IMPLICATIONS OF ASYMPTOMATIC *PLASMODIUM FALCIPARUM* INFECTIONS FOR FUTURE SYMPTOMATIC MALARIA INFECTION AND ONWARD TRANSMISSION IN WESTERN KENYA

Kelsey M. Sumner

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the Gillings School of Global Public Health.

Chapel Hill
2021

Approved by:

Brian W. Pence

Steven M. Taylor

Wendy Prudhomme-O'Meara

Jessie K. Edwards

Michael Emch

ABSTRACT

Kelsey M. Sumner: Implications of asymptomatic *Plasmodium falciparum* infections for future symptomatic malaria infection and onward transmission in Western Kenya
(Under the direction of Brian W. Pence)

Despite efforts to reduce the malaria burden in high transmission areas like Western Kenya, malaria has persisted, making it important to identify remaining malaria reservoirs. Asymptomatic infections could be sustaining transmission, but their role as a reservoir has not been quantified. Furthermore, the relationship between asymptomatic malaria and future symptomatic infection is poorly understood.

Using a 29-month cohort of 268 participants residing in Western Kenya, this dissertation aimed to (1A) investigate the hazard of symptomatic *P. falciparum* when exposed to asymptomatic malaria versus no infection; (1B) establish the odds of symptomatic infection when exposed to new compared to previously acquired infections; and (2) estimate the relative contributions of asymptomatic and symptomatic human infections to successful onward mosquito transmission events.

With a frailty Cox model, aim 1A found that infection with asymptomatic malaria compared to being uninfected greatly increased the short-term, 1-month hazard of symptomatic malaria [hazard ratio: 2.61, 95% CI: 2.05 to 3.33], regardless of parasite density or participant age, but the association weakened as the follow-up period was expanded. Next, using amplicon deep sequencing to determine genetically distinct malaria infections (haplotypes) acquired over time, aim 1B identified that, compared to

infections with only recurrent haplotypes, incident infections with only new haplotypes had higher odds of symptomatic malaria [odds ratio (OR): 3.24, 95% CI: 1.20 to 8.78] but infections with both new and recurrent haplotypes had comparable odds of symptomatic malaria [OR: 0.64, 95% CI: 0.15 to 2.65]. Using amplicon deep sequencing of human and mosquito samples as well as probabilistic modelling, aim 2 observed that, compared with symptomatic infections, asymptomatic infections more than doubled the odds of transmission to a mosquito [OR 2.66, 95% CI: 2.05 to 3.47] and were the likely source of 94.6% (95% CI: 93.1 to 95.8%) of mosquito infections.

These findings indicate that asymptomatic infections increase the 1-month hazard of symptomatic malaria, are more common in incident infections with previously seen haplotypes, and are major contributors to mosquito infections. Taken together, this research provides a rationale to include asymptomatic infections as a part of malaria reduction interventions in high transmission regions.

ACKNOWLEDGMENTS

First and foremost, I would like to thank the Webuye study participants for their long-term participation in this study. I appreciate the hard work of the field technicians in Webuye and their engagement with the study participants: I. Khaoya, L. Marango, E. Mukeli, E. Nalianya, J. Namae, L. Nukewa, E. Wamalwa, and A. Wekesa. In addition, I am thankful for the study coordinating skills of J. Kipkoech Kirui; he has played a key role in the project's management. L. Abel and A. Obala have added entomological expertise to this project. I appreciate J. Mangeni for study development and discussion.

I am indebted to E. Freedman for both creating and implementing the laboratory sample processing for this project. I thank the assistance with laboratory sample and data processing by V. Liao, A. Nantume, S. Kim, and J. Saelens. I also appreciate the analytical advice I received from A. Wesolowski, J. Petersen, and C. Markwalter.

I appreciate the funding resources for this project. I received a Carolina Global Initiative Pre-Dissertation Travel Award to fund my travel to the dissertation research site, where I was able to meet with the local team, shadow sample collection, and advance dissertation question development. The dissertation project was supported by the National Institute of Allergy and Infectious Diseases (R21AI126024 and R01AI146849). While completing this work, I was also supported by GlaxoSmithKline.

My dissertation committee has been an integral part of this project, and I am thankful for their support and advice. Thank you to S. Taylor who, being a professor at

Duke University, has not had the official role as my advisor but has acted as one nonetheless. I thank him for always having his door open to talk through concepts, pushing me to expand my work with sensitivity analyses, and having a supportive laboratory environment in which to complete my PhD. I thank my advisor, B. Pence, for mentoring me despite my research being in a different subject area. Thank you for learning my malaria jargon, adding me to your advisee group, and always offering kind words of advice. I am appreciative of W. Prudhomme-O'Meara including me in this project, data set, and analyses that she has worked to implement for over 10 years. I appreciate her having me as a part of the team and am thankful of her analytical expertise. I also thank M. Emch for being a part of my dissertation committee, offering a geographical perspective for the analyses, and always being open to talk through concepts. I thank J. Edwards for joining my committee later in the process, offering helpful advice for the time-to-event analyses, and being a supportive committee member. Finally, thank you to S. Meshnick who was an excellent advisor for my master's and PhD at UNC-Chapel Hill. He pushed me to always question results and helped me grow as a scientist. I thank him for his mentorship even during his illness.

Finally, thank you to my family and friends for their support during the PhD process. Thank you especially to A. Keller-Silver, J. Sumner, C. Sumner, A. Sumner, W. Sumner, J. Masten, B. Masten, M. Masten, and A. M. Masten for reading my work and supporting me throughout the program. I am especially thankful of my husband, A. A. Masten, for his unwavering encouragement, bringing me hot chocolate when I stayed up late writing or studying, and teaching me to be a better scientific communicator through many presentation practices. Thank you all for your support!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACT	Artemisinin-based combination therapy
AIC	Akaike information criteria
AL	Artemether-Lumefantrine
CI	Confidence interval
DAG	Directed acyclic graph
DBS	Dried blood spot
gDNA	Genomic deoxyribonucleic acid
HR	Hazard ratio
HS-RDT	Highly-sensitive rapid diagnostic test
IQR	Interquartile range
ITN	Insecticide-treated net
MDA	Mass drug administration
MOI	Multiplicity of infection
NE	Not evaluated
OR	Odds ratio
PFAMA1	*Plasmodium falciparum* apical membrane antigen-1
PFCSP	*Plasmodium falciparum* circumsporozoite protein
qPCR	Real-time, quantitative polymerase chain reaction
RDT	Rapid diagnostic test
SE	Standard error
SNP	Single-nucleotide polymorphism

CHAPTER I: SPECIFIC AIMS

Over 70% of Kenya's population lives in a high transmission area of *Plasmodium falciparum.*[1,2] Prevalence is highest in Western Kenya where parasites are found in up to 80% of residents.[3] Malaria reduction efforts are ongoing;[4,5] but, high *P. falciparum* prevalence persists, highlighting the need to identify the region's malaria reservoirs.[6,7]

Asymptomatic malaria infections could be reservoirs for sustained malaria transmission and are prevalent in Western Kenya.[6,8,9] Previous work has longitudinally assessed asymptomatic malaria incidence, intra-host immunology, parasite density, and mosquito biting exposure;[10–13] but, the amount asymptomatically-infected humans transmit malaria to mosquitoes remains unresolved. It is also unclear whether asymptomatic infections impact the severity of subsequent symptomatic infections within the same host, with asymptomatic infections increasing the risk of future symptomatic infections in some studies and decreasing the risk in others.[14–23] Understanding the transmission dynamics and disease progression of asymptomatic malaria is important for understanding its reservoir role.

Our overall objective was to explore the natural history of asymptomatic *P. falciparum* infection and its relationship to future symptomatic infection and transmission to mosquitoes. Our central hypothesis was that preceding, new asymptomatic infections would increase the short-term risk of future symptomatic infection and be a large source of malaria transmission compared to symptomatic malaria. To investigate this hypothesis, we used a longitudinal cohort in Webuye, Western Kenya, a hyperendemic

malaria site. We collected blood samples from human participants asymptomatic (monthly) and symptomatic (as suspected) for malaria as well as mosquitoes from participant households (weekly). We aimed to:

Aim 1A: Investigate the hazard of symptomatic *P. falciparum* when exposed to asymptomatic infections compared to no malaria infections. Asymptomatic infections had malaria parasites in blood samples and no malaria symptoms. Symptomatic infections had malaria parasites in blood samples and at least one malaria symptom (ex. fever, headache, nausea). We assessed the hazard of symptomatic *P. falciparum* infection in humans across the time-varying exposure (asymptomatic malaria infection *vs.* no malaria infection). We hypothesized that participants with asymptomatic infections during the study period would have a higher short-term hazard of future symptomatic infection compared to those that were uninfected.

Aim 1B: Establish the odds of symptomatic *P. falciparum* infection compared to asymptomatic infection when exposed to new infections compared to recurrent or persistent infections. Using only blood samples from humans with asymptomatic and symptomatic malaria, we performed amplicon deep sequencing of two polymorphic *P. falciparum* gene segments to create haplotypes (distinct sequences of a parasite gene target) that represented genetically-distinct malaria infections. We determined genetic relatedness of an individual's infections by comparing haplotypes across infections. We analyzed the odds of symptomatic compared to asymptomatic malaria across infections with new versus recurrent or persistent haplotypes. We hypothesized we would observe greater odds of symptomatic infection compared to asymptomatic infection when a new haplotype was present in the infection.

Aim 2: Estimate the relative contributions of asymptomatic and symptomatic *P. falciparum* human infections to successful onward mosquito transmission events in a high transmission area. Using blood samples from (i) humans with asymptomatic and symptomatic malaria and (ii) female *Anopheles* mosquitoes from those participants' households, we sequenced two parasite gene targets and counted the number of shared parasite haplotypes between a human infection and mosquito found in the human's household. A probability of human-to-mosquito transmission value was calculated based on the time, distance, and haplotypes shared between human and mosquito samples; this value was compared across humans with asymptomatic to symptomatic infections. We hypothesized that asymptomatic humans compared to symptomatic humans would have a higher probability of malaria transmission to mosquitoes.

Using data collected with an established malaria research team in Webuye,[24–33] we determined if asymptomatic infections were providing a protective effect against symptomatic infection or becoming symptomatic themselves over time. We also established if asymptomatic humans highly contributed to mosquito infection and onward transmission. Findings inform the utility of active test and treat strategies for finding asymptomatic malaria in endemic regions.

CHAPTER II: BACKGROUND

P. falciparum Overview and Global Burden

Malaria is a large global health problem with over 200 million cases reported annually.[34] Most cases occur in sub-Saharan Africa, which is the focal area for transmission of the most deadly malaria species, *Plasmodium falciparum.*[2,34] Over 70% of Kenya's population lives in a high transmission area of *P. falciparum*, leading to a high number of malarial deaths.[1] In fact, Kenya is one of 29 countries that contributed to 95% of malarial deaths that occurred world-wide in 2019.[34] In Kenya, the risk of malaria is highest in the Western region where *P. falciparum* parasite prevalence is greater than 300 cases per 1000 population.[1]

Many vector control and preventative therapy measures have been implemented in hyperendemic areas like Western Kenya. In sub-Saharan Africa, access to insecticide-treated bed nets increased from 5 to 68% from 2000 to 2019, and indoor residual spraying has remained a protective vector control strategy.[34] In mainly West Africa, pregnant women and children have received prophylactic antimalarials through seasonal malaria chemoprevention programs.[34] These preventative efforts are reducing malaria cases across Africa, but case reductions have plateaued since 2010,[34,35] with efforts presumably having a smaller effect. Many regions are still experiencing a high number of malaria cases and, for some countries, increasing case counts. This malaria persistence could be due to intervention efforts not effectively reducing the malaria

reservoir, which describes individuals and infections that disproportionately contribute more to parasite transmission. As intervention efforts push regions toward elimination, it becomes increasingly important to identify these malaria reservoirs.

P. falciparum Life Cycle

Identification of reservoirs for malaria transmission requires knowledge of the disease's life cycle. *P. falciparum's* life cycle involves an asexual stage in humans where commitment to the sexual gametocyte stage occurs and a sexual replication stage in the vector, female *Anopheles* mosquitoes, where replication between male and female gametocytes occurs.[36]

To assess human-to-mosquito malaria transmission, we followed the malaria parasite at different stages in its life cycle. For study participants, we captured asexual and sexual malaria parasites during the blood stage of infection using dried blood spots. The blood stage of the parasite life cycle is responsible for causing clinical symptom development in infected humans.[37] For *P. falciparum* there is a hypothesized incubation period of 9 to 14 days between a person being bitten by a mosquito and the development of symptoms;[38] however, this incubation period is not well supported with a study finding symptoms developed up to 12 weeks after the initial asymptomatic infection was recorded.[16] In some case studies, untreated infections remained asymptomatic for 6 months to 13 years.[39] Using the dried blood spots collected during the parasite's blood stage in humans, we identified parasite haplotypes (distinct sequences of a gene target) as a measure of parasite genetic diversity. We detected parasite haplotypes from humans consisting of asexual and sexual (gametocyte) parasites.

When female *Anopheles* mosquitoes bite a malaria-infected person, they ingest malaria gametocytes. To follow gametocyte transmission to mosquitoes, we collected mosquitoes from participants' households. Over an average of 9 to 18 days,[37] ingested gametocytes sexually reproduce in the mosquitoes' midgut to produce ookinetes and then oocysts that later burst to release sporozoites into the mosquitoes' salivary glands. These sporozoites could then be transferred to another person via a mosquito bite.[40,41] Multiple genetically-distinct malaria parasites could reside in a mosquito's salivary glands and mosquitoes could take multiple blood meals, leading to human infection with multiple malaria parasite haplotypes.[42] We detected parasite haplotypes in female *Anopheles* mosquito abdomens and matched those to parasite haplotypes in humans to capture human-to-mosquito malaria transmission and quantify how asymptomatic and symptomatic infections in humans act as reservoirs for transmission.

Asymptomatic Malaria and Mosquito Transmission

Asymptomatic malaria has been hypothesized to be a significant source, or reservoir, of sustained malaria transmission.[9,36,43–46] Asymptomatic malaria infections could be pre-symptomatic infections that have not yet progressed to symptomatic malaria or infections that never progress to symptomatic malaria due to a partial, anti-disease immune response.[47] Twenty-four percent of sub-Saharan Africa's population is estimated to harbor an asymptomatic infection,[35] and asymptomatic infection prevalence correlates with transmission intensity; high transmission areas have higher asymptomatic infection prevalence.[48–51] In high transmission regions, asymptomatic malaria also likely occurs at a higher prevalence than symptomatic malaria.[35,52–54] This is seen in Western Kenya where asymptomatic infections were found in 38 to 50% of

school-aged children living in hyperendemic areas.[6,8]

High prevalence of asymptomatic malaria is an enormous issue because the same asymptomatic infection could persist in individuals for years.[39] This long infection duration allows many opportunities for a mosquito to bite an infected individual and transmit malaria to someone else. Many factors influence the likelihood of this route of malaria transmission, such as mosquito gametocyte ingestion, malaria parasite density (asexual and sexual), human host factors, multiclonal infections, and immunological responses.[55–58]

The frequency of *P. falciparum* transmission from asymptomatic or symptomatic humans to mosquitoes has been measured using cross-sectional and experimental studies, but results have been conflicting.[59–63] For example, an experimental study in Western Kenya assessed membrane feeding of *Anopheles gambiae* mosquitoes on asymptomatic and symptomatic infected humans and found that asymptomatic infections passed more gametocytes to mosquitoes than symptomatic infections.[61] This was observed in the higher oocyst rate in the mosquitoes that fed on asymptomatic humans (12%) compared to those that fed on symptomatic humans (0.6%).[61] In contrast, a study in Brazil discovered that parasites from symptomatic infections were transmitted to mosquitoes more often than those from asymptomatic infections, but asymptomatic infections were more prevalent and persisted for a longer duration.[59] It is important to note that these studies had small sample sizes, participants who recently took antimalarial treatment or contrived mosquito feeding systems, making results not easily generalizable. Modeling studies have tried to estimate the amount of human-to-malaria transmission that could occur,[64,65] suggesting that submicroscopic individuals

(with parasite densities below the threshold for microscopy detection) contribute to 20 to 50% of human-to-mosquito malaria transmissions in low-transmission settings[65] and >50% of transmissions in high-transmission settings.[66,67] Many cross-sectional and experimental studies have also looked at the amount of human-to-mosquito transmission for sub-microscopic infections.[66,68–77] However, these modelling, cross-sectional, and experimental studies focused on submicroscopic infections, which are not always asymptomatic. Similarly, asymptomatic infections are not always submicroscopic. Thus, the amount of onward transmission caused by asymptomatic compared to symptomatic malaria has not been sufficiently assessed and asymptomatic malaria's role as a reservoir has not been clarified.

Asymptomatic Malaria and Human Health

The prevalence of asymptomatic malaria varies across groups depending on human host factors that influence susceptibility to symptom development. Sex has been observed as an effect modifier for asymptomatic and symptomatic infections with males experiencing higher prevalence of both infection types.[78] Age has also been reported as an important factor for malaria infection as young children have the highest risk of symptomatic or severe infection.[79–81] Additional factors like transmission intensity and infection with multiple genetically-distinct malaria parasites (multiclonality) could explain susceptibility to symptomatic malaria. Multiclonal infections in older children in high transmission areas have been associated with a reduced risk of symptomatic malaria,[82,83] whereas multiclonal infections in younger children in low transmission areas have been associated with an increased risk of symptomatic malaria.[84–86]

Because of the high likelihood of being infected with multiple malaria parasites at

the same time, it is imperative to assess multiclonality in high malaria transmission areas like Kenya.[70,87–89] Multiclonality could be an indicator for how haplotype diversity influences development of symptomatic infections, with higher haplotype diversity protecting against symptomatic progression.[90,91] Repeated exposure to haplotypes has been hypothesized to create partial anti-disease immunity to those haplotypes, causing lower parasite densities and fewer symptomatic infections.[92–95] This theory has been supported by studies that observed an association between new infections and an increased risk of symptomatic illness, finding that an asymptomatic infection rarely became symptomatic.[20,96–99] Yet, another study observed 47% of children with asymptomatic infections later developed symptomatic illness from the initial asymptomatic infection.[16] We used amplicon deep sequencing technology to assess genetically-distinct haplotypes within and between infections during the study period and determine if an individual harbors the same parasites in asymptomatic compared to symptomatic infections.

Chronic asymptomatic infections could have negative health consequences.[100] Chronic asymptomatic malaria has been associated with malnutrition;[101,102] low platelet count and hemoglobin levels in children;[103] anemia;[101,102,104,105] adverse pregnancy outcomes such as preterm birth, low birth weight, and mortality;[104,106] invasive bacterial infections (non-typhoid salmonellae);[107] and cognitive impairment.[108] These conditions suggest that it could be beneficial to treat asymptomatic infections.

