CURRICULUM VITAE

Jimmy A. Vareta, PhD

Department of Epidemiology and Human Genetics

University of Maryland School of Medicine

jimmyvareta@gmail.com

EDUCATION

University of Maryland Baltimore, MD
PhD Molecular Epidemiology, December 2023

Michigan State University, East Lansing, MI
MSc., Clinical Laboratory Science, 2016.
Certificate, Epidemiology, 2016.

Nelson Mandela University, Port Elizabeth, Republic of South Africa
BTech., Biomedical Technology (*cum laude*), 2013.

Malawi College of Health Sciences, Lilongwe, Malawi
Diploma, Biomedical Sciences, 2003.

HONORS AND AWARDS

- MasterCard Foundation master's degree Scholarship, Michigan State University (2014).
- National Institute of Health (NIH) Fogarty D43 PhD Scholarship (2017).
- American Society of Tropical Medicine and Hygiene (ASTMH) 2021 Annual Meeting travel award.

RELEVANT EXPERIENCE

Blantyre Malaria Project (BMP), Laboratory Technician, Kamuzu University of Health Sciences, Blantyre Malawi (2008 – 2011, 2013 – 2014, and 2016 – 2017).
Project: "Pathogenesis of Cerebral Malaria in Children in Malawi."

- Performed in vitro *P. falciparum* and human endothelial cells culture.
- Assisted in conducting autopsy of children to understand the pathology of cerebral malaria.
- Performed PCR and RT-PCR to detect and genotype *P. falciparum* asexual parasites and gametocytes in field samples.

- Performed ELISA to measure plasma and cerebrospinal fluid levels of various antigens and antibodies.
- Assisted in establishing a Molecular Core Laboratory (under BMP) at College of Medicine, Kamuzu University of Health Sciences, Malawi.
- Managed biological samples shipments and freezer inventory system.
- Assisted in quality control of assays and scheduling required maintenance of laboratory equipment.

Drug Resource Enhancement Against HIV&AIDS (DREAM Program), Clinical Laboratory Technician, Malawi (2006 – 2007).
Project: "Evaluating Effectiveness of Different Combinations of anti-Retroviral Drugs in Patients with HIV&AIDS."

- Performed Hematology tests including full blood counts and CD4 counts using flow Cytometry.
- Performed liver function (aspartate aminotransferase and alanine aminotransferase) and kidney function tests (albumin and creatinine) using spectrophotometry.
- Performed HIV viral load using real-time quantitative PCR.
- Assisted in data management.

Wellcome Trust Research Laboratories, Laboratory Technician, Blantyre, Malawi (2003 - 2007).
Project: "Randomized Double-blind Placebo Controlled Clinical Trial evaluating effectiveness of Azithromycin as Prophylaxis to Prevent Preterm Labor in Pregnant Women in Southern Malawi."

- Performed malaria microscopy and serological assays to detect sexually transmitted diseases in pregnant women.
- Performed bacterial culture and sensitivity using blood samples (blood cultures), cerebrospinal fluid, ascitic fluid and other body fluids.
- Performed *Mycobacterium tuberculosis* cultures.
- Drawing blood from sheep to use as an ingredient for bacterial culture media (blood and chocolate agar).

TEACHING AND MENTORING

- Graduate teaching assistant in Introductory SAS, Statistics for Epidemiologists and Regression Analysis courses at University of Maryland Baltimore (UMB).
- Assisted in training undergraduate Medical Laboratory Technology students from Kamuzu University of Health Sciences during laboratory rotations.
- Mentored students on rotations at Malaria Research Laboratory at UMB including Natalie Hortsman (MPH student at Johns Hopkins University) and Otuto Nwagbata (Medical student at UMB).

PUBLICATIONS

Peer-reviewed journal articles

1. Milner DA, **Vareta J**, Valim C, *et al*. Human cerebral malaria and *Plasmodium falciparum* genotypes in Malawi. *Malar J*. 2012; 11:35.
2. Sisya TJ, Kamn'gona RM, **Vareta JA**, *et al*. Subtle changes in *Plasmodium falciparum* infection complexity following enhanced intervention in Malawi. *Acta Tropica*. 2015; 142:108-114.
3. Fulakeza J, McNitt S, **Vareta J**, *et al*. Comparison of msp genotyping and a 24 SNP molecular assay for differentiating Plasmodium falciparum recrudescence from reinfection. *Malar J*. 2019;18(1):84.
4. Thakur KT, **Vareta J**, Carson KA, *et al.* Cerebrospinal fluid Plasmodium falciparum histidine-rich protein-2 in pediatric cerebral malaria. *Malar J*. 2018;17(1):125.
5. **Vareta, J**., Buchwald, A.G., Barrall, A. *et al.* Submicroscopic malaria infection is not associated with fever in cross-sectional studies in Malawi. *Malar J* **19,** 233 (2020).

Manuscripts in progress

1. **Vareta J**., *et al*., Genotyping *Plasmodium falciparum* mature gametocytes using amplicon deep sequencing.
2. **Vareta J**., *et al*., Association between host and parasite factors and gametocyte production in *P. falciparum* infection.

RELEVANT SKILLS

Data Analysis
- Proficient in Biostatistics.
- Proficient in SAS, STATA, SPSS, and R statistical software.
- Proficient in longitudinal data analysis.
- Proficient in Data Science and Machine Learning.
- Proficient in SQL.

Bioinformatics
- Proficient in bioinformatics data analysis skills.
- Proficient in Unix scripting language and Python programming language.

Laboratory techniques
- Tissue cell culture (*Plasmodium falciparum* and human endothelial cells).
- Phlebotomy.
- Flow cytometry.
- DNA and RNA isolation, PCR, qRT-PCR.
- Clinical/Diagnostic microbiology.
- DNA Sequencing.
- Tissue cell culture (*Plasmodium falciparum* and human endothelial cells).
- Western blot

ABSTRACT

Dissertation Title: Impact of Host and Parasite Factors on Gametocyte Production in *Plasmodium falciparum*

Jimmy A. Vareta, Doctor of Philosophy, 2023

Dissertation Directed By: Miriam K. Laufer, MD, MPH[a]

Shannon Takala-Harrison, PhD[b]

[a] Director, Malaria Research Program, Center for Vaccine Development and Global Health, University of Maryland School of Medicine

[b] Head, Genomic Epidemiology Unit, Malaria Research Program, Center for Vaccine Development and Global Health, University of Maryland School of Medicine

Stalled progress in reducing the malaria burden over the past few years suggests the need to develop new interventions to augment existing ones, including interventions aimed at interrupting gametocyte transmission from humans to the mosquito vector. To develop and effectively apply interventions that target gametocytes, there is a need to understand patterns of gametocytemia observed between individuals. Gametocytemia varies by host age, season, symptom status, antimalarial drugs use, and complexity of infection; however, the underlying mechanisms of this variation are not fully understood.

Complexity of infection may modulate gametocytemia in *P. falciparum*; however, mechanisms of how clone composition influences gametocyte production are not clear. Addressing this gap requires a genotyping assay that can detect and estimate relative clone frequency of gametocytes and asexual parasites in infections.

The dissertation aims were two-fold: 1) To develop an amplicon sequencing assay to genotype *P. falciparum* mature gametocytes; and 2) To evaluate the impact of host and parasite factors and gametocyte production in *P. falciparum* infections.

We identified a polymorphic region of the *pfs230* gene as a marker to distinguish *P. falciparum* mature gametocyte clones. When evaluating the impact of host and parasite factors on gametocyte production and gametocytemia, we found that more complex and high parasite density infections were more likely to produce and harbor gametocytes. The proportion of infections that produced gametocytes were similar between age-groups and between symptomatic and asymptomatic individuals, but children and asymptomatic individuals were more likely to harbor gametocytes than adults and symptomatic individuals, respectively. These findings suggest that complexity of infection and parasite density may increase gametocyte production, but additional factors such as host immunity and duration of infection may contribute to the presence or absence of gametocytes after initiation of gametocyte production. Coupled with the development of the gametocyte genotyping assay which will be an important tool for studies aimed at understanding dynamics of gametocyte production in polyclonal infections, understanding the impact of host and parasite factors on gametocyte production and gametocytemia will help explain variation in gametocytemia observed between

individuals. This knowledge could inform development and effective deployment of transmission interrupting interventions.

Impact of Host and Parasite Factors on Gametocyte Production in *Plasmodium falciparum*.

by
Jimmy A. Vareta

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, Baltimore in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACT – artemisinin-based combination therapy

CCC – concordance correlation coefficient

cDNA – complementary deoxyribose nucleic acid

DNA – deoxyribose nucleic acid

GCR – gametocyte conversion rate

GDV1 – gametocyte development protein 1

IPT – intermittent preventive treatment

IRS – indoor residual spraying

ITN – insecticides-treated bed nets

LOD – limit of detection

PCR – polymerase chain reaction

RBC – red blood cell

RNA – ribonucleic acid

SAC – school age children

WHO – world health organization

CHAPTER I. BACKGROUND

Introduction to malaria

Malaria is a devastating vector-borne disease caused by different species of an apicomplexan parasite of the genus *Plasmodium.* Five species of *Plasmodium* cause disease in humans: *Plasmodium falciparum, P. vivax*, *P. ovale, P. malariae* and *P. knowlesi* which infects monkeys, but human cases have been reported in South-East Asia. *P. falciparum* is responsible for the most severe forms of the disease and majority of deaths, especially in sub-Saharan Africa where this species accounts for over 99% of all malaria cases. *P. vivax* is widely distributed globally but causes milder disease and has a lower fatality rate than *P. falciparum.* These two *Plasmodium* species are of the highest public health concern [1].

Apart from the public health burden of malaria in many countries around the world, the socioeconomic impact of this disease continues to be devastating in malaria endemic regions, especially in sub-Saharan Africa where the burden is highest [2-4]. Governments in malaria endemic countries spend a substantial amount of money every year to fund malaria control and prevention programs while other equally important sectors remain underfunded. For example, between 2010 and 2021, governments in malaria endemic countries contributed approximately 33% of the total funding for malaria control and elimination [1]. Furthermore, people in these countries lose their

daily livelihoods because of absenteeism from work and caring for sick relatives. In the 1950s the World Health Organization (WHO) rolled out a global malaria eradication program based on indoor residual spraying of dichlorodiphenyltrichloroethane and mass drug administration, but the program was abandoned with limited gains mainly due to parasite resistance to chloroquine, mosquito vector resistance to insecticides, and lack of funding and political will [5]. Recently, there have been renewed calls to eradicate malaria worldwide [6]. However, achieving eradication will require development of novel multifaceted malaria interventions.

In 2021, the WHO estimated that there were 247 million malaria cases and 619,000 malaria deaths worldwide. Approximately 95% of the cases and 96% of the deaths occurred in sub-Saharan Africa where the disease disproportionately affects children under the age of 5 years and pregnant women: 76% of reported malaria deaths occurred in children under the age of 5 years old and 32% of pregnant women are exposed to the malaria parasite every year [1]. Although there has been extensive roll out of interventions to reduce malaria incidence in recent years, the burden of the disease remains very high. From 2000 to 2015, there was a steady decline in the number of malaria cases worldwide from 245 million down to 230 million. This decline raised hopes that the fight against malaria could be won; however, since 2016, the number of malaria cases increased from 232 million to 247 million in 2021. This increase in the number of cases occurred mostly in the WHO Africa region. Similarly, the number of deaths due to malaria declined steadily from 897,000 in 2000 down to 568,000 in 2019. Since 2019, the number of malaria deaths increased to 619,000 in 2021. Malaria deaths in children

worldwide declined between 2000 to 2015 from 87.3% of total malaria deaths to 76.8 %, but there has been no change in rates of deaths in this age group since 2016 [1].

P. falciparum life cycle

The life cycle of *P. falciparum* takes place in humans (asexual stage reproduction) and mosquitoes (sexual stage reproduction). The parasite is transmitted when an infected female *Anopheles* mosquito bites a human host and injects sporozoites into the peripheral blood circulation. Sporozoites migrate to the liver where they invade hepatocytes and develop into liver schizonts. After 6 – 14 days, liver schizonts rupture and release merozoites into the peripheral blood circulation. Merozoites invade red blood cells (RBCs) and develop into ring forms six hours after entry and mature into trophozoites 24 hours later. Trophozoites undergo cell division and develop into schizonts containing 6 – 30 merozoites [7]. Forty-eight hours after initial invasion, schizonts rupture releasing merozoites that infect more RBCs **(Figure 1)**. Exponential replication of asexual parasites; sequestration of parasitized RBCs to brain endothelial cells and other organs; and rupture of infected RBCs releasing parasite proteins and immunological response of the human host to these proteins is responsible for the pathology of the disease [7, 8]. Individuals develop partial immunity after multiple encounters with *P. falciparum* parasites. This partial immunity wanes after some time and individuals are not protected when they are exposed to the parasite again. Over time and after multiple exposures to the parasite, individuals may develop immunity that prevents clinical malaria [9]. This clinical immunity increases with age, and therefore, in malaria endemic regions, host age is commonly used as a surrogate for *P. falciparum* clinical immunity.

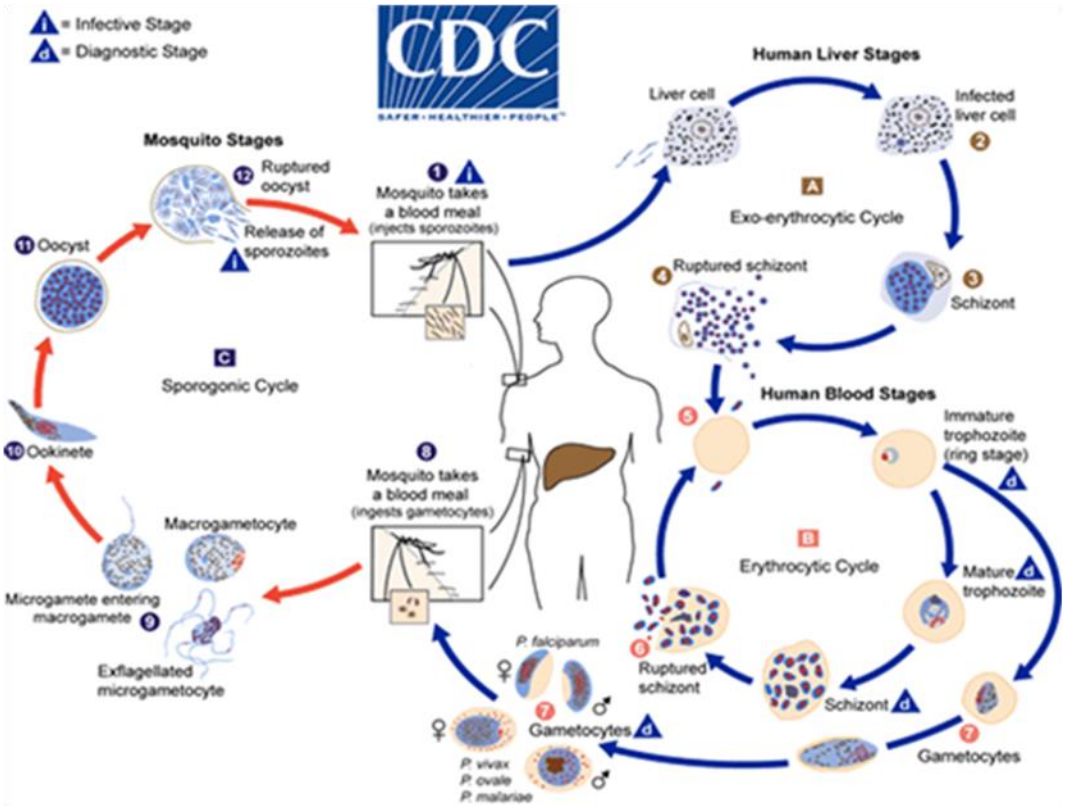

Figure 1: The *P. falciparum* life cycle [CDC]**.** Female *Anopheles* mosquito injects sporozoites during blood feeding. Sporozoites move to the liver where they undergo schizogony and release merozoites into blood circulation and infects RBCs.

The ability of the asexually replicating parasites to differentiate into transmissible forms is responsible for the spread of malaria and parasite survival [10]. During the *P. falciparum* erythrocytic cycle in humans, a subset of parasites differentiates to produce female and male gametocyte forms which are eventually transmitted to mosquitoes during blood feeding [11]. The ratio of female to male gametocytes is usually biased toward female, with 2 – 5 or more female gametocytes to 1 male gametocyte [12]. *P. falciparum* gametocytes take about 10 – 12 days to develop from committed schizonts to mature gametocytes [13]. During this period, gametocytes go through five distinct stages of development (Stage I – V). Immature gametocytes (stage I) sequester in bone marrow and spleen where they develop [14-16] into stages II - IV before they are subsequently

released back into peripheral blood as stage V mature gametocytes **(Figure 2)**. Circulating mature gametocytes are ingested by female *Anopheles* mosquitoes during a blood meal. Once ingested by the mosquito, the gametocytes undergo gametogenesis (sexual reproduction) in the mosquito midgut. Changes in the environment inside the mosquito midgut, such as a drop in temperature, increase in pH, and presence of xanthuremic acid induces gametogenesis [17, 18]. Each male gametocyte undergoes cell division to produce eight microgametes and each female gametocyte form produces one macrogamete. Inside the midgut, microgametes exflagellate and fertilize with macrogametes to form zygotes that subsequently develop into motile ookinetes and cross the midgut of the mosquito to form oocysts. These oocysts mature and burst to release sporozoites that migrate to salivary glands of the mosquito [13]. The presence of sporozoites in salivary glands enable mosquitoes to be infectious once they bite humans.

Gametocytes detection in peripheral blood is important for identification of reservoirs of transmission that can be targeted to reduce transmission of the malaria parasite from humans to mosquitoes. Previously, gametocyte detection was based on microscopy; only recently have molecular detection methods been widely used in field studies. Molecular detection of gametocytes relies on the detection and quantification of transcript abundance of a gene that is exclusively expressed in mature gametocytes. Some molecular based methods currently being used for gametocyte detection target the *pfs25* gene (PF3D7_1031000), which is highly expressed in female gametocytes [19]. More recently, the *ccp4* (PF3D7_0903800; highly expressed in female gametocytes) and *mget* genes (PF3D7_1469900; highly expressed in male gametocyte) have been used to distinguish between the two populations [20]. Unlike microscopy-based detection of

female and male gametocytes which relies on slight differences in morphology, molecular detection of female and male gametocytes has allowed accurate quantification of gametocyte sex ratio which plays a major role in sexual reproduction in the mosquito host [21]. Microscopy-based detection of gametocytes is less sensitive compared to molecular detection with some field studies reporting gametocyte prevalence estimates by molecular methods that are three times higher than estimates based on microscopy [19, 22]. Use of the more sensitive molecular methods for gametocytes detection has allowed identification of asymptomatic submicroscopic infections with gametocytes which are more prevalent in school age children (SAC) in malaria endemic regions and are a reservoir for onward transmission of malaria[19, 23].

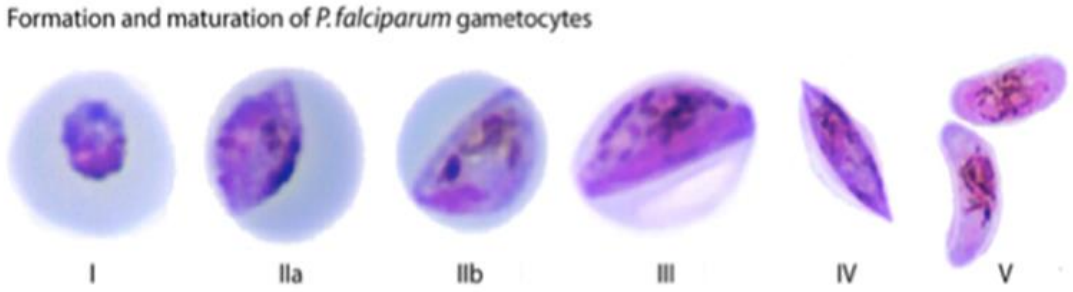

Figure 2: Five developmental stages of *P. falciparum* gametocytes. Stages I to stage IV are present in bone marrow and spleen where they develop, and stage 5 are mature gametocytes circulating in peripheral blood [11].

Threats to Current Malaria Control and Prevention Measures

Malaria control and prevention measures recommended by the WHO include artemisinin-based combination therapies (ACTs) which are used as a first-line antimalarial drug, vector control measures which comprises indoor spraying of insecticides (IRS) and insecticides-treated bed nets (ITNs), seasonal malaria

chemoprevention in some areas, and intermittent preventive treatment for children and pregnant women (IPT). Recently, the WHO has recommended 2 vaccines to prevent malaria in children: RTS, S/AS01 in 2021 and R21/Matrix-M in October 2023. Although ACTs, IRS and ITNs have registered some success in reducing the burden of the disease, reported *P. falciparum* resistance to artemisinin derivatives and partner drugs in South-East Asia and recently in several countries in Africa (Eritrea, Rwanda and Uganda), and vector resistance to insecticides reported in many malaria endemic settings threaten to derail the progress that has been made to reduce the malaria burden and also threatens malaria elimination efforts [24-30]. The threats posed by parasite resistance to ACT and mosquito vector resistance to insecticides coupled with stalled progress in reducing the burden of malaria suggest that new malaria interventions are urgently needed to augment existing control and prevention measures.

Transmission interrupting interventions

Stagnation of progress in reducing the malaria burden in the past few years and threats posed by *P. falciparum* resistance to ACTs and malaria vector resistance to insecticides has stimulated research into the development of interventions aimed at interrupting transmission of the malaria parasite from humans to the mosquito vector. These interventions are key to reducing the burden of malaria in high transmission areas, and are also crucial in malaria elimination and eradication efforts in areas with low burden of the disease [31]. One of the targets of transmission interrupting interventions is the gametocyte. Gametocytes are transmitted from the human host to the mosquito vector to allow

competition of the parasite life cycle; therefore, this parasite stage is a perfect target for transmission interrupting interventions [11]. Interruption of transmission of the parasite from humans to mosquitoes can be achieved by the use of drugs that can kill mature gametocytes circulating in peripheral blood and transmission blocking vaccines that provide immune responses that inhibit gametogenesis in the mosquito host [32]. To develop effective interventions that target gametocytes, and to design viable strategies to roll out these interventions to maximize their effectiveness, it is imperative to understand variations in gametocytemia between individuals, as estimates of gametocytemia are useful when determining human-to-mosquito transmission potential of the malaria parasite in populations.

