
Thailand. Clin Infect Dis.