Conversely, asymptomatic infections could have protective health benefits for the individual against future symptomatic infection; however, the relationship between asymptomatic malaria and future symptomatic illness is still poorly understood. Some

studies observed that having a history of asymptomatic malaria decreased the risk of subsequent symptomatic illness.[17,19–21] Other reports discovered that an asymptomatic infection actually increased the risk of symptomatic malaria.[14–16,18] A few studies observed the direction of effect to differ even within its study population by age groups or transmission intensities.[22,23] The contradictory findings could be due to differences in study populations, designs, and diagnostic techniques. Yet more information is needed to discern how asymptomatic malaria infections are associated with future symptomatic disease development.

Significance

With our study's 14 to 29-month longitudinal design and highly-sensitive amplicon deep sequencing technology, we were able to build upon previous studies of the natural history of asymptomatic malaria. We aimed to fill the gap in knowledge on whether asymptomatic infections were associated with future symptomatic illness and if new genetically-distinct infections were associated with symptom development. We also determined if asymptomatic infections were transmitted to mosquitoes more frequently than symptomatic infections and, thus, serving as a reservoir for sustained malaria transmission. In order to do this, we investigated the incidence of asymptomatic and symptomatic *P. falciparum* infections in human and mosquito samples collected in Webuye, Kenya over 29 months. Study results inform the utility of active test and treat strategies for reducing asymptomatic malaria infections in high transmission regions.

Innovation

To clarify the impact of asymptomatic malaria on intra-host morbidity and onward transmission to mosquitoes, in each of the aims we used enhanced methodology and

study design compared to previous work.

In aim 1A, we used a novel approach to capture how asymptomatic malaria varied over 29 months across all ages (range: 1 to 85 years), creating a more complete view of infection dynamics compared to previous work. Many prior studies only included children,[14–16,18,21,23] missing information about infection dynamics across all ages. Past studies also failed to measure true asymptomatic exposure length by using infrequent cross-sectional surveys to identify asymptomatic infections,[15,16,18,19,21–23] coding exposure binarily as always being exposed after an asymptomatic infection occurred[20] or only following infections for 9 or 30 days.[14,15] We more precisely captured asymptomatic malaria exposure using a time-varying method that allowed participants to change exposure status throughout follow-up.[109]

In aim 1B, we expanded upon previous work by following a diverse parasite population for 14 months with frequent sampling and high-resolution parasite genotyping (amplicon deep sequencing). This approach overcame limitations in previous studies that had small sample sizes with brief follow-up,[97–99] infrequent sampling,[86,99] genotyping approaches with high failure rates,[20] and an inability to capture multiclonal genotypes.[20,86,96–99] Specifically, genotyping approaches that use PCR-restricted fragment length polymorphism to detect size variants,[20,86,96–99] capture only 30% of the unique clones present compared to amplicon deep sequencing.[110] Due to this inability to capture all clones present,[111] prior work employing these approaches failed to accurately classify haplotypes as new, recurrent or persistent. Using fine-scale genotypes created by the more sensitive amplicon deep sequencing method,[110] we were able to more definitively partition the distinct effects of new, recurrent or persistent

haplotypes within infections, which has previously not been done.[16,20,86,96–99]

In aim 2, we built upon previous studies by capturing participant-to-mosquito transmission longitudinally, in a larger study population, and in a natural setting with mosquitoes collected within participants' households. Prior studies measured participant-to-mosquito malaria transmission using experimental direct or membrane feeding by laboratory-reared mosquitoes,[59–63] and therefore could not capture variance in the feeding behaviors of vectors[112] or natural trajectories of infections;[113] the feeding methods used in these studies failed to represent numerous participant-, mosquito-, and parasite-related factors that are critical to transmission, limiting their generalizability. We quantified asymptomatic malaria transmission to mosquitoes in a more natural, generalizable setting using mosquitoes caught in participants' households post-overnight feeding.

The innovative approaches employed across the three dissertation aims allowed us to learn more about the natural history of asymptomatic malaria. Using a novel time-varying exposure coding method to measure asymptomatic malaria exposure in both children and adults, we gained more precise estimates of asymptomatic malaria exposure and its influence on subsequent symptomatic illness. Using frequent sampling and amplicon deep sequencing methodology, we were able to more finely differentiate between new, recurrent, and persistent haplotypes and their association with infection dynamics. Finally, we were the first study to estimate participant-to-mosquito malaria transmission using amplicon deep sequencing and probabilistic modeling, offering a generalizable quantification of participant-to-mosquito malaria transmission.

CHAPTER III: METHODS

Study Site

The longitudinal prospective research study was conducted in Webuye, a town in Bungoma East sub-county in Western Kenya. Webuye is in a rural region where most residents work in small-scale agricultural practices. In 2009, most of the population had access to improved water (88.9%) and sanitation (96.6%), but there was little access to electricity (4.5%) and paved roads (6.0%).[114] The majority of the sub-county's population lives below the poverty line (52.9%).[114] Malaria is a large health issue in the sub-county, with malaria listed as the cause of 89.2% of patients' first outpatient visits to a health facility.[114] Overall, malaria transmission is high and endemic in Webuye.[1] Transmission has two seasonal peaks following the long rainy season (May to June) and the short rainy season (September to October). *P. falciparum* is the main parasite species in the region.[1] The major *Anopheles* mosquito species are *An. gambiae*, *An. arabiensis*, *An. funestus*, and *An. merus.*[1]

Study Population

A longitudinal cohort of individuals residing in 38 households across 3 villages (Kinesamo, Maruti, and Sitabicha) was established in June 2017 and followed until November 2019. The three villages were selected based on their similar high levels of malaria prevalence observed in a previous cross-sectional study in the area [Mean *P. falciparum* prevalence in 2013: Kinesamo (18.4%), Maruti (20.8%), Sitabicha (22.8%)].[32] Within each village, households were enrolled by randomly selecting the index

household from among those in the previous cohort and then enrolling additional households radially until 12 households were enrolled per village. Two enrolled households were replaced with neighboring households when the entire household moved out of the study area during follow-up. Participants were excluded from the study if they were less than 1-year-old, refused any component of the study or sample collection or completed less than 2 months of follow-up.

Participant Sample Collection and Surveillance

For each participant, demographic and behavioral questionnaires were administered and dried blood spot samples collected every month (**Figure 3.1**). Dried blood spots were collected by pricking the participant's finger and placing three blood spots on filter paper. The questionnaires included malaria-related information experienced in the past month, such as travel history, insecticide-treated net (ITN) use, malaria-like illness, and antimalarial use. Study participants were invited to contact the study team when experiencing malaria-like symptoms; when this occurred, a team member visited the participant's household to record symptoms, collect a dried blood spot, and test for *P. falciparum* malaria using a rapid diagnostic test (RDT) (Carestart © Malaria HRP2 *Pf* from Accessbio).[115] If the RDT produced negative results, the participant was referred with the test results to the nearest health facility. If the RDT produced positive results, the participant received Artemisinin-based Combination Therapy (ACT) from the local pharmacy at no charge. Microscopy was not performed.

Mosquito Collection and Identification

One morning each week, the study team visited participant households to collect indoor resting mosquitoes via vacuum aspiration with Prokopacks.[116] Participants were

asked to leave windows and doors closed until the study team arrived. Mosquitoes were stored in collection cups in boxes with icepacks until they were transported to the laboratory in Webuye. Mosquitoes were killed using chloroform and sorted by genus (*Anopheles* or *Culex*) and sex (male or female). The female *Anopheles* were speciated and dissected between the head and abdomen.

Participant and Mosquito Sample Processing

Participant dried blood spot samples and female *Anopheles* mosquito abdomens were shipped to Duke University in Durham, North Carolina, where they were molecularly processed to determine *P. falciparum* infection status and haplotypes. The dried blood spots and mosquito parts were distributed into 384-well plates with one punched dried blood spot per participant per well and one mosquito abdomen per mosquito per well. *P. falciparum* parasite genomic deoxyribonucleic acid (gDNA) was Chelex-extracted from each of the samples. Each sample was tested in duplicate for *P. falciparum* parasites using real-time, quantitative Polymerase Chain Reaction (qPCR) and the *P. falciparum* 364r assay. A human β-tubulin assay was also included in the qPCR as an internal control. Samples were defined as *P. falciparum*-positive if: (i) both replicates amplified *P. falciparum* and both Ct values were < 40 or (ii) 1 replicate amplified *P. falciparum* and Ct value was < 38.

Dual-indexed libraries were prepared for amplicon deep sequencing of the ~300 nucleotide polymorphic *P. falciparum* parasite gene targets apical membrane antigen-1 (*pfama1*) and circumsporozoite protein (*pfcsp*) on an Illumina miSeq platform.[117] Using Trimmomatic, CutAdapt, and BBmap,[118–120] sample reads were mapped to the 3D7 reference sequences for *pfama1* and *pfcsp* and primers removed.[121,122] Sample read

pairing, haplotype calling, and chimera removal were performed using DADA2 version 1.8.[123] This process outputted haplotypes (distinct sequences of the *pfama1* and *pfcsp* gene targets) to be used as a measure of parasite genetic diversity. Haplotypes were censored if: (i) supported by < 250 reads within the sample; (ii) supported by < 3% of the sample's total read depth; (iii) deviation from the expected nucleotide length of 300 for *pfama1* or 288 for *pfcsp*; or (iv) a minority haplotype distinguished by a one single-nucleotide polymorphism (SNP) difference from another haplotype within the sample that had a read depth > 8 times the read depth of the minority haplotype.[124] We also removed a haplotype from the overall population if it was defined by a single variant position that was only variable within that haplotype. All genetic sequences are available through the National Center for Biotechnology Information GenBank (BioProject Number PRJNA646940).

Data Analysis

The data analysis approach was largely motivated by the high amount of malaria transmission in the study site. Malaria endemicity caused a high number of asymptomatic and symptomatic infection events, a diverse malaria parasite population, and frequent turnover of malaria infections. To capture a more complete picture of asymptomatic infection dynamics, we utilized amplicon deep sequencing and multi-level modelling methodology. All analyses were conducted using R (versions 3.6.1 or 4.0.2).[125,126]

Aim 1A Analysis

The main exposure assessed was individual-level exposure to asymptomatic *P. falciparum* infection compared to no *P. falciparum* infection. Exposure status was

ascertained at each monthly follow-up visit and allowed to vary each month using a method proposed by Hernán *et al*.[109] The method treated each monthly follow-up visit as a new study entry, recalculating the time to symptomatic malaria using the monthly follow-up visit date as the origin.[109] Participant follow-up was imputed for the first consecutive missed monthly visit during each follow-up period using the exposure status in the previous month.[9] The main outcome assessed was time to symptomatic malaria infection (in days). A symptomatic infection was *P. falciparum*-positive by both RDT and qPCR in a participant with at least one symptom consistent with malaria during a sick visit. Repeated symptomatic infections were included.

We compared how the short and long-term hazard of symptomatic illness differed across exposure to asymptomatic malaria compared to no malaria infections, expanding participant follow-up from 1 to 3, 6, 12, and 29 months. First, we compared time to symptomatic malaria stratified by asymptomatic exposure using Kaplan-Meier curves and the log-rank test. Next, we assessed the hazard of symptomatic malaria across asymptomatic exposure controlling for confounding covariates identified in a directed acyclic graph (DAG) analysis and participant-level correlation with a frailty Cox proportional hazards model. Effect measure modification by age and sex was assessed by stratifying the multivariate frailty Cox proportional hazards model by age category or sex. Pre-symptomatic and post-treatment sensitivity analyses were conducted to assess differences in model results.

We also investigated how the detectability of asymptomatic infections affected the 1-month, short-term hazard of symptomatic malaria. We compared multivariate frailty Cox proportional hazards model results across varying parasite density thresholds

as well as malaria diagnostic limits of detection.

Aim 1B Analysis

For every person, we classified parasite haplotypes created by amplicon deep sequencing in each of their infections as: (i) new, a haplotype not previously observed in that person; (ii) recurrent, one previously observed in that person but not in the most recent DBS sample; or (iii) persistent, a haplotype previously observed in the most recent DBS sample. For the main exposure, we categorized each infection using the above haplotype classifications, assigning categories independently for *pfama1* and *pfcsp* haplotypes. For the main outcome, each *P. falciparum* infection was categorized as asymptomatic or symptomatic.

Using a multi-level logistic regression model, we assessed odds of symptomatic malaria as a function of new, recurrent or persistent haplotypes in: (i) incident infections where none of the haplotypes in the infection were previously observed in the participant's most recent DBS or (ii) persistent infections where at least one haplotype persisted between consecutive DBS collections occurring within 30 days. The logistic regression model included a participant-level random intercept and controlled for confounding covariates identified in a DAG. We evaluated effect measure modification by age on the multiplicative scale.

Aim 2 Analysis

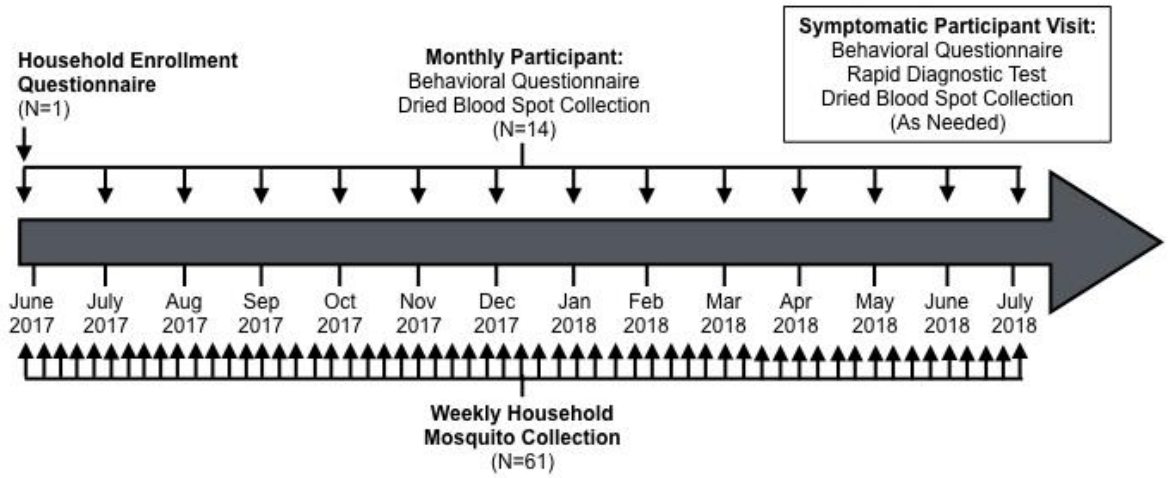
The main exposure was classification of an infection as asymptomatic or symptomatic. To first assess the likelihood of transmission by symptomatic status within participants, we compared the proportion of mosquitoes that shared a haplotype between a participant's asymptomatic and symptomatic infections. We computed the

proportion of participant-mosquito pairings that shared at least one haplotype across participants that had at least one asymptomatic and one symptomatic infection. We assessed the statistical significance of differences in these proportions between asymptomatic and symptomatic infections using a multi-level logistic regression model that included random intercepts at the participant and household levels and controlled for confounding covariates.

For a more comprehensive measure of transmission across all participants, we created a probabilistic model to estimate the probability that a shared haplotype between a participant and a mosquito represented a *P. falciparum* transmission event. The probability of transmission estimate was based on three distinct features: (i) the time interval between the participant's infection and mosquito collection; (ii) the distance between the household of the participant and the household where the mosquito was collected; and (iii) the prevalence and number of parasite haplotypes shared. We compared values between participants with asymptomatic and symptomatic infections using a multi-level logistic regression model with random intercepts at the participant and household levels as well as confounding covariates as identified in a DAG analysis. Sensitivity analyses were conducted to compare coding choices for the probabilistic combination of time, distance, and haplotypes.

The contribution to the infectious reservoir made by asymptomatic infections was calculated using the odds ratio estimate obtained from the probabilistic method for participant-to-mosquito transmission.

Figure 3.1. Follow-up visit and sample collection design.
The first 14 months of participant follow-up and mosquito collections are illustrated below. This follow-up design was repeated for up to 29 months for aim 1A and 14 months for aims 1B and 2.



CHAPTER IV: IMPACT OF ASYMPTOMATIC *PLASMODIUM FALCIPARUM* INFECTION ON THE RISK OF SUBSEQUENT SYMPTOMATIC MALARIA IN A LONGITUDINAL COHORT IN KENYA

Introduction

Asymptomatic *Plasmodium falciparum* infections are common across sub-Saharan Africa, with a meta-analysis estimating that 24% of the continent's children harbored an asymptomatic infection in 2015.[35] When such infections persist chronically they can have adverse effects on the host[101,102,104,105] as well as serve as a reservoir for malaria transmission.[61,62]

Because of the potential for adverse effects from asymptomatic infections, it is important to learn more about the natural history of asymptomatic *P. falciparum* and its implications for future symptom development. Previous studies observed that a history of asymptomatic malaria decreased the risk of subsequent symptomatic infection,[17,19–21] while others reported that exposure to asymptomatic malaria actually increased the risk of symptomatic illness.[14–16,18] A few studies observed the direction of effect even differed within its study population based on age or transmission intensity.[22,23] Many of these studies had infrequent cross-sectional surveys for asymptomatic infections,[16,18,19,21–23] short follow-up periods,[14,15] or assessment of only children,[14–16,18,21,23] possibly missing informative asymptomatic events across a range of ages and, thus, not fully capturing the natural history of asymptomatic malaria.

We investigated the natural history of asymptomatic *P. falciparum* infections in a high-transmission setting using a 29-month longitudinal cohort of people aged 1 to 85

years in Western Kenya. Using monthly active case detection of asymptomatic infections and passive capture of symptomatic events, we investigated the likelihood of symptomatic malaria following an asymptomatic *P. falciparum* infection. Because asymptomatic malaria could be indicative of a pre-symptomatic state, we hypothesized that, compared to uninfected people, those with an asymptomatic infection would have a higher short-term hazard of symptomatic malaria.

Materials and Methods

Study Population, Sample Collection, and Sample Processing

From June 2017 to November 2019 we followed a cohort of 268 people aged 1 to 85 years living in 38 households in a rural setting in Webuye, Western Kenya.[127] For each person, asymptomatic *P. falciparum* infections were detected monthly by active surveillance through collecting questionnaires and dried blood spot (DBS) samples for post-hoc molecular parasite detection. Symptomatic *P. falciparum* infections were detected using passive surveillance by testing people with self-reported symptoms with a malaria rapid diagnostic test (RDT) (Carestart© Malaria HRP2 *Pf* from Accessbio) and collecting a DBS.[115] People with positive RDT results were treated with Artemether-Lumefantrine (AL).

DBS were processed to detect *P. falciparum* infections by extracting genomic DNA (gDNA) from DBS and then tested in duplicate for *P. falciparum* parasites using a duplex real-time PCR (qPCR) assay targeting the *P. falciparum pfr364* motif and human β-tubulin gene.[128,129] Samples were defined as *P. falciparum*-positive if: (i) both replicates amplified *P. falciparum* and both Ct values were < 40 or (ii) 1 replicate amplified *P. falciparum* and the Ct value was < 38. Parasite density was estimated

against a standard curve included on each qPCR reaction plate consisting of templates at known parasite densities ranging from 1 to 2,000 parasites/μL.