Variation in gametocytemia between infections.

Understanding variation in gametocytemia between individuals is critical for the development and prudent application of interventions that target transmission of the malaria parasite from humans to mosquitoes. Gametocyte carriage follow certain demographic patterns: among others, it varies by age, season, antimalarial drug use, and symptoms status [11, 19, 33-38]. Surveillance data from malaria endemic countries have shown higher proportion of infection with gametocytes in school age children (SAC) compared to adults and younger children [19, 22, 39]. Other studies have shown higher gametocyte carriage in asymptomatic compared symptomatic individuals; in individuals taking antimalarial drugs or who have recently taken antimalarial drugs compared to those not taking antimalarial drugs; in anemic individuals compared to individuals with normal hemoglobin levels; in

individuals with chronic *P. falciparum* infections compared to individuals with incident infection; and during the rainy season compared to the dry season [35-38, 40, 41]. However, causes of these variations in gametocytemia between individuals are not clearly understood. Variation in gametocytemia could be due to differences in sexual commitment of blood stage parasites or could be due to factors such as host immune responses that may interfere with gametocyte production and maturation.

Molecular mechanisms of sexual commitment in *P. falciparum*

During the intraerythrocytic stage in humans, malaria parasites either invest in asexual replication or gametocytes production. In *P. falciparum,* gametocyte production starts with a small proportion of asexual parasites committing to become gametocytes. Sexual commitment is believed to be initiated before the schizont stage [42]. Each merozoite released from a committed schizont after rupture eventually develops into a schizont which produces either male or female gametocytes [42]. Sexual commitment is induced by *P. falciparum* gametocyte development protein 1(GDV1) which antagonizes heterochromatin protein 1 (HP1) [43, 44]. HP1 suppresses the activation of *P. falciparum* ApiAP2-G (*Pf*AP2-G), a master transcription regulator of genes that are involved in gametocyte development [45, 46]. AP2-G is one of the 20 – 27 proteins that make up the apicomplexan AP2 family of DNA binding proteins. These proteins function as transcription factors that regulate expression of genes responsible for development in apicomplexan parasites [47]. In *P. falciparum,* AP2-G plays an important role in sexual commitment by upregulating the expression of genes required for gametocyte

development (including *Pfs16*, *Pfg27/25*, *Pfg14.744*, *Pfg14.745*, and *Pfg14.748*) [45, 48, 49] **(Figure 3)**. The *ap2-g* gene is highly expressed in schizonts and gametocyte-committed rings, and it is used as a molecular marker for sexual commitment [41, 47, 50].

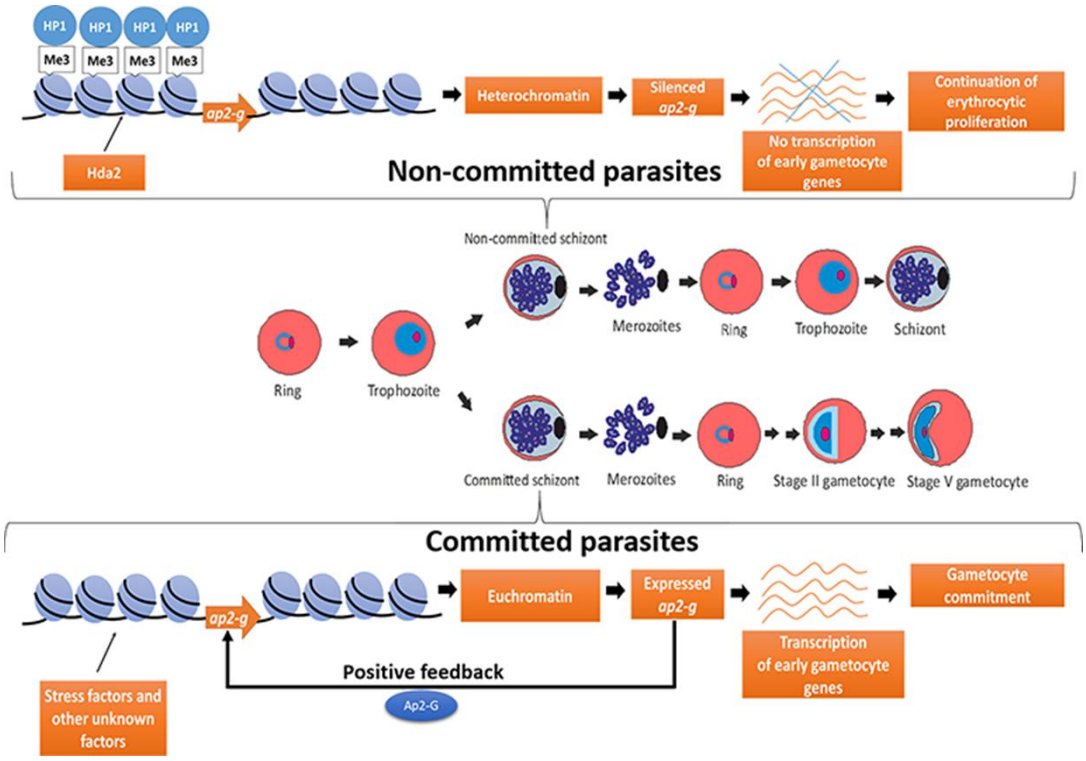

Figure 3: Mechanisms underlying *ap2-g* gene expression and sexual commitment in *P. falciparum* [51].

Expression of the *ap2-g* gene is limited to 1 – 30% of all asexual parasites in an infection [48], suggesting that not all asexual parasites in an infection commit to produce gametocytes. It has been suggested that gametocyte production could be a stochastic process; however, there is abundant evidence contradicting this suggestion [52-54]. *P. falciparum* isolates from different patients vary in gametocytes conversion rates (GCR) – the proportion of asexual parasites in circulation that are committed to develop into

gametocytes [52]. In experimental studies using *P. falciparum* clones from laboratory isolates, distinct parasite clones have different GCRs [53]. Furthermore, some isolates that have been in culture for a long period of time lose their ability to produce gametocytes because of deletions in a locus spanning the GDV1 gene on chromosome 9 of the genome [54, 55]. Also, reduced levels of lysophosphatidylcholine, a common component of plasma, increases gametocyte production *in-vitro* [56]. This evidence suggests that gametocyte production is not a stochastic process, but instead is influenced by host and parasite genetic factors. It remains unclear whether the influence of these factors on sexual commitment explains the variation in gametocytemia observed between individuals.

Measuring gametocyte production

Gametocytemia is commonly used as a surrogate for gametocyte production. Although gametocytemia is useful when determining transmission potential of an infection, it does not necessarily accurately capture gametocyte production [57]. GCR has been proposed as a more direct metric for estimating gametocyte production [58]. However, this metric is difficult to measure in natural infection as it requires at least two consecutive samples with strict timing. Recently, it has been demonstrated that expression levels of the *ap2-g* gene are positively associated with GCR [41, 59]. Thus, *ap2-g* gene expression is a reasonable surrogate for estimation of gametocyte production, and its use in studies investigating the influence of host and parasite factors on

gametocyte production could help in understanding variation in gametocytemia observed between individuals.

Influence of host immunity on gametocyte production

Although molecular mechanisms underlying sexual commitment and gametocyte development have recently been elucidated [43-46], host factors that influence these processes in natural *P. falciparum* infection are not clearly understood and remain an important knowledge gap. Filling this knowledge gap could help in understanding the variation in gametocytemia observed between individuals in *P. falciparum* infections. Previous studies have evaluated the association between host factors, such as immunity or host age (a proxy for host immunity) and gametocytemia. In an *in-vitro* study using sera from Gambian children exposed to malaria and non-exposed Europeans, addition of serum from exposed children to *P. falciparum* cultures substantially reduced gametocyte production [60]. Similar results have been observed in animal model studies that used *P. yoelii* and *P. chabaudi* [61-63]. Gametocytemia may also be influenced by immune responses acting against asexual parasites by reducing the number of asexual parasites that commit to develop into gametocytes [64, 65].

There is abundant evidence demonstrating that there are immune responses that act against immature (stage I – IV) and mature gametocytes (stage V). Immune responses against immature gametocytes may reduce the number of developing gametocytes which subsequently impact gametocytemia [64]. A study conducted in Thailand has shown that sera from malaria patients reduces the number of stage I gametocytes and hinders their

maturation [66]. A recent study showed that exported proteins produced by immature gametocytes (stage I - III) elicit immune responses, suggesting that naturally acquired immune responses act against immature gametocytes [64]. Other studies have shown that CD4 + T cells respond to antigens on the surface of mature gametocytes suggesting that cellular immunity may play an important role in the clearance of mature gametocyte in peripheral blood circulation [67, 68]. Likewise, immune responses against mature gametocytes shorten their circulation time, thus impacting their density in peripheral blood [11], as evidenced by a study which was conducted in Indonesia demonstrating that individuals with a history of malaria had lower gametocytemia compared to individuals with limited exposure to the parasite [66]. Furthermore, a study conducted in Gabon demonstrated that surface proteins on RBCs infected with mature gametocytes elicit an immune response [69]. These findings suggest that immune responses act against immature gametocyte during development as well as circulating mature gametocytes, and this may impact gametocytemia independent of asexual parasite density. Further evidence of the role of immunity on gametocytemia is demonstrated by the association between host age and gametocyte carriage. Host age, which is a proxy for immunity, has been shown to influence gametocytemia in natural *P. falciparum* infection. Cross-sectional studies conducted in Malawi and Burkina Faso demonstrated that SAC had increased gametocytemia compared to younger children and older adults [19, 22]. Findings from these studies also demonstrate that naturally acquired immunity against immature and mature gametocytes may influence gametocytemia in *P. falciparum*. Notwithstanding, these studies were either conducted *in-vitro* under controlled conditions or measured gametocytemia which may not an accurately reflect gametocyte production which

captures the proportion of gametocytes that have survived immune responses against immature and mature gametocytes [45, 48, 59].

Duration of infection and gametocyte production

Malaria surveillance using both microscopy and molecular diagnosis in endemic countries has found a high prevalence of asymptomatic *P. falciparum* infections. A surveillance study conducted at the end of the dry season in 2012 (September-October) and at the end of the rainy season in 2013 (April-May) in three districts (Blantyre rural, Chikwawa and Thyolo) in southern Malawi found that asymptomatic infections made up 88% of all *P. falciparum* infections detected [23]. Since these infections are detected in otherwise healthy individuals, they are mostly untreated and persist for months or years. Use of molecular methods for gametocyte detection in community surveillance studies have shown that a substantial proportion of submicroscopic asymptomatic infections harbor gametocytes; therefore, these infections are considered important reservoirs of onward transmission of *P. falciparum* [19, 70]. Evidence of the importance of these infections in transmission was demonstrated in a longitudinal study of incident and chronic infections in children in Burkina Faso where they showed that all individuals with chronic infection developed gametocytes during follow-up but only 35% of incident infections developed gametocytes [41]. In the same study, using membrane feeding assays, it was shown that chronic infections had higher mosquito infectivity compared to incident infections, indicating that these low parasitemia infections are infectious to mosquitoes and are an important transmission reservoir. It is thought that high

gametocyte carriage in these asymptomatic long duration infections is due to continuous gametocyte production over long periods [40]; however, it is not clearly understood whether these chronic infections are more likely to commit to produce gametocytes than incident infections.

Influence of polyclonal infections on gametocyte production.

P. falciparum infections in high transmission settings are typically polyclonal [71-75]. Polyclonal infections can result from bites from multiple mosquitoes each infected with different parasite clones, or from a bite from a single mosquito that is infected with multiple parasite clones [58]. The impact of polyclonal infections on gametocyte production in natural infections has not been clearly delineated. According to evolutionary theory, competition of co-infecting parasite clones in a polyclonal infection could influence investment in asexual replication or sexual differentiation in multiple ways [76]. For example, two clones may compete for resources to survive (i.e., competing for RBCs) or nonspecific immune responses against a dominant clone in an infection impacting densities of co-infecting minority clones and the dominant clone will invest more in asexual replication and minority clones will invest more in sexual reproduction [77]. These clone interactions could influence gametocyte production and gametocytemia in infected individuals. However, very little research has been done to understand how these clone interactions influence gametocyte production in *P. falciparum* natural infections. There are contradictory theories on how clone competition may impact sexual reproduction by individual clones in polyclonal infections. It is

possible for stressed minority clones to invest more in sexual reproduction as demonstrated in studies where antimalarial drug treatment and anemia are associated with high gametocytemia [11, 35, 37, 78], and in *in vitro P. falciparum* cultures where it has been shown that stressed parasites produce more gametocytes [11]. In contrast, majority clones in an infection may also invest more in sexual reproduction. This is demonstrated in *in vivo* studies using *P. chabaudi* infections in mouse models that have shown that majority clones in polyclonal infections have competitive advantage over minority clones and invest more in sexual reproduction [76, 79, 80].

Studies aimed at understanding reproductive investment in the context of polyclonal infection have mostly been done in mouse models. These controlled laboratory studies mostly used two distinct clones of the parasite which is not a true reflection of the complexity of infection observed in high transmission settings where up to 8 clones can be observed in one infection [76]. In addition, findings from these studies cannot be generalized without confirmatory evidence from field studies. Many field studies that have attempted to answer this question have looked at the association between complexity of infection (COI) and gametocytemia [36, 40, 41]. Evaluating these associations cannot answer this question as these studies have only shown a positive association between COI and gametocyte carriage but have not demonstrated the contribution of each clone on gametocytemia in an infection. To definitively answer this question, a well-designed longitudinal study coupled with a gametocyte genotyping method that can detect and estimate relative clone frequency in both asexual parasites and gametocytes in a polyclonal infection is needed.

Gametocyte genotyping techniques that have previously been developed are mostly based on length polymorphism markers [81-83]. These length polymorphism markers cannot estimate relative clone frequency which is important in identification of majority or minority clones in an infection [81, 82]. With recent advances in short-read sequencing technologies, amplicon deep sequencing of polymorphic regions of antigen genes is a method of choice for measuring COI and relative clone frequency in *P. falciparum* infections. This approach has been successfully used to genotype blood stage *P. falciparum* parasites in field settings [84-86]. However, the approach has not been utilized to genotype gametocytes in *P. falciparum* infections. Use of amplicon based genotyping method to genotype concurrently asexual parasites and mature gametocytes could be an important tool to understand gametocyte production dynamics in the context of polyclonal infections.

Study Objectives and Public Health Significance

The aims of this dissertation project were two-fold: 1) To develop an assay to genotype *P. falciparum* gametocytes using amplicons sequencing of a polymorphic region in a gene exclusively expressed in mature gametocytes; and 2) To evaluate the impact of host and parasite factors on gametocyte production in *P. falciparum* infection in the context of a longitudinal cohort study conducted in Malawi between 2019 and 2020. Development of a genotyping method based on amplicon deep sequencing is crucial in understanding how clone composition influence gametocyte production in *P. falciparum*. Specifically, the genotyping assay can be used in longitudinal studies to track clones in an

infection to determine which clones are more likely to produce gametocytes. Understanding the contribution of host and parasite factors that influence gametocyte production and gametocytemia will help to explain these metrics. This knowledge will be crucial for understanding the biology of gametocytes which may inform development and deployment of transmission interrupting interventions by identifying groups of individuals who are likely to carry gametocytes and potentially allow application of interventions prior to individuals becoming infectious.

Innovation

This will be the first study to develop an amplicon deep sequencing assay to genotype mature gametocytes in *P. falciparum* infection which can be used to understand gametocyte production dynamics in polyclonal infections. Unlike *P. falciparum* gametocyte genotyping methods based on length polymorphism markers, this amplicon sequencing based genotyping assay will be more likely to capture minority clones and will allow estimation of the relative frequency of each clone among gametocytes in the infection. Furthermore, the amplicon sequencing approach will increase throughput through multiplexing of samples.

In this proposed dissertation project, I used *P. falciparum* positive blood samples collected from a well characterized community-based household surveillance cohort conducted under the Malawi International Center of Excellence for Malawi Research (ICEMR) to understand the association between host and parasite factors and variation in gametocyte production and gametocytemia. The ICEMR study was aimed at

characterizing human to mosquito transmission of *P. falciparum* over time and how this is impacted by application of universal insecticide treated bed nets. Data generated from this dissertation and membrane feeding and immunological data generated by research colleagues in Malawi can be linked to accurately investigate dynamics of gametocyte production and gametocytemia in individuals over time. Furthermore, the data could be used to identify factors that influence gametocyte production in natural *P. falciparum* infection which could help explain variation in gametocyte carriage and transmission between individuals. This information will inform deployment of interventions aimed at interrupting transmission of the malaria parasite from humans to the mosquito host.

CHAPTER II. STUDY DESIGN AND METHODS

The ICEMR Malawi household-based clustered cohort study

Study design

The *P. falciparum* positive samples that were used in this dissertation research were collected as part of the Malawi ICEMR household cohort study (ICEMR, NIH grant number: U19A1089683) conducted within catchment areas of Namanolo and Ntaja health centers in Balaka and Machinga districts, respectively, in southern Malawi between April 2019 and March 2020. The two districts are hot and dry rural areas with high malaria transmission.

Human subjects research approvals were obtained from the University of Maryland School of Medicine`s and Michigan State University`s Institutional Review Board, and the Kamuzu University of Health Sciences Research Ethics Committee.

Participants

The ICEMR cohort study included 96 households in each health center (Namanolo and Ntaja) catchment areas with the goal of detecting infection, disease, and potential for transmission. Briefly, each of the two sites was divided into four quadrants

and within each quadrant four index households were sequentially sampled from the population census done in 2018. The five nearest households were identified to form a cluster of households. Index houses were excluded if there was less than 100 meters between the cluster boundary and proceeding clusters. Index households and five neighboring households were offered enrollment with replacement of any refusals or migrations with the next nearest households. All members of each household were invited to participate in the study. The study was an open cohort where new household members were invited to participate, thus length of follow-up for each individual varied. Consent was sought from the head of the households, and after consent was granted, children and adults were recruited to participate in the study and provided individual level consent. Demographics and other information were collected from individuals upon recruitment. Monthly surveys were conducted in all individuals to collect information on bed net use, treatment, and symptoms. All study participants were followed with active case detection visits every 4 – 6 weeks yielding a maximum of ten visits over twelve months or until they were lost to follow up. Passive case detection was conducted by a study nurse at the health center. Study participants presenting for evaluation of illnesses were interviewed by the study nurse, but diagnosis, evaluation and treatment were conducted by the routine health center providers according to the standard of care. Malaria treatment was provided to those diagnosed with malaria. Samples at active and passive case detection visits were collected by finger prick. Blood was collected into Eppendorf tubes, 50 µL was pipetted on to filter paper to create two blood spots and 100 µL of blood was placed in RNA Protect cell reagent (Qiagen, CA). Filter paper samples were stored at room temperature with desiccants for up to 2 weeks and then stored at -20 °C until *P. falciparum* detection

and RNA protect samples were refrigerated for up to 2 weeks and then transitioned to -80 °C freezer until nucleic acid extraction.

P. falciparum detection and quantification

P. falciparum detection and quantification of the ICEMR blood samples was done in the ICEMR molecular core laboratory at Kamuzu University of Health Sciences in Blantyre, Malawi. Genomic DNA was extracted from filter paper blood samples from all visits and *P. falciparum* 3D7 laboratory strain standards of known parasite density (6000, 600, 60 and 6 parasites/ µL) using the dried blood spots protocol in the Qiamp DNA MiniKit kit (Qiagen, CA). *P. falciparum* detection and quantification was based on quantitative real-time PCR targeting the 18s rRNA gene [87]. Briefly, 5 µL of DNA extracted from each sample and the standards was used in a real-time qPCR using TaqMan Universal Mastermix (Applied Biosystems) and probe on a Step One Plus Real-Time PCR System (Applied Biosystems). Parasite density (parasites per µL) was estimated via linear regression applied to cycle threshold (Ct) values and parasite density of the *P. falciparum* 3D7 laboratory strain standards. The following primers and probe were used: Forward Primer at 25µm: 5' – GTA ATT GGA ATG ATA GGA ATT TAC AAG GT – 3'; Reverse Primer at 25µm: 5' – TCA ACT ACG AAC GTT TTA ACT GCA AC – 3'; and Probe at 25µm: (5'-[FAM] TGC CAG CAG CCG CGG TAA TTC [TAMRA]-3').

Gametocyte detection and quantification

Stored RNA Protect stabilized blood from all *P. falciparum* positive samples was sent to the University of Maryland School of Medicine (UMSOM) malaria research laboratory for gametocyte detection and quantification, *ap2-g* gene expression RT-PCR, and amplicon deep sequencing. At UMSOM, the samples were thawed for DNA and RNA extraction: 200 µL was used for RNA extraction using RNeasy Plus 96 kit (Qiagen, CA) and RNA was eluted in 50 µL of Low EDTA TE (1X); and the remaining sample (~150 µL) was used for DNA extraction using QIAamp DNA Blood Kit (Qiagen, CA) and DNA was eluted in 40 µL of Low EDTA TE (1X). Extracted RNA and DNA was stored at -80 °C until analysis. All RNA samples underwent Human Tubulin one step RT-PCR to determine if the RNA was successfully extracted. Details about primers and probe, reaction mix volumes, and cycling conditions for Human Tubulin RT-PCR are in the Appendix **(Table 8)**.

Gametocytes quantification for *P. falciparum* positive samples was based on quantification of the *ccp4* (female gametocytes) and *mget* (male gametocytes) genes transcripts using RT-qPCR as described previously [20]. In brief, 5 µL of RNA extracted from *P. falciparum* positive samples using RNeasy Plus 96 kit and serially diluted standards generated from target RNA copies transcribed *in vitro* using MEGAShortscript T7 high yield transcription kit (Invitrogen) and quantified using Qubit 3 fluorometer (ThermoFisher) was used in a multiplex one step RT-qPCR with Luna Universal Probe One-Step RT-qPCR Kit (NEB) and probes on a Light Cycler 96 Real-time PCR machine (Roche). Reaction mix and cycling conditions are described in Appendix Table A. All RT-qPCR reactions were performed in duplicate. If there were discordant results in the

duplicate reactions, the RT-qPCR was repeated, and if still discordant, a third run was conducted as a tie breaker. *ccp4* and *mget* copy number per mL in test samples was estimated via a linear regression using Ct values of the serially diluted standards and their respective *ccp4* and *mget* copy number per mL. To obtain female and male gametocyte density, estimated test sample *ccp4* copy number per mL was multiplied by a factor of 4 (4 RNA copies of per female gametocyte cell) and estimated test sample *mget* copy number was multiplied by a factor of 9.8 (9.8 RNA copies per male gametocyte cell) [20].