Exposure and Outcome Ascertainment

The main exposure was an asymptomatic *P. falciparum* infection during monthly active case detection assessments, defined as *P. falciparum*-positive by qPCR in a person lacking symptoms. People who were *P. falciparum*-negative by qPCR during monthly visits were considered uninfected. Participant follow-up was imputed for the first consecutive missed monthly visit during each follow-up period. The exposure status in the previous month was assumed to be the exposure status of the missed monthly visit.[9] If a person missed two or more consecutive monthly visits, they were considered lost to follow-up and censored at the time of the imputed monthly visit.

The main outcome assessed was days to symptomatic malaria infection. We defined symptomatic *P. falciparum* infection as the presence of at least one symptom consistent with malaria during a sick visit and *P. falciparum*-positive by both RDT and qPCR. Sensitivity analyses for which alternate definitions of symptomatic malaria were used are reported in the supplement. Events occurring within 14 days of receipt of AL for a symptomatic infection were censored. Participants were allowed to enter and leave the study throughout the study period. At the end of the study period, all participants were censored.

Hazard of Symptomatic Malaria Analysis

Across all participants, we estimated the hazard of subsequent symptomatic malaria when infected with asymptomatic malaria compared to being uninfected at monthly visits. The hazard of symptomatic malaria was calculated within multiple follow-

up periods: (i) 1 month; (ii) 3 months (iii) 6 months; (iv) 12 months; and (v) 29 months (entire study period). For each follow-up period exceeding 1 month, exposure status was ascertained at every monthly visit and allowed to vary each month using a method proposed by Hernán *et al*.[109] The method treated each monthly follow-up visit as a new study entry, recalculating the time to symptomatic malaria or censoring using each monthly follow-up visit date as the origin and attributing the exposure in that month as the exposure status from that month up until the event or censoring occurred (**Figure 4.1**). Results predicted the likelihood of symptomatic illness within 1 month, 3 months, 6 months, etc. when harboring an asymptomatic infection versus being uninfected during each follow-up period. This exposure coding method was chosen due to its ability to capture the exposure at multiple time points with less risk of misclassification or left truncation bias compared to alternative time-varying coding approaches (**Table S4.1**).

Statistical Modeling

We first estimated the time to symptomatic malaria using Kaplan-Meier curves and the log-rank test. Differences in median time to symptomatic malaria were also compared across select covariates using the Wilcoxon Rank Sum test with continuity correction for dichotomous variables or the Kruskal-Wallis test for polytomous variables. The Bonferroni correction was applied to all table *p*-values to account for repeated measures during the 29 months of follow-up.

In order to account for anticipated confounders of the relationship between asymptomatic infection and symptomatic malaria, we next computed a multivariate frailty Cox proportional hazards model (**Equation 4.1**).

$$\frac{h_1(t)_i}{h_0(t)_i} = \exp(\alpha_i + \beta_1 Aymptomatic\ infection_{im} + \beta_2 Age5to15_i + \beta_3 Ageover15_i + \beta_4 Female_i + \beta_5 Regular\ bed\ net\ usage_i + \beta_6 Village{:}\ Maruti + \beta_7 Village{:}\ Sitabicha + \epsilon_i) \quad \textbf{(4.1)}$$

The model controlled for the following confounders as determined by a directed acyclic graph (**Figure S4.1**): age (<5 years, 5-15 years, >15 years), sex, and regular bed net usage (averages > 5 nights a week sleeping under a bed net – yes, no). To account for differences in malaria prevalence across the three villages, we also included a covariate in the model to represent each village. We allowed the main exposure to vary each month based on the monthly follow-up visit infection status (m). A random effect at the participant level (α_i) accounted for potential correlated outcomes due to multiple observations per person. The proportional hazards assumption was assessed using Kaplan-Meier curves and Schoenfeld residual plots.

We tested for effect measure modification by age and sex by stratifying the multivariate frailty Cox proportional hazards model by age category (<5 years, 5-15 years, >15 years) or sex. Hazard ratios and 95% confidence intervals of the main effect of interest were compared across stratified models. The log-likelihood ratio test compared a Cox proportional hazards model with an interaction term between the potential modifier and main exposure to **Equation 4.1**.

We computed two additional time-to-event models of the relationship between asymptomatic infection and subsequent symptomatic illness using alternate subsets of events. Firstly, because asymptomatic infections could represent incipiently symptomatic (i.e. “pre-symptomatic”) infections, we conducted an analysis in which all monthly follow-up visits occurring within 14 days prior to a symptomatic infection were

excluded, reducing the possibility that pre-symptomatic infections could be misclassified as asymptomatic. The analysis was conducted using **Equation 4.1** for the 1, 3, 6, 12, and 29-month follow-up periods. Secondly, to investigate this relationship following the definitive clearance of parasites, we conducted an analysis in which participants did not enter the analysis until 14 days following AL treatment for a symptomatic episode. This analysis focused only on the 1-month hazard of symptomatic malaria and used the frailty Cox proportional hazards model described in **Equation 4.1**. Differences in the post-treatment and main analysis data sets were compared using the Pearson's χ^2 test with Bonferroni correction for 29 months of follow-up. All statistical analyses were performed using R (version 4.0.2)[125] with the packages tidyverse,[130] survminer,[131] survival,[132] coxme,[133] lme4,[134] and ggalluvial.[135] Statistical significance was assessed at an α level of 0.05.

Detectability of Asymptomatic Infections

To investigate how detectability of asymptomatic infections influenced results, we assessed how the parasite density of asymptomatic infections changed the 1-month hazard of symptomatic malaria. To do so, asymptomatic infections defined as above were further classified as meeting a series of thresholds of parasite densities: any density, >1, > 10, > 100, > 500, and > 1000 parasites/μL. These classifications were assigned in a non-mutually exclusive fashion to asymptomatic infections, and then the 1-month likelihood of symptomatic malaria relative to uninfected people was modeled using the Cox proportional hazards model in **Equation 4.1**. Because parasite density influences infections' detectability by common diagnostics, we next projected the likelihood of detection by available diagnostics onto these asymptomatic infections and

estimated the risk of progression to symptomatic malaria as a function of likely detectability. To do so, we compared the hazard of symptomatic malaria among people who were uninfected to those with asymptomatic infections that would have been detectable by: (i) qPCR (>0 parasites/μL), (ii) highly-sensitive (HS)-RDT (≥ 1 parasites/μL), (iii) conventional RDT (≥ 100 parasites/μL), or (iv) light microscopy (≥ 500 parasites/μL). Each asymptomatic infection was classified as detected or undetected by each diagnostic's threshold, and then we used the Cox proportional hazards model in **Equation 4.1.**

Ethical Considerations

The study was approved by institutional review boards of Moi University (2017/36), Duke University (Pro00082000), and the University of North Carolina at Chapel Hill (19-1273). All participants provided written informed consent, and those over age 8 provided additional assent.

Results

For 29 months, we followed 268 participants from three villages in Western Kenya. After excluding participants with less than two months of follow-up, the analysis data set consisted of 257 participants with a median of 222 days (interquartile range (IQR): 89 to 427) of follow-up. Overall, 5379 person-months at risk were observed with 1842 (34.2%) person-months of asymptomatic malaria exposure. Exposure status frequently changed for participants and remained constant for only 16 (6.2%) people across follow-up (**Figure 4.2A**). We recorded 266 symptomatic malaria events. Participants had a median of 1 (IQR: 0, 2) symptomatic infection during follow-up. Median time to symptomatic malaria when asymptomatically-infected (173, IQR: 49,

399) was shorter than when uninfected (230, IQR: 98, 402), as well as shorter for participants aged 5-15 years or living in the village Maruti (**Table 4.1**). However, comparison of Kaplan-Meier curves did not indicate a significant difference in the median time to symptomatic malaria by asymptomatic infection for the full 29 months (*p*-value = 0.100 by log-rank test) (**Figure 4.2B).** Results for secondary case definitions for symptomatic malaria were overall similar and are provided in the supplement (**Tables S4.2 and S4.3**).

Short-term Effect of Asymptomatic Malaria Exposure

In a univariate frailty Cox proportional hazards model, compared to uninfected people, the 1-month crude Hazard Ratio (HR) of symptomatic malaria for participants with asymptomatic infections was 2.69 [95% confidence interval (CI): 2.12 to 3.43]. This association was similar in a model controlling for covariates [adjusted HR: 2.61, 95% CI: 2.05 to 3.33] (**Table 4.2, Figure 4.3A**) as well as when using alternative modeling approaches and case definitions for symptomatic malaria (**See supplemental information**). This relationship between asymptomatic malaria and subsequent symptomatic illness was not modified by age (*p*-value = 0.447 by log-likelihood ratio test). Asymptomatic infections were associated with increased likelihoods of subsequent symptomatic malaria in all age categories: < 5 years [HR: 3.77, 95% CI: 2.02 to 7.04], 5-15 years [HR: 2.45, 95% CI: 1.79 to 3.35], and > 15 years [HR: 2.55, 95% CI: 1.57 to 4.15] (**Table S4.4**). In contrast, sex did modify this relationship (*p*-value = 0.006 by log-likelihood ratio test) (**Table S4.4**), whereby the risk of symptomatic malaria following asymptomatic infection was lower for males [HR: 1.76, 95% CI: 1.24 to 2.50] compared to females [HR: 3.71, 95% CI: 2.62 to 5.24] (**Figure S4.2**).

In subset analyses, compared to uninfected people, the hazard of symptomatic malaria was increased in those with asymptomatic infections by more than 1.7 times [HR: 1.77, 95% CI: 1.26 to 2.47] when limited to those with events more than 14 days after exposure ascertainment (**Figure 4.3A**) and more than 2.5 times [HR: 2.54, 95% CI: 1.76 to 3.67] when limited to those following antimalarial treatment (**Figure 4.3A**). Effect measure modification by neither participant age nor sex was observed in these subset analyses (**Table S4.4**). The post-treatment analysis, which sampled a population with at least one symptomatic infection, had significantly more participants that identified as female and were 5-15 years-old compared to the full, main analysis data set (**Table S4.5**).

Long-term Effect of Asymptomatic Malaria Exposure

Next, we assessed the relationship between asymptomatic infection and subsequent symptomatic malaria using prolonged follow-up. Extending the follow-up period led to a diminution in the hazard ratios of symptomatic malaria comparing those asymptomatically-infected versus uninfected over the following 3 [HR: 1.64, 95% CI: 1.40 to 1.94], 6 [HR: 1.38, 95% CI: 1.20 to 1.58], 12 [HR: 1.12, 95% CI: 1.00 to 1.25], and 29-month [HR: 1.11, 95% CI: 1.01 to 1.22] time periods (**Table 4.2, Figure 4.3B**). In the 29-month analysis, this relationship was modified by participant age (*p*-value < 0.001 by log-likelihood ratio test) with the strongest relationship between asymptomatic infection and future symptomatic malaria in children < 5 years [HR: 1.38, 95% CI: 1.05 to 1.81], second-strongest in children 5-15 years [HR: 1.16, 95% CI: 1.02 to 1.32], and weakest in adults > 15 years [HR: 0.96, 95% CI: 0.81 to 1.13] (**Table S4.4**). Consistent with the 1-month analysis, we observed modification by sex in some models, with

females having higher risk for symptomatic disease (**Table S4.4**). Results from sensitivity analyses are recorded in the supplement (**Table S4.4**).

Short-term Effect of Detectability of Asymptomatic Infections

Owing to the consistently elevated short-term risk of symptomatic malaria in people with asymptomatic infections, we investigated the effect of parasite density in these infections on the likelihood of subsequent symptomatic malaria with 1 month. Compared to uninfected people, the risk of symptomatic malaria was significantly increased by asymptomatic infections of all parasite densities, with the highest adjusted hazard for those with densities > 1000 parasites/μL [HR 3.99, 95% CI 2.41 to 6.62] (**Figure 4.4A**). To the density of each asymptomatic infection, we applied a threshold of detectability for parasite diagnostics and estimated the 1-month hazard of subsequent symptomatic malaria as a function of detectability by qPCR, HS-RDT, RDT, and light microscopy. In this analysis, relative to uninfected people, the hazard of symptomatic malaria was increased among people with asymptomatic infections detectable by qPCR [HR: 1.70, 95% CI: 1.33 to 2.17], HS-RDT [HR: 2.11, 95% CI: 1.63 to 2.72], RDT [HR: 1.93, 95% CI: 1.36 to 2.74], and light microscopy [HR: 2.41, 95% CI: 1.56 to 3.71] (**Figure 4.4B**).

Discussion

Using a 29-month longitudinal cohort based in a high malaria transmission region of Kenya, we investigated the influence of asymptomatic *P. falciparum* infections on the risk of symptomatic malaria. In the short term, compared to uninfected individuals, people with asymptomatic infections had a more than 2-fold increased risk of symptomatic malaria within 1 month irrespective of patient age. Additionally, this

elevated risk of symptomatic malaria was associated with asymptomatic infections at densities detectable by all parasite detection approaches. As follow-up time was expanded, the association between asymptomatic infection status and the hazard of subsequent symptomatic malaria remained overall positive but weakened. Collectively, our finding that detection of an asymptomatic *P. falciparum* infection confers an elevated risk of future symptomatic malaria supports the routine treatment of infections even in the absence of symptoms.

We observed that asymptomatic infections were associated with a 2.6-times increased risk of symptomatic malaria within 1 month. This association was consistent when using multiple case definitions for symptomatic malaria and in pre-symptomatic and post-treatment sensitivity analyses. Previous studies that detected asymptomatic infections using microscopy also reported an increased short-term hazard of symptomatic illness among children within 9 to 30 days after having an asymptomatic malaria infection.[14,15] We built upon these studies by detecting asymptomatic infections using qPCR, a highly sensitive method with a low limit of detection,[129] in participants of all ages and similarly found that asymptomatic infections had a high probability of being quickly followed by symptomatic illness. The increased short-term hazard could be due to acquisition of new blood-stage parasites[20] or progression of the original infection to symptoms;[14] however, because this positive association was upheld in sensitivity analyses that accounted for misclassification of asymptomatic infections as pre-symptomatic, the increased hazard observed was not driven solely by pre-symptomatic infections. We observed an increased short-term hazard of symptomatic malaria after asymptomatic infection regardless of participant age, which suggests that even adults

with low density asymptomatic infections have an increased 1-month, short-term hazard of symptomatic malaria.

This elevated risk of symptomatic malaria was present for asymptomatic infections at any parasite density and detectability by field diagnostics. Notably, this increased hazard of symptomatic malaria grew when harboring asymptomatic infections with higher parasite densities. Our results were similar to previous work that also observed higher parasite densities increased the odds of subsequent symptomatic illness.[14,85] This association has been used to prioritize high density, easily detectable infections for treatment; however, this prioritization is not representative of long-term infection dynamics, as low parasite density infections have been observed to later develop into higher density ones.[9,136] Additionally, we found that asymptomatic infections at all parasite densities, even those detectable by qPCR, were at higher risk of symptomatic illness. Thus, our results do not support the notion that parasite detectability should be used to risk stratify people for treatment.

We found that, compared to males, asymptomatically-infected females had a statistically significant higher hazard of symptomatic malaria shortly following an asymptomatic infection. This was in contrast to previous work in high malaria transmission regions that observed higher general malaria burden among males;[78,137,138] however, no previous longitudinal studies of time-to-symptomatic malaria have compared the hazard of symptomatic illness stratified by asymptomatically-infected males and females, with most studies only including sex as a covariate in models. This stratification of asymptomatic infections by sex could have important biological implications for the natural history of asymptomatic infections, as recent

research has suggested that females clear asymptomatic infections faster than males.[139] More research is needed to determine potential mechanisms for sex-based differences in the hazard of symptomatic malaria following an asymptomatic infection.

We observed that the elevated risk for symptomatic disease associated with asymptomatic infection weakened as the follow-up length expanded from 3 to 29 months. In fact, the hazard of symptomatic illness was only significantly increased for up to 6 months following an asymptomatic infection. This phenomenon has been previously observed as a methodological flaw for hazard ratios, whereby the magnitude of the average hazard ratio decreases as follow-up time increases,[140] and possibly explains previous conflicting studies with variable long-term follow-up: some previous work found an increased risk of future symptomatic illness within 9-12 months when infected with asymptomatic malaria[16,18] but those with greater than 1 year of follow-up observed a decreased risk.[19–21,141] Thus, the follow-up period is a critical factor for assessing the relationship between asymptomatic infection and the future hazard of symptomatic malaria. We addressed this potential issue by presenting a series of average hazard ratios over a range of follow-up periods and encourage future studies to employ this approach.

We used a novel approach to capture how asymptomatic malaria varied over time. Most previous work used an intention-to-treat approach for asymptomatic infections identified in cross-sectional surveys;[16–19,21–23] however, this method can misclassify person-time if the exposure frequently changes, as happens with asymptomatic infections in high transmission areas. For previous studies with more frequent asymptomatic sampling, the projects had short follow-up periods (9-30

days)[14,15] that obscured the true duration of asymptomatic infection, or coded the exposure as always being exposed after an asymptomatic infection occurred.[20] We more precisely recorded asymptomatic malaria exposure using a time-varying method proposed by Hernán *et al*.[109] that allowed participants to change exposure status throughout follow-up, capturing a more complete view of infection dynamics with lower risk of exposure misclassification. This new method produced an effect estimate predictive of future risk regardless of prior exposure making it also less prone to left truncation bias, which can occur with methods that create additive measures of months of exposure. The Hernán *et al*.[109] method has been used in previous studies of cardiovascular or kidney disease,[142–144] but has never before been used to study malaria. We urge more studies to incorporate frequent longitudinal sampling of asymptomatic infections into time-to-symptomatic malaria analyses as well as to include new time-varying exposure methodology.

This study had some limitations. Asymptomatic infections were only captured at monthly follow-up visits, missing transient asymptomatic infections between visits. By allowing participant exposure to vary over time, we assumed exchangeability between the exposed and unexposed groups. This was mitigated by the observation that approximately 94% of the study population changed exposure status at least once during follow-up. Finally, we detected asymptomatic infections only by qPCR and estimated parasite densities only using molecular methods. However, we estimated the influence of detectability of asymptomatic malaria using commonly-accepted thresholds for clinical diagnostics.

In conclusion, using a novel exposure coding method and frequent sampling of

both children and adults over 29 months, we found that asymptomatic *P. falciparum* infections had a high likelihood of being shortly followed by symptomatic illness across all ages and parasite densities. These results suggest interventions focus on treating and reducing asymptomatic malaria in high-transmission settings.

Table 4.1. Covariate distribution across symptomatic malaria events in 29 months of follow-up

	Total person-months* (N, %)	**Person-months ending in symptomatic infections**** (N, %)	**Median time to symptoms for entire study** (days, IQR)	***P*-value comparing time to symptoms**
Main exposure				**<0.001[a]**
No infection	3537 (65.8)	1580 (65.7)	230 (98, 402)	-
Asymptomatic infection	1842 (34.2)	826 (34.3)	173 (49, 399)	-
Age				**0.015[b]**
< 5 years	812 (15.1)	329 (13.7)	226 (82, 435)	-
5-15 years	2279 (42.4)	1319 (54.8)	199 (70, 379)	-
> 15 years	2288 (42.5)	758 (31.5)	244 (97, 426)	-
Sex				0.779[a]
Male	2360 (43.9)	1190 (49.5)	229 (86, 420)	-
Female	3019 (56.1)	1216 (50.5)	202 (76, 384)	-
Regular bed net usage[#]				1.000[a]
No	1425 (26.5)	730 (30.3)	210 (82, 386)	-
Yes	3954 (73.5)	1676 (69.7)	217 (80, 403)	-
Village				**<0.001[b]**
Kinesamo	1854 (34.5)	876 (36.4)	233 (89, 418)	-
Maruti	1681 (31.3)	745 (31.0)	174 (64, 350)	-
Sitabicha	1844 (34.3)	785 (32.6)	231 (90, 421)	-

Abbreviations: IQR, interquartile range
[#]Regular bed net usage was a person averaging > 5 nights a week sleeping under a bed net.
*Total person-months indicates the total number of monthly follow-up visits ending in a symptomatic infection or censoring for full 29-month follow-up.
**Symptomatic infections were defined using the primary case definition where a participant was *P. falciparum*-positive by both RDT and qPCR as well as had at least one symptom consistent with malaria during a sick visit.
[a] Wilcoxon Rank Sum test with continuity correction and Bonferroni correction for repeated measures.
[b] Kruskal-Wallis test with Bonferroni correction for repeated measures.
Significant estimates are bolded.

Table 4.2. Predicted hazard of symptomatic malaria across follow-up periods

	1-month HR (95% CI)	3-month HR (95% CI)	6-month HR (95% CI)	12-month HR (95% CI)	29-month HR (95% CI)
Main exposure					
No infection	Ref	Ref	Ref	Ref	Ref
Asymptomatic infection	**2.61 (2.05, 3.33)**	**1.64 (1.40, 1.94)**	**1.38 (1.20, 1.58)**	1.12 (1.00, 1.25)	**1.11 (1.01, 1.22)**
Age					
< 5 years	Ref	Ref	Ref	Ref	Ref
5-15 years	1.37 (0.90, 2.08)	1.61 (1.00, 2.61)	**1.99 (1.07, 3.71)**	2.37 (0.97, 5.77)	**2.52 (1.26, 5.01)**
> 15 years	**0.56 (0.36, 0.88)**	0.74 (0.46, 1.21)	0.83 (0.44, 1.53)	0.88 (0.37, 2.08)	0.97 (0.51, 1.84)
Sex					
Male	Ref	Ref	Ref	Ref	Ref
Female	0.93 (0.70, 1.24)	0.84 (0.61, 1.16)	0.80 (0.53, 1.20)	0.68 (0.38, 1.21)	**0.63 (0.40, 0.99)**
Regular bed net usage[#]					
No	Ref	Ref	Ref	Ref	Ref
Yes	1.00 (0.70, 1.43)	0.81 (0.55, 1.20)	0.70 (0.43, 1.16)	0.59 (0.29, 1.21)	**0.52 (0.30, 0.89)**
Village					
Kinesamo	Ref	Ref	Ref	Ref	Ref
Maruti	1.08 (0.77, 1.52)	1.11 (0.75, 1.64)	1.14 (0.69, 1.88)	1.13 (0.56, 2.31)	1.09 (0.64, 1.85)
Sitabicha	0.72 (0.49, 1.05)	0.80 (0.53, 1.21)	0.76 (0.45, 1.29)	0.73 (0.35, 1.51)	0.70 (0.40, 1.23)

Abbreviations: CI, confidence interval; HR, adjusted hazard ratio; IQR, interquartile range; Ref, reference
[#]Regular bed net usage was defined as a person averaging > 5 nights a week sleeping under a bed net.
*Total person-months indicates the total number of monthly follow-up visits ending in a symptomatic infection or censoring.
**Symptomatic infections were defined using the primary case definition where a participant was *P. falciparum*-positive by both RDT and qPCR as well as had at least one symptom consistent with malaria during a sick visit.
[a] Wilcoxon Rank Sum test with continuity correction and Bonferroni correction for repeated measures.
[b] Kruskal-Wallis test with Bonferroni correction for repeated measures.
Significant estimates are bolded.

Figure 4.1. Schematic of how asymptomatic exposure status was ascertained for one participant's follow-up using the Hernán *et al.* method.
The method treated each monthly follow-up visit as a new study entry for the participant, recalculating the time to symptomatic malaria using the monthly follow-up visit date as the origin. The exposure status for each monthly follow-up visit became the exposure status for the follow-up period. The follow-up period ended if the participant had a symptomatic infection or was censored due to the study ending or becoming lost to follow-up. One hypothetical participant's follow-up across 29 months is illustrated here.

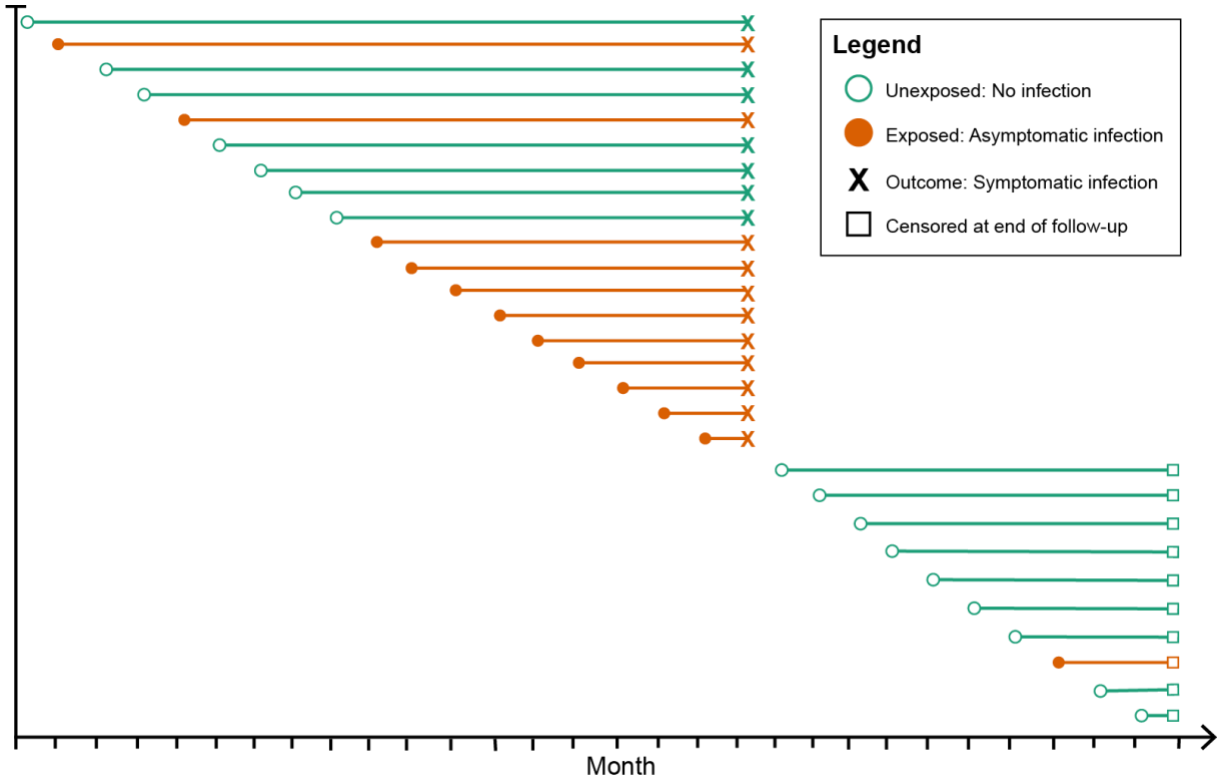


Figure 4.2. Asymptomatic malaria versus no infection exposure classification over time.
4.2A: The proportion of participants who had either an asymptomatic infection (orange) or no infection (green) at each monthly visit is indicated by the bars. The ribbons connecting the bars illustrate the proportion of participants who moved exposure status from month to month. Orange ribbons indicate the proportion of participants with asymptomatic infections and green the proportion with no infection. **4.2B:** A Kaplan-Meier survival curve assessing median time to symptomatic malaria is illustrated across the full 29 months of follow-up stratified by cumulative asymptomatic malaria exposure.

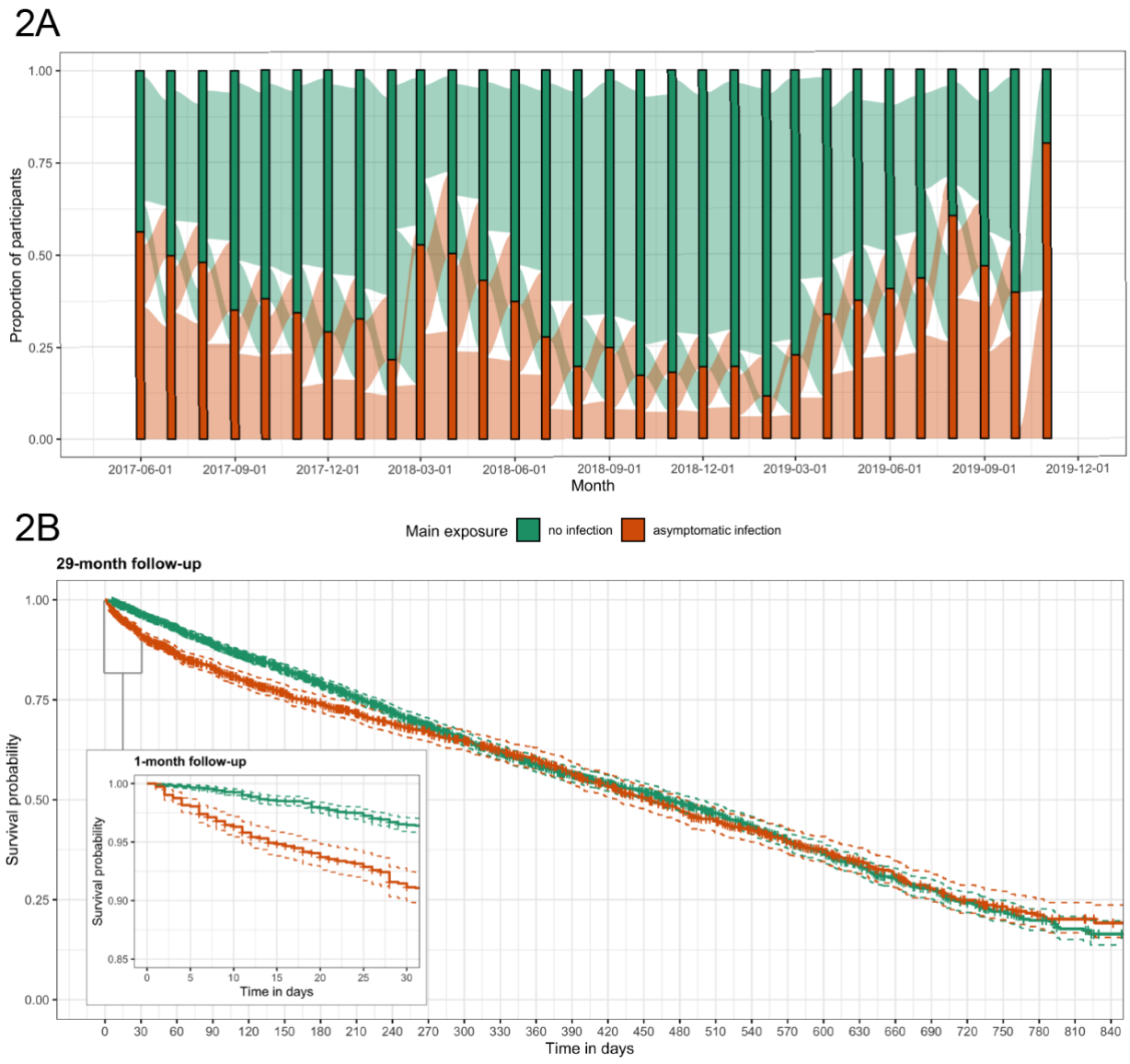


Figure 4.3. Hazard of symptomatic malaria after exposure to asymptomatic infections compared to no malaria infection over time.
4.3A: Frailty Cox proportional hazards model results comparing exposure to asymptomatic malaria infections versus no infection over time and hazard of symptomatic malaria in the main model using the primary outcome coding as well as across exposure and outcome sensitivity analyses. The pre-symptomatic analysis model removed monthly follow-up visits that occurred within 14 days prior to a symptomatic malaria infection. The post-treatment analysis model was restricted to only participants who had at least one symptomatic malaria infection and follow-up began at the monthly visit at least 14 days post study-administered antimalarial treatment for each person's initial symptomatic infection. All models controlled for covariates participant age, sex, bed net usage, and village. **4.3B:** Models were run with differing follow-up for each participant ranging from 1 to 29 months and controlled for covariates participant age, sex, bed net usage, and village.

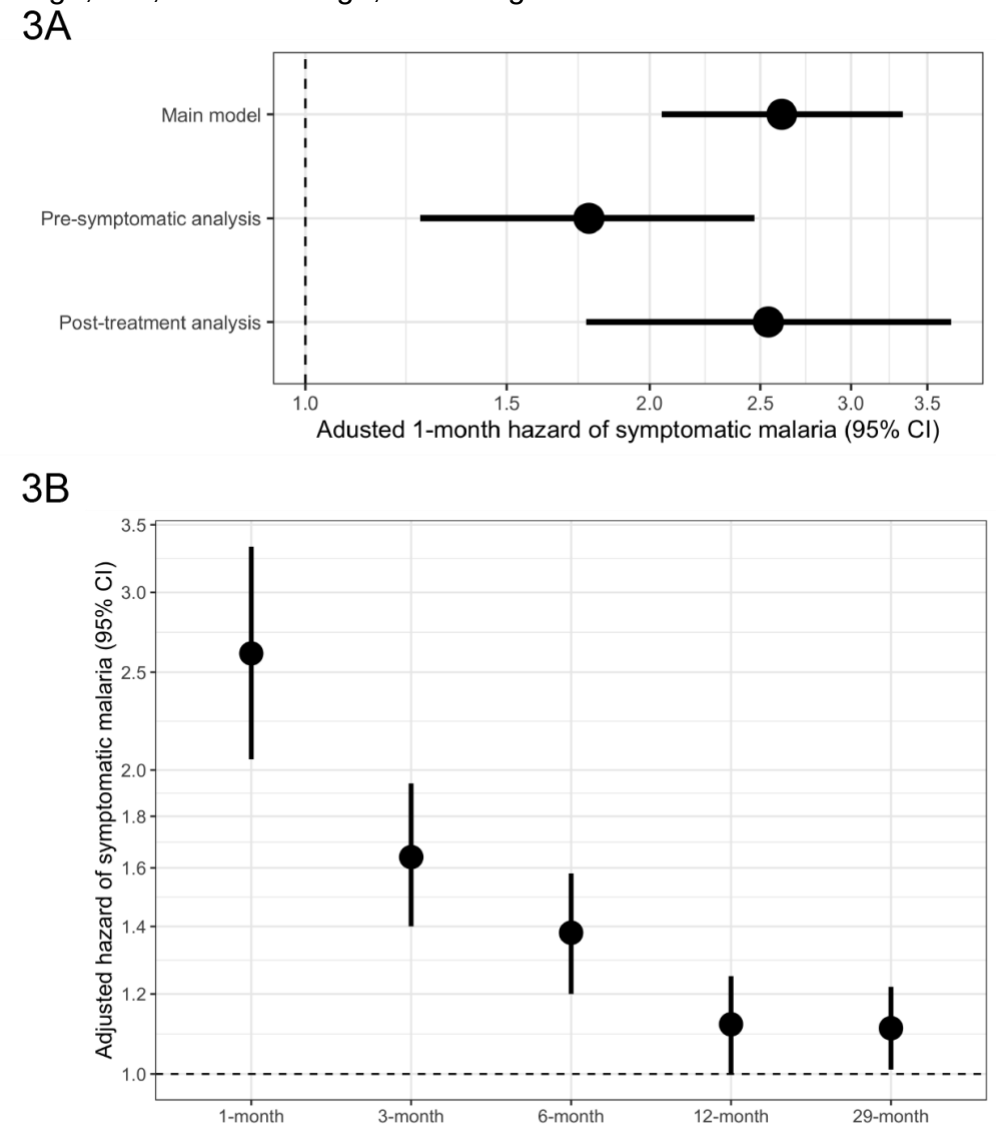
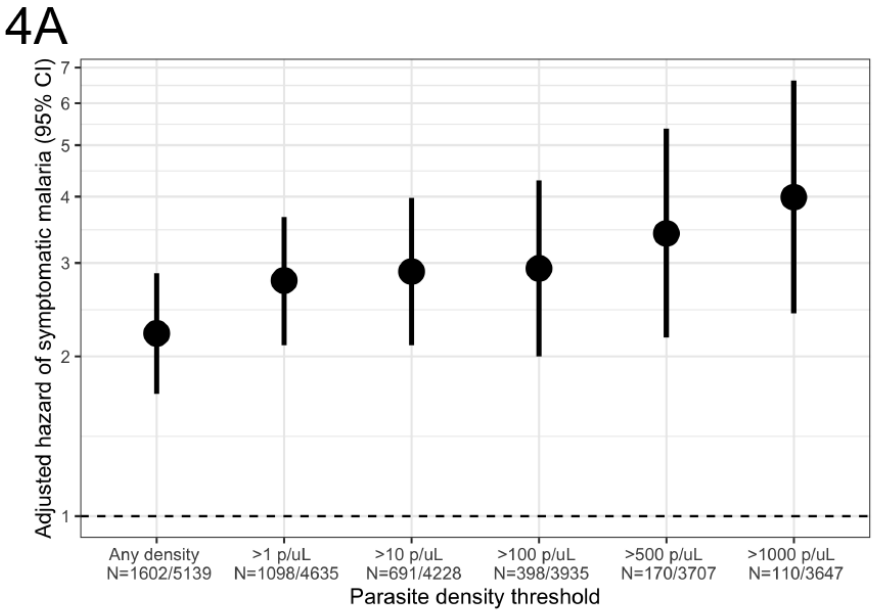
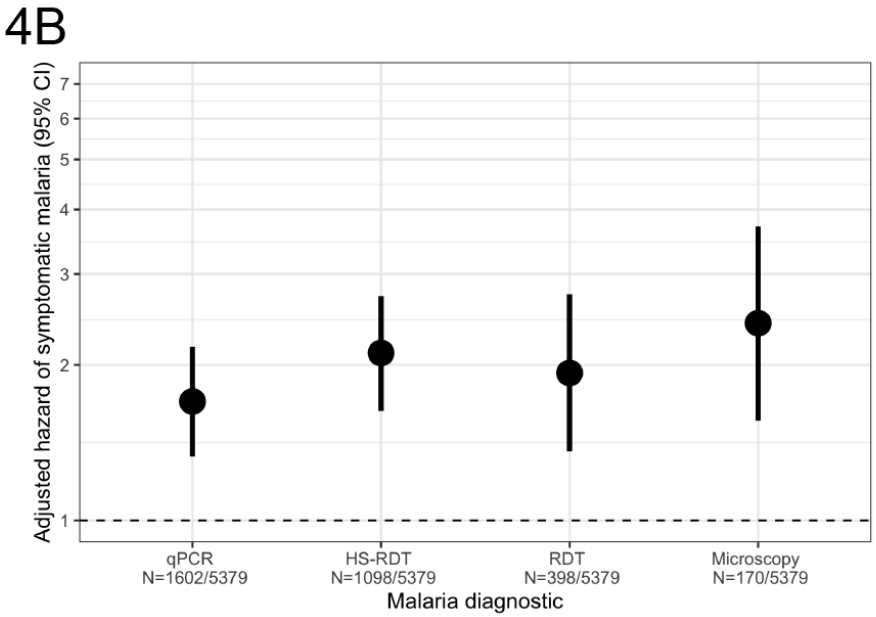


Figure 4.4. Detectability of asymptomatic malaria infections and short-term hazard of symptomatic malaria.
4.4A: Frailty Cox proportional hazards model results for the short-term, 1-month hazard of symptomatic malaria comparing models restricted to asymptomatic infections with varying parasite density thresholds in parasites/μL (p/μL) across malaria diagnostics. All models compared asymptomatic malaria exposure to no infection exposure and controlled for covariates participant age, sex, bed net usage, and village. **4.4B:** Frailty Cox proportional hazards model results for the 1-month hazard of symptomatic malaria comparing models with asymptomatic infections with varying detectability across malaria diagnostics. Models compared asymptomatic malaria exposure to having no infection over time and controlled for covariates participant age, sex, bed net usage, and village.