Aim 1 Methods

Identification of candidate gametocyte genotyping markers

Molecular detection of mature gametocytes target genes that are primarily expressed in mature gametocytes. Detection of these mature gametocyte genes in RT-PCR of RNA extracted from individuals infected with *P. falciparum* indicate presence of mature gametocytes in an infection. To identify a genotyping marker for mature gametocyte, I compiled a list of genes that are primarily expressed in *P. falciparum* mature gametocytes using published *P. falciparum* stage specific single cell RNA-seq data [88]. Using this list, I searched a *P. falciparum* "microhaplotype" dataset that contains 4465 genomic regions (200 bp) that show high genetic diversity [89] to select regions with expected heterozygosity of 0.5 and above as candidate gametocyte genotyping markers. To improve genotyping resolution, overlapping candidate marker

regions were combined (when possible) to obtain larger amplicons of a size amenable to coverage by Illumina MiSeq paired-end reads.

Haplotype resolution of candidate genotyping markers

To evaluate the potential haplotype resolution of candidate genotyping markers, the number of haplotypes was estimated at the genotyping markers loci from whole genome sequences (WGS) data generated from 205 Malawian *P. falciparum* field isolates. The Malawian *P. falciparum* isolates were collected in a cohort study that was conducted at Mfera Health Center in Chikwawa district between 2014 and 2015. The cohort study has been described previously [90]. Briefly, 120 symptomatic individuals with uncomplicated *P. falciparum* malaria confirmed by histidine-rich protein-2 rapid diagnostic test and microscopy, and without human immunodeficiency virus were followed monthly for 2 years. At each clinic visit, blood was collected on filter paper, on a slide, and in a Lithium Heparin tube for parasite culture. Part of the blood in Lithium Heparin was whole genome sequenced. Approval for human subject research was granted by the University of Maryland School of Medicine and Michigan State University Institutional Review Boards and Kamuzu University of Health Sciences Health Sciences Research Committee. For comparison, the number of haplotypes in 2 loci from commonly used blood stage amplicon sequencing markers (*ama-1* and *sera-2*) [84, 91] were also estimated from the same isolates.

Specificity of candidate genotyping markers

To determine the specificity of the candidate gametocyte genotyping markers for detection of mature gametocytes (compared to asexual parasite stages), I extracted RNA from cultured ring stage parasites (10^5 rings/µL) of the MRA–1183 *P. falciparum* SenTh026.04 strain (a non-gametocyte producing line obtained from BEI Resources) [92] and mature gametocytes (10 gametocytes/µL) cultured from NF54 strain (a gametocyte producing laboratory strain) using the RNeasy Mini kit (Qiagen, CA). Complementary DNA (cDNA) was synthesized from the RNA using QuantiTect Reverse Transcription kit (Qiagen, CA). PCR with primers targeting the candidate marker was performed on the cDNA using the KAPA HiFi HotStart ReadyMix PCR kit (Kapa Biosystems, Wilmington, MA). Primers were designed using Primer-BLAST [93] (Appendix Table B). A higher concentration of MRA-1183 rings compared to NF54 gametocytes was used to reflect the higher asexual parasite density (compared to gametocyte density) often observed in natural *P. falciparum* infections [39], and to determine if any low-level expression of the candidate marker gene in rings could overestimate mature gametocyte genotypes. A region of the *pfs25* gene, a gametocyte detection marker known to be highly specific to female gametocytes [9], was amplified as a positive control for gametocyte specificity. All RT-PCRs were done with a reverse-transcriptase negative control to rule out DNA contamination in extracted RNA and a PCR negative control to rule out contamination of reagents. PCR products from the RT-PCR were visualized using 3% agarose gel electrophoresis.

Quantitative ability of genotyping assay in experimental mature gametocyte mixtures of distinct *P. falciparum* isolates

Culture of distinct P. falciparum isolates and gametocyte induction

To identify distinct *P. falciparum* isolates to use for mature gametocyte mixtures to mimic a polyclonal infection, I compared DNA sequences spanning the candidate genotyping markers regions in 3 clinical isolates (MLW5, MLW227 and MLW389) and a laboratory strain of *P. falciparum* (NF54). Two distinct isolates that had DNA sequence differences at the candidate gametocyte genotyping marker regions were MLW5 and NF54 cultured to obtain mature gametocytes and MRA-1183 *P. falciparum* SenTh026.04 isolate was also cultured to obtain ring stage parasites (rings). Cultures were maintained in RPMI-1640 media supplemented with 5% human serum at 5% hematocrit and incubated at 37 ºC in an atmosphere of 5% carbon dioxide (CO_2), 5% Oxygen (O_2) and 90% Nitrogen (N_2). Parasite growth was evaluated by monitoring parasitemia every 24 hours using microscopy of Giemsa-stained smears. Cultures with a high percentage of rings was treated with 5% sorbitol to obtain synchronous cultures for gametocyte induction and RNA and DNA extraction.

To induce gametocyte production in MLW5 and NF54, cultures were allowed to grow until the parasitemia was between 4 – 10%. Thereafter, cultures were diluted to a parasitemia of 0.7% if mostly ring stage parasites; 0.6% if mixed; and 0.5% if mostly trophozoites. Diluted cultures were maintained in complete media without gentamicin and the media was changed every day for 15 days preferably at the same time of the day. Seven days after induction, cultures were checked for early-stage gametocytes (stage II

&III) using microscopy. On days 13, 14, and 15 post induction, cultures were checked for the presence of mature gametocytes.

Mature gametocytes from the distinct isolates and MRA-1183 rings were quantified and serially diluted in human whole blood and stored in RNA Protect (Qiagen, CA). DNA was extracted from mature gametocytes from distinct *P. falciparum* isolates using the QIAamp DNA blood kit (Qiagen, CA). RNA was purified using the RNeasy Plus 96 kit (Qiagen, CA) according to the manufacturer`s instructions.

Mature gametocyte mixtures

RNA extracted from the mature gametocyte stocks (100 gametocytes/µL) from each of the two distinct strains (MLW5 and NF54) were mixed in duplicate in different proportions to mimic RNA extracted from an infection with two unique gametocyte genotypes **(Table 1)**. DNA extracted from mature gametocyte stocks of the same two strains was also mixed at the same proportions described for the RNA mixtures.

Table 1: NF54 and MLW5 mature gametocyte mixed at different ratios. RNA extracted from the mixtures was genotyped to estimate the number of gametocyte clones and their relative clone frequency. DNA extracted from the same mixtures was genotyped to estimate the number of clones among all parasite stages in the infection and their relative frequency.

Mature gametocyte distinct clones	Clone proportion in mixture
NF54: MLW5	1:0
NF54: MLW5	10:1
NF54: MLW5	5:1
NF54: MLW5	2:1
NF54: MLW5	1:1
NF54: MLW5	1:2
NF54: MLW5	1:5
NF54: MLW5	1:10
NF54: MLW5	0:1

NF54 = *P. falciparum* NF54 laboratory strain
MLW5 = Malawian *P. falciparum* field isolate

Sequencing library preparation

A primary PCR was performed on cDNA from the mixtures of mature gametocytes using primers (including 5' overhang sequences) targeting the gametocyte genotyping marker. A reverse transcriptase negative control and PCR negative and positive controls were included in all PCR runs. To allow multiplexing by sample, a secondary PCR (i.e., index PCR) was performed with primers targeting overhang sequences on primary PCR primers to attach unique barcodes and sequencing adapters (IDT for Illumina Nextera DNA UD Indexes Set A, B, C and D, San Diego, CA) **(Figure 4)**. The primary RT–PCR and index PCR primer sequences, reaction mixes, and cycling conditions are provided in Appendix Table B and C. Pooled barcoded amplicons were purified using the AMPure XP bead kit (Beckman, IN) and underwent sequencing using 300 bp paired-end reads on an Illumina MiSeq (Illumina). Sequencing runs included an Enterobacteria phage PhiX control (Illumina, PhiX Control v3) for quality control. The same primary and index PCR was performed on genomic DNA extracted from the mature gametocyte mixtures.

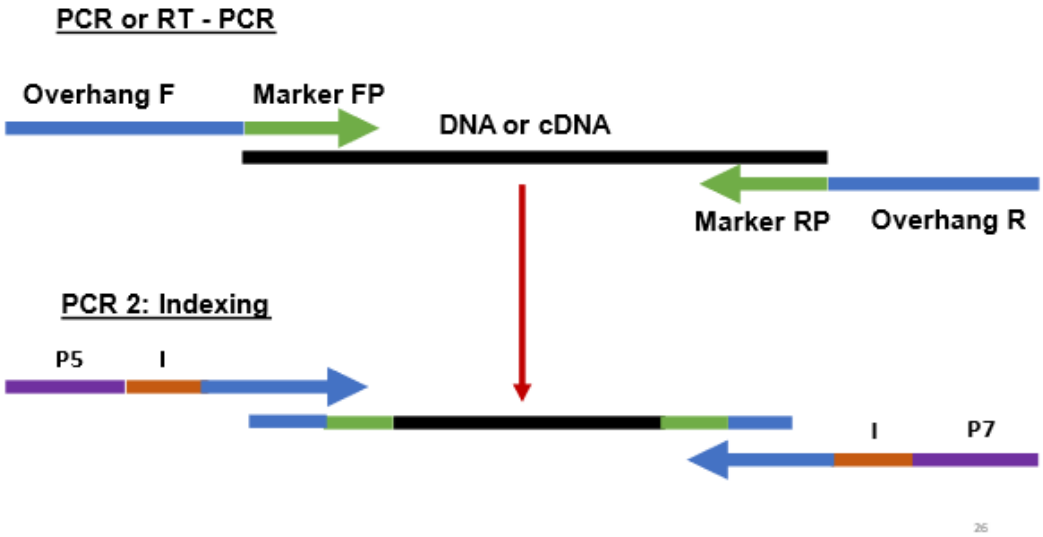

Figure 4: Two step PCR to generate amplicons for sequencing. The first PCR (DNA) and RT-PCR (RNA) targets the genotyping markers sequence and primers contain overhang sequences at their 5` ends. The second PCR targets the amplicons generated from PCR 1. Primers for the second PCR are complementary to the overhang sequences from PCR/RT-PCR 1 and unique indexes and sequencing adapters are attached to their 5` ends.

Processing of sequencing reads

The SeekDeep [94] software package with default settings was used to demultiplex sequencing reads; trim adapter, overhang, primer and barcode sequences from the reads; assess read quality and remove poor quality reads; and call haplotypes. Trimmed paired-end reads were merged and clustered based on unique haplotypes after excluding likely chimeric reads and reads with insertions and deletions. Each unique haplotype generated from RT-PCR amplicons represents a distinct gametocyte clone and each unique haplotype generated from PCR amplicons represents a distinct clone amongst all parasite life stages within the infection (in this case the experimental gametocyte mixture). Clone frequency was estimated based on the proportion of

sequencing reads corresponding to a given haplotype out of the total number of reads from all clones in the infection/mixture.

Limit of detection of the genotyping marker

A total of 101 *P. falciparum* gametocyte positive blood samples from the ICEMR Malawi household clustered cohort study (described above) selected non-randomly to represent a range of gametocyte densities were used to determine the limit of detection (LOD) of the genotyping assay. RT-PCR targeting the candidate gametocyte genotyping marker was performed on RNA extracted from the gametocyte positive samples and amplicons were sequenced as described above.

Genotyping DNA and RNA in field samples

The genotyping assay was applied to genotype DNA and RNA from 594 gametocyte-positive infections from the ICEMR cohort study (described above) to compare parasite haplotypes present in DNA (all circulating parasite stages) and RNA (circulating gametocytes) from the same infections. From among these infections, 209 RNA samples were successfully genotyped and 161 had corresponding DNA genotypes from the same infection, for a total of 161 paired DNA and RNA samples corresponding to 102 individuals.

Data analysis

To evaluate whether the genotyping approach can estimate gametocyte clone frequency in a quantitative manner, a concordance correlation coefficient (CCC) [95] was used to estimate the agreement between observed haplotype frequencies in gametocyte mixtures after sequencing and expected haplotype frequencies based on the ratios used to prepare mature gametocyte mixtures. The CCC in our study estimates agreement between the expected and observed haplotype abundance, with perfect concordance corresponding to the 45° line in a scatterplot **(Figure 5)**. The CCC range from -1 to 1: -1 representing perfect discordance between the two measurements; and 1 representing perfect agreement between the measurements. A CCC of ≥ 0.9 is considered good agreement and the cutoff used in this study.

A probit regression model was used to determine the LOD of the gametocyte genotyping assay in field isolates by estimating male and female gametocyte densities at which the probability of a sample being successfully genotyped is 0.95. All statistical analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, North Carolina) and R (version 4.2.3) statistical software packages.

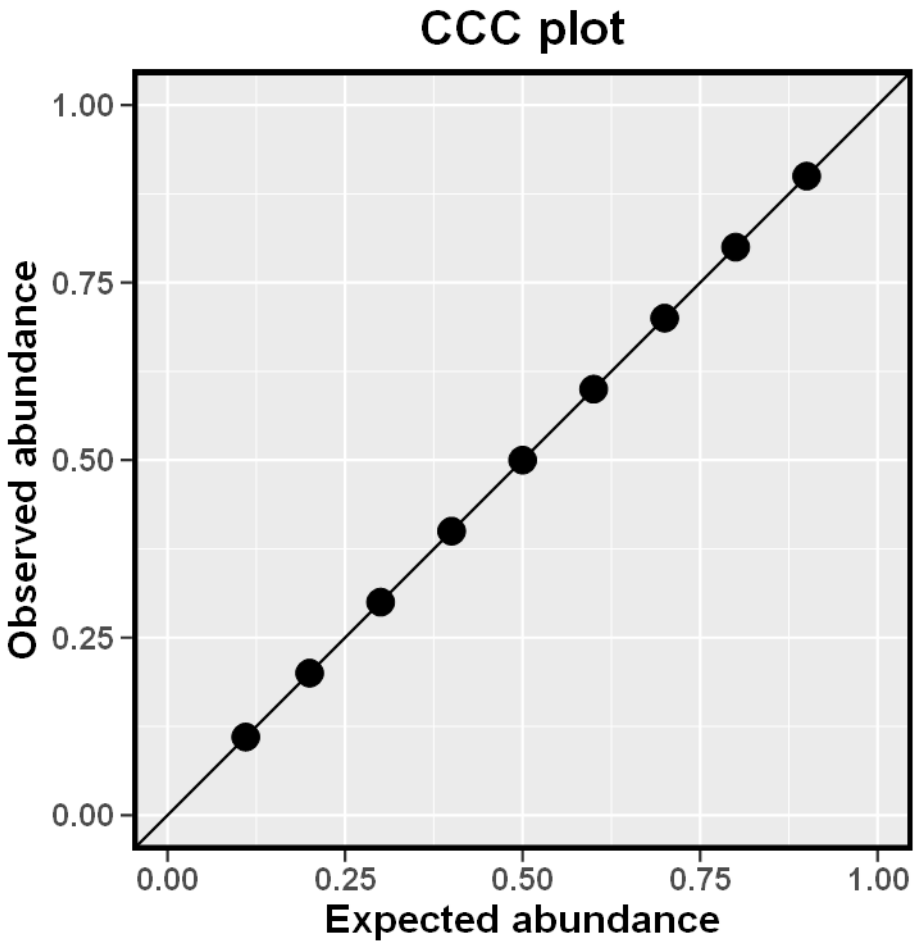

Figure 5: A scatter plot demonstrating perfect agreement between expected and observed haplotype abundance. All measurements (observed and expected) fall on the 45° line indicating a concordance correlation coefficient (CCC) of 1.

Aim 2 methods.

Study design and population

This specific aim used *P. falciparum* blood samples collected as part of the Malawi ICEMR (NIH grant number: U19A1089683) household-clustered cohort study described above. The ICEMR cohort study recruited 962 individuals aged > 6 months in 96 households from each health center catchment area. All individuals contributed 7,976

samples, of which 2138 were found to be *P. falciparum* positive. Of these, 1771 underwent gametocyte detection and quantification, *ap2-g* gene expression RT-PCR and *P. falciparum* genotyping (367 samples were not shipped to UMSOM).

Gametocyte Detection and Quantification

To detect and quantify *P. falciparum* gametocytes, RNA extracted from all *P. falciparum* positive samples underwent RT-qPCR using primers targeting the *ccp4* (female gametocyte marker) and *mget* (male gametocyte marker) genes as described above [20].

ap2-g Gene Transcript Abundance.

Two delta-delta CT (2^-ΔΔCT) relative quantification RT-PCR was used to quantify *ap2-g* gene transcript abundance in *P. falciparum* positive samples. Briefly, cDNA was synthesized from remaining RNA (after gametocyte detection) and RNA extracted from ring stage parasites from a strain of *P. falciparum* that does not produce gametocytes and expresses very little *ap2-g* gene (MRA-1183 SenTh026.04) [92] using QuantiTect Reverse Transcription Kit (Qiagen, CA). The cDNA from the test samples and from MRA-1183 was used in a PCR targeting the *ap2-g* gene and the *sbp1* gene (ring stages marker) as reference on a Light Cycler 96 Real-time PCR machine (Roche). The following specific primers were used to target the *ap2-g* gene transcripts: forward primer 5`-AAC AAC GTT CAA TAA ATA AGG-3` and reverse primer 5`-ATG TTA ATG TTC CCA AAC AAC CG-3` [45]. The following specific primers were used to target the *sbp1*

reference gene: Forward primer: 5` -GCA AAA CAA GCC GTA CAT GTT G – 3` and reverse primer: TTG CTA GGT AAT ATC CTT TTC TTT TTC C – 3` [50].

Parasite Genotyping by Amplicon Sequencing

To estimate COI, DNA extracted from *P. falciparum* positive samples was genotyped using the assay described in chapter 3 of this dissertation [Vareta JA. *et al*., under revision for *Malaria Journal*]. In brief, a PCR was performed targeting a polymorphic region in the *pfs230* gene. A second PCR was done to add unique barcodes and sequencing adapters to amplicons from the first PCR to allow for multiplexing of samples for sequencing. Barcoded amplicons from all samples were pooled and sequenced using an Illumina MiSeq (Illumina). SeekDeep [94] was used to demultiplex samples, cluster unique reads, and determine the number of haplotypes (number of clones) present in each infection as described above.

Definitions

The primary outcome of interest was gametocyte production quantified by measuring transcript abundance of the *ap2-g* gene in ring stages of *P. falciparum* positive samples. *ap2-g* transcript abundance was defined as fold change in expression of the *ap2-g* gene in test samples compared to its expression in ring stage parasites in a non-gametocyte producing control strain of *P. falciparum* (MRA-1183 SenTh026.04) [92]. *P. falciparum* positive samples were classified as having ring stage parasites if they had

measurable *sbp1* gene expression via RT-PCR. Another outcome of interest was gametocytemia which was quantified by summing female gametocytes density measured by expression level of *ccp4* gene and male gametocyte density measured by expression levels of *mget* gene. Predictors of interest included COI, age, season when sample was collected, symptoms, and history of antimalarial treatment. COI was defined as the number of distinct haplotypes observed per infection based on genotyping results from deep sequencing of amplicons from PCR targeting the *pfs230* gene. Due to the distribution of COI variable, COI was categorized into infections with one clone (clonal infection); infections with 2 to 3 clones; and infections with more than 3 clones. The age of study participants was categorized into three groups: young children (< 6 years of age); school-age children (SAC) (6 - 15 years old); and adults (≥16 years). Season when the samples were collected was categorized as dry season (May - November) and rainy season (December - April). History of antimalarial treatment was defined as reported antimalarial drug treatment in the past month prior to each visit (yes/no). Symptoms of malaria was defined as reported symptoms (any symptoms including fever in the last 48 hours) at the health center and axillary temperature of ≥ 37.5 at the time of visit (yes/no).

Data analysis

The proportion of ring stage infections that had *ap2-g* gene expression and the proportion of infections containing gametocytes were computed within strata of predictors and covariates. Bivariate analyses of predictors and *ap2-g* gene expression (yes/no) and gametocyte carriage (yes/no) were performed using unadjusted logistic

regression while accounting for clustering with household and individual. Comparison of *ap2-g* transcript abundance and gametocyte density within strata of predictors was done using unadjusted mixed effects linear regression while accounting for clustering within household and individual. In the final adjusted models, two multilevel models of the association between predictors and outcomes were performed: 1) A mixed effects logistic regression model to evaluate the association between predictors and *ap2-g* gene expression (expressed or not) and gametocyte carriage (gametocytes present or not); and 2) a mixed effects linear model to evaluate the association between predictors and *ap2-g* gene transcript abundance in infections that expressed *ap2-g* and gametocyte density in infections that carried gametocytes. *ap2-g* gene transcript abundance and gametocyte density in the linear mixed effects model were log transformed to fulfill normality assumption of continuous outcome variables. Both models accounted for clustering at household and individual level and were adjusted for gender, gametocyte sex ratio (in gametocyte carriage and gametocyte density models) and parasite density. All statistical analyses were done in SAS 9.4 statistical software (SAS Institute, Cary, NC) at alpha level of 0.05.

CHAPTER III. GENOTYPING *PLASMODIUM FALCIPARUM* GAMETOCYTES USING AMPLICON DEEP SEQUENCING

ABSTRACT

Background

Understanding the dynamics of gametocyte production in polyclonal *Plasmodium falciparum* infection requires a genotyping method that detects distinct gametocyte clones and estimates their relative frequencies. Here, a marker was identified and evaluated to genotype *P. falciparum* mature gametocytes using amplicon deep sequencing.

Methods

A data set of polymorphic regions of the *P. falciparum* genome was mined to identify a gametocyte genotyping marker. To assess marker resolution, number of unique haplotypes in the marker region was estimated from 205 Malawian *P. falciparum* whole genome sequences. Specificity of the marker for detection of mature gametocytes clones was evaluated using reverse transcription-polymerase chain reaction of RNA extracted from NF54 mature gametocytes and rings from a non-gametocyte-producing strain of *P. falciparum*. Amplicon deep sequencing was performed on experimental mixtures of

mature gametocytes from two distinct parasite clones, as well as gametocyte positive *P. falciparum* field isolates to evaluate the quantitative ability and determine the limit of detection of the genotyping approach.