CHAPTER V: EXPOSURE TO DIVERSE *PLASMODIUM FALCIPARUM* GENOTYPES SHAPES THE RISK OF SYMPTOMATIC MALARIA IN INCIDENT AND PERSISTENT INFECTIONS: A LONGITUDINAL MOLECULAR EPIDEMIOLOGIC STUDY IN KENYA

Introduction

Plasmodium falciparum causes over 200 million clinical malaria cases annually.[145] Many of these infections occur in young children, who are more likely to develop symptomatic malaria compared to adults.[19,78,146] This age-dependent risk of symptomatic disease is thought to develop following repeated exposure to *P. falciparum* that produces adaptive, disease-controlling immune responses.[92–95,147] The targets and mechanisms of this naturally-acquired immunity remains largely obscure.

In the absence of generalizable, measurable immune correlates, the contours of functional clinical immunity to parasites and disease have been inferred from patterns of disease risk and parasite genetics. Specifically, the adaptive immunity theory has been supported by studies reporting that symptomatic malaria is often associated with the presence of parasite genotypes that were unobserved in prior infections;[20,86,96–99] this suggests that symptomatic malaria results from new infections that can exploit gaps in immunologic memory. These prior studies, though, have been limited in scope and follow-up,[86,97–99] resolution of genotyping approach,[20,86,96–99] and an inability to partition effects of parasite genotypes between newly-acquired and persistent infections, which collectively limit the generalizability of findings. Furthermore, most[20,86,97–99] have interrogated neutral parasite genes that do not clearly encode targets of functional

immunity, which further limits causal inference of immunologic mechanisms. A clearer understanding of the influence of parasite genetic diversity on disease risk would inform the development of polyvalent vaccines.

To explore how the specific *P. falciparum* infections acquired over time influence the risk of symptomatic malaria, we investigated the association between *Plasmodium falciparum* genotypes and an individual's risk of symptomatic infection using a 14-month longitudinal cohort in Western Kenya. To do so, we classified each person's infections as harboring novel, recurrent or persistent parasites on the basis of amplicon deep sequencing of two diverse parasite genes that encode targets of known functional immunity at the liver (circumsporozoite protein, *pfcsp*) and blood (apical membrane antigen-1, *pfama1*) stages, and analyzed associations between haplotype categories and the odds of developing symptomatic malaria. We hypothesized that, compared with infections harboring parasite genotypes previously observed within a person's prior infections, infections harboring hitherto-unobserved haplotypes would be associated with increased likelihood of symptomatic malaria.

Materials and Methods

Study Population and Sample Collection

From June 2017 to July 2018, a longitudinal cohort was followed in Webuye, Western Kenya, consisting of members 1 to 85 years of 38 randomly-selected households.[127] Asymptomatic *P. falciparum* infections were detected by active case detection using monthly dried blood spot (DBS) collection, in which parasites were detected by real-time PCR (qPCR; see below). Symptomatic malaria infections were detected by passive surveillance whereby participants experiencing malaria-like

symptoms contacted the study team, were tested for malaria using a rapid diagnostic test (RDT; Carestart© Malaria HRP2 *Pf* from Accessbio),[115] and had a DBS collected. RDT-positive participants were treated with Artemether-Lumefantrine.

Sample Processing

Molecular sample processing has been previously described. Genomic DNA (gDNA) was extracted from DBS using Chelex-100 and tested in duplicate for *P. falciparum* parasites using a duplex qPCR assay targeting the *P. falciparum pfr364* motif and human β-tubulin gene.[128,129] *P. falciparum*-positive samples were genotyped across gene segments encoding *pfama1* and *pfcsp* using PCR amplification and sequencing on an Illumina MiSeq platform.[117,148] Read processing and haplotype inference were as previously described;[149] briefly, reads were quality-filtered and mapped to the 3D7 reference sequences for *pfama1* and *pfcsp.*[121,122,148] Haplotype inference was performed on mapped reads using DADA2 (version 1.8) as implemented in R (version 4.0.2)[123,125] and resulting haplotypes were further filtered to reduce haplotype false discovery risk using previously-validated criteria.[124] The output was a catalog of all *pfcsp* and *pfama1* unique haplotypes in each qPCR-positive infection for each person. Sequences are available through GenBank (PRJNA646940).

Exposure and Outcome Assessment

For every person, we classified each parasite haplotype in each of their infections as: (i) new, a haplotype not previously observed in that person; (ii) recurrent, one previously observed in that person but not in the most recent DBS sample; or (iii) persistent, a haplotype previously observed in the most recent DBS sample irrespective of whether the sample was a routine monthly sample or a symptomatic event (**Figure**

5.1). For the main exposure, we categorized each infection using the above haplotype classifications, and therefore each infection could contain a mixture of new, recurrent or persistent haplotypes; these categories were assigned independently for *pfama1* and *pfcsp* haplotypes.

For the main outcome, each *P. falciparum* infection was categorized as asymptomatic or symptomatic. An asymptomatic infection was *P. falciparum*-positive by qPCR in a person lacking symptoms and seen during monthly follow-up visits. A symptomatic infection was *P. falciparum*-positive by both RDT and qPCR in a participant with at least one symptom consistent with malaria during a sick visit. Infections were excluded from the outcome ascertainment process if they occurred within 14 days of taking Artemether-Lumefantrine for a symptomatic infection or were the person's first infection during the study (which would harbor only new haplotypes). To assess potential for pre-symptomatic infections, across all symptomatic infections, we compared the time since the preceding asymptomatic infection between infections with and without persistent haplotypes using a Kruskal-Wallis χ^2 test.

We assessed odds of symptomatic malaria as a function of haplotypes in two distinct types of malaria infections: (i) incident infections where none of the haplotypes in the infection were previously observed in the participant's most recent DBS or (ii) persistent infections where at least one haplotype persisted between consecutive DBS collections, excluding infections where participants had a symptomatic infection, were prescribed antimalarials, and had another infection with persistent haplotypes within 30 days of the initial infection.

Comparing Odds of Symptomatic Malaria Among Incident Infections

We first assessed the odds of symptomatic malaria across incident infections. We conducted a multi-level logistic regression comparing the odds of having a symptomatic compared to an asymptomatic infection across people infected with (i) only new haplotypes; (ii) new and recurrent haplotypes; or (iii) only recurrent haplotypes (**Equation 5.1**).

$$\ln(Symptomatic\ malaria_i) = \alpha_i + \beta_1 Age_i + \beta_2 Number\ of\ prior\ infections_i + \beta_3 Transmission\ season_i + \beta_4 Multiplicity\ of\ infection_i + \beta_5 Haplotype\ comparison_i + \epsilon_i \quad \textbf{(5.1)}$$

The model included a participant-level random intercept and controlled for confounding covariates identified in a DAG (**Figure S5.1**): participant age (≤ 15 or > 15 years), number of prior malaria infections the person suffered during the study (≤ 3 or > 3 infections), transmission season (≤ 50 or > 50 mosquitoes collected in the prior 14 days across study site), and multiplicity of infection (≤ 2 or > 2 haplotypes). Thresholds for categorization were determined by functional form assessment. Differences in model covariates stratified by symptomatic status were compared using the Pearson's χ^2 test. *p*-values were adjusted using the Bonferroni correction for the maximum number of incident infections a participant suffered (N=6).

We evaluated effect measure modification by age on the multiplicative scale by computing multi-level logistic regression models stratified by age category (≤ 15 or > 15 years), with covariates as above. Direction of effect and 95% confidence intervals were compared across age-stratified models. Additionally, the log-likelihood ratio test compared output from an adjusted multi-level logistic regression with an interaction term

between age and the haplotype categories to results from the model in **Equation 5.1**.

Comparing Odds of Symptomatic Malaria Among Persistent Infections

We next evaluated the association between persistent infections and symptomatic malaria risk by focusing on infections harboring persistent haplotypes, which were defined as haplotypes also observed in testing immediately prior to the episode. To do this, we investigated, among only persistent infections, the relationship between presence of new compared to persistent haplotypes and odds of symptomatic malaria. First, we identified infections with persistent haplotypes and classified them based on presence or absence of additional haplotypes: (i) only persistent; (ii) new and persistent; (iii) recurrent and persistent; or (iv) new, recurrent, and persistent. Across these four categories, we compared the number of days since previous infection using the Kruskal-Wallis χ^2 test.

After assessing time differences between persistent haplotype categories, we restricted the data set to only infections with persistent haplotypes occurring within 30 days and computed a multi-level logistic regression. Using the same random intercept and covariates as **Equation 5.1**, the model compared the odds of developing symptomatic malaria across people infected with (i) mixed haplotypes (persistent haplotypes + new or recurrent haplotypes) or (ii) only persistent haplotypes. Differences in model covariates stratified by symptomatic status were investigated using the Pearson's χ^2 test with *p*-values adjusted using the Bonferroni correction for 6 infections. Effect measure modification by age was evaluated as described above. All statistical analyses were performed using R (version 4.0.2).[125]

Ethical Considerations

The study was approved by the ethical review boards of Moi University (2017/36), Duke University (Pro00082000) and the University of North Carolina at Chapel Hill (19-1273). All study participants provided consent to participate in the study.

Results

Haplotype Classification and Infection Categorization

Over 14 months, we identified 902 asymptomatic and 137 symptomatic *P. falciparum* infections (**Figure 5.2A**). After parasite sequencing, we obtained genotypes for 861 *P. falciparum* infections among 239 people, with a range from 1 to 14 infections during the study period (mean: 3.8). From these data, events meeting criteria for analysis as outcomes consisted of 109 *pfcsp* haplotypes in 622 infections (534 asymptomatic and 88 symptomatic) across 186 people; 435 (69.9%) harbored new haplotypes, 320 (51.4%) harbored recurrent haplotypes, and 213 (34.2%) had persistent haplotypes (**Figure 5.2B**). A plurality of infections (27.2%) harbored only new haplotypes (N=169/622). Results for *pfama1* are recorded in the supplement (**Figure S5.2**).

Assessing the potential for pre-symptomatic infections across all symptomatic infections (N=88), persistent *pfcsp* haplotypes (N=37) were not as commonly found in symptomatic infections as new or recurrent haplotypes (N=51); however, the time interval in days since the last infection starting from a symptomatic infection was significantly shorter for infections with persistent *pfcsp* haplotypes (median: 12, range: 2 to 63) than for those without persistent haplotypes (median: 39, range: 4 to 327) (*p*-value <0.001 by Kruskal-Wallis χ^2 test) (**Figure 5.3**). The same relationship was seen

for *pfama1* haplotypes (**Figure S5.3**).

Analysis of Symptomaticity in Incident Infections

We first assessed if the presence of new haplotypes influenced odds of symptomatic malaria among 409 incident (358 asymptomatic and 51 symptomatic) infections. Symptomatic infections were more likely to consist of only new haplotypes, occur in children, arise during the high malaria transmission season, and have a lower multiplicity of infection (**Tables 5.1 and S5.1**). Compared to infections composed of only recurrent *pfcsp* haplotypes, odds of symptomatic malaria were not significant in those with both new and recurrent haplotypes [OR: 0.64, 95% CI: 0.15 to 2.65] but significantly higher for those harboring only new haplotypes [OR: 3.24, 95% CI: 1.20 to 8.78] (**Figure 5.4A**). Results were similar but not statistically significant for *pfama1* (**Figure S5.4**). In age-stratified models, the effect was similar in children ≤ 15 years [OR: 3.01, 95% CI: 1.02 to 8.84] and adults > 15 years [OR: 4.00, 95% CI: 0.44 to 36.08], indicating, along with similarity in model fit between models with and without an interaction term for age (*p*-value = 0.996 by log-likelihood ratio test), that age did not modify the effect of haplotype classification on symptoms (**Figure 5.4B**).

Analysis of Symptomaticity in Persistent Infections

Persistent *pfcsp* haplotypes were identified in 213 infections and categorized into: (i) only persistent (N=57); (ii) new and persistent (N=76); (iii) recurrent and persistent (N=29); or (iv) new, recurrent, and persistent (N=51) (**Figure 5.2B**). Across all categories, the time interval since the previous infection ranged from 2 to 96 days (median: 28) (**Figure 5.5A**). Although we observed a cluster of infections consisting solely of persistent haplotypes with very small intervals, overall, the number of days

since previous infection did not differ between haplotype categories (*p*-value = 0.249 by Kruskal-Wallis χ^2 test). Results were similar for *pfama1* (**Figure S5.5**).

In order to test whether the acquisition of new or recurrent haplotypes affected the odds of symptomatic compared to asymptomatic malaria among people with a background of persistent parasite haplotypes, we restricted the data set to only persistent infections that occurred within 30 days. Meeting this 30-day criteria, we assessed 139 infections with at least one persistent *pfcsp* haplotype across 109 asymptomatic and 30 symptomatic infections. Symptomatic infections were more likely to consist of only persistent haplotypes and have a lower multiplicity of infection (**Tables 5.2 and S5.2**). Compared to infections with only persistent *pfcsp* haplotypes, the acquisition of additional haplotypes (either new or recurrent) was not associated with symptomatic disease [OR: 0.77, 95% CI: 0.21 to 2.75] (**Figure 5.5B**). Results were similar using *pfama1* (**Figure S5.6**). Owing to small sample sizes, we could not assess effect measure modification of these associations by age.

Discussion

In a high-transmission setting in Western Kenya, incident *P. falciparum* infections composed of parasite haplotypes that were hitherto unobserved within an individual increased that person's odds of symptomatic malaria. In contrast, the appearance of new haplotypes in a person who was already infected with persistent haplotypes was not associated with increased odds of symptoms. Collectively, our results are consistent with a model of anti-disease immunity in which genetically-distinct parasites can overcome immunity and cause disease in incident infections, but this ability is attenuated by the presence of persistent, tolerated parasites.

Compared to infections with haplotypes a person has experienced previously, we found incident infections with only new haplotypes were associated with increased odds of symptomatic malaria over 3-fold. These results are consistent with the phenomena that partial variant-specific immunity is acquired over time to provide anti-disease protection, and extend the findings of prior studies that report an increased risk of symptomatic malaria when infected with novel haplotypes.[20,86,96–99] Notably, our findings resulted from approaches that overcame limitations in these studies, including small sample sizes with brief follow-up,[97–99] infrequent sampling,[86,99] genotyping approaches with high failure rates,[20] and an inability to capture multiclonal genotypes.[20,86,96–99] Specifically, genotyping approaches that use PCR-restricted fragment length polymorphism to detect size variants,[20,86,96–99] capture only 30% of the unique clones present compared to amplicon deep sequencing.[110] Due to this inability to capture all clones present in complex infections common to high-transmission areas,[111] these approaches fail to accurately classify haplotypes as new or recurrent. Using fine-scale genotypes created by the more sensitive amplicon deep sequencing method,[110] we were able to more definitively partition the distinct effects of new or recurrent haplotypes within incident infections. Our results, which are predicated on 14 months of follow-up of a diverse population with frequent sampling and high-resolution parasite genotyping, suggest that symptomatic malaria amongst frequently-infected residents of a high-transmission setting is associated with the acquisition of blood-stage parasites to which a person has been hitherto unexposed.

Surprisingly, this increased risk of symptomatic disease with new parasite haplotypes was attenuated when new haplotypes were mixed with recurrent ones. This

could not be attributed to the “persistence” of these recurrent haplotypes, because this analysis was restricted to incident infections, suggesting the acquisition of “known” (i.e. recurrent) parasite strains may mediate the disease-causing effects of new haplotypes; this could result from cross-reacting immune recognition of recurrent parasites that either enhances parasite clearance or attenuates immune activation,[147,150] competition between haplotypes that reduces pathogenesis[151] or alternate mechanisms. Also surprising, and in contrast to a prior report,[20] we observed an increased risk of symptomatic malaria when new haplotypes were present in both adults and children. The ability to register this effect is likely due to the use of a more sensitive genotyping method that could capture diverse clones in polygenomic infections, which are more common in adults. The presence of this risk in adults supports an age-independent mechanism for this phenomenon, despite the common assumption that by reaching adulthood one has acquired durable immunity to diverse parasites.

Interestingly, in contrast to incident infections, persistent infections were not likely to be symptomatic when supplemented by new or recurrent parasite haplotypes. The presence of persistent haplotypes in persistent infections suggests a state of immune tolerance may be maintained and limit responses to superinfections.[150] Alternately, the pre-existing blood-stage infection with persistent haplotypes may limit efficiency of establishing and maintaining super-infections, as observed in a murine model;[141] this would be consistent with the original meaning of ‘premunition’, wherein contemporaneous infection confers resistance to superinfection.[152] Finally, asymptomatic infections with persistent haplotypes could be pre-symptomatic and, thus, not greatly impacted by the acquisition of new or recurrent haplotypes. Regardless of

mechanism, these results illustrate the importance of distinguishing between recurrent and persistent haplotypes in incident and persistent infections, which has previously not been done.[16,20,86,96–99] Moreover, future work could assess variability in haplotype within-host competition,[151] virulence of specific haplotypes,[153] and host immune responses to more directly measure how new, recurrent, and persistent haplotypes affect symptomatic malaria risk.

Broadly, our results highlight not only the role that incident and persistent infections have on reducing odds of symptomatic disease, but also the critical influence of parasite genetic diversity on this relationship. In population-based studies, reduced transmission as reflected by decreases in passively-detected cases can increase and shift the severity of disease,[154] possibly by reduced acquisition of anti-disease immunity in childhood. In our study, such anti-disease immunity manifested in incident infections such that symptomaticity was prevented by the presence of recurrent haplotypes. Because these recurrent haplotypes require exposure to prior diverse infections, reduced exposure would increase the likelihood that incident infections are composed of new haplotypes and therefore likely to manifest symptoms. However, if reduced transmission is accompanied by reductions in parasite genetic diversity as has been reported in several settings,[155,156] even with fewer prior infections the per-infection likelihood that a parasite will harbor recurrent haplotypes would remain high and thereby attenuate symptoms. Future studies could explore if specific haplotypes at disease-mediating loci differentially modify the risk of malaria and thereby furnish targets for surveillance.