Results

A 400 bp region of the *pfs230* gene was identified as a gametocyte genotyping marker. A larger number of unique haplotypes was observed at the *pfs230* marker (53) compared to *sera-2* (23) and *ama-1* (18) markers in field isolates from Malawi. RNA and DNA genotyping accurately estimated gametocyte and total parasite clone frequencies when evaluating agreement between expected and observed haplotype frequencies in gametocyte mixtures, with concordance correlation coefficient of 0.97 [95% CI: 0.92 – 0.99] and 0.92 [95% CI: 0.83 – 0.97], respectively. The detection limit of the genotyping method for male gametocytes was 1.64 gametocytes/µl [95% CI: 1.12 – 2.88] and for female gametocytes was 19.42 gametocytes/µl [95% CI: 13.23 – 36.06], owing to greater expression of the marker gene in male compared to female gametocytes.

Conclusions

A region of the *pfs230* gene was identified as a marker to distinguish *P. falciparum* gametocytes. Amplicon deep sequencing of this marker can be used to estimate the number and relative frequency of parasite clones among mature gametocytes within *P. falciparum* infections. This gametocyte genotyping marker will be an important tool for studies aimed at understanding dynamics of gametocyte production in polyclonal *P. falciparum* infections.

BACKGROUND

During the erythrocytic stage of the malaria parasite in humans, a fraction of asexual parasites differentiates into gametocytes [11]. Gametocytes are the sexual form of the parasite transmitted from humans to mosquitoes and are an important target for interventions that interrupt human-to-mosquito transmission [10, 11]. Polyclonal *Plasmodium falciparum* infections are prevalent in high transmission settings [72, 74, 75, 96]. As multiple different asexual parasite clones are circulating and multiplying in the peripheral bloodstream, the contribution of each clone to the total gametocytemia and the factors that impact the relative gametocytemia of different clones are not clear. Understanding the dynamics of gametocytemia corresponding to different clones in the context of polyclonal *P. falciparum* infections will help identify parasite and/or host factors that influence gametocyte production during natural infection. Such knowledge will fill a gap in our understanding of gametocyte biology, which may inform development of malaria prevention strategies targeting human-to-mosquito transmission of the parasite.

Identifying the specific parasite clones that have produced gametocytes in a polyclonal *P. falciparum* infection requires a genotyping assay that distinguishes clones in both asexual parasites and mature gametocytes. Development of such an assay requires identification of a polymorphic region of a gene primarily expressed in mature gametocytes to serve as a genotyping marker. Identification of gametocyte-specific markers for genotyping has been challenging, as many genes are not exclusively expressed in gametocytes [88], and those that are, often display relatively limited genetic diversity [97]. Genotyping of gametocyte-specific markers will allow estimation of clone

composition among gametocytes (by genotyping RNA) and all parasite blood stages (by genotyping genomic DNA) in polyclonal *P. falciparum* infections. Previous gametocyte genotyping methods have utilized length-polymorphism markers in genes encoding proteins such as *osmiophilic body protein* (*pfg377*) and *6-cysteine protein P230* (*pfs230)* [33, 81-83, 98]. Such length-polymorphism markers are prone to amplification bias towards smaller fragments and rely on gel electrophoresis or capillary electrophoresis to determine fragment sizes, resulting in limited sensitivity, low throughput and resolution, and inability to quantify relative clone frequencies [85, 99, 100].

With increased access to short-read sequencing technologies, researchers are performing deep sequencing of small amplicons representing regions of high heterozygosity in *P. falciparum* antigen genes to estimate infection complexity and parasite genetic relatedness, and to track malaria transmission dynamics [75, 85, 99, 101-105]. The large read depth of amplicon deep sequencing allows for an increased ability to detect low frequency clones (as low as 1%) and quantification of relative clone frequencies in an infection based on the proportion of reads corresponding to a given haplotype [85, 86]. Furthermore, multiplexed sequencing increases throughput with the capability of sequencing up to 384 samples in one sequencing run using the Illumina MiSeq platform. Although amplicon deep sequencing is now readily used to genotype *P. falciparum* infections in field settings, it has not been adapted to genotype gametocytes in *P. falciparum* infections. Therefore, this study was designed to develop an amplicon deep sequencing approach to genotype *P. falciparum* mature gametocytes in natural infections.

METHODS

Identification of the candidate gametocyte genotyping markers

Genes expressed primarily in *P. falciparum* mature gametocytes were identified based on stage specific single cell RNA-sequencing data [88]. Polymorphic regions of these genes were identified by mining a published *P. falciparum* microhaplotypes dataset containing 4465 highly polymorphic genomic regions of 200 bp in length [89] for characterization as potential candidate gametocyte genotyping markers. To improve genotyping resolution, overlapping candidate marker regions were combined (when possible) to obtain larger amplicons of a size amenable to coverage by Illumina MiSeq paired-end reads. One region that was the most diverse among the identified polymorphic regions was selected as the candidate gametocyte genotyping marker. To evaluate potential haplotype resolution of the candidate genotyping marker, the number of unique haplotypes was estimated at the marker locus from whole genome sequences (WGS) data generated from 205 Malawian *P. falciparum* field isolates [90, 106]. For comparison, the number of unique haplotypes at two loci widely used for blood stage amplicon sequencing (*ama-1* and *sera-2*) [84, 91] was also estimated from the same isolates.

Specificity of the candidate genotyping markers for detection of mature gametocytes

To determine the specificity of the candidate gametocyte genotyping markers for detection of mature gametocytes (compared to asexual parasite stages), RNA was extracted from cultured ring stage parasites (10^5 rings/µL) of the MRA–1183 *P.*

falciparum SenTh026.04 strain (a non-gametocyte producing line obtained from BEI Resources) [92] and mature gametocytes (10 gametocytes/µL) cultured from NF54 strain (a gametocyte producing laboratory strain) using RNeasy Mini kit (Qiagen, CA). Complementary DNA (cDNA) was synthesized from the RNA using QuantiTect Reverse Transcription kit (Qiagen, CA). PCR with primers targeting the candidate marker was performed on the cDNA using the KAPA HiFi HotStart ReadyMix PCR kit (Kapa Biosystems, Wilmington, MA). Primers were designed using Primer-BLAST [93] (Appendix **Table 9 and 10**). A higher concentration of MRA-1183 rings compared to NF54 gametocytes was used to reflect the higher asexual parasite density (compared to gametocyte density) often observed in natural *P. falciparum* infections [39], and to determine if any low-level expression of the candidate marker gene in rings could overestimate mature gametocyte genotypes. A region of the *pfs25* gene, a gametocyte detection marker known to be highly specific to female gametocytes [9], was amplified as a positive control for gametocyte specificity. All RT-PCRs were done with a reverse-transcriptase negative control to rule out DNA contamination in extracted RNA. PCR products from the RT-PCR were visualized using agarose gel electrophoresis.

Quantitative ability of genotyping assay in experimental gametocyte mixtures of distinct *P. falciparum* isolates

Mature gametocyte mixtures

Two distinct strains of *P. falciparum* (a Malawian field isolate (MLW5) and the NF54 strain) were cultured to obtain mature gametocytes. These strains differed at 2 sites

within the length of the candidate genotyping marker sequence. Mature gametocytes from each of the two parasite strains were counted and mixed with uninfected human whole blood to obtain a stock of 100 gametocytes/μL. RNA was extracted from mature gametocytes from each strain stock using the extraction method described above. Genomic DNA was also extracted from the same gametocyte stocks using a QIAamp DNA blood kit (Qiagen, CA). RNA extracted from the mature gametocyte stocks from each of the two strains were mixed in duplicate in different proportions (NF54: MLW5=1:0, 10:1, 5:1, 2:1, 1:1, 1:2, 1:5, 1:10, and 0:1) to mimic RNA extracted from an infection with two unique gametocyte genotypes. DNA extracted from mature gametocyte stocks of the same two strains was also mixed at the same proportions described for the RNA mixtures.

Sequencing library preparation

A primary PCR was performed on cDNA from the mixtures of mature gametocytes using primers (including 5' overhang sequences) targeting the gametocyte genotyping marker. A reverse transcriptase negative control and PCR negative and positive controls were included in all PCR runs. To allow multiplexing by sample, a secondary PCR (i.e., index PCR) was performed with primers targeting overhang sequences on primary PCR primers to attach unique barcodes and sequencing adapters (IDT for Illumina Nextera DNA UD Indexes Set A, B, C and D, San Diego, CA). The primary RT–PCR and index PCR primer sequences, reaction mixes, and cycling conditions are provided in Additional file (Table A1 and A2). Pooled barcoded amplicons were purified using the AMPure XP bead kit (Beckman, IN) and underwent sequencing

using 300 bp paired-end reads on an Illumina MiSeq (Illumina). Sequencing runs included an Enterobacteria phage PhiX control (Illumina, PhiX Control v3) for quality control. The same primary and index PCR was performed on genomic DNA extracted from mature gametocyte mixtures.

Processing of sequencing reads

The SeekDeep [94] software package with default settings was used to demultiplex sequencing reads; trim adapter, overhang, primer and barcode sequences from the reads; assess read quality and remove poor quality reads; and call haplotypes. Trimmed paired-end reads were merged and clustered based on unique haplotypes after excluding likely chimeric reads and reads with insertions and deletions. Each unique haplotype generated from RT-PCR amplicons represents a distinct gametocyte clone and each unique haplotype generated from PCR amplicons represents a distinct clone amongst all parasite life stages within the infection (in this case the experimental gametocyte mixture). Clone frequency was estimated based on the proportion of sequencing reads corresponding to a given haplotype out of the total number of reads from all clones in the infection/mixture.

Limit of detection of the genotyping marker

Study samples

Samples for determination of the limit of detection (LOD) of the genotyping assay were obtained from a household-based cohort study conducted as part of the Malawi

International Center of Excellence for Malaria Research (ICEMR). The study was conducted within the catchment areas of the Namanolo and Ntaja health centers in Balaka and Machinga districts, respectively, in southern Malawi from April 2019 to March 2020. The cohort study included 96 households, with all members of each household invited to participate in the study. Consent was sought from the head of each household, followed by individual level consent from other adults and assent from children >14 years of age. The cohort included 962 individuals who were followed actively every 4-6 weeks, yielding a maximum of ten visits over the 10-month study. Passive case detection was conducted by a study nurse at the health center, and malaria treatment provided by health center providers according to national guidelines. Blood samples were collected by finger prick at each active and passive case detection visit, with ~50 µL blotted onto filter paper (stored at room temperature with desiccant) for PCR detection of *P. falciparum* and 100 µL preserved in RNA Protect cell reagent (Qiagen, CA) (stored at -80 °C) for gametocyte detection.

Gametocyte detection

Gametocyte quantification for *P. falciparum*-positive samples was based on expression of the *ccp4* (female gametocyte) and *mget* (male gametocyte) genes, assessed using RT-qPCR as described previously [20]. In brief, RNA was extracted from each sample using the RNeasy Plus 96 kit (Qiagen). Five µL of RNA, and serially diluted standards, were used in a multiplex one-step RT-qPCR using the Luna Universal Probe One-Step RT-qPCR Kit (NEB) and probes and run on a Light Cycler 96 Real-time PCR machine (Roche). Standards were generated from target RNA copies transcribed *in vitro* using the MEGAShortscript T7 high yield transcription kit (Invitrogen) and were

quantified using Qubit 3 fluorometer (ThermoFisher). All RT-qPCR reactions were performed in duplicate, with a third PCR run as a tie breaker to resolve discordant results. Copy number per mL for each gene was estimated using the average cycle threshold (Ct) of each marker in samples and a linear regression equation of Ct values of the serially diluted standards and their respective copy number per mL. To obtain female and male gametocyte density, estimated test sample *ccp4* copy number per mL was multiplied by a factor of 4 (4 RNA copies per female gametocyte cell) and estimated test sample *mget* copy number was multiplied by a factor of 9.8 (9.8 RNA copies per male gametocyte cell) [20].

A total of 101 *P. falciparum* gametocyte-positive blood samples were selected to represent a range of gametocyte densities for determination of the limit of detection (LOD) of the genotyping assay. RT-PCR targeting the candidate gametocyte genotyping marker was performed on RNA extracted from the gametocyte-positive samples and amplicons were sequenced as described above.

Genotyping DNA and RNA in field samples

The genotyping assay was applied to genotype DNA and RNA from 594 gametocyte-positive infections from the ICEMR cohort study (described above) to compare parasite haplotypes present in DNA (all circulating parasite stages) and RNA (circulating gametocytes) from the same infections. From among these infections, 209 RNA samples were successfully genotyped and 161 had corresponding DNA genotypes

from the same infection, for a total of 161 paired DNA and RNA samples corresponding to 102 individuals.

Data analysis

To evaluate whether the genotyping approach can estimate gametocyte clone frequency in a quantitative manner, a concordance correlation coefficient (CCC) [95] was used to estimate the agreement between observed haplotype frequencies in gametocyte mixtures after sequencing and expected haplotype frequencies based on ratios used to prepare mature gametocyte mixtures. A probit regression model was used to determine the LOD of the gametocyte genotyping assay in field isolates by estimating male and female gametocyte densities at which the probability of a sample being successfully genotyped is 0.95. Statistical analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, North Carolina) and R (version 4.2.3).

RESULTS

Selection of the gametocyte genotyping marker and its specificity to mature gametocytes

Two polymorphic regions in a gene encoding *pfs230* (Pf3D7_0209000, coordinates: 373786 – 374186 and 375686- 376086) and one polymorphic region in a gene encoding *pfg377* (*Pf*3D7_1250100, coordinates: 2050053 - 2050353) were identified as candidate mature gametocyte genotyping markers. When evaluating the diversity of the three candidate genotyping markers in 205 WGS data generated from Malawian *P. falciparum* isolates, the candidate markers had similar or greater unique

haplotype number (*pfs230*: 51 and 53, *pfg377*:24) compared to *sera-2* and *ama-1* blood stage marker loci [84, 91], which had 23 and 18 unique haplotypes in the same WGS data set, respectively **(Table 2)**. RT–PCR of the candidate markers in NF54 mature gametocytes and MRA–1183 *P. falciparum* SenTh026.04 ring stage parasites showed amplification of NF54 mature gametocytes and no amplification of MRA-1183 ring stage parasites, despite ring stage parasitemia being four orders of magnitude greater than the gametocytemia in the test sample, suggesting that the markers are specific to mature gametocytes **(Figure 6)**. A 400 bp polymorphic region of the gene encoding *pfg377* (Pf3D7_0209000, coordinates:375686 - 376086) out the three identified markers was selected as a marker for the gametocyte genotyping assay because it detected more haplotypes in WGS data set generated from Malawian *P. falciparum* isolates compared to the other two markers, and the gene spanning this marker is expressed in both male and female gametocyte, with greater expression in males than females, allowing for high sensitivity compared to the *pfg377* gene marker (Pf3D7-1250100) which is expressed in female gametocytes only [107].

Table 2: Number of variable sites and unique haplotypes detected by three candidate gametocyte genotyping markers and commonly used blood stage short amplicon markers (*sera-2* and *ama-1*) in whole genome sequences from 205 *P. falciparum* field isolates collected in Malawi.

Gene ID	Gene Name	Marker Start	Marker End	Variable Sites	Number of haplotypes	Marker Size
Pf3D7_1250100	*pfg377*	2050053	2050353	5	24	300
Pf3D7_0209000	*Pfs230*	373786	374186	6	51	400
Pf3D7_0209000	*pfs230*	375686	376086	11	53	400
Pf3D7_0207900	*sera-2*	320763	321020	5	23	257
Pf3D7_1133400	*ama-1*	1294056	1294456	5	18	400

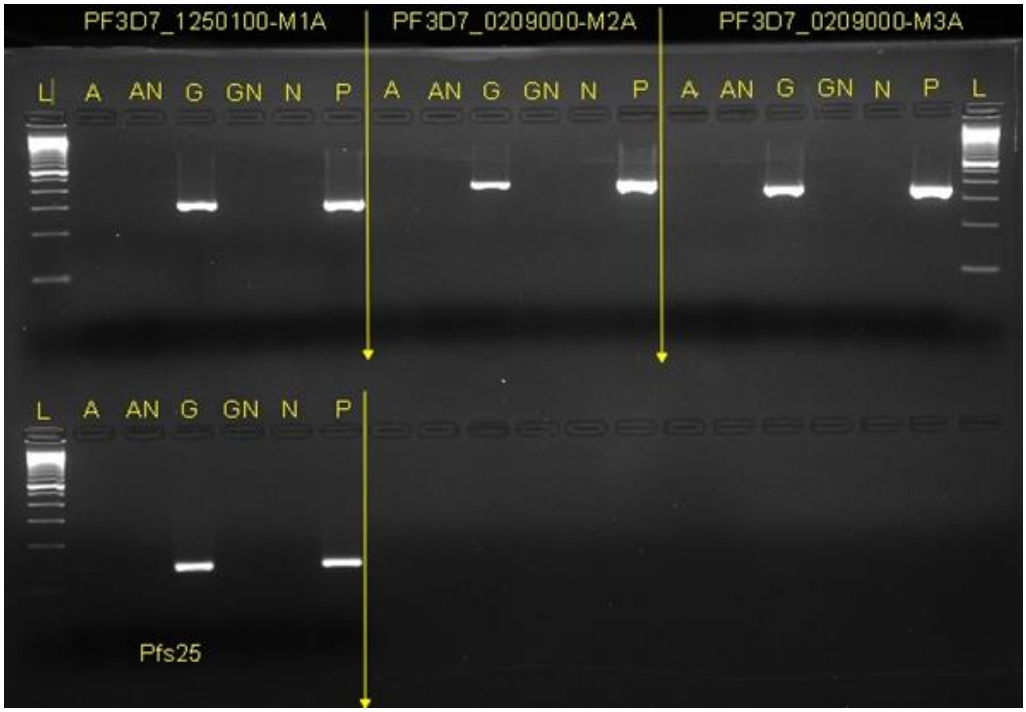

Figure 6: Specificity of the *pfg377* and *pfs230* genes genotyping markers to mature gametocytes. Lane A in the *pfg377*, *pfs230* and *pfs25* panel represents RT-PCR performed on RNA extracted from a high-parasite density culture (100,000 parasites/µL) of a non-gametocyte-producing *P. falciparum* strain MRA-1183, with corresponding RT negative control (AN). Lane G represents the same RT-PCRs on RNA extracted from NF54 gametocytes (10 gametocytes/µL) with corresponding RT negative (GN). Lane N and P represent the PCR negative and positive controls, respectively. Lane L represents 100bp DNA ladder.

Estimating clone frequency among gametocytes and total parasites.

The *pfs230* genotyping marker locus was amplified from DNA and RNA extracted from mature gametocyte mixtures, and the amplicons were sequenced. Sequencing coverage ranged from 4000 to 163316 reads prior to read filtering. After filtering erroneous reads, sequencing coverage ranged from 2819 to 37445 reads. The observed proportions of reads of the two isolates in the DNA and RNA gametocyte

mixtures after filtering are shown in **Figure 7A**. The expected proportion of reads (proportions used to prepare gametocyte mixtures) for each isolate in the mixtures was compared to its corresponding observed proportion of reads. Agreement between expected and observed haplotype frequencies was high in both RNA (CCC = 0.97;95% CI: 0.92 – 0.99) and DNA (CCC = 0.92;95% CI: 0.83 – 0.97) mixtures (**Figure 7B and 7C**). These results suggest that the genotyping assay accurately estimates relative clone frequency of gametocytes (when using RNA) and total parasite clone frequency (when using DNA) in polyclonal infections. We detected a minority clone in the Malawian field isolate that was previously undetected in the whole genome sequencing data generated from this isolate, a further testament of the increased sensitivity of amplicon deep sequencing for detection of minority clones (unknown haplotype in **Figure 7A**).

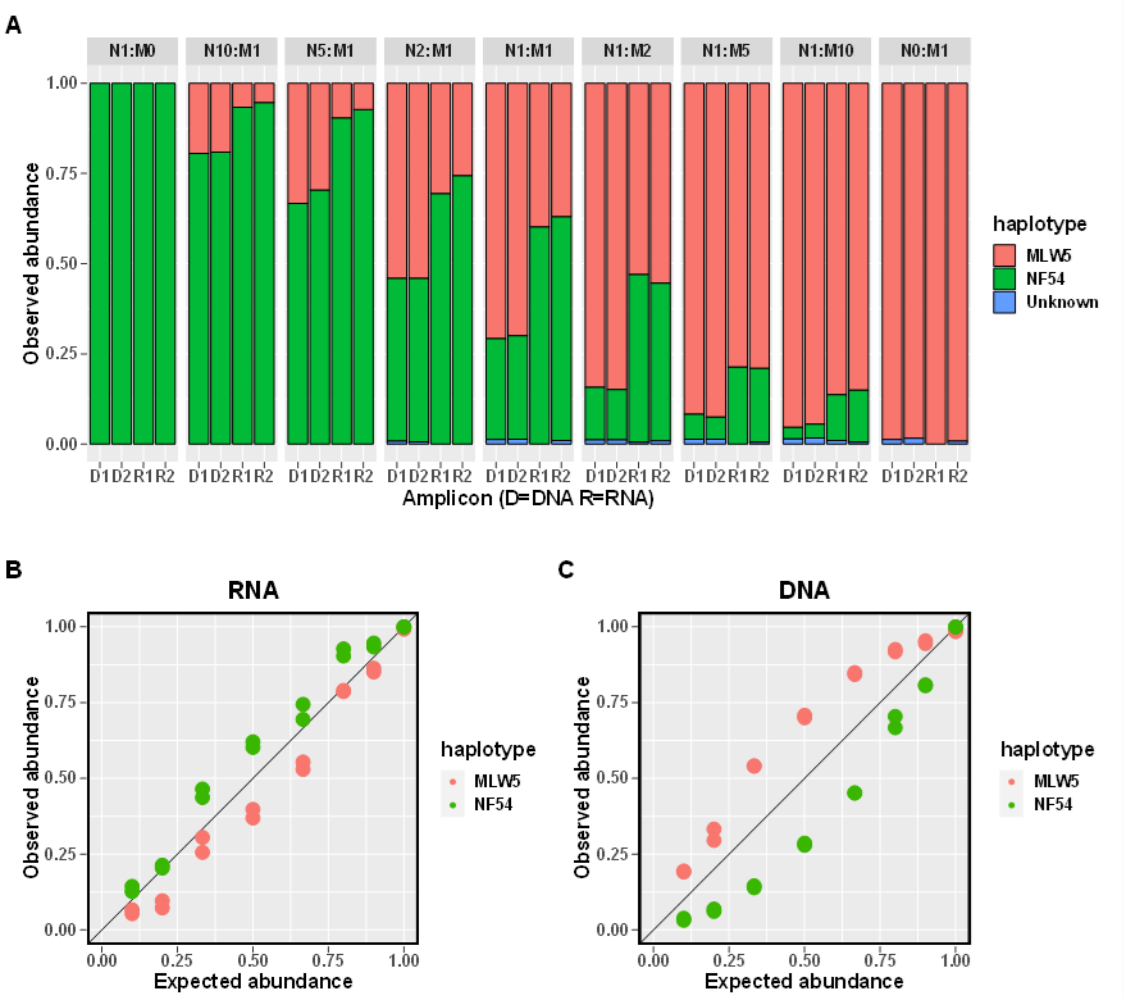

Figure 7: Agreement between expected clone abundance (gametocyte mixture ratios) and observed abundance (clone frequency as estimated by the proportion of reads corresponding to a given clone) of RNA and DNA extracted from NF54 and MLW5 mature gametocytes mixtures. **A)** Stacked bar graphs depicting observed clone frequency after amplicon sequencing of mature gametocyte mixtures of NF54 (N) and MLW5 (M) strains DNA (D1 and D2 are replicates) and RNA (R1 and R2 are replicates) were mixed in the following ratios: 1:0, 10:1, 5:1, 2:1, 1:1, and the reverse. **B and C)** The observed RNA and DNA clone frequencies for each clone are plotted against the expected frequencies (black dotted diagonal line) and were found to be highly concordant (RNA CCC = 0.97; DNA CCC = 0.92).

Limit of detection of the genotyping assay

To determine the LOD of the genotyping assay, cDNA from 101 gametocyte positive samples underwent PCR followed by sequencing. After applying SeekDeep extractor filtering to the raw reads, two populations of reads were observed: one ranging from 0 to 15 per sample and another ranging from 10000 to 100000 per sample (Figure 8). Samples with filtered reads ranging from 0 to 15 were assigned as genotyping failures and samples with filtered reads ranging from 10000 to 100000 were assigned as successfully genotyped. The distribution of total and filtered read counts by sample and by female and male gametocytes density are shown in **Figure 8 A-D**.

Out of 101 RNA samples sequenced, 48 (47%) were successfully genotyped. Fourteen unique haplotypes were observed among the 48 successfully genotyped samples: 27 samples (26.7%) with a single gametocyte haplotype; 18 (17.8%) with 2 gametocyte haplotypes; and 3 (3%) with 3 gametocyte haplotypes. Median male gametocyte density in samples that were successfully genotyped was 0.67 gametocyte/µL [IQR: 0.21 – 1.02] compared to 0.02 gametocytes/µl [IQR: 0.008 – 0.09] in samples that failed sequencing. Median female gametocyte density in samples that were successfully genotyped was 5.58 gametocyte/µL [IQR: 1.52 – 14.02] compared to 0.32 gametocytes/µL [IQR: 0.03 – 1.45] in samples that failed sequencing. Using probit regression, the LOD of the genotyping method based on male gametocyte density was estimated to be 1.64 gametocytes/µL [95% CI: 1.12 – 2.88] and based on female gametocyte density, 19.42 gametocytes/µl [95% CI: 13.23 – 36.06].

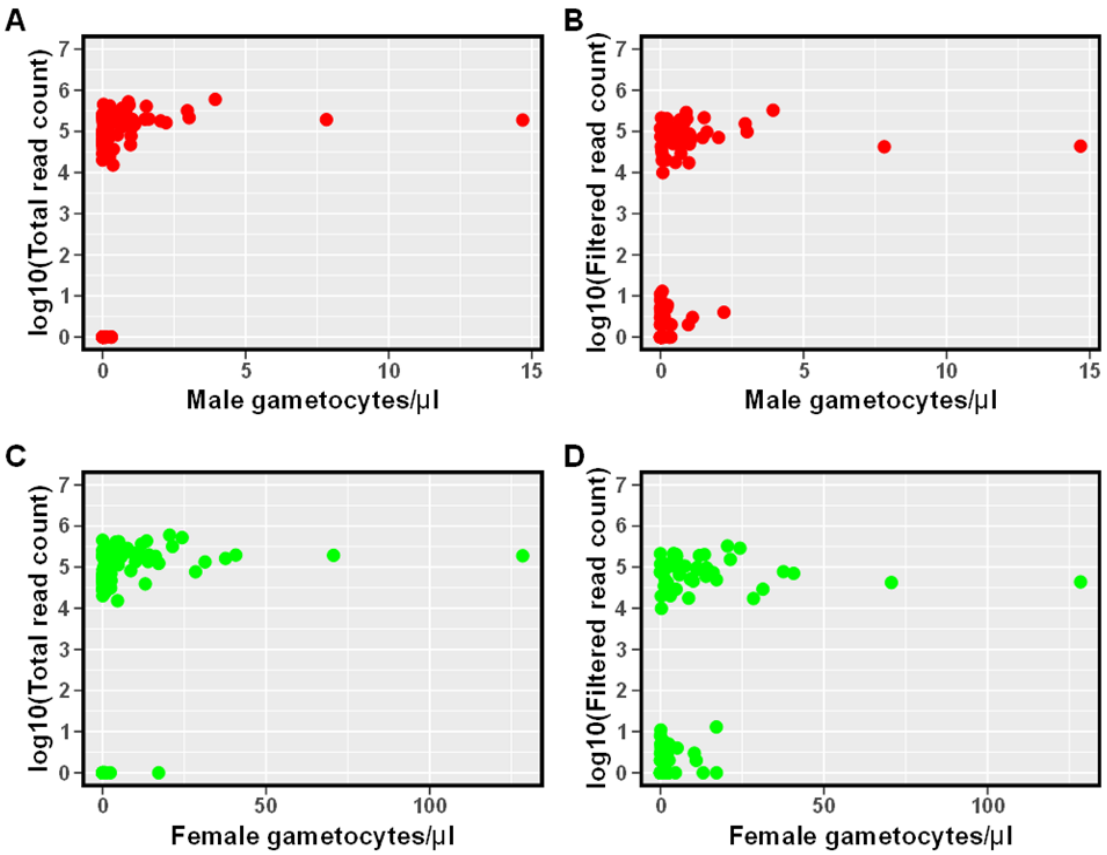

Figure 8: Scatter plots of total and filtered reads by male and female gametocyte density. **(A and B)** Scatter plots of raw and filtered read count by male gametocytes density. **(C and D)** Scatter plots of raw and filtered read count by female gametocytes density.

Concordance of haplotypes between DNA and RNA genotyping

Amplicon sequencing with the *pfs230* marker was performed on DNA and RNA extracted from 594 gametocyte-positive *P. falciparum* infections from the ICEMR cohort study to determine the haplotypes present among all circulating parasite stages (DNA genotyping) and among circulating gametocytes (RNA genotyping) within each infection. From these infections, 209 RNA samples were successfully genotyped, with genotyping success consistent with the estimated LOD of the assay, and 161 infections had both RNA

and DNA genotypes available. Eighty-eight (54.7%) of the 161 infections were polyclonal. Thirty-nine unique haplotypes were observed among the RNA and DNA genotypes. Fifteen infections (9.3%) had a discordant haplotype identified in the RNA sample that was not observed in its paired DNA sample. The mean COI among all parasite stages (DNA samples) was 3.36 [range:1 - 8] and in gametocytes (RNA samples) was 1.67 [range:1 – 4] (Table 3). The prevalence of each haplotype was similar in DNA and RNA samples (Figure 9), with the three most prevalent haplotypes being the same between all parasites (DNA) and gametocytes (RNA).

Table 3: Performance of the gametocyte genotyping assay in 161 *P. falciparum* gametocyte positive paired DNA and RNA field samples.

	Paired DNA/RNA samples (n=161 infections, 102 individuals)
Unique haplotypes in all samples (DNA/RNA)	39
Unique haplotypes among all parasite stages (DNA samples)	31
Unique haplotypes among gametocytes (RNA samples)	28
Polyclonal infections	88 (54.7%)
Infections with an RNA haplotype not observed among DNA haplotypes	8 (4.9%)
Mean COI among all parasite stages (DNA samples)	3.36 [range: 1 – 8]
Mean COI among gametocytes (RNA samples)	1.67 [range: 1 – 4]

COI = complexity of infection

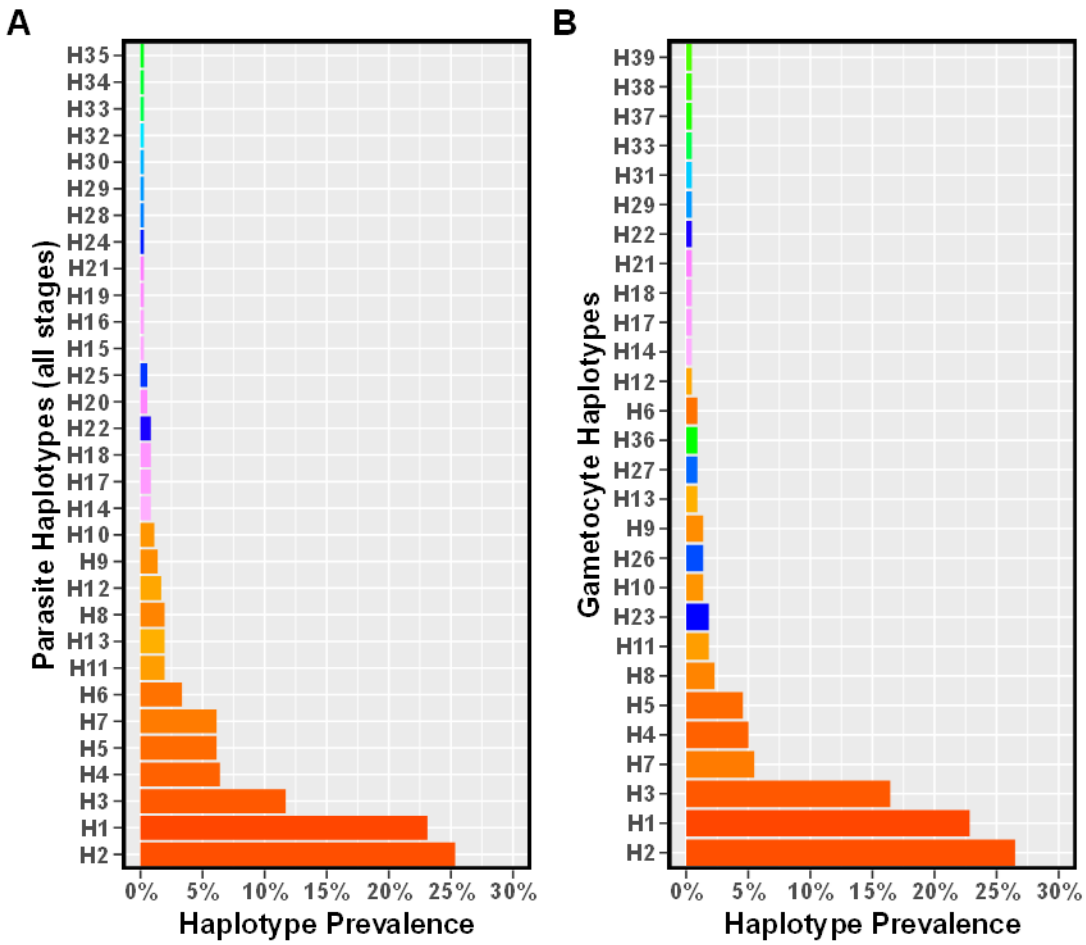

Figure 9: Frequency of unique haplotypes in 161 gametocyte-positive *P. falciparum* infections among **A)** all parasite stages in the infection (DNA samples) and **B)** gametocytes in the infection (RNA samples). Haplotypes are shown on the y-axis, with each haplotype represented by a unique-colored bar. The prevalence of the haplotype among the samples is shown on the x-axis.

DISCUSSION

In this study, we developed an assay to genotype mature gametocytes in *P. falciparum* infections through deep sequencing of RT-PCR amplicons generated from a polymorphic region of the *pfs230* gene. The genotyping assay is specific to mature gametocytes, is quantitative, and has higher haplotype resolution than known *P.*

falciparum blood stage amplicon-based genotyping markers. Furthermore, the assay performs well in field samples with low gametocyte density. The approach provides a useful tool to advance our understanding of how infection complexity and clone composition impact gametocyte production in *P. falciparum* infections. This knowledge will provide further insights into parasite transmission biology, which may inform development of transmission-interrupting interventions.

Notwithstanding the limited genetic diversity of many genes highly expressed in mature gametocytes [97], we identified a polymorphic region of the *pfs230* gene as a marker to genotype gametocytes in *P. falciparum* infections. Since the genotyping marker was selected from a published data set of microhaplotypes extracted from WGS data from *P. falciparum* isolates representative of global malaria isolates [89], gametocyte genotyping using this marker should be applicable in a range of malaria-endemic settings. The specificity of the genotyping marker to mature gametocytes was an important consideration in the assay development. Most genes that are highly expressed in mature gametocytes are also expressed at low levels in other asexual parasite stages, including ring stages [88]. Therefore, low-level expression of a marker gene in ring stage parasites, if detectable by RT-PCR, may overestimate the number of clones represented among gametocytes (because clones present among asexual stages will also be detected), especially in infections with high parasitemia [20]. Here, our RT-PCR results showed that even at high ring stage parasitemia (10^5 rings/μL), there was no detectable expression of the marker gene. This finding demonstrates that the gametocyte genotyping marker is specific to mature gametocytes.

In our WGS data from 205 Malawian *P. falciparum* isolates, our new gametocyte genotyping marker was shown to have higher haplotype resolution than two widely used blood-stage amplicon sequencing markers, *sera-2* and *ama-1*. We observed a smaller number of unique haplotypes at both our new genotyping marker locus (53/205 isolates) and at the *sera-2* (23/205) and *ama-1* (18/205) loci in these Malawian isolates compared to prior studies that examined fewer samples [86]. The Malawian *P. falciparum* WGS data were generated from samples collected during a longitudinal cohort study and represent isolates collected from the same individuals over time (the 205 isolates were sampled from 95 individuals) [90, 106]. As a result, the sampled isolates likely do not represent independent infections and probably underrepresent the diversity that would be observed in independent samples from this parasite population. However, because all markers were evaluated using the same data set, we would expect diversity at all markers to be equally underrepresented, i.e., the relative diversity should not be biased.

Because of high read depth when using amplicon deep sequencing (read depth range was 2,819 to 37,445 reads, in our study), the ability to detect minor clones in polyclonal infections is improved [86]. High concordance (CCC >0.90) between observed and expected clone proportions for both RNA and DNA in the gametocyte mixtures suggest that the sequencing read depth was large enough to detect minor clones with frequency of 10% (the smallest dilution tested). We suspect that the lowest detectable frequency of minor clones is <10% based on our ability to detect a minor clone within the Malawian field isolate included in the experimental mixtures that has a frequency of <2% and was not detected using WGS data with lower average read depth (~150x). The ability to generate quantitative estimates of parasite clone frequency among

blood stage parasites and gametocytes within an infection may allow assessment of associations between parasite clone frequency (e.g., whether a given clone is a majority or minority clone within an infection) and parasite investment in sexual reproduction (although the contribution of gametocytes to total blood stage clone frequencies may need to be accounted for in such analyses). Furthermore, the ability to detect unique clones and their relative frequency could also allow identification of specific clones that more often produce gametocytes. Such parasites could then be characterized further to understand genetic determinants of increased gametocyte production.

A major challenge of genotyping gametocytes in field settings is that most infections have low gametocyte density, which can affect the number of samples that are successfully genotyped and the detection of minority clones in polyclonal infections [108]. This problem is compounded in submicroscopic infections, which are prevalent in many malaria endemic settings [109]. As we note, less than half of the gametocyte-positive field samples in our study were successfully genotyped, with genotyping success consistent with the LOD estimated for our assay. Although the LOD of our assay would likely allow genotyping of submicroscopic infections, it is not as sensitive as the RT-qPCR used to detect and quantify gametocytes in our study, a finding similar to that observed for other gametocyte genotyping markers [10]. We estimated that our assay can successfully genotype infections with male gametocyte density as low as 1.64 gametocytes/μL; however, the limit of detection based on female gametocyte density was higher at 19.42 gametocytes/μL. Prior gene expression studies have indicated that the *pfs230* gene has a 5-fold higher expression in male gametocytes than females [107]. However, *P. falciparum* infections are often female-biased [11, 12], with a ratio of

female to male gametocytes on the order of 2 – 5 females to 1 male or even higher depending on transmission intensity [110, 111]. The fact that many *P. falciparum* infections contain more female gametocytes may offset, to an extent, the lower expression of the genotyping marker gene in female gametocytes; however, where possible, downstream analyses should adjust results by gametocyte sex ratio.

We genotyped DNA and RNA extracted from the same infections to determine if observed haplotypes among all sampled parasite stages (DNA genotyping) are concordant with haplotypes observed among gametocytes (RNA genotyping). In a small proportion of infections (9.3%), haplotypes were observed in the RNA fraction (representing gametocytes) that were not observed in the DNA fraction (representing all parasites in the infection). This type of discordance has also been observed with other gametocyte genotyping markers, but in a 4-6 times larger proportion of infections than observed in our study, depending on the marker [10]. This discordance may reflect infections where gametocytes make up a comparatively small fraction of the total parasites and are below the limit of detection in the context of the overall infection represented by the DNA fraction.

CONCLUSIONS

In this study, we developed a gametocyte genotyping assay based on deep sequencing of an amplicon marker within the *pfs230* gene. This high-resolution genotyping approach can detect minority clones within *P. falciparum* infections and can estimate relative clone frequency at low gametocyte densities. This genotyping assay can

be used to answer important questions about parasite transmission biology, including the influence of polyclonal infections on gametocyte production and dynamics of gametocytes within *P. falciparum* infections.

CHAPTER IV. IMPACT OF HOST AND PARASITE FACTORS ON GAMETOCYTE PRODUCTION IN *PLASMODIUM FALCIPARUM*

ABSTRACT

Background

Plasmodium falciparum gametocytemia varies between individuals; however, it is not known if this variation is due to the influence of host and parasite factors on the initiation of gametocyte production or on developing and mature gametocytes themselves or both. In this study, we evaluated the association between host and parasite factors and gametocyte production and gametocytemia in *P. falciparum* infections to distinguish between these possibilities.

Methods

RNA extracted from *P. falciparum* positive blood collected as part of a household-clustered cohort study conducted in southern Malawi from April 2019 to March 2020 underwent RT-qPCR to measure expression of the *ccp4* and *mget* genes to detect and quantify gametocytes and RT-PCR to measure expression of the *ap2-g* gene as a marker of gametocyte production. Amplicon sequencing of a polymorphic region of the *pfs230*

was performed on genomic DNA extracted from the *P. falciparum* positive samples to estimate infection complexity (i.e., the number of unique parasite clones in an infection). Mixed effects logistic regression was used to evaluate the association between predictors (host age, complexity of infection, season, symptoms, history of antimalarial drugs use, and parasite density) and two outcomes: gametocyte production and gametocytemia, while accounting for host sex and gametocyte sex ratio. Mixed effects linear regression was used to evaluate the association between the predictors and *ap2-g* gene transcript abundance and gametocyte density in infections with detectable *ap2-g* gene expression and detectable gametocytes, respectively.

Results

A total of 1108 samples positive for *P. falciparum* parasites of any stage were identified, and 1047 of these samples were positive for ring stage parasites. *ap2-g* gene expression was detectable in 41% (492/1047) of infections with ring stage parasites, and gametocytes were detected in 53.6% (594/1108) of samples with any parasite stage. In the final adjusted models, infections with >3 clones had greater odds of expressing *ap2-g* compared to single clone infections [OR 2.52, 95% CI 1.41 – 4.50], but infection complexity was not significantly associated with the presence of gametocytes. Children <6 years old had greater odds of harboring gametocytes [OR 2.13, 95% CI 1.02 – 4.48] compared to adults. Symptomatic infections had lower odds of harboring gametocytes compared to asymptomatic infections [OR 0.17, 95% CI 0.11 – 0.29], but the proportion of infections with detectable *ap2-g* gene expression did not differ significantly between symptomatic and asymptomatic infections. *ap2-g* gene transcript abundance was 65.3%

lower in children <6 years of age compared to adults. In contrast, gametocyte density was 293% higher in children <6 years of age compared to adults.

Conclusions

Our findings show that factors associated with gametocyte production in *P. falciparum* also contribute to variation in gametocyte carriage, but additional factors may contribute to variation in gametocyte carriage after initiation of gametocyte production. While *ap2-g* gene expression is necessary for gametocyte production, additional factors such as immunity and duration of infection may contribute to the presence or absence of gametocytes after initiation of gametocyte production.

BACKGROUND

Stalled progress towards malaria control and elimination has stimulated research into the development of interventions aimed at interrupting transmission of the malaria parasite from the human host to the mosquito vector. The gametocyte, the parasite stage that is transmitted from humans to mosquitoes, is a key target for transmission interrupting interventions [1, 112, 113]. Gametocytemia is useful metric in determining human-to-mosquito transmission potential. Therefore, to develop and apply transmission interrupting interventions effectively, it is critical to understand variation in gametocytemia observed between individuals.

During the erythrocytic stage in humans, the malaria parasite either invests in asexual replication or gametocyte production. Factors that modulate the tradeoff between asexual replication and gametocyte production could explain variation in gametocytemia observed between individuals. In *P. falciparum*, gametocyte production starts with sexual commitment where a fraction of asexual parasites differentiates to develop into gametocytes [11]. Recent studies examining the molecular mechanisms underlying gametocyte production have identified *P. falciparum* ApiAP2-G protein (*Pf*AP2-G) as a master transcription regulator of sexual commitment [42, 45, 46]. *Pf*AP2-G upregulates expression of genes required for gametocyte development [48, 49]. The *ap2-g* gene is expressed in schizonts and gametocyte-committed rings and is used as a molecular marker to distinguish sexually committed rings from non-committed rings which are morphologically indistinguishable [41, 47, 48, 50]. After sexual commitment, immature gametocytes circulate briefly in peripheral blood and subsequently sequester in the bone marrow and spleen where they progress through five distinct developmental stages [14-

16] before they are released back into the peripheral circulation as mature gametocytes which are accessible to female *Anopheles* mosquitoes during blood feeding [11].