The study had some limitations. While amplicon deep sequencing was a

sensitive method for identifying different malaria infections,[110] it might not have captured all genetically-distinct infections that occurred during the study. To account for this, we compared results across two unlinked parasite gene targets, *pfama1* and *pfcsp*. We did not observe malaria infections that participants acquired before the study; misclassifying a haplotype as new when it might have been present in an individual prior to the study would bias results towards the null. Additionally, persistent infections were possibly pre-symptomatic. Future studies could have more frequent longitudinal sampling to distinguish between asymptomatic and pre-symptomatic infections at a finer scale.

In conclusion, these results indicate that infections harboring novel haplotypes increased the likelihood of symptomatic malaria in incident infections, but not when acquired in the presence of persistent infections. Future research could explore at the immunological level how the impact of new haplotypes changes the risk of symptomatic malaria when compared to recurrent or persistent haplotypes.

Table 5.1. Distribution of symptomatic status across covariates for incident infections with new and recurrent *pfcsp* haplotypes

	Asymptomatic infections (N=358)	Symptomatic infections (N=51)	*P*-value
Haplotype category, N (%)			<0.001[a]
Only new	133 (37.2)	36 (70.6)	
New and recurrent	134 (37.4)	5 (9.8)	
Only recurrent	91 (25.4)	10 (19.6)	
Age, N (%)			0.016[a]
≤ 15 years	213 (59.5)	42 (82.4)	
> 15 years	145 (40.5)	9 (17.6)	
Number of prior malaria infections[#], N (%)			1.000[a]
≤ 3	252 (70.4)	40 (78.4)	
> 3	106 (29.6)	11 (21.6)	
Transmission season*, N (%)			0.004[a]
Low	245 (68.4)	22 (43.1)	
High	113 (31.6)	29 (56.9)	
Multiplicity of infection, N (%)			0.022[a]
1-2 *pfcsp* haplotypes	200 (55.9)	40 (78.4)	
> 2 *pfcsp* haplotypes	158 (44.1)	11 (21.6)	

Abbreviations: IQR, interquartile range; NE, not evaluated
[#] During their participation in the cohort prior to the event.
* Low: ≤ 50 mosquitoes collected in the two weeks prior; High: > 50 mosquitoes
[a] Pearson's χ^2 test with Bonferroni correction for repeated measures for 6 infections

Table 5.2. Distribution of symptomatic status across covariates for consecutive infections with persistent and new or recurrent *pfcsp* haplotypes

	Asymptomatic infections (N=109)	Symptomatic infections (N=30)	*P*-value
Haplotype category, N (%)			0.002[a]
Mixed types of haplotypes[^]	86 (78.9)	13 (43.3)	
Only persistent haplotypes	23 (21.1)	17 (56.7)	
Age, N (%)			1.000[a]
≤ 15 years	76 (69.7)	23 (76.7)	
> 15 years	33 (30.3)	7 (23.3)	
Number of prior malaria infections[#], N (%)			0.801[a]
≤ 3	65 (59.6)	23 (76.7)	
> 3	44 (40.4)	7 (23.3)	
Transmission season*, N (%)			0.755[a]
Low	70 (64.2)	14 (46.7)	
High	39 (35.8)	16 (53.3)	
Multiplicity of infection, N (%)			<0.001[a]
1-2 *pfcsp* haplotypes	29 (26.6)	21 (70.0)	
> 2 *pfcsp* haplotypes	80 (73.4)	9 (30.0)	

Abbreviations: IQR, interquartile range; NE, not evaluated

[^] Infections with mixed types of haplotypes harbored persistent haplotypes + at least one new or recurrent haplotype.

[#] During their participation in the cohort prior to the event.

* Low: ≤ 50 mosquitoes collected in the two weeks prior; High: > 50 mosquitoes

[a] Pearson's χ^2 test with Bonferroni correction for repeated measures for 6 infections

Figure 5.1. Haplotype categorization throughout participant follow-up.
Two hypothetical scenarios illustrate how the malaria haplotypes that participants acquired over time were categorized as new, recurrent or persistent.

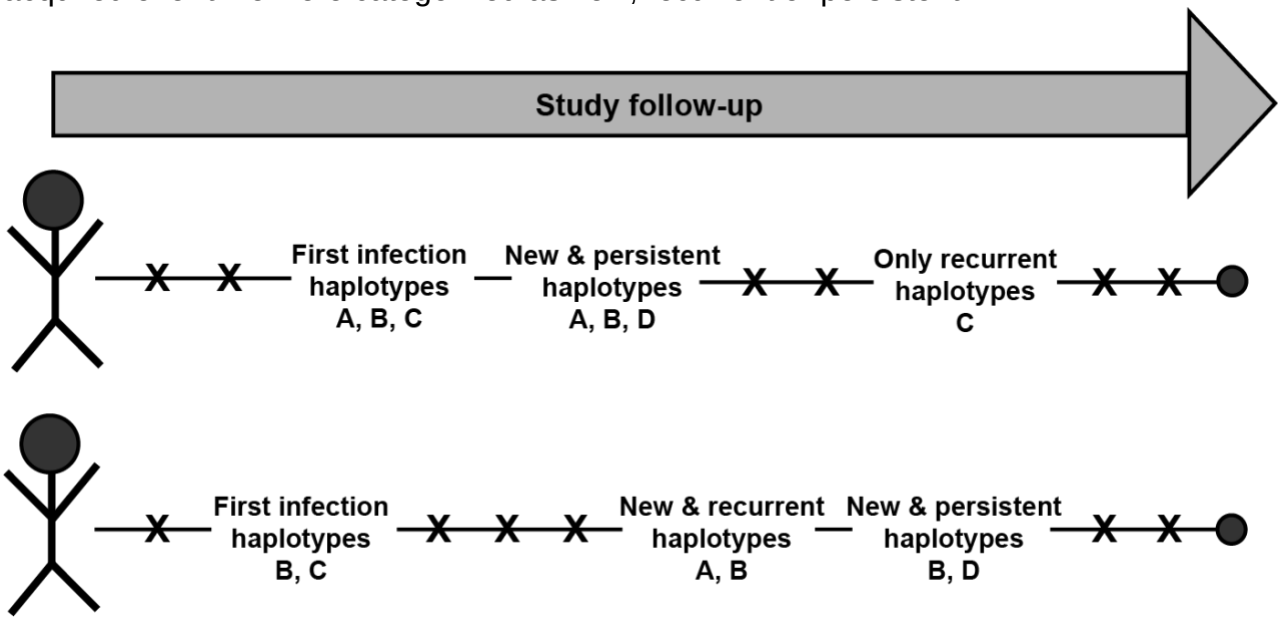


Figure 5.2. Total number of *P. falciparum* infection types and categorization of *pfcsp* haplotypes within these infections.
5.2A: Asymptomatic and symptomatic *P. falciparum* positive samples were captured during 14 months of sampling. Symptomatic infections were captured during as-needed sick visits and asymptomatic infections during monthly visits. A person's initial infection is light grey. Subsequent infections for that person were used for outcome ascertainment (dark grey). **5.2B:** Overlap of *pfcsp* haplotype categories across all symptomatic and asymptomatic *P. falciparum* infections (N=622). Numbers indicate the number of infections that had haplotypes within each category: new, recurrent or persistent.

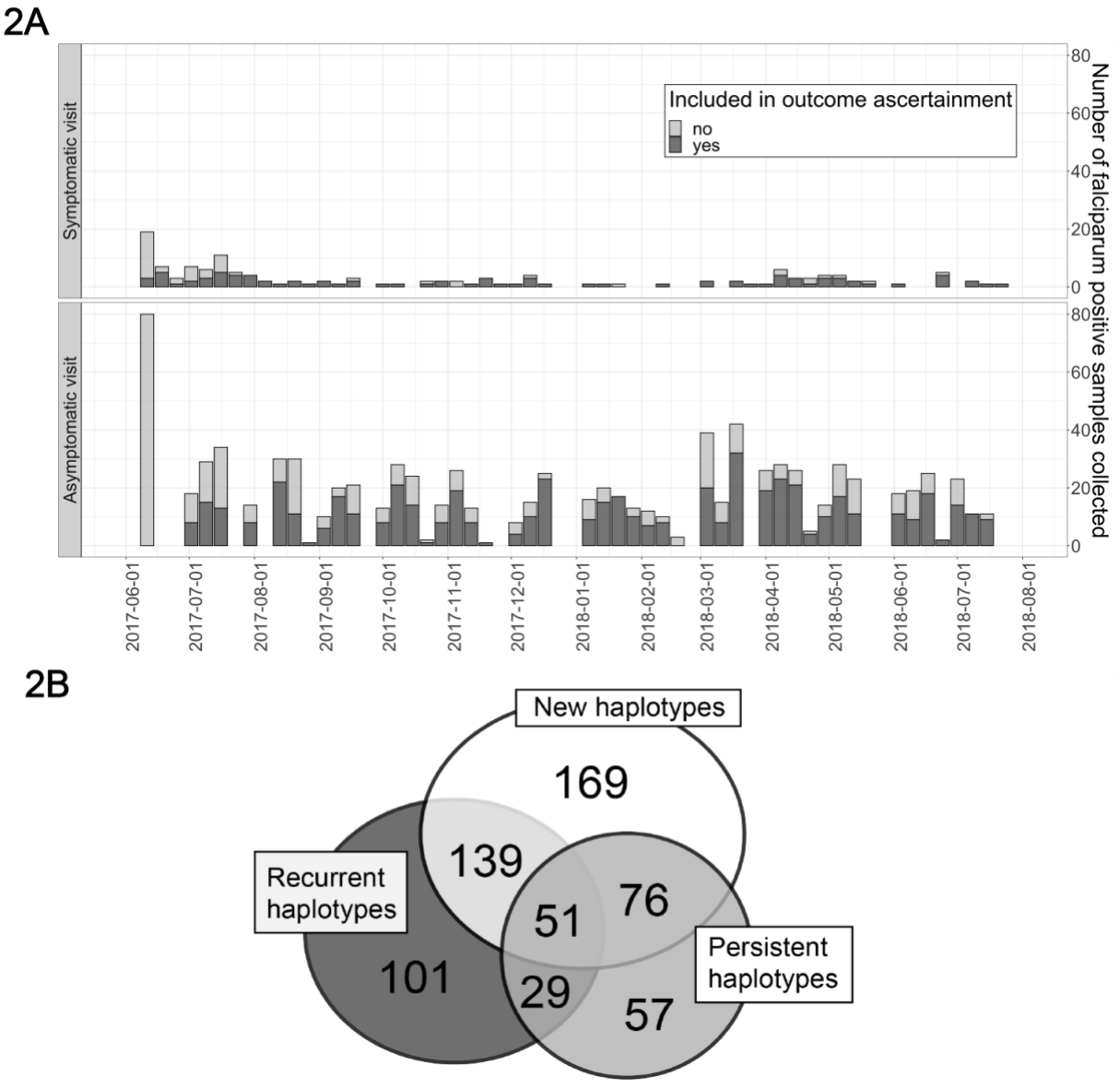


Figure 5.3. Comparison of time since previous infection for symptomatic malaria infections with or without persistent *pfcsp* haplotypes.
Distribution of the number of days since previous infection across all symptomatic malaria infections stratified by whether or not the infection had *pfcsp* haplotypes persist from the asymptomatic to symptomatic infection.

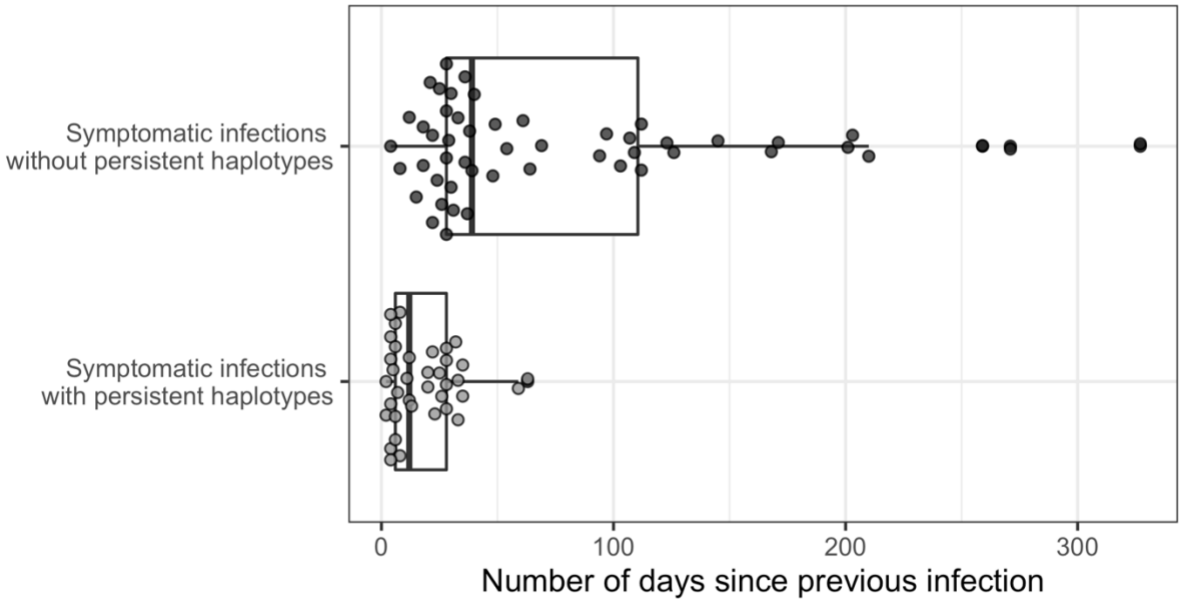


Figure 5.4. Incident infections: Comparison of odds of symptomatic malaria between infections harboring new versus recurrent *pfcsp* haplotypes.
5.4A: Multi-level logistic regression results for the odds of symptomatic malaria comparing (i) only new versus only recurrent and (ii) new and recurrent versus only recurrent *pfcsp* haplotypes. Dots indicate point estimate of the odds ratio, and lines the 95% confidence intervals. **5.4B:** Assessment of effect measure modification on symptomatic disease by age. Adjusted multi-level logistic regression models comparing the odds of developing symptomatic malaria between (i) only new versus only recurrent and (ii) new and recurrent versus only recurrent haplotypes were computed conditioned on age category.

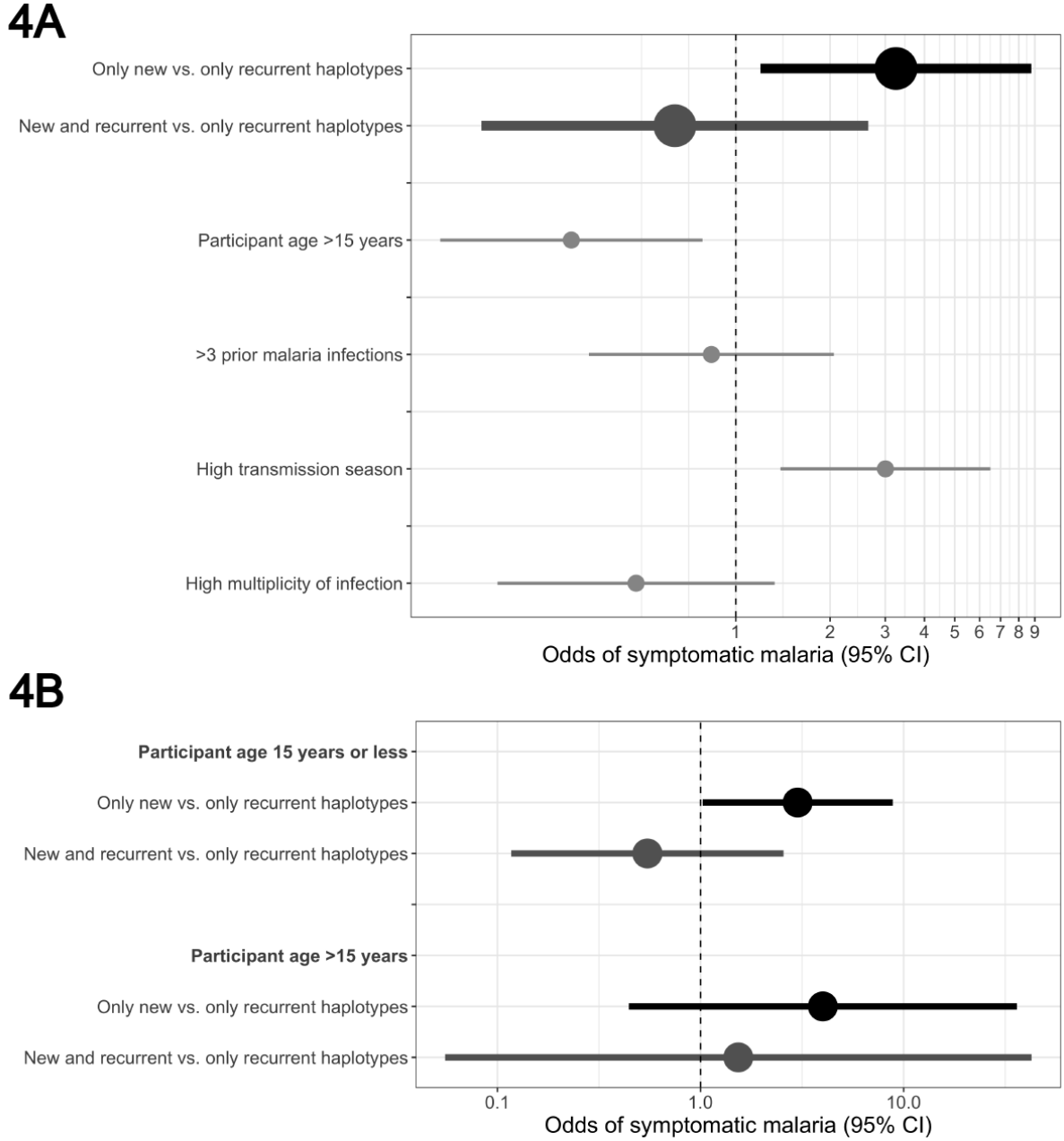
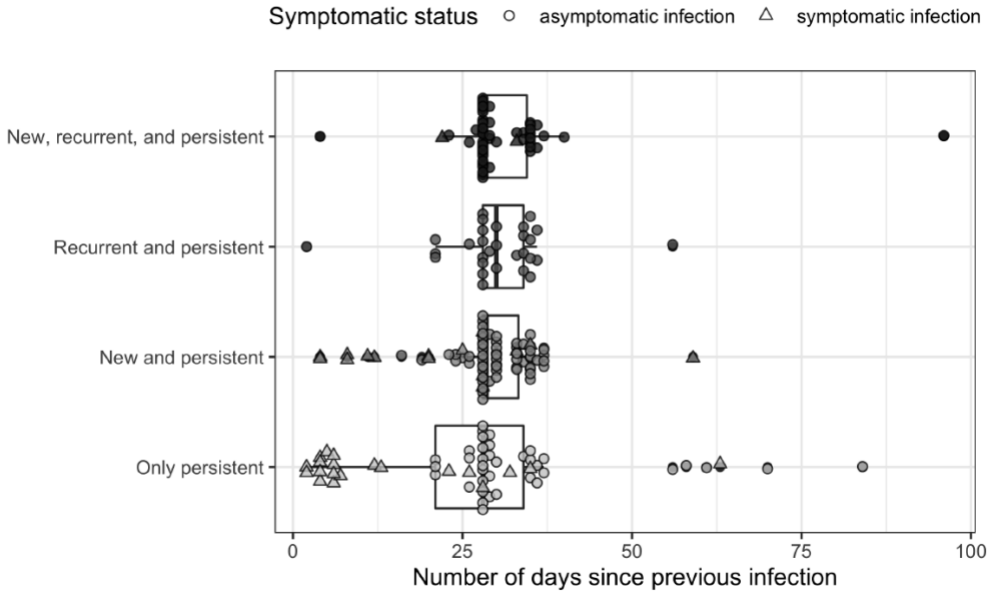
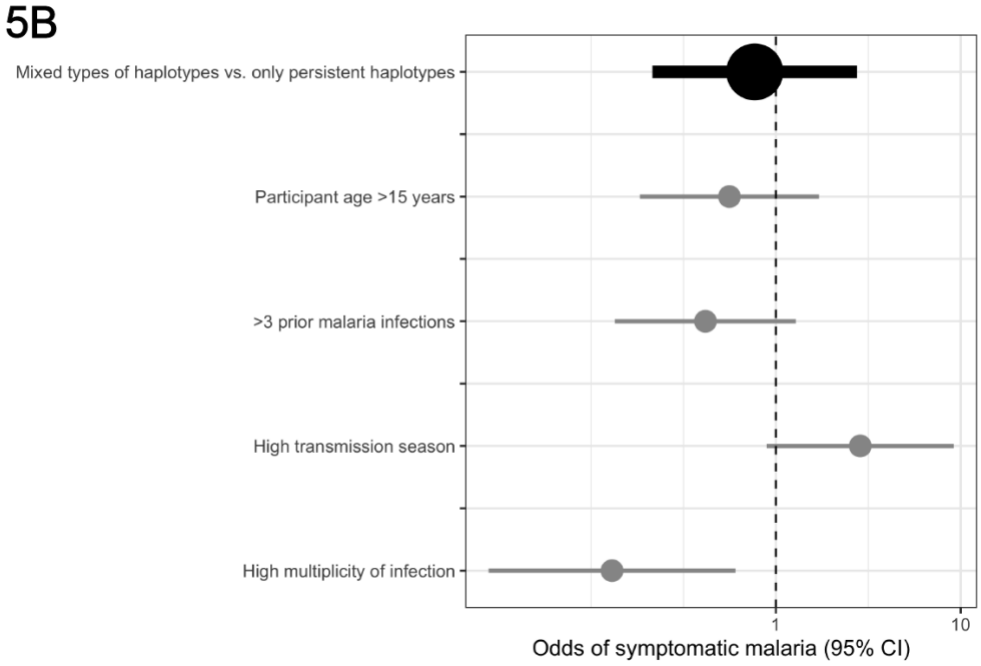


Figure 5.5. Persistent infections: Comparison of odds of symptomatic malaria between infections harboring mixed versus only persistent *pfcsp* haplotypes. **5.5A:** Distribution of the number of days since previous infection for malaria infections with persistent *pfcsp* haplotypes. Infections were categorized into: (i) only persistent; (ii) new and persistent; (iii) recurrent and persistent; and (iv) new, recurrent, and persistent. Asymptomatic infections were represented by circles and symptomatic ones by triangles. **5.5B:** Adjusted multi-level logistic regression results for the odds of symptomatic malaria comparing consecutive infections with mixed types of haplotypes versus only persistent haplotypes. Dots represent odds ratios and lines the corresponding 95% confidence intervals.

5A



5B



CHAPTER VI: GENOTYPING COGNATE *PLASMODIUM FALCIPARUM* IN HUMANS AND MOSQUITOES TO ESTIMATE ONWARD TRANSMISSION OF ASYMPTOMATIC INFECTIONS[1]

Introduction

Despite sustained malaria prevention efforts, progress in malaria control has stalled since 2010, with 228 million malaria episodes in 2018.[145] This persistence could result from a failure to target and mitigate infections in individuals or populations that disproportionally contribute to malaria transmission, so called malaria reservoirs. Sustained *Plasmodium falciparum* transmission despite case reductions could result from asymptomatic *P. falciparum* infections.[9,36,43–46] Asymptomatic infections are defined as the presence of parasites in the blood at any density in the absence of malaria-like symptoms[52] and typically represent either a state prior to development of symptoms (i.e. pre-symptomatic)[52] or one in which symptoms are attenuated due to non-sterilizing adaptive immunity.[157] Asymptomatic infections include both submicroscopic and microscopically patent infections which have different capacities for infecting mosquitoes.[158] Because asymptomatic infections are sub-clinical and therefore often remain untreated,[39] asymptomatically infected people can remain infectious to mosquitoes for prolonged periods and fuel onward transmission despite control

[1] This chapter previously appeared as an article in *Nature Communications*. The original citation is as follows: Sumner KM, Freedman E, Abel L, Obala A, Pence BW, Wesolowski A, Meshnick SR, Prudhomme-O'Meara W, Taylor SM. "Genotyping cognate *Plasmodium falciparum* in humans and mosquitoes to estimate onward transmission of asymptomatic infections," *Nature Communications* 12 (February 2021): 909.

measures.[17,99]

The relative contribution of asymptomatic infections to overall malaria transmission is incompletely understood. Several studies have compared the transmission potential of asymptomatic and symptomatic *P. falciparum* infections to mosquitoes and have generally confirmed that such infections are transmissible.[59–63] However, the small sample sizes and use of experimental approaches with artificial membrane feeding by laboratory-reared mosquitos limit generalizability by failing to capture variations in human activity, vector complexity and behavior, and parasite biology that influence transmissibility in natural settings. Such controlled feeding studies are critical to understand the fundamental biology of parasite transmission, and studies in natural, uncontrolled settings are necessary to confidently extend these insights to understand how they shape disease epidemiology. It is particularly critical to understand the impact of these infections in complex high-transmission settings, in which asymptomatic infections are highly prevalent but not commonly prioritized in transmission-reduction efforts. Such efforts include enhanced testing, treatment, and prevention on either mass or focal scales, and these tools can be employed more efficiently and rationally with a better understanding of the relative transmissibility of asymptomatic *P. falciparum* infections.

We investigated the contribution of asymptomatic *P. falciparum* infections to successful mosquito infection in a 14-month longitudinal cohort of 239 people in Western Kenya, a hyperendemic area where asymptomatic infections are common.[6,8] In these households, we collected cognate infections in both people and indoor-resting Anopheline mosquitoes, under the premise that, owing to the endophilic and

endophagic preferences for feeding by the principal vectors *Anopheles gambiae* and *A. funestus*, household transmission would be both measurable and substantial. Building upon previous studies, our approach combines empirical data collection of naturally fed mosquitoes, parasite genotyping using amplicon deep sequencing, and probabilistic modelling to estimate the transmissibility from people to mosquitoes of asymptomatic relative to symptomatic *P. falciparum* infections. We hypothesized that, compared to symptomatic infections, asymptomatic infections would be a larger source of infected mosquitoes.

Methods

Study Population and Data Collection

A longitudinal cohort of households across three villages (Kinesamo, Maruti, and Sitabicha) in Bungoma county, Kenya was established in June 2017 and followed until July 2018. The three villages were selected based on their high malaria prevalence in a previous cross-sectional study.[32] All household members in participating households over the age of 1 year were offered enrollment. Sample collection details have been reported.[127] For each participant, demographic and behavioral questionnaires were administered and dried blood spot (DBS) samples collected every month. The DBS were tested for *P. falciparum* parasites using real-time PCR post-hoc (see below), and therefore parasites detected in asymptomatic people were not treated. Participants contacted the study team at any time when experiencing symptoms consistent with malaria, at which time they were tested for malaria using a rapid diagnostic test (RDT) (Carestart Malaria HRP2 *Pf* from Accessbio)[115] and, if positive, treated with Artemether-Lumefantrine. DBS were also collected at the time of RDT testing. One morning each

week, indoor resting mosquitoes were collected from participant households using vacuum aspiration with Prokopacks.[116] From these collections, female *Anopheles* mosquitoes were identified morphologically and transected to separate the abdomen from the head and thorax.

Participant and Mosquito Sample Processing

Genomic DNA (gDNA) was extracted from mosquito abdomens and DBS samples using a Chelex-100 protocol.[128] gDNA from each DBS and mosquito was tested in duplicate using a duplex TaqMan real-time PCR (qPCR) assay targeting the *P. falciparum pfr364* motif and the human β-tubulin gene.[129] Samples were defined as *P. falciparum*-positive if: (i) both replicates amplified *P. falciparum* and both Ct values were < 40 or (ii) 1 replicate amplified *P. falciparum* and Ct value was < 38. *P. falciparum*-positive samples were genotyped across variable segments of genes encoding the apical membrane antigen-1 (*pfama1*) and circumsporozoite protein (*pfcsp*) as previously described,[148] with some additional steps taken for low parasite density samples (see **Supplementary Information**). Ultimately, dual-indexed libraries were prepared for both gene targets, then pooled and sequenced on an Illumina MiSeq platform.[117]

Sequencing reads were filtered based on read length and Phred quality scores and mapped to the 3D7 reference sequences for *pfama1* and *pfcsp*.[118–122,148] We performed haplotype inference on mapped reads using DADA2 (version 1.8) as implemented in R (version 3.6.1).[123,126] These putative haplotypes were then further filtered in order to mitigate the risk of false discovery by removing haplotypes from a sample that met any of the following criteria: (i) supported by < 250 reads within the sample; (ii) supported by < 3% of the sample's total read depth; (iii) deviation from the

expected nucleotide length of 300 for *pfama1* or 288 for *pfcsp*; or (iv) a minority haplotype distinguished by a one single-nucleotide polymorphism (SNP) difference from another haplotype within the sample that had a read depth > 8 times the read depth of the minority haplotype.[124] Finally, we removed a haplotype from the overall population if it was defined by a single variant position that was only variable within that haplotype (see **Supplementary Information, Figures S6.1-S6.4**). All genetic sequences are available through the National Center for Biotechnology Information GenBank (BioProject Number PRJNA646940).

Exposure Assessment

The main exposure was the classification of an infection as asymptomatic or symptomatic. We defined an asymptomatic infection as a *P. falciparum* infection detected by qPCR during active case detection in a participant lacking symptoms. We defined a symptomatic infection as a *P. falciparum* infection detected by both RDT and qPCR during passive case detection in a participant with at least one malaria-like symptom. To reduce potential for exposure misclassification, individual asymptomatic and symptomatic infections were excluded from the analysis if they occurred within 14 days of taking study-prescribed antimalarials for a symptomatic infection.

Within-participant Modeling of Transmissibility

To assess the likelihood of transmission by symptomatic status, we compared the proportion of mosquitoes that shared a haplotype between a participant's asymptomatic and symptomatic infections. To do so, we included only participants that had at least one asymptomatic and one symptomatic infection. We then paired each participant's infection events with all mosquitoes that were collected within 3 kilometers

as well as between 7 days prior to and 14 days following the participant infection, in order to constrain the search space for plausible transmission events to within time and distance parameters that are consistent with parasite and mosquito biology. For each infection, we computed the proportion of participant-mosquito pairings that shared at least one haplotype, and did so separately using either *pfcsp* or *pfama1* haplotypes.

We assessed the statistical significance of differences in these proportions between asymptomatic and symptomatic infections using a multi-level logistic regression model (**Equation 6.1**):

$$\ln\left(Transmission\ event_{ij}\right) = \alpha_i + \alpha_j + \beta_1 Asymptomatic\ infection_{ij} + \beta_2 Parasite\ density_{ij} + \beta_3 Mosquito\ abundance_{ij} + \epsilon_{ij} \quad \textbf{(6.1)}$$

This model included a random intercept at the participant level (α_i) and one at the household level (α_j) to account for repeated measures of participants clustered in households throughout the study. The model included covariates for parasite density (in parasites/μL in the participant samples) and mosquito abundance (expressed as the total number of female *Anopheles* mosquitoes collected within the week following the participant infection as <75 mosquitoes or ≥75 mosquitoes). The cutoffs for the number of mosquitoes chosen to represent mosquito abundance was determined by a functional form assessment and known malaria seasonality.

Probabilistic Modeling of Transmission Across All Participants

For a more comprehensive measure of transmission across all participants, we created a probabilistic model to estimate the probability that a shared haplotype between a participant and a mosquito represented a *P. falciparum* transmission event [$P(TE_{all})$]. $P(TE_{all})$ was estimated for the pairing of each infected participant with an

infected mosquito on the basis of three distinct features: (i) the time interval between the participant's infection and mosquito collection [$P(TE_t)$], (ii) the distance between the household of the participant and the household where the mosquito was collected [$P(TE_d)$], and (iii) the prevalence and number of parasite haplotypes shared [$P(TE_h)$]. For each pairing, we calculated these terms and then multiplied them to estimate $P(TE_{all})$. $P(TE_{all})$ values were computed independently using *pfama1* or *pfcsp* haplotypes.

Probability of Participant-to-mosquito Transmission Over Time

$P(TE_t)$ was defined as the probability of participant-to-mosquito parasite transmission as a function of the time interval between specimen collections. The rationale for this term was that participant-to-mosquito transmission could only occur within a certain time window based on the mosquito lifespan and parasite life cycle.[37,159] $P(TE_t)$ was assigned as 1 if a mosquito was collected within a 21-day window of the participant infection, spanning 7 days before the participant infection and 14 days after (**Figure S6.5**). This range only allowed participant-to-mosquito malaria transmission to be captured, following infections from participants to mosquito abdomens. If the mosquito was collected outside of this window, $P(TE_t)$ was set to 0. A sensitivity analysis was conducted to assess how differences in the time window chosen affected results, expanding the time window to allow mosquitoes to be collected up to 30 days after the participant's infection (**Figure S6.6**).

Probability of Participant-to-mosquito Transmission Over Distance

$P(TE_d)$ was defined as the probability of participant-to-mosquito parasite transmission as a function of Euclidean distance between specimen collections. This

term was included to restrict participant-mosquito pairs to only be considered as a possible transmission event within a reasonable distance for a mosquito to fly. $P(TE_d)$ was calculated using a modified negative exponential distribution previously observed in a study tracking *Anopheles* mosquito movement (**Equation 6.2; Figure S6.7**):[160]

$$P(TE_d) = e^{-3d} \quad \textbf{(6.2)}$$

For example, by 0.66 kilometers from the participant, which was the maximum distance blood fed *Anopheles* mosquitoes were observed to fly in a Kilifi study,[161] the probability of transmission was already low (14%) and at 3 kilometers it had dropped to 0% entirely. A sensitivity analysis was conducted to evaluate how changing the distance between specimen collection to allow specimen collection at a distance greater than 3 kilometers influenced results (**Figure S6.8**). We also compared the number of *pfcsp* haplotypes shared within 3 kilometers compared to at a distance of greater than 3 kilometers (**Figure S6.9**).

Probability of Participant-to-mosquito Transmission Over Haplotypes

$P(TE_h)$ was defined as the probability that a shared haplotype represented a participant-to-mosquito parasite transmission as a function of the number of shared haplotypes and the population prevalence of each shared haplotype. The premise of the calculation of $P(TE_h)$ was that the probability that haplotype sharing represented a transmission event increased with a higher number of haplotypes shared as well as the rarity of those haplotypes across the study population. $P(TE_h)$ was calculated independently for *pfama1* and *pfcsp* haplotypes using **Equation 6.3**:

$$P(TE_h) = (1 - \prod_{n=1}^{s} PopPrev_n{}^{1/3})(\frac{s}{MOI_i}) \quad \textbf{(6.3)}$$

where:

1. s indicates the number of haplotypes of a gene target (*pfcsp* or *pfama1*) that are shared between components of the pair;
2. $PopPrev$ indicates the *prevalence of the haplotype across the entire study population*, calculated by dividing the number of samples with that haplotype by the total number of samples in the study. $PopPrev$ is calculated from 1 to s, where s is the total number of shared haplotypes between the participant and mosquito pair. $PopPrev$ for each haplotype is rescaled by taking the cubed root, as it is highly right skewed; and
3. MOI_i is the participant's multiplicity of infection (MOI), represented by the number of *unique gene* haplotypes observed in the participant's infection (i).

We applied the term $(\frac{s}{MOI_i})$ in order to mitigate the risk of biasing $P(TE_h)$ towards larger values in participants with high MOI values. If no haplotypes were shared between the participant and mosquito pair, $P(TE_h)$ = 0. $P(TE_h)$ was calculated independently for *pfcsp* and *pfama1*. A sensitivity analysis was conducted comparing $P(TE_h)$ calculated independently for each gene target to $P(TE_h)$ calculated using both gene targets (**Figure S6.10**).

Probability of Participant-to-mosquito Transmission Over All Variables

Individual terms for $P(TE_t)$, $P(TE_d)$, and $P(TE_h)$ were combined into a final estimate of the probability of transmission [$P(TE_{all})$]. For participant-mosquito pairs that had probability values > 0 of $P(TE_t)$, $P(TE_d)$, and $P(TE_h)$, $P(TE_{all})$ was calculated using **Equation 6.4**:

$$P(TE_{all}) = P(TE_t) * P(TE_d) * P(TE_h) \quad \textbf{(6.4)}$$

If $P(TE_t) > 0$ and $P(TE_d) > 0$ but $P(TE_h) = 0$, $P(TE_{all}) = 0$. $P(TE_t)$, $P(TE_d)$, and $P(TE_h)$ were rescaled to the range 0 to 1 to be comparable when multiplying.

Statistical Analysis

To estimate the probability of a participant-to-mosquito transmission event using the probabilistic method for participant-to-mosquito transmission [$P(TE_{all})$] across time, distance, and the haplotypes shared, we compared values between participants with asymptomatic and symptomatic infections using a multi-level logistic regression model (**Equation 6.5**).

$$\ln(Transmission\ event_{ij}) = \alpha_i + \alpha_j + \beta_1 Asymptomatic\ infection_{ij} + \beta_2 Parasite\ density_{ij} + \beta_3 Age5to15_{ij} + \beta_4 Age15over_{ij} + \beta_5 Mosquito\ abundance_{ij} + \beta_6 Kinesamo\ village + \beta_7 Sitabicha\ village + \epsilon_{ij} \quad \textbf{(6.5)}$$

We included a random intercept at the participant level (i) to account for repeated measures for participants who experienced multiple malaria infections (asymptomatic or symptomatic) throughout the study. To consider different transmission intensities between households, we included a random intercept at the household level (j). We controlled for confounding covariates that we identified in a DAG (**Figure S6.11**) and performed functional form assessments on continuous variables prior to inclusion (See **Supplementary Information, Tables S6.3-S6.5**). The final model included covariates for village, parasite density in the participant samples in parasites/μL (linear), participant age at study enrollment (categorized: <5 years, 5-15 years, >15 years), and mosquito abundance (expressed as the total number of female *Anopheles* mosquitoes collected within the week following the participant infection as <75 mosquitoes or ≥75 mosquitoes). To reduce skew for the multi-level model, parasite density was centered

and rescaled to have a mean of 0.