Not all parasites within an infection express the *ap2-g* gene, with one study demonstrating expression in 1 – 30% of asexual parasites [48]. These results indicate that not all asexual parasites in an infection produce gametocytes at a given time. It has been suggested that gametocyte production could be a stochastic process; however, there is contradictory evidence to this hypothesis [52-54]. *P. falciparum* isolates from different patients vary in gametocytes conversion rates (GCR) – the proportion of asexual parasites in circulation that are committed to develop into gametocytes [52]. In experimental studies using *P. falciparum* laboratory strains, distinct parasite strains have different GCRs [53]. Furthermore, some isolates that have been in culture for long period of time lose their ability to produce gametocytes because of deletions in a locus spanning the gametocyte development protein 1 (GDV1) gene on chromosome 9 of the *P. falciparum* genome [54, 55]. Also, reduced levels of lysophosphatidylcholine, a common component of plasma, have been shown to increase gametocyte production [56, 114]. Combined, this evidence suggests that malaria parasite investment in asexual replication or gametocyte production is not a stochastic process but is likely dependent on host and parasite genetic factors.

Gametocytemia in *P. falciparum* infections follow certain demographic patterns with higher gametocyte carriage rates observed in infections from school age children (SAC), anemic individuals, asymptomatic individuals, and individuals taking antimalarial drugs, as well as in polyclonal and high parasitemia infections [11, 19, 33-36]. However, it is not known if these variations are due to the influence of host and parasite factors on

the initiation of gametocyte production and/or on the developing and mature gametocytes themselves. With the complex development of gametocytes which take approximately 10 – 12 days to reach maturity [13], and the impact of immune responses against both immature and mature gametocytes [64, 67, 68], gametocytemia at a given time point only captures gametocytes that have survived during development in deep tissues and in peripheral blood [37]. Therefore, the influence of host and parasite factors on gametocyte production and gametocyte carriage may not be the same. In this study, we evaluated the association between host age, complexity of infection, season, symptoms, history of antimalarial drug use, and parasite density and gametocyte production in *P. falciparum* infection based on *ap2-g* gene expression, a marker of parasite sexual commitment. Using the same samples, we also evaluated the association between these factors and gametocytemia to determine if these factors impact gametocyte production and gametocytemia in the same way. Thereafter, we evaluated the association between these factors and *ap2-g* gene transcript abundance and gametocyte density in infections with detectable *ap2-g gene* expression and infections that had detectable gametocytes, respectively.

METHODS

Study design and population

This analysis used *P. falciparum* blood samples collected as part of the Malawi International Center of Excellence for Malaria Research (ICEMR, NIH grant number: U19A1089683) household-clustered cohort study that was conducted within the

catchment areas of health centers in two high transmission rural settings: Namanolo, Balaka and Ntaja, Machinga, from April 2019 to March 2020. The ICEMR cohort study included 96 households in each health center catchment areas with the goal of detecting infection, disease, and potential for transmission. Briefly, each of the two sites was divided into four quadrants and within each quadrant four index households were sequentially sampled from the population census done in 2018. The five nearest households were identified to form a cluster of households. Index houses were excluded if there was less than 100 meters between the cluster boundary and proceeding clusters. Index households and five neighboring households were offered enrollment with replacement of any refusals or migrations with the next nearest households. All members of each household were invited to participate in the study. The study was an open cohort where new household members were invited to participate. After consent was obtained from adults and assent from children over 14 years of age, all eligible children and adults were recruited to participate in the study. The cohort included 962 individuals, aged >6 months, who contributed 7,976 blood samples that underwent 18s ribosomal RNA qPCR [87]. Demographics and other information were collected from individuals upon recruitment. Monthly surveys were conducted in all individuals to collect information on bed net use, history of antimalarial treatment, and symptoms. All study participants were followed with active case detection visits every 4 – 6 weeks yielding a maximum of ten visits over twelve months or until they were lost to follow up. Passive case detection was conducted by a study nurse at the health center. Study participants presenting for evaluation of illnesses were interviewed by the study nurse, but diagnosis, evaluation and treatment were conducted by health center providers according to the standard of care.

Malaria treatment was provided to those diagnosed with malaria. Samples at active and passive case detection visits were collected by finger prick into Eppendorf tubes; 50 µL was pipetted on to a filter paper to create two blood spots and 100 µL of blood was placed in RNA Protect cell reagent (Qiagen, CA). Filter paper samples were stored at 4 °C with desiccants for up to 2 weeks and then stored at -20 °C until *P. falciparum* detection by qPCR and RNA protect samples were also stored at 4 °C for up to 2 weeks and then transitioned to -80 °C freezer until nucleic acid extraction.

Laboratory procedures

18S ribosomal RNA gene qPCR

P. falciparum detection and quantification of the ICEMR blood samples was done in the ICEMR molecular core laboratory at Kamuzu University of Health Sciences in Blantyre, Malawi. Genomic DNA was extracted from filter paper blood samples from all visits and *P. falciparum* 3D7 laboratory strain standards of known parasite density (6000, 600, 60 and 6 parasites/ µL) using the dried blood spot protocol of the QIAamp DNA MiniKit (Qiagen, CA). *P. falciparum* detection and quantification was based on quantitative real-time PCR targeting the 18s rRNA gene [87]. Briefly, 5 µL of DNA extracted from each sample and the standards was used in a real-time qPCR using TaqMan Universal Mastermix (Applied Biosystems) and probe on a Step One Plus Real-Time PCR System (Applied Biosystems). Parasite density (parasites per µL) was estimated via linear regression applied to Ct values and parasite density of the *P. falciparum* 3D7 laboratory strain standards. The following primers and probe were used:

Forward Primer at 25µm: 5' – GTA ATT GGA ATG ATA GGA ATT TAC AAG GT – 3'; Reverse Primer at 25µm: 5' – TCA ACT ACG AAC GTT TTA ACT GCA AC – 3'; and Probe at 25µm: (5'-[FAM] TGC CAG CAG CCG CGG TAA TTC [TAMRA]-3').

DNA and RNA Extraction

Stored RNA Protect stabilized blood from samples that were positive for 18S ribosomal RNA gene qPCR was sent to the University of Maryland School of Medicine (UMSOM) malaria research laboratory for gametocyte detection and quantification, *ap2-g* gene expression RT-PCR, and amplicon deep sequencing of the *pfs230* gene. At UMSOM, the samples were thawed for DNA and RNA extraction: 200 µL was used for RNA extraction using RNeasy Plus 96 kit (Qiagen, CA) and RNA was eluted in 50 µL of Low EDTA TE (1X); and the remaining sample (~150 µL) was used for DNA extraction using QIAamp DNA Blood Kit (Qiagen, CA) and DNA was eluted in 40 µL of Low EDTA TE (1X). Extracted RNA and DNA was stored at -80 °C until analysis.

Gametocyte Detection and Quantification

To detect and quantify *P. falciparum* gametocytes, extracted RNA from *P. falciparum* positive samples underwent RT-qPCR using primers targeting the *ccp4* (female gametocyte marker) and *mget* (male gametocyte marker) genes as described previously [20]. In brief, 5 µL of RNA template and serially diluted standards generated from target RNA copies transcribed *in vitro* using MEGAShortscript T7 high yield transcription kit (Invitrogen) and quantified using Qubit 3 fluorometer (ThermoFisher)

was used in a multiplex one step RT-qPCR with Luna Universal Probe One-Step RT-qPCR Kit (NEB) and probes on a Light Cycler 96 Real-time PCR machine (Roche). All RT-qPCR reactions were performed in duplicate. If there were discordant results in the duplicate reactions, the RT-qPCR was repeated, and if still discordant after rerun, a third run was conducted as a tie breaker. *ccp4* and *mget* copy number per mL in test samples was estimated via linear regression using Ct values of the serially diluted standards and their respective *ccp4* and *mget* copy number per mL. To obtain female and male gametocyte density, estimated test sample *ccp4* copy number per mL was multiplied by a factor of 4 (4 RNA copies of per female gametocyte cell) and estimated test sample *mget* copy number was multiplied by a factor of 9.8 (9.8 RNA copies per male gametocyte cell)[20].

ap2-g Gene Transcript Abundance.

Two delta-delta CT (2^-ΔΔCT) relative quantification RT-PCR was used to quantify *ap2-g* gene transcript abundance in RNA samples. Briefly, cDNA was synthesized from remaining RNA (after gametocyte detection) and RNA extracted from ring stage parasites from a strain of *P. falciparum* that does not produce gametocytes and expresses very little *ap2-g* gene (MAL-1183 SenTh026.04) [92] using QuantiTect Reverse Transcription Kit (Qiagen, CA). Five µL of cDNA from the test samples (*P. falciparum* positive samples) and from MRA-1183 was used in a PCR with SYBR Green PCR Master Mix (Applied Biosystems) on a Light Cycler 96 Real-time PCR machine targeting the *ap2-g* gene and the *sbp1* gene (ring stages marker) as reference. The following specific primers were used to target *ap2-g* gene transcripts: forward primer 5`-

AAC AAC GTT CAA TAA ATA AGG-3` and reverse primer 5`-ATG TTA ATG TTC CCA AAC AAC CG-3` [45]. The following specific primers were used to target the *sbp1* gene: Forward primer: 5` -GCA AAA CAA GCC GTA CAT GTT G – 3` and reverse primer: TTG CTA GGT AAT ATC CTT TTC TTT TTC C – 3` [50].

Parasite Genotyping by Amplicon Sequencing

To estimate COI, extracted DNA was genotyped using the assay developed previously in chapter 3 of this dissertation [Vareta JA. *et al*., under revision for *Malaria Journal*]. In brief, a PCR was performed targeting a polymorphic region of *pfs230* gene. A second PCR was done to add unique barcodes and sequencing adapters to amplicons from the first PCR to allow for multiplexing of samples for sequencing. Barcoded amplicons from all samples were pooled and sequenced using an Illumina MiSeq (Illumina). SeekDeep [94] was used to demultiplex samples, cluster unique reads, and determine number of haplotypes (number of clones) present in each infection as described above.

Definitions

Gametocyte production was quantified by measuring transcript abundance of the *ap2-g* gene in ring stages of *P. falciparum* positive samples. Gametocyte production was evaluated in two ways: 1) whether an infection had detectable *ap2-g* gene expression; and 2) *ap2-g* gene transcript abundance in those infections with detectable *ap2-g* gene

transcripts. *ap2-g* gene transcript abundance was defined as fold change in expression of the *ap2-g* gene in test samples compared to its expression in ring stage parasites in a non-gametocyte producing control strain of *P. falciparum* (MRA-1183 SenTh026.04) [92]. Infections that were positive for *sbp1* gene were defined as harboring ring stage parasites. Gametocyte carriage was quantified by summing female and male gametocytes density, measured by expression levels of the *ccp4* and *mget* genes, respectively. Gametocyte carriage was also evaluated in two ways: 1) whether an infection had detectable *ccp4* or *mget* genes expression; and 2) sum of *ccp4* and *mget* transcripts abundance (converted to gametocyte density) in infections that had detectable gametocytes. Predictors of interest included COI, host age, season when sample was collected, symptoms, history of antimalarial treatment and parasite density. COI was defined as the number of distinct haplotypes observed within an infection based on genotyping results from deep sequencing of *pfs230* gene amplicons. Due to the distribution of the COI variable (range: 1 – 8), COI was categorized into infections with a single clone; infections with 2 to 3 clones; and infections with more than 3 clones. The age of study participants was categorized into three groups: young children (< 6 years of age); school-age children (SAC) (6 - 15 years old); and adults (≥16 years). Season when the samples were collected was categorized as dry season (May - November) and rainy season (December - April). History of antimalarial treatment was defined as reported antimalarial drug treatment in the past month prior to each visit. Symptoms of malaria was defined as reported symptoms at the health center and axillary temperature of ≥ 37.5 at the time of visit. Parasite density was measured as the number of malaria parasites per μL using quantification of *P. falciparum* 18S ribosomal RNA gene as described above.

Data analysis

The proportion of ring stage infections with detectable *ap2-g* gene expression and the proportion of infections with detectable gametocytes were computed within strata of predictors and covariates. Bivariate analyses of predictors and infections with detectable *ap2-g* gene expression and detectable gametocytes (yes/no) were performed using unadjusted mixed effects logistic regression while accounting for clustering within household and individual. Bivariate comparison of *ap2-g* gene transcript abundance and gametocyte density within strata of predictors was done using unadjusted linear mixed effects regression while accounting for clustering within household and individual. In the final adjusted models, two multilevel models of the association between predictors and outcomes were performed: 1) A mixed effects logistic regression model to evaluate the association between predictors and detectable *ap2-g* gene expression and gametocyte carriage; and 2) a mixed effects linear model to evaluate the association between predictors and *ap2-g* gene transcript abundance in infections with detectable *ap2-g* gene expression and gametocyte density in infections with detectable gametocytes. The *ap2-g* gene transcript abundance and gametocyte density in the linear mixed effects model were log transformed to satisfy the normality assumption of continuous outcome variables. Both models accounted for clustering at household and individual level and were adjusted for gender, gametocyte sex ratio (in gametocyte carriage and gametocyte density models) and parasite density. All statistical analyses were done in SAS 9.4 statistical software (SAS Institute, Cary, NC) at alpha level of 0.05.

RESULTS

Overall, 7976 samples were collected from 986 individuals in 198 households during the study period. A total of 2138 samples (26.8%) tested positive for *P. falciparum* by 18s ribosomal RNA gene qPCR conducted in Malawi. RNA was extracted from 1,771 *P. falciparum* positive samples that were shipped to UMSOM for analysis. After *ccp4* and *mget* RT-qPCR and *ap2-g* gene RT-PCR, 663 samples that were negative for both *ccp4* and *mget* expression and *sbp1* expression were considered false positives for *P. falciparum* and were excluded from the analysis, leaving 1108 total samples. Gametocytes were detected in 53.6% (594/1108) of the samples and *sbp1* expression was detected in 94.5% (1047/1108) of the samples. Forty-one percent (429/1047) of samples that expressed the *sbp1* gene had detectable *ap2-g* gene expression and all samples with detectable *ap2-g* gene expression had detectable *sbp1* gene expression. DNA was extracted from 1,492 samples, with 279 samples not having sufficient material for DNA extraction after RNA extraction. Among the samples that underwent DNA extraction, 866 (58 %) were successfully genotyped via deep sequencing of the *pfs230* amplicon.

Association between *ap2-g* gene expression and gametocyte carriage

We first evaluated the association between *ap2-g* gene expression and gametocyte carriage. A significantly higher proportion of infections with detectable *ap2-g* gene expression had detectable gametocytes (59.9%) compared to infections with no detectable *ap2-g* gene expression (49.6%) (p-value = 0.0086). In the final model accounting for clustering within household and individual and adjusting for all predictors

and covariates, a significantly greater odds of gametocyte carriage was observed in infections with detectable *ap2-g* gene expression (OR=2.73; 95% CI: 1.38 – 4.08) compared to infections without detectable *ap2-g* gene expression. No association was observed between *ap2-g* gene transcript abundance and gametocyte density in infections with detectable *ap2-g* gene expression.

Factors associated with gametocyte production

The distribution of infections with detectable *ap2-g* gene expression by categories of potential predictors and covariates is shown in **Table 4**. Unadjusted comparison show significant associations between age-group, season, symptoms, and COI and presence of infections with detectable *ap2-g* gene expression. Associations between study site, participant sex, and history of antimalarial drug use and presence of infections with detectable *ap2-g* gene expression were not statistically significant. In the final adjusted model controlling for all predictors and host sex, the odds of an infection expressing *ap2-g* gene were significantly greater in infections with >3 clones [OR: 2.52; 95% CI:1.41 – 4.50] compared to monoclonal infections and with increased parasite density (for a 2-fold increase in parasite density, the odds of an infection having detectable *ap2-g* gene expression increased by 63 %). The odds of an infections expressing *ap2-g* gene were not significantly different between age-groups; between the rainy and the dry season; between symptomatic and asymptomatic individuals; and between individuals with and without a history of antimalarial drug use **(Table 5)**.

Table 4: Bivariate association between infections with detectable *ap2-g* gene expression and gametocytes and predictors of interest and covariates.

	Ring stage positive [a] n=1047	Detectable *ap2-g* gene expression [b] n=429		*P. falciparum* positive n=1108	Gametocyte positive[c] n=594	
		n(row%)	p value	(col %)	n(row%)	p value
Site						
Namanolo	393	159 (40.5)		420	236 (56.2)	
Ntaja	654	270 (41.3)	0.4377	688	358 (52.0)	0.1666
Age group						
< 6 years	189	97 (51.3)	<0.0001	195	103 (52.8)	0.0738
6 – 15 years	622	277 (44.5)	<0.0001	655	390 (59.5)	<0.0001
16 + years	236	55 (23.3)	Ref	258	101 (39.2)	Ref
Sex						
Female	558	221 (39.6)		587	304 (51.8)	
Male	489	208 (42.5)	0.2828	521	290 (55.7)	0.1906
Season						
Dry	472	216 (45.8)		494	288 (58.3)	
Rainy	575	213 (37.0)	0.0077	614	306 (49.8)	0.0139
Symptoms						
No	724	246 (34.0)		779	493 (63.3)	
Yes	323	183 (56.7)	<0.0001	329	101 (30.7)	<0.0001
History Antimalarial drugs[d]						
No	1000	413 (41.3)		1057	568 (53.7)	
Yes	47	16 (34.0)	0.2062	51	26 (51.0)	0.7027
COI[e]						
Clonal	349	130 (37.3)		378	188 (49.7)	
2 – 3 clones	328	151 (46.0)	0.0224	333	185 (55.6)	0.1096
>3 clones	110	69 (62.7)	<0.0001	111	82 (73.9)	0.0008

[a] ring stage positive defined as samples that had detectable *sbp1* gene expression.
[b] *ap2-g* positive defined as samples that were positive for *ap2-g* gene in all samples that were *sbp1* positive after RT-PCR.
[c] Gametocyte positive defined as samples that were positive for *ccp4* or *mget* genes after qPCR.
[d] 3 samples were missing history of antimalarial drugs status.
[e] 286 samples failed genotyping. COI = complexity of Infection.
p values from unadjusted mixed effects logistic regression accounting for clustering within household and individual.

Factors associated with gametocyte carriage

The distribution of gametocyte carriage by categories of potential predictor variables and covariates is shown in **Table 4**. Unadjusted comparisons show significant associations between age-group, season, symptoms, and COI and presence of gametocytes. Associations between study site, host sex, and history of antimalarial drug use and detectable gametocytes were not statistically significant. In the final adjusted

model controlling for all predictors (including host sex and gametocyte sex ratio), the odds ratio of an infection carrying gametocytes versus not was 2.13 [95% CI: 1.02 – 4.48] and 1.88 [95% CI: 1.07- 3.33] in children <6 years of age and school age children compared to adults, respectively. Lower odds of carrying gametocytes were observed in symptomatic compared to asymptomatic individuals [OR = 0.17; 95% CI: 0.10 – 0.29]. For a 2-fold increase in parasite density, the odds of an infection carrying gametocytes increased by 6%. In the final model, infection complexity, season, and history of antimalarial drug use were not significantly associated with the presence of gametocytes, although the association with complexity of infection was borderline **(Table 5)**.

Table 5: Odds ratio of associations between predictors and *ap2-g* expression (dichotomous) and gametocyte carriage using mixed effects logistic regression model.

	Detectable *ap2-g* gene expression[a]		**Gametocyte positive[d]**	
	***ap2-g* gene expression yes/no**	**Adjusted multilevel model[b] OR [95% CI][c]**	**Gametocytes yes/no**	**Adjusted multilevel model OR [95% CI]**
COI[e]				
Clonal	130/219	1.00 (Ref)	188/190	1.00 (Ref)
2 – 3 clones	151/177	1.27 [0.85 – 1.90]	185/148	1.12 [0.73 – 1.70]
>3 clones	69/41	2.52 [1.41 – 4.50]	82/29	1.85 [0.94 – 3.64]
Age group				
<6 years old	97/92	1.06 [0.55 – 2.06]	103/92	2.13 [1.02 – 4.48]
6 – 15 years old	277/345	1.10 [0.66 – 1.83]	390/265	1.89 [1.07 – 3.33]
≥16 years old	55/181	1.00 (Ref)	101/157	1.00 (Ref)
Season				
Dry	216/256	1.00 (Ref)	288/206	1.00 (Ref)
Rainy	213/362	0.81 [0.56 – 1.17]	306/308	0.79 [0.55 – 1.15]
Symptoms				
No	246/478	1.00 (Ref)	493/286	1.00 (Ref)
Yes	183/140	0.80 [0.49 – 1.29]	101/228	0.17 [0.10 – 0.29]
History of antimalarial drugs				
No	413/587	1.00 (Ref)	568/489	1.00 (Ref)
Yes	16/31	0.53 [0.21 -1.34]	26/25	0.77 [0.32 – 1.89]
Parasite density[g]				
Parasite density	429	1.63 [1.49 – 1.78]	594	1.068 [1 – 1.15]

[a] Detectable *ap2-g* expression = samples positive for *ap2-g* gene in *sbp1* positive samples after RT-PCR.
[b] Mixed effects logistic regression accounting for clustering within household and individual and controlling for predictors, host sex, and gametocyte sex ratio.
[c] Odds ratio, 95% confidence interval.
[d] Detectable gametocyte defined as samples that were positive for *ccp4* or *mget* genes after qPCR.
[e] COI = complexity of infection.
[f] Log parasite density.

Factors associated with *ap2-g* gene transcript abundance and gametocyte density

Comparisons of *ap2-g* gene transcript abundance and gametocyte density among samples with detectable *ap2-g* gene expression and detectable gametocytes, respectively, by predictors and covariates are described in **Table 6**. Unadjusted comparisons indicate significantly higher median *ap2-g* gene transcript abundance in adults compared to school age children or children <6 years of age and in asymptomatic compared to symptomatic individuals. Median gametocyte density was significantly higher in children <6 years of age compared to adults, in females compared to males, and during the rainy compared to the dry season. In the final adjusted model controlling for all predictors, sex, and gametocyte sex ratio (in gametocyte density model), *ap2-g* gene transcript abundance was 65.3% lower in children <6 years of age compared to adults. In contrast, gametocyte density was 293% higher in children <6 years of age compared to adults. Although *ap2-g* gene transcript abundance was similar during the rainy and dry season, gametocyte density was 55.7% higher during the rainy than dry season. *ap2-g* gene transcript abundance and gametocyte density were 60% and 49.7% lower in symptomatic compared to asymptomatic infections, respectively. Also, parasite density was positively associated with *ap2-g* gene transcript abundance: for a 1% increase in parasite density, *ap2-g* gene transcript abundance increased by 99%. However, the association between parasite density and gametocyte density was not statistically significant **(Table 7)**.

Table 6: Bivariate association between predictors and *ap2-g* transcript abundance, gametocyte density and parasite density.

	***ap2-g* transcript abundance[a] (n = 492)**		**Gametocyte density (n = 594)**		**Parasite density (n = 1108)**	
	Median [IQR][b]	**p value**	**Median [IQR][c]**	**p value**	**Median [IQR][d]**	**p value**
Age group						
< 6 years	4.82 [0.50 – 29.04]	0.0057	2.24 [0.41 – 10.06]	0.0069	20.37 [1.95 – 791.13]	<0.0001
6 – 15 years	11.09 [2.38 – 54.00]	0.7192	0.66 [0.15 – 3.69]	0.8086	37.56 [3.57 – 448.37]	<0.0001
16 + years	20.09 [1.34 – 95.01]	Ref	0.72 [0.13 – 3.04]	Ref	4.93 [1.63 – 29.47]	Ref
Sex						
Female	7.40 [2.18 – 50.03]		1.07 [0.20 – 5.13]		11.19 [2.08 – 167.50]	
Male	10.35 [1.03 – 50.68]	0.9263	0.64 [0.14 – 3.40]	0.3518	23.04 [2.38 – 393.38]	0.3967
Season						
Dry	9.36 [2.51 – 43.51]		0.58 [0.13 – 3.51]		15.62 [2.30 – 239.24]	
Rainy	8.60 [0.93 – 58.68]	0.7298	1.06 [0.23 – 5.20]	0.0224	13.8 [2.11 – 270.68]	0.3517
Symptoms						
No	23.59 [3.25 – 58.68]		0.94 [0.18 – 4.57]		10.49 [1.94 – 82.12]	
Yes	4.98 [0.56 – 35.04]	<0.0001	0.57 [0.13 – 3.05]	0.1131	156.13 [3.72 – 7910.2]	<0.0001
History of antimalarial drugs						
No	8.55 [2.02 – 50.77]		0.75 [0.16 – 4.32]		60.59 [7.7 – 781.42]	
Yes	27.92 [6.71 – 45.99]	0.2141	1.36 [4.09 – 3.91]	0.4995	50.11 [10.79 – 1054.1]	0.4619
COI[e]						
Clonal	8.27 [0.91 – 69.07]	Ref	0.77 [0.14 – 5.46]	Ref	24.42 [3.94 – 970.63]	Ref
2 – 3 clones	7.49 [0.89 – 41.36]	0.9828	0.89 [0.20 – 4.19]	0.4873	107.31 [22.30 – 1409.61]	0.0005
>3 clones	10.10 [3.41 – 39.53]	0.3461	1.01 [0.20 – 5.50]	0.6723	278.92 [61.06 – 660.84]	<0.0300

[a] *ap2-g* gene transcript abundance = fold change in *ap2-g* gene expression in test compared to MRA-1183 control.
[b] Interquartile range.
[c] Gametocyte density (per µL) = sum of female gametocyte density (*ccp4* gene) and male gametocyte density (*mget*).
[d] parasites per microliter (µL) as quantified by 18S Ribosomal RNA qPCR
[e] COI = complexity of infection.

Table 7: Percent change and standard errors of the association between predictors and *ap2-g gene* transcript abundance and gametocyte density using mixed effects linear models.

	% change *ap2-g* gene transcript abundance[a] [SE][b]	p-value	% change gametocyte density[c] [SE]	p-value
COI				
Clonal	Ref		Ref	
2 – 3 clones	- 42.37 [0.29]	0.0584	- 18.56 [0.25]	0.4068
>3 clones	- 47.99 [0.38]	0.0913	- 12.61 [0.3530]	0.7026
Age group				
<6 years old	- 65.11 [0.46]	0.0243	293.50 [0.45]	0.0026
6 – 15 years old	- 6.54 [0.40]	0.8663	44.77 [0.3416]	0.2798
≥16 years old	Ref		Ref	
Season				
Dry	Ref		Ref	
Rainy	- 3.19 [0.27]	0.8995	55.74 [0.21]	0.0345
Symptoms				
No	Ref		Ref	
Yes	- 60.42 [0.35]	0.0090	- 49.69 [0.32]	0.0303
History of antimalarial drugs				
No	Ref		Ref	
Yes	221.52 [0.70]	0.0951	20.08 [0.56]	0.7455
Parasite density				
Parasite density[d]	99.87 [0.06]	0.0312	100.02 [0.04]	0.5074

[a] *ap2-g* transcript abundance measured as fold change in expression
[b] SE=standard error
[c] Gametocyte density measured by
[d] Parasite density (μL) log transformed
Accounting for clustering within household and individual and adjusting for sex and gametocyte sex ratio.

DISCUSSION

Our results indicate that factors associated with gametocyte production in *P. falciparum* also contribute to variation in gametocyte carriage, but additional factors may contribute to variation in gametocyte carriage after initiation of gametocyte production. We found that infections with multiple clones were more likely to produce gametocytes compared to monoclonal infections. In addition, the odds of gametocyte carriage increased with parasite density. Although the proportion of infections that produced gametocytes was similar between age-groups and between symptomatic and

asymptomatic individuals, children and asymptomatic individuals were more likely to carry gametocytes than adults and symptomatic individuals, respectively. These findings suggest that COI and parasite density play a crucial role in initiation of gametocyte production, but host immunity and duration of infection may influence gametocyte carriage after sexual commitment.

Findings from previous studies have shown that polyclonal *P. falciparum* infections are more likely to carry gametocytes than clonal infections [36, 40, 115, 116]. Here, we show that polyclonal infections are also associated with higher gametocyte production based on *ap2-g* gene expression. Although not statistically significant after controlling for parasite density, polyclonal infections were also more likely to carry gametocytes, confirming findings from previous studies. It is important to note that in bivariate associations, we observed a statistically significant association between COI and gametocyte carriage. Non-significant results in our adjusted model compared to other studies could be due to differences in: covariates included in our final model and in the final models from previous studies; sensitivity of the genotyping methods used; and categorization of COI. The positive association between COI and parasite density and *ap2-g* gene expression observed in this study suggest that these factors may trigger gametocyte production and the association observed between these factors and gametocyte carriage could be a result of these factors influence on sexual commitment. The influence of COI and parasite density on gametocyte production could result from competition for resources between clones in complex infections and/or exponential growth in parasitemia that could be detrimental to the host that may subsequently trigger the parasites to invest more in sexual reproduction as a survival mechanism. This

hypothesis agrees with theoretical studies that have suggested high reproductive investment due to clone competition and high parasitemia [117].

Notably, we did not observe any association between host age and gametocyte production after adjusting for parasite density in our final model in this population, but gametocyte carriage was higher in children <6 years old and school age children than adults. In our bivariate analysis of age and gametocyte production we found that children were more likely to produce gametocytes than adult. Since host age is a proxy of naturally acquired immunity, these findings together suggest that naturally acquired immunity impacts parasitemia which drives gametocyte production in the parasite, but once gametocytes production is initiated, stronger naturally acquired immunity against immature and mature gametocytes in adults may result in lower rates of gametocyte carriage in this age-group compared to children who have less acquired immunity.

We observed higher gametocyte carriage in infections from asymptomatic compared to symptomatic individuals, which agrees with findings from previous studies that have shown higher gametocyte carriage in chronic compared to incident infections [40, 41]. The findings could be explained by differences in the duration of these infections in the human host. Unlike symptomatic infections, which are mostly incident, short duration and treated, asymptomatic infections are mostly low density chronic and often untreated and persist longer in an individual. These infections continuously produce gametocytes at lower lever over longer periods, thus the higher gametocyte carriage observed. Another explanation could be that symptomatic incident infections are investing more in asexual replication than asymptomatic infections [40, 41].

Our finding of no association between history of antimalarial drug use and gametocyte production and carriage disagrees with a previous study that showed increased *ap2-g* gene expression and gametocyte carriage with intake of antimalarial drugs [37]. This discrepancy could be due to the differences in the way history of antimalarial treatment was measured. In our study, participants were asked if they had taken antimalarial drugs in the past month, without documentation of the specific timing of when drugs were taken. The study that observed increased gametocyte production following treatment monitored individuals at 12-hour intervals after treatment as part of efficacy studies. The way we measured antimalarial drug intake could have captured individuals who had taken antimalarial drugs weeks prior, which may not be an informative time frame for our outcome of interest.

Among infections with detectable gametocyte production, *ap2-g* expression was higher in adults compared to children <6 years of age, while in contrast, children <6 years of age had higher gametocyte density than adults. The contrast between the *ap2-g* gene transcript abundance and gametocyte density between these age groups was unexpected. The explanation for this finding could be that there is high *ap2-g* transcript abundance by parasites expressing *ap2-g* in adults, but developing and mature gametocytes in this age group are impacted by stronger immunity resulting in low gametocyte density. Factors that contribute to higher *ap2-g* gene transcript abundance in adults compared to children need to be explored in future studies.

Because of the long follow-up intervals (1 month) in the cohort study, we could not establish a temporal relationship between sexually committed rings at one visit with mature gametocytes present at a subsequent visit. Gametocyte development from sexual

commitment to mature gametocytes takes approximately 10-12 days [118]. It is possible that gametocytes observed after one month were not from the asexual parasites from a prior visit but a new infection two weeks later. By analyzing the data cross-sectionally, we evaluated the association between predictors and gametocyte production and between predictors and gametocyte carriage separately. However, analysis of the association between the predictors and gametocyte production using *ap2-g* gene expression and gametocyte carriage separately still shows that these metrics are affected differently by host and parasite factors. Recent studies have shown that the parenchyma of the bone marrow and the spleen are a reservoir of asexual parasite. This indicates that some sexual commitment may occur exclusively in these tissues [119, 120]. Our sampling of the peripheral blood did not detect parasites in the bone marrow and spleen. This could have introduced a bias if the extravascular parasites were systematically different from the parasites sampled in the peripheral circulation with regards to known or unknown variables. The extent to which extravascular commitment contributes to gametocytemia is currently unknown; however, this study still provides new insights into gametocyte production that occurs in the vasculature.

Recent studies have shown that the parenchyma of the bone marrow and the spleen are a reservoir of asexual parasites. This indicates that some sexual commitment may occur exclusively in these tissues [119, 120]. Our sampling of the peripheral blood did not detect parasites in the bone marrow and spleen. This could have introduced a bias if the extravascular parasites were systematically different from the parasites sampled in the peripheral circulation with regards to known or unknown variables. The extent to which extravascular commitment contributes to gametocytemia is currently unknown;

however, this study still provides new insights into gametocyte production that occurs in the vasculature.

CONCLUSIONS

Findings from this study provide evidence that factors associated with gametocyte production in *P. falciparum* also contribute to variation in gametocyte carriage, but additional factors may also contribute to gametocyte carriage after initiation of gametocyte production. These findings add to the existing body of knowledge on *P. falciparum* gametocyte biology which could help explain variations in gametocytemia observed between individuals in *P. falciparum* infections. This knowledge could inform development and effective deployment of transmission interrupting interventions to reduce malaria burden in endemic regions.

CHAPTER V. DISCUSSION

The main driving force of this work has been our keen interest to understand host and parasite factors that contribute to variation in gametocytemia between individuals in *P. falciparum* infections. In this dissertation work, we developed an assay to genotype *P. falciparum* gametocytes using amplicon deep sequencing as a tool to understand the dynamics of gametocyte production in the context of polyclonal infections. Thereafter, using *P. falciparum* positive blood samples from a household clustered cohort study conducted in southern Malawi from April 2019 to March 2020, we evaluated the impact of host and parasite factors on gametocyte production and gametocytemia to identify factors that contribute to variation in gametocytemia observed between individuals.

Understanding the dynamics of gametocyte production in polyclonal infections requires a genotyping assay that distinguishes both asexual parasites and gametocyte clones and can estimate their relative frequencies in an infection. Here, we developed an assay to genotype gametocytes in *P. falciparum* through sequencing of RT-PCR amplicons generated from a polymorphic region of the *pfs230* gene which is primarily expressed in mature gametocytes. The assay is specific to mature gametocytes, is quantitative, and has higher haplotype resolution than known *P. falciparum* blood stage amplicon-based genotyping markers. The genotyping assay we have developed here can be used in many ways to advance our understanding of gametocyte biology. It can be

used to determine which clones in a polyclonal infection are more likely to produce gametocytes by tracking asexual parasite and gametocyte clones over time. Answering this question could help delineate parasite genetic factors involved in gametocyte production. Through tracking of individual clones in an infection, WGS of a clone with a high gametocyte production phenotype can be compared to WGS of a clone with a low gametocyte production phenotype to identify genes responsible for differences in gametocyte production. The assay can definitively address the important question as to how long mature gametocytes survive in peripheral blood by tracking appearance and disappearance of distinct clones of mature gametocytes over time. The assay can also be used to genotype blood stage parasite given that we have shown that the diversity of the *pfs230* gene marker is comparable to or greater than that of other commonly used blood stage markers. All these applications will help to advance our understanding of *P. falciparum* transmission biology.

With the complex development of *P. falciparum* gametocytes which take 10 – 12 days to mature [13], gametocyte carriage at a given time point only captures gametocyte survival [37]. As many studies have shown variation in gametocyte carriage between individuals [11, 19, 34, 35, 41], we wanted to determine if the variation in gametocyte carriage observed between individuals is due to the influence of host and parasite factors on gametocyte production or gametocyte carriage or both. Using *P. falciparum* positive blood samples collected from Malawi, we went on to show that COI and parasite density play a crucial role in initiation of gametocyte production. The influence of these factors on gametocyte production is also reflected in gametocyte carriage, but host immunity and duration of infection may influence gametocyte carriage after sexual commitment. These

results suggest that parasite genetic factors and within-host competition may influence gametocyte production, but host factors largely impact developing and mature gametocytes. The findings are important in the understanding of the biology gametocytes, as this adds to the existing body of knowledge on factors that influence gametocyte production and carriage.

Gametocytemia is a useful metric to determine the transmission potential of malaria in populations; therefore, understanding variations in gametocytemia is important when deploying transmission interrupting interventions. Higher gametocyte carriage in children and asymptomatic individuals observed in this study indicate that these populations are important reservoirs of transmission and can be intentionally targeted with transmission interrupting interventions. Higher gametocyte carriage in asymptomatic individuals highlight the contribution of long duration infections to human-to-mosquito transmission of the malaria parasite. Although gametocyte density is associated with host infectiousness and likelihood of parasite transmission to the mosquito vector [11], asymptomatic chronic infections with low gametocyte density may constitute a larger proportion of existing infections in the population and should be targeted to reduce malaria transmission. Despite having lower gametocytemia, these chronic infections have been shown to be highly infectious to mosquitoes [41] and are common in high transmission settings [23]. Therefore, malaria screening campaigns using molecular detection methods to identify and treat these infections in the community or mass drug administration campaigns could be viable options to eliminate this transmission reservoir.

Findings from this dissertation work are important because they add to the existing body of knowledge about *P. falciparum* transmission biology which could inform development and deployment of interventions that reduce transmission of the malaria parasite from humans to mosquitoes. The gametocyte genotyping assay developed herein will help in answering important questions about *P. falciparum* transmission biology. As observed in this study and other studies, COI is positively associated with gametocyte production. It is believed that this association is due to clone competition in polyclonal infection that increases gametocyte production. However, the contribution of each clone to gametocytemia in an infection has not been determined due to lack of a viable genotyping method that can longitudinally track asexual parasites clones to determine if they eventually produce gametocytes. Therefore, the genotyping assay developed here will be an important tool to understand gametocyte production dynamics in polyclonal infections.

This dissertation work opens the door for several exciting further investigations. The gametocyte genotyping assay developed herein is an important tool for studies aimed at understanding gametocyte production dynamics in polyclonal infections. The assay can be used in longitudinal studies to understand how clone composition influence gametocyte production by tracking majority and minority asexual parasites clones in an infection to determine which clone will produce more gametocytes. Another exciting investigation will be to study why *ap2-g* gene transcript abundance is higher in adults than children when *ap2-g* gene is expressed but gametocyte density is low in this age group compared to children gametocytes carriage is high. The low gametocyte density in adults could be explained by stronger immune response to gametocytes, but we could

only speculate why *ap2-g* gene transcript abundance was high in adults when *ap2-g* gene was expressed in all age groups. It is possible that once gametocyte production is initiated, greater naturally acquired immunity increases expression of the *ap2-g* gene to enhance gametocyte development but the same immunity acts against immature and gametocytes, thus impacting gametocytemia. This hypothesis needs to be explored in future studies.

As research efforts are focusing on developing interventions that interrupt human-to-mosquito transmission of the malaria parasite in endemic regions, clear understanding of the biology of transmission stages of the malaria parasite will be crucial in mitigating challenges encountered when designing these interventions. In this body of work, we have developed an assay to genotype gametocytes in *P. falciparum* which can be used as a tool to understand parasite genetic factors that influence gametocyte production in natural infection. Furthermore, we have identified factors that contribute to variation in gametocytemia between individuals which will be crucial in identifying targets of transmission interrupting interventions and adding to the general understanding of gametocyte biology. The findings from this dissertation research and future studies based on our findings will add to the existing body of knowledge on malaria transmission biology. This will in turn inform development and prudent application of malaria interventions to interrupt transmission of the malaria parasite from the human host to the mosquito vector.

APPENDIX

Table 8: *ccp4*, *mget* and human tubulin genes qRT-PCR reaction mix and cycling conditions.

MGET/CCp4 Reaction Mix					
	Reagent	**Conc**	**1 rxn (µL)**	**Final conc**	
	RNAse free water		0.95		
	2X NEB Luna Master mix	2X	10	1X	
	CCp4 primers	10µM (F + R)	1.8	0.3µM	
	CCp4 probe	10µM	0.4	0.3µM	
	MGET primers	10µM (F + R)	0.45	0.23µM	
	MGET probe	10µM	0.4	0.3µM	
	Luna WarmStart Enzyme	20X	1	1X	
	RNA template		5		
		Total	20		
Human Tubulin Reaction mix					
	Reagent	**Conc**	**1 rxn (µL)**	**Final conc**	
	RNAse free water		2.5		
	2X NEB Luna Master mix	2X	5	1X	
	HumanTuBB_F2	10µM	0.15	0.15µM	
	HumanTuBB_R2	10µM	0.15	0.15µM	
	HumanTuBB_Hex	10µM	0.2	0.20µM	
	Luna WarmStart Enzyme	20X	0.5	1X	
	RNA template		1.5		
		Total	10		
CCp4/MGET Cycling conditions					
		Temperature	**Time**	**Cycles**	
	Preincubation	55 °C	900s	1	
	Preincubation	95 °C	60s	1	
	Amplification	95 °C	10 seconds		
		60 °C	60 seconds	45	
Human Tubulin Cycling conditions					
		Temperature	**Time**	**Cycles**	
	Preincubation	55 °C	900s	1	
	Amplification	95 °C	60s		
		95 °C	10 seconds		
		60 °C	60 seconds	45	

Table 9: RT-PCR and PCR primer sequences for *pfs230*-M3A genotyping marker with overhang adapter sequences for library preparation.

Primer Name	**Overhang Adaptor Sequence (5`-3`)**
pfs230-M3A_Foward *pfs230*-M3A_Reverse	TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G
	Primer Sequences (5`-3`)
	CTG AAA ACG ACA ATG AAT ATG AGT TGG GAA CAA TTG GTG GAA CA
	Full oligos (5`-3`)
	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGAAAACGACAATGAATATGAGT GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGGGAACAATTGGTGGAACA
Primer Name	**Primer Sequences (5`-3`)**
Pfs25_Forward Pfs25_Forward	GAA ATC CCG TTT CAT ACG CTT G AGT TTT AAC AGG ATT GCT ATC TAA

Table 10: Reverse transcription, PCR and index PCR reaction preparation and cycling conditions.

Reverse transcription reaction preparation and cycling conditions					
DNA elimination					
	1 reaction				
Components	**Volume(µl)**				
gDNA wipeout buffer (7X)	2				
RNAse Free Water	7				
Template (RNA)	5				
Total	14				
complementary DNA synthesis					
	1 reaction				
Components	**Volume(µl)**				
Quantscript reverse transcriptase	1				
Quantscript RT buffer (5X)	4				
RT primer mix	1				
Template (DNAse treated RNA)	14				
Total	**20**				
Cycling conditions					
	Temperature	**Time**			
DNA Elimination	42 ℃	2 minutes			
Reverse Transcription	42 ℃	30 minutes			
Reverse transcriptase inactivation	95 ℃	3 minutes			
Primary PCR reaction preparation and cycling conditions					
PCR reaction mix					
		1 reaction			
Components	**Concentration**	**Volume(µl)**	**Final concentration**		
PCR grade water		5.5	**N/A**		
2X Kapa Ready Mix	**2X**	12.5	**1X**		
Forward Primer	**(5µM)**	1	**(0.2µM)**		
Reverse Primer	**(5µM)**	1	**(0.2µM)**		
Template (cDNA or genomic DNA)		5			
Total		**25**			
Cycling conditions					
Step	**Temperature**	**Duration**	**Cycles**		
Initial denaturation	95 ℃	3 minutes	1		
Denaturation	98 ℃	20 Sec			
Annealing temperatures	62 ℃	30 sec	25 cycles		
Extention	72 ℃	30 sec			
Final extention	72 ℃	1 minute	1		
Index PCR reaction mix and cycling conditions (adding indexes)					
PCR reaction mix					
		1 reaction			
Components	**Concentration**	**Volume(µl)**	**Final concentration**		
PCR grade water		6.5	**N/A**		
2X Kapa Ready Mix	**2X**	12.5	**1X**		
Indexes		1			
Template (amplicon)		5			
Total		**25**			
Cycling conditions					
Step	**Temperature**	**Duration**	**Cycles**		
Initial denaturation	95 ℃	3 minutes	1		
Denaturation	98 ℃	20 Sec			
Annealing temperatures	62 ℃	30 sec	8 cycles		
Extention	72 ℃	30 sec			
Final extention	72 ℃	1 minute	1		

REFERENCES

1. WHO. *World Malaria Report 2022*. 2022 [cited 2023 May 22, 2023]; Available from: https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022.

2. Sarma, N., et al., *The Economic Burden of Malaria: Revisiting the Evidence.* Am J Trop Med Hyg, 2019. **101**(6): p. 1405-1415.

3. Gallup, J.L. and J.D. Sachs, *The economic burden of malaria.* Am J Trop Med Hyg, 2001. **64**(1-2 Suppl): p. 85-96.

4. Sachs, J. and P. Malaney, *The economic and social burden of malaria.* Nature, 2002. **415**(6872): p. 680-5.

5. Najera, J.A., M. Gonzalez-Silva, and P.L. Alonso, *Some lessons for the future from the Global Malaria Eradication Programme (1955-1969).* PLoS Med, 2011. **8**(1): p. e1000412.

6. Roberts, L. and M. Enserink, *Malaria. Did they really say ... eradication?* Science, 2007. **318**(5856): p. 1544-5.

7. Cowman, A.F., et al., *Malaria: Biology and Disease.* Cell, 2016. **167**(3): p. 610-624.

8. Cowman, A.F., D. Berry, and J. Baum, *The cellular and molecular basis for malaria parasite invasion of the human red blood cell.* J Cell Biol, 2012. **198**(6): p. 961-71.

9. Bousema, T., et al., *Asymptomatic malaria infections: detectability, transmissibility and public health relevance.* Nat Rev Microbiol, 2014. **12**(12): p. 833-40.

10. van der Watt, M.E., J. Reader, and L.M. Birkholtz, *Adapt or Die: Targeting Unique Transmission-Stage Biology for Malaria Elimination.* Front Cell Infect Microbiol, 2022. **12**: p. 901971.

11. Bousema, T. and C. Drakeley, *Epidemiology and infectivity of Plasmodium falciparum and Plasmodium vivax gametocytes in relation to malaria control and elimination.* Clin Microbiol Rev, 2011. **24**(2): p. 377-410.

12. Tadesse, F.G., et al., *Gametocyte Sex Ratio: The Key to Understanding Plasmodium falciparum Transmission?* Trends Parasitol, 2019. **35**(3): p. 226-238.

13. Josling, G.A. and M. Llinas, *Sexual development in Plasmodium parasites: knowing when it's time to commit.* Nat Rev Microbiol, 2015. **13**(9): p. 573-87.

14. Joice, R., et al., *Plasmodium falciparum transmission stages accumulate in the human bone marrow.* Sci Transl Med, 2014. **6**(244): p. 244re5.

15. Aguilar, R., et al., *Molecular evidence for the localization of Plasmodium falciparum immature gametocytes in bone marrow.* Blood, 2014. **123**(7): p. 959-66.

16. Pelle, K.G., et al., *Transcriptional profiling defines dynamics of parasite tissue sequestration during malaria infection.* Genome Med, 2015. **7**(1): p. 19.

17. Billker, O., A.J. Miller, and R.E. Sinden, *Determination of mosquito bloodmeal pH in situ by ion-selective microelectrode measurement: implications for the regulation of malarial gametogenesis.* Parasitology, 2000. **120 (Pt 6)**: p. 547-51.

18. Billker, O., et al., *Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito.* Nature, 1998. **392**(6673): p. 289-92.

19. Coalson, J.E., et al., *High prevalence of Plasmodium falciparum gametocyte infections in school-age children using molecular detection: patterns and predictors of risk from a cross-sectional study in southern Malawi.* Malar J, 2016. **15**(1): p. 527.

20. Meerstein-Kessel, L., et al., *A multiplex assay for the sensitive detection and quantification of male and female Plasmodium falciparum gametocytes.* Malaria Journal, 2018. **17**(1): p. 441.

21. Robert, V., et al., *Effect of gametocyte sex ratio on infectivity of Plasmodium falciparum to Anopheles gambiae.* Trans R Soc Trop Med Hyg, 1996. **90**(6): p. 621-4.

22. Ouédraogo, A.L., et al., *Age-dependent distribution of Plasmodium falciparum gametocytes quantified by Pfs25 real-time QT-NASBA in a cross-sectional study in Burkina Faso.* Am J Trop Med Hyg, 2007. **76**(4): p. 626-30.

23. Walldorf, J.A., et al., *School-Age Children Are a Reservoir of Malaria Infection in Malawi.* PLoS One, 2015. **10**(7): p. e0134061.

24. Noedl, H., et al., *Evidence of artemisinin-resistant malaria in western Cambodia.* N Engl J Med, 2008. **359**(24): p. 2619-20.

25. Dondorp, A.M., et al., *Artemisinin resistance in Plasmodium falciparum malaria.* N Engl J Med, 2009. **361**(5): p. 455-67.

26. Amaratunga, C., et al., *Artemisinin-resistant Plasmodium falciparum in Pursat province, western Cambodia: a parasite clearance rate study.* Lancet Infect Dis, 2012. **12**(11): p. 851-8.

27. Mbengue, A., et al., *A molecular mechanism of artemisinin resistance in Plasmodium falciparum malaria.* Nature, 2015. **520**(7549): p. 683-7.

28. Straimer, J., et al., *Drug resistance. K13-propeller mutations confer artemisinin resistance in Plasmodium falciparum clinical isolates.* Science, 2015. **347**(6220): p. 428-31.

29. Stica, C., et al., *Characterizing the molecular and metabolic mechanisms of insecticide resistance in Anopheles gambiae in Faranah, Guinea.* Malar J, 2019. **18**(1): p. 244.

30. Uwimana, A., et al., *Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study.* Lancet Infect Dis, 2021. **21**(8): p. 1120-1128.

31. Ross, A. and N.M. Brancucci, *On a mission to block transmission.* Elife, 2018. **7**.

32. Chawla, J., J. Oberstaller, and J.H. Adams, *Targeting Gametocytes of the Malaria Parasite Plasmodium falciparum in a Functional Genomics Era: Next Steps.* Pathogens, 2021. **10**(3).

33. Nwakanma, D., et al., *High gametocyte complexity and mosquito infectivity of Plasmodium falciparum in the Gambia.* Int J Parasitol, 2008. **38**(2): p. 219-27.

34. Dinko, B., et al., *Gametocyte Development and Carriage in Ghanaian Individuals with Uncomplicated Plasmodium falciparum Malaria.* Am J Trop Med Hyg, 2018. **99**(1): p. 57-64.

35. Nacher, M., et al., *Decreased hemoglobin concentrations, hyperparasitemia, and severe malaria are associated with increased Plasmodium falciparum gametocyte carriage.* J Parasitol, 2002. **88**(1): p. 97-101.

36. Sondo, P., et al., *Plasmodium falciparum gametocyte carriage in symptomatic patients shows significant association with genetically diverse infections, anaemia, and asexual stage density.* Malar J, 2021. **20**(1): p. 31.

37. Portugaliza, H.P., et al., *Plasmodium falciparum sexual conversion rates can be affected by artemisinin-based treatment in naturally infected malaria patients.* EBioMedicine, 2022. **83**: p. 104198.

38. Oduma, C.O., et al., *Increased investment in gametocytes in asymptomatic Plasmodium falciparum infections in the wet season.* BMC Infect Dis, 2021. **21**(1): p. 44.

39. Ouédraogo, A.L., et al., *The plasticity of Plasmodium falciparum gametocytaemia in relation to age in Burkina Faso.* Malar J, 2010. **9**: p. 281.

40. Andolina, C., et al., *Plasmodium falciparum gametocyte carriage in longitudinally monitored incident infections is associated with duration of infection and human host factors.* Sci Rep, 2023. **13**(1): p. 7072.

41. Barry, A., et al., *Higher gametocyte production and mosquito infectivity in chronic compared to incident Plasmodium falciparum infections.* Nat Commun, 2021. **12**(1): p. 2443.

42. Bruce, M.C., et al., *Commitment of the malaria parasite Plasmodium falciparum to sexual and asexual development.* Parasitology, 1990. **100 Pt 2**: p. 191-200.

43. Filarsky, M., et al., *GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing.* Science, 2018. **359**(6381): p. 1259-1263.

44. Rea, E., K.G. Le Roch, and R. Tewari, *Sex in Plasmodium falciparum: Silence Play between GDV1 and HP1.* Trends Parasitol, 2018. **34**(6): p. 450-452.

45. Kafsack, B.F., et al., *A transcriptional switch underlies commitment to sexual development in malaria parasites.* Nature, 2014. **507**(7491): p. 248-52.

46. Brancucci, N.M.B., et al., *Heterochromatin protein 1 secures survival and transmission of malaria parasites.* Cell Host Microbe, 2014. **16**(2): p. 165-176.

47. Balaji, S., et al., *Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains.* Nucleic Acids Res, 2005. **33**(13): p. 3994-4006.

48. Poran, A., et al., *Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites.* Nature, 2017. **551**(7678): p. 95-99.

49. Josling, G.A., et al., *Dissecting the role of PfAP2-G in malaria gametocytogenesis.* Nat Commun, 2020. **11**(1): p. 1503.

50. Prajapati, S.K., et al., *The transcriptome of circulating sexually committed Plasmodium falciparum ring stage parasites forecasts malaria transmission potential.* Nat Commun, 2020. **11**(1): p. 6159.

51. Che Julius, N., F.d.A.R. Thiago, and P. Gabriele, *The Biology of Malaria Gametocytes*, in *Current Topics in Malaria*, J.R.-M. Alfonso, Editor. 2016, IntechOpen: Rijeka. p. Ch. 7.

52. Chin, W. and W.E. Collins, *Comparative studies of three strains of Plasmodium falciparum isolated by the culture method of Trager and Jensen.* Am J Trop Med Hyg, 1980. **29**(6): p. 1143-6.

53. Graves, P.M., R. Carter, and K.M. McNeill, *Gametocyte production in cloned lines of Plasmodium falciparum.* Am J Trop Med Hyg, 1984. **33**(6): p. 1045-50.

54. Bhasin, V.K. and W. Trager, *Gametocyte-forming and non-gametocyte-forming clones of Plasmodium falciparum.* Am J Trop Med Hyg, 1984. **33**(4): p. 534-7.

55. Eksi, S., et al., *Plasmodium falciparum gametocyte development 1 (Pfgdv1) and gametocytogenesis early gene identification and commitment to sexual development.* PLoS Pathog, 2012. **8**(10): p. e1002964.

56. Brancucci, N.M.B., et al., *Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite Plasmodium falciparum.* Cell, 2017. **171**(7): p. 1532-1544.e15.

57. Schneider, P. and S.E. Reece, *The private life of malaria parasites: Strategies for sexual reproduction.* Mol Biochem Parasitol, 2021. **244**: p. 111375.

58. Carter, L.M., et al., *Stress and sex in malaria parasites: Why does commitment vary?* Evol Med Public Health, 2013. **2013**(1): p. 135-47.

59. Usui, M., et al., *Plasmodium falciparum sexual differentiation in malaria patients is associated with host factors and GDV1-dependent genes.* Nat Commun, 2019. **10**(1): p. 2140.

60. Buckling, A. and A.F. Read, *The effect of partial host immunity on the transmission of malaria parasites.* Proc Biol Sci, 2001. **268**(1483): p. 2325-30.

61. Ono, T., et al., *Gametocytogenesis induction by Berenil in cultured Plasmodium falciparum.* Exp Parasitol, 1993. **77**(1): p. 74-8.

62. Motard, A., et al., *Immunization with the malaria heat shock like protein hsp70-1 enhances transmission to the mosquito.* Int Immunol, 1995. **7**(1): p. 147-50.

63. Smalley, M.E. and J. Brown, *Plasmodium falciparum gametocytogenesis stimulated by lymphocytes and serum from infected Gambian children.* Trans R Soc Trop Med Hyg, 1981. **75**(2): p. 316-7.

64. Dantzler, K.W., et al., *Naturally acquired immunity against immature Plasmodium falciparum gametocytes.* Sci Transl Med, 2019. **11**(495).

65. Koepfli, C., et al., *Blood-Stage Parasitaemia and Age Determine Plasmodium falciparum and P. vivax Gametocytaemia in Papua New Guinea.* PLoS One, 2015. **10**(5): p. e0126747.

66. Tonwong, N., et al., *Natural infection of Plasmodium falciparum induces inhibitory antibodies against gametocyte development in human hosts.* Jpn J Infect Dis, 2012. **65**(2): p. 152-6.

67. Riley, E.M., et al., *Cellular and humoral immune responses to Plasmodium falciparum gametocyte antigens in malaria-immune individuals. Limited response*

to the 48/45-kilodalton surface antigen does not appear to be due to MHC restriction. J Immunol, 1990. **144**(12): p. 4810-6.

68. Good, M.F., et al., *Human T clones reactive to the sexual stages of Plasmodium falciparum malaria. High frequency of gamete-reactive T cells in peripheral blood from nonexposed donors.* J Immunol, 1987. **138**(1): p. 306-11.

69. Gebru, T., et al., *Recognition of Plasmodium falciparum mature gametocyte-infected erythrocytes by antibodies of semi-immune adults and malaria-exposed children from Gabon.* Malar J, 2017. **16**(1): p. 176.

70. Ouédraogo, A.L., et al., *Substantial contribution of submicroscopical Plasmodium falciparum gametocyte carriage to the infectious reservoir in an area of seasonal transmission.* PLoS One, 2009. **4**(12): p. e8410.

71. Patgiri, S.J., et al., *Characterization of drug resistance and genetic diversity of Plasmodium falciparum parasites from Tripura, Northeast India.* Sci Rep, 2019. **9**(1): p. 13704.

72. Roh, M.E., et al., *High Genetic Diversity of Plasmodium falciparum in the Low-Transmission Setting of the Kingdom of Eswatini.* J Infect Dis, 2019. **220**(8): p. 1346-1354.

73. Nkhoma, S.C., et al., *Intra-host dynamics of co-infecting parasite genotypes in asymptomatic malaria patients.* Infect Genet Evol, 2018. **65**: p. 414-424.

74. Touray, A.O., et al., *Diversity and Multiplicity of P. falciparum infections among asymptomatic school children in Mbita, Western Kenya.* Sci Rep, 2020. **10**(1): p. 5924.

75. Lalremruata, A., et al., *Species and genotype diversity of Plasmodium in malaria patients from Gabon analysed by next generation sequencing.* Malar J, 2017. **16**(1): p. 398.

76. Wargo, A.R., et al., *Transmission stage investment of malaria parasites in response to in-host competition.* Proc Biol Sci, 2007. **274**(1625): p. 2629-38.

77. Mideo, N., *Parasite adaptations to within-host competition.* Trends Parasitol, 2009. **25**(6): p. 261-8.

78. Churcher, T.S., et al., *Predicting mosquito infection from Plasmodium falciparum gametocyte density and estimating the reservoir of infection.* Elife, 2013. **2**: p. e00626.

79. Pollitt, L.C., et al., *Competition and the evolution of reproductive restraint in malaria parasites.* Am Nat, 2011. **177**(3): p. 358-67.

80. Mideo, N. and T. Day, *On the evolution of reproductive restraint in malaria.* Proc Biol Sci, 2008. **275**(1639): p. 1217-24.

81. Wampfler, R., et al., *Novel genotyping tools for investigating transmission dynamics of Plasmodium falciparum.* J Infect Dis, 2014. **210**(8): p. 1188-97.

82. Menegon, M., et al., *Genotyping of Plasmodium falciparum gametocytes by reverse transcriptase polymerase chain reaction.* Mol Biochem Parasitol, 2000. **111**(1): p. 153-61.

83. Abdel-Wahab, A., et al., *Dynamics of gametocytes among Plasmodium falciparum clones in natural infections in an area of highly seasonal transmission.* The Journal of Infectious Diseases, 2002. **185**(12): p. 1838-1842.

84. Early, A.M., et al., *Detection of low-density Plasmodium falciparum infections using amplicon deep sequencing.* Malar J, 2019. **18**(1): p. 219.

85. Lerch, A., et al., *Development of amplicon deep sequencing markers and data analysis pipeline for genotyping multi-clonal malaria infections.* BMC Genomics, 2017. **18**(1): p. 864.

86. Lerch, A., et al., *Longitudinal tracking and quantification of individual Plasmodium falciparum clones in complex infections.* Sci Rep, 2019. **9**(1): p. 3333.

87. Hermsen, C.C., et al., *Detection of Plasmodium falciparum malaria parasites in vivo by real-time quantitative PCR.* Mol Biochem Parasitol, 2001. **118**(2): p. 247-51.

88. Howick, V.M., et al., *The Malaria Cell Atlas: Single parasite transcriptomes across the complete Plasmodium life cycle.* Science, 2019. **365**(6455).

89. Tessema, S.K., et al., *Sensitive, Highly Multiplexed Sequencing of Microhaplotypes From the Plasmodium falciparum Heterozygome.* The Journal of Infectious Diseases, 2020.

90. Buchwald, A.G., et al., *Association Between Age and Plasmodium falciparum Infection Dynamics.* Am J Epidemiol, 2019. **188**(1): p. 169-176.

91. Miller, R.H., et al., *A deep sequencing approach to estimate Plasmodium falciparum complexity of infection (COI) and explore apical membrane antigen 1 diversity.* Malar J, 2017. **16**(1): p. 490.

92. Park, D.J., et al., *Sequence-based association and selection scans identify drug resistance loci in the Plasmodium falciparum malaria parasite.* Proc Natl Acad Sci U S A, 2012. **109**(32): p. 13052-7.

93. Ye, J., et al., *Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction.* BMC Bioinformatics, 2012. **13**: p. 134.

94. Hathaway, N.J., et al., *SeekDeep: single-base resolution de novo clustering for amplicon deep sequencing.* Nucleic acids research, 2018. **46**(4): p. e21-e21.

95. Lin, L.I., *A concordance correlation coefficient to evaluate reproducibility.* Biometrics, 1989. **45**(1): p. 255-68.

96. Lopez, L. and C. Koepfli, *Systematic review of Plasmodium falciparum and Plasmodium vivax polyclonal infections: Impact of prevalence, study population characteristics, and laboratory procedures.* PLoS One, 2021. **16**(6): p. e0249382.

97. Escalante, A.A., et al., *Polymorphism in the gene encoding the Pfs48/45 antigen of Plasmodium falciparum. XI. Asembo Bay Cohort Project.* Molecular and Biochemical Parasitology, 2002. **119**(1): p. 17-22.

98. Paganotti, G.M., et al., *Genetic complexity and gametocyte production of Plasmodium falciparum in Fulani and Mossi communities in Burkina Faso.* Parasitology, 2006. **132**(5): p. 607-614.

99. Zhong, D., et al., *Molecular approaches to determine the multiplicity of Plasmodium infections.* Malaria journal, 2018. **17**(1): p. 172-172.

100. Jones, S., et al., *Should Deep-Sequenced Amplicons Become the New Gold Standard for Analyzing Malaria Drug Clinical Trials?* Antimicrob Agents Chemother, 2021. **65**(10): p. e0043721.

101. Daniels, R., et al., *A general SNP-based molecular barcode for Plasmodium falciparum identification and tracking.* Malar J, 2008. **7**: p. 223.

102. Neafsey, D.E., et al., *Genetic Diversity and Protective Efficacy of the RTS,S/AS01 Malaria Vaccine.* The New England journal of medicine, 2015. **373**(21): p. 2025-2037.

103. Gruenberg, M., et al., *Amplicon deep sequencing improves Plasmodium falciparum genotyping in clinical trials of antimalarial drugs.* Scientific Reports, 2019. **9**(1): p. 17790.

104. Lin, J.T., et al., *Using Amplicon Deep Sequencing to Detect Genetic Signatures of Plasmodium vivax Relapse.* J Infect Dis, 2015. **212**(6): p. 999-1008.

105. Hathaway, N.J., et al., *SeekDeep: single-base resolution de novo clustering for amplicon deep sequencing.* Nucleic Acids Research, 2017. **46**(4): p. e21-e21.

106. Shah, Z., et al., *Whole-genome analysis of Malawian Plasmodium falciparum isolates identifies possible targets of allele-specific immunity to clinical malaria.* PLoS Genet, 2021. **17**(5): p. e1009576.

107. Lasonder, E., et al., *Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression.* Nucleic Acids Res, 2016. **44**(13): p. 6087-101.

108. Wampfler, R., et al., *Strategies for detection of Plasmodium species gametocytes.* PLoS One, 2013. **8**(9): p. e76316.

109. Okell, L.C., et al., *Submicroscopic infection in Plasmodium falciparum-endemic populations: a systematic review and meta-analysis.* J Infect Dis, 2009. **200**(10): p. 1509-17.

110. Delves, M.J., et al., *Male and female Plasmodium falciparum mature gametocytes show different responses to antimalarial drugs.* Antimicrob Agents Chemother, 2013. **57**(7): p. 3268-74.

111. Talman, A.M., et al., *Gametocytogenesis: the puberty of Plasmodium falciparum.* Malar J, 2004. **3**: p. 24.

112. Ashley, E.A., et al., *Spread of artemisinin resistance in Plasmodium falciparum malaria.* N Engl J Med, 2014. **371**(5): p. 411-23.

113. Toé, K.H., et al., *Increased pyrethroid resistance in malaria vectors and decreased bed net effectiveness, Burkina Faso.* Emerg Infect Dis, 2014. **20**(10): p. 1691-6.

114. Abdi, A.I., et al., *Plasmodium falciparum adapts its investment into replication versus transmission according to the host environment.* Elife, 2023. **12**.

115. Mobegi, V.A., et al., *Genome-wide analysis of selection on the malaria parasite Plasmodium falciparum in West African populations of differing infection endemicity.* Mol Biol Evol, 2014. **31**(6): p. 1490-9.

116. Lamptey, H., et al., *The prevalence of submicroscopic Plasmodium falciparum gametocyte carriage and multiplicity of infection in children, pregnant women and adults in a low malaria transmission area in Southern Ghana.* Malar J, 2018. **17**(1): p. 331.

117. Koella, J.C. and R. Antia, *Optimal Pattern of Replication and Transmission for Parasites with Two Stages in Their Life Cycle.* Theoretical Population Biology, 1995. **47**(3): p. 277-291.

118. Josling, G.A., K.C. Williamson, and M. Llinás, *Regulation of Sexual Commitment and Gametocytogenesis in Malaria Parasites.* Annu Rev Microbiol, 2018. **72**: p. 501-519.

119. Venugopal, K., et al., *Plasmodium asexual growth and sexual development in the haematopoietic niche of the host.* Nat Rev Microbiol, 2020. **18**(3): p. 177-189.

120. De Niz, M., et al., *Plasmodium gametocytes display homing and vascular transmigration in the host bone marrow*. Sci Adv, 2018. **4**(5): p. eaat3775.