Prior to modeling, differences in model covariates across symptomatic status were assessed using the Wilcoxon Rank Sum test with continuity correction for continuous variables and the Pearson's χ^2 test for categorical variables. Differences in MOI across sample types were calculated using the Pairwise Wilcoxon Rank Sum test with the Bonferroni correction for multiple comparisons. All *p*-values obtained from the bivariate tests were adjusted using the Bonferroni correction to account for repeated measures across participants of up to 14 infections, which was the maximum number of infections observed in any participant during study follow-up.

We conducted a sensitivity analysis of different coding choices for the probabilistic combination of time, distance, and haplotypes [$P(TE_{all})$]. We re-computed the logistic regression model specified above testing variable cutoffs for the binary coding of the outcome variables. Cutoffs for what was considered a participant-to-mosquito malaria transmission ranged from 0.00 to 0.55 due to sparse data restrictions above 0.55. The model contained the same covariates and random effects as the model in **Equation 6.5**.

Contribution to Infectious Reservoir of Asymptomatic Infections

The contribution to the infectious reservoir made by asymptomatic infections was calculated using the odds ratio estimate obtained from the probabilistic method for participant-to-mosquito transmission [$P(TE_{all})$] across time, distance, and the haplotypes shared. The odds ratio for the binary coding of $P(TE_{all})$ was used, where any value of $P(TE_{all})$ indicated a participant-to-mosquito transmission event. The contribution to the infectious reservoir (C_{at}) was estimated using **Equation 6.6:**[62,68]

$$C_{at} = \frac{P_{at}I_a}{P_{at}I_a + P_{st}I_s} * (100) \quad \textbf{(6.6)}$$

where:

1. P_{at} represents the proportion of all infections that were asymptomatic (a) during each month (t) of follow-up;
2. P_{st} represents the proportion of all infections that were symptomatic (s) during each month (t) of follow-up;
3. I_a indicates the likelihood a mosquito was infected by someone with an asymptomatic (a) infection. I_a was calculated using the odds ratio obtained from the multi-level logistic regression model that estimated the probability of participant-mosquito transmission with $P(TE_{all})$ coded binarily: $OR = \frac{I_a}{I_s}$. Model random effects and covariates were previously described in **Equation 6.5.** I_a did not vary across months; and
4. I_s represents the likelihood a mosquito was infected by someone with a symptomatic (s) infection. I_s was calculate as follows: $I_s = 1 - I_a$. I_s did not vary across months.

A cumulative value of C_{at} was calculated across the entire follow-up period and represented by C_a. The upper and lower limits of the 95% confidence interval for C_a were calculated using the upper and lower limits for the 95% confidence interval from the estimated odds ratio. All statistical analyses were performed using R (version 3.6.1).[126]

Ethical Considerations

All adult participants provided written informed consent prior to participation. Participants between the ages of 1 and 18 years old were included if their parent or

legal guardian provided written informed consent. Verbal assent was also obtained from children between 8 and 18 years. The study was approved by the ethical review boards of Moi University (2017/36), Duke University (Pro00082000) and the University of North Carolina at Chapel Hill (19-1273).

Results

From June 2017, we enrolled 268 participants across 3 villages in Bungoma County, Kenya, in the cohort study; after excluding participants with either zero *P. falciparum* infections or less than 2 months of follow-up, the analysis data set consisted of 239 participants across 38 households who were visited monthly for active case detection of asymptomatic infections and as-needed for passive case detection of symptomatic infections. In these participants across 14 months, we recorded 137 symptomatic *P. falciparum* infections during 501 sick visits and 902 asymptomatic *P. falciparum* infections during 2312 routine visits (**Figure 6.1**). From their households, we collected 1494 female *Anopheles* mosquitoes; of 1450 mosquito abdomens with gDNA available, we identified 203 *P. falciparum*-positive mosquitoes.

These 1242 real-time PCR-positive *P. falciparum* infections (N=902 asymptomatic infections, N=137 symptomatic infections, and N=203 infected mosquito abdomens) were genotyped for the *P. falciparum* parasite genes encoding apical membrane antigen-1 *(pfama1)* and circumsporozoite protein *(pfcsp)* using PCR amplification, amplicon deep sequencing, and a validated haplotype inference program with strict quality-filtering criteria.[123] *Pfcsp* and *pfama1* were selected owing not to phenotypes associated with their protein products but rather to their sequence diversity, which enables capture of diverse parasite strains and matching strains between

hosts.[148] Results for *pfama1* haplotypes are reported in the supplement. For *pfcsp*, we obtained analyzable haplotypes that passed our custom quality filtering for 1046 samples (84.2%), across which we identified 229 unique *pfcsp* haplotypes. These haplotypes harbored variants at 72 nucleotide positions in the sequenced segment of *pfcsp*; variants at 37 (51.4%) of these positions were previously reported (**Figure S6.2**).[162–164] Many haplotypes were observed across all three sample types, but some haplotypes were private to asymptomatic infections, symptomatic infections or mosquitoes (**Figures 6.2, S6.12**). Between sample types, the median *pfcsp* multiplicity of infection (MOI) was higher for mosquitoes (6, Interquartile range [IQR]: 4 to 9) compared to either symptomatic infections (1, IQR: 1 to 3, *p*-value < 0.001 by Pairwise Wilcoxon Rank Sum Test) or asymptomatic infections (3, IQR: 1 to 7, *p*-value < 0.001) (**Figure 6.2**).

We used these *P. falciparum* haplotypes as identifiers by which to estimate parasite transmission from people to mosquitoes by computing pairwise metrics of parasite haplotype sharing between infected participants and mosquitoes. We first analyzed a subset of 65 participants who suffered both asymptomatic and symptomatic infections, from whom we paired all 225 infected events (N=143 asymptomatic and N=82 symptomatic infections) with infected mosquitoes that were collected: (i) between 7 days before and 14 days after the event, and (ii) within 3 kilometers of the participant's household. This yielded 1565 participant-mosquito pairs for the 225 events; this subset of participants and events was similar to the overall population (**Table S6.1**). For each event, we computed the proportion of participant-mosquito pairings in which at least 1 *pfcsp* haplotype was shared between the mosquito and the participant. In a multi-level

logistic regression model controlling for parasite density and mosquito abundance, compared to their symptomatic infections, their asymptomatic infections had higher odds of sharing a parasite haplotype with infected mosquitoes [OR: 2.56, 95% CI: 1.36 to 4.81] (**Figure 6.3**). Results were similar but not statistically significant in parallel analyses using *pfama1* [OR: 1.30, 95% CI: 0.63 to 2.69] (**Figures S6.13-S6.14**), indicating that, compared to when the individuals suffered symptomatic infections, their asymptomatic infections were more likely to result in successful parasite transmission to mosquito vectors.

In order to more comprehensively analyze transmission across all participants irrespective of their infection counts, we extended our assessment of transmission using pairings of all infections in participants and mosquitoes using a probabilistic model of transmission. Across all samples over 14 months, there were 159,285 potential pairings of infected participants and mosquitoes, and after applying the aforementioned temporal and geographic distance constraints to these pairings to remove those with implausible transmission potential (**Figure 6.4**), the final analysis data set consisted of 3727 participant-mosquito pairs. These comprised 198 participants and 182 mosquitoes that were drawn from 37 households across all 3 villages. Among these 3727 pairings, mosquitoes paired with asymptomatic participants (N=3012) outnumbered those paired with symptomatic participants (N=715). Compared to those including asymptomatic infections, pairings including symptomatic infections had higher parasite densities (*p*-value < 0.001 by Wilcoxon Rank Sum test), were more likely to occur in a participant under 5 or over 15 (*p*-value < 0.001 by Pearson's χ^2 test) (**Tables 6.1 and S6.2**) and typically occurred during periods with larger mosquito abundance (*p*-value < 0.001 by

Pearson's χ^2 test). Across the all pairings, the median number of haplotypes shared within a participant-mosquito pair was 1 (range: 0 to 8, IQR: 0 to 2) for asymptomatic and 0 (range: 0 to 7, IQR: 0 to 1) for symptomatic infections.

For each of these 3727 pairings, we computed the probability that a shared haplotype between a participant and mosquito represented a transmission event [$P(TE_{all})$] as a function of three indices: (i) temporal distance [$P(TE_t)$], (ii) geographic distance [$P(TE_d)$], and (iii) the prevalence and quantity of shared haplotypes between samples [$P(TE_h)$] (**Figure 6.4**). The rationale for this approach was to assign a probability to each pair that reflected the level of confidence that the pair represented a participant-to-mosquito transmission event. The probability increased for pairs that were closer in space or time and for those which shared a higher number of haplotypes or haplotypes that were comparatively rare across samples. We aggregated all three terms into a compound estimate of a probable transmission event [$P(TE_{all})$]; across all pairings (N=3727), the median $P(TE_{all})$, was 0.05 (IQR: 0.00 to 0.15), and among only those pairings with at least 1 shared haplotype (N=2278), the median $P(TE_{all})$ was 0.12 (IQR: 0.06 to 0.21). $P(TE_{all})$ represented a relative likelihood that a human and mosquito pair that shared parasite haplotypes represented a transmission event and should be interpreted relative to other $P(TE_{all})$ values.

We compared our estimates of transmission for each pairing between those with asymptomatic and symptomatic infections using a multi-level logistic regression on $P(TE_{all})$, controlling for parasite density, participant age, mosquito abundance, and village. Using *pfcsp* haplotypes for haplotype indices, compared to symptomatic infections, asymptomatic infections had 50% higher odds of being matched to a

mosquito infection [OR: 1.50, 95% CI: 1.07 to 2.10] (**Figure 6.5**). In parallel analyses using *pfama1*, we observed a similar increase in the odds of transmission to mosquitoes from asymptomatic compared to symptomatic infections [OR: 1.22, 95% CI: 0.82 to 1.82] (**Figure S6.15**). We re-computed regression models after dichotomizing our estimated $P(TE_{all})$ values at various cutoffs from 0.00 to 0.55, reflecting the range of $P(TE_{all})$ values and increasing stringency for defining a transmission event (**Figure 6.5**). Across this broad range of $P(TE_{all})$ definitions, asymptomatic infections had consistently higher odds of onward parasite transmission. When we defined a transmission event as any non-zero $P(TE_{all})$ value, reflecting sharing between participant and mosquito of any number and quality of haplotypes, compared to symptomatic infections, asymptomatic infections more than doubled the odds of transmission to a mosquito [OR: 2.66, 95% CI: 2.05 to 3.47].

Finally, we used this measurement to estimate the contribution of each type of infection to onward transmission across our population as a function of the monthly proportion of all infections that were asymptomatic, which varied from 73.4% to 97.4% between months. Using these, we estimated that monthly contributions to mosquito infections by asymptomatic infections varied from 88.0% to 99.0% (**Figure 6.5**), and averaged across all months in our high and perennial transmission setting, asymptomatic infections were the source of 94.6% (95% CI: 93.1 to 95.8%) of mosquito infections.

Discussion

In this longitudinal epidemiological and entomological cohort in Western Kenya, we investigated the relative contributions to onward *P. falciparum* transmission of

asymptomatic compared to symptomatic *P. falciparum* infections. To do so, we analyzed parasite haplotypes in people and mosquitoes using a probabilistic model to directly estimate participant-to-mosquito malaria transmission. We report that, compared to symptomatic people, those with asymptomatic infections had more than double the odds of transmission to mosquitoes. Owing to this as well as the high prevalence of asymptomatic infections, we estimated that asymptomatic infections were the source of nearly all *P. falciparum* parasites infecting mosquitoes. Our findings provide an explicit rationale to target asymptomatic *P. falciparum* infections as a component of transmission-reducing programs.

Across 14 months of observation in a high-transmission setting, asymptomatic *P. falciparum* infections were the major source of onward malaria transmission. Specifically, relative to symptomatic infections, asymptomatic infections had 2.66-fold odds of probable malaria transmission to mosquitoes. Our findings are consistent with results from smaller studies, which suggested that asymptomatic infections were more likely to transmit to mosquitoes than symptomatic infections.[61,62] Those studies used experimental feeding on infected blood by laboratory-reared mosquitoes to measure transmission, and therefore could not capture variance in the feeding behaviors of vectors[112] or the natural trajectories of infections.[113] Our findings build upon these previous studies by capturing participant-to-mosquito transmission longitudinally, in a larger study population, and in a natural setting with mosquitoes collected within participants' households. Notably, we also observed this positive association between asymptomatic infections and transmission among the overall cohort using the alternate, unlinked parasite genotyping locus of *pfama1* as well as among a subset of participants

who suffered both asymptomatic and symptomatic infections during the study period. Although our study does not enable the identification of a clear mechanism for this association, a lack of symptoms may allow a longer duration of infection, and thereby enable both the development of gametocytes as well as more opportunities to be bitten by and transmit malaria to mosquitoes.[39]

Our approach used probabilistic modelling of genotypes captured by amplicon deep sequencing to estimate *P. falciparum* transmission. Prior studies of participant-to-mosquito malaria transmission using alternate approaches[59–63] have incompletely captured the complexity of natural systems, which limits their generalizability. Mosquito feeding experiments employing either direct skin or membrane feeding fail to represent numerous participant-, mosquito-, and parasite-related factors that are critical to transmission. These critical factors include variance among mosquito vectors in biting preferences,[165] behaviors,[112] and success;[166] among parasites in replication rates[113] and gametocyte production;[167] and among participants in exposure to vectors[168] and care-seeking behavior.[169] Similarly, this complexity also confounds the use of gametocyte prevalence or density as a proxy for transmission,[170] which may more precisely define which infections can rather than do transmit. Other studies have used modelling approaches to estimate how transmission dynamics could change in a more realistic setting, finding that submicroscopic infections are a large source of malaria spread;[65,136] however, the studies did not examine how transmission differed by symptomaticity.

This approach to measure participant-to-mosquito transmission offers a scalable tool that can be adapted to diverse settings. Consistent with prior reports from high-

transmission settings,[148,171–174] we found high diversity of *pfama1* and *pfcsp* haplotypes in our study site, likely the result of strong balancing selection on these loci exerted by immune pressure. This large number of unique haplotypes allowed us to both identify matches between participant-mosquito pairs as well as weight the relevance of those matches for potential transmission events based on the quantity and rarity of shared haplotypes. Importantly, amplicon deep sequencing enabled this approach with its technical ability to capture minority variants within mixed infections[110] and scalability in a large field study.[175] More precise estimations of individual transmission events as well as mapping of transmission chains, may require novel approaches using higher-dimensional genotyping combined with analytic models that resolve polygenomic infections. Our results highlight how integrated genetic and computational approaches can be implemented in large field studies to leverage parasite genetic diversity for investigating fundamental features of parasite epidemiology.

Using this approach, we observed that the median number of *pfcsp* haplotypes (or MOI) was much higher in mosquito infections (6) than in asymptomatic (3) or symptomatic (1) human infections (**Figure 6.2**). This high median MOI in mosquito abdomens is surprising given that wild-caught[176] and membrane-fed[177] Anopheline mosquitoes typically have < 5 oocysts, suggesting that the high amount of genetic diversity that we observed was likely harbored by a very small number of oocysts in the collected mosquitoes. This could have resulted from the transmission to mosquitoes of cryptic haplotypes that were undetectable in asexual human infections, as has been reported with both *P. falciparum* and *P. vivax,*[178] although both sample types were processed analogously and were subjected to identical haplotype quality filtering

criteria. On a related note, partial immune recognition of expressed circumsporozoite protein or apical membrane antigen-1 variants, which are expressed in the liver or blood stage respectively, may have served to differentially limit the densities of certain variants below the limits of detection in human infections while allowing passage to and propagation in mosquitoes. Finally, given evidence that *Anopheles gambiae* can take multiple bloodmeals per gonotropic cycle,[179–181] and that this behavior may be enhanced by an existing sporozoite infection of the mosquito,[182] these oocysts may represent an accumulation of parasites acquired over multiple feedings on multiple days from multiple infected humans, which collectively would enhance the diversification of midgut parasites.

The finding that asymptomatic *P. falciparum* infections are the primary source of infections in mosquito vectors provides an explicit rationale to target these infections in order to reduce transmission in highly-endemic settings. Across sub-Saharan Africa asymptomatic *P. falciparum* infections are highly prevalent:[49,50,52,183] one meta-analysis estimated a continent-wide prevalence in 2015 of 24% among just children aged 2 to 10 years.[35] Asymptomatic infections have been targeted in prior studies either by testing defined geographic or demographic groups (i.e. active case detection) or by foregoing testing and implementing mass-drug administration (MDA) of antimalarials.[184] Both active case detection and MDA have proven effective or been implemented in low-transmission, pre-elimination settings, where they have been recommended as interventions to accelerate progress to elimination.[185] In contrast, high-transmission settings like ours rely on bed net use, access to care, use of rapid diagnostics, and treatment with artemisinin-based combination therapies (ACT) for control.[145] Despite the

adoption of all of these interventions in our study site, asymptomatic infections remained the major source of mosquito infections, suggesting the need for enhanced interventions. The efficacy of such interventions in high-transmission settings on the asymptomatic reservoir specifically – and on transmission reduction more generally – may be feasibly testable with novel tools to estimate transmission using serologic[186] or parasite genetic[155] measures.

Our study had several limitations. Symptomatic infections were quickly diagnosed and treated with effective therapy under our protocol which likely reduced the duration of these infections and therefore limited their transmission potential. This access to diagnosis and treatment is higher than is generally available across sub-Saharan Africa,[145] though recent reports indicate gradual improvement in quality clinical management.[187] Conversely, we may have under-detected asymptomatic infections and therefore over-represented symptomatic infections, owing either to the sparse monthly sampling for asymptomatic infections or the inability to capture transmission from symptomatic infections during their asymptomatic or pre-symptomatic phase. We expect that this would serve mainly to bias our analyses towards the null by providing relatively more opportunities for symptomatic infections to match to mosquitoes. Similarly, mosquito sampling was necessarily sparser than human sampling, precluding absolute measurement of all transmission events but allowing for relative estimations to onward transmission. We had no direct measurement of gametocytes due to the types of sample collection, precluding a direct analysis of their participation in transmission; however, we adjusted models for asexual parasite density, which has been suggested as a proxy for gametocyte density.[158] We only measured transmission directly within

households, and cannot capture events occurring in other settings; this limitation is mitigated by the known nocturnal feeding preference of local vectors. Finally, many infections in participants and mosquitoes had low parasite densities, which increases the risk of haplotype false discovery.[124] To mitigate this risk, we enforced stringent haplotype censoring based on read quality and haplotype abundance consistent with prior studies.[70,124,188]

In our longitudinal study of paired participant and mosquito *P. falciparum* infections, compared to people with symptomatic malaria infections, those with asymptomatic infections were more than twice as likely to successfully transmit *P. falciparum* to *Anopheles* mosquitoes. Future studies can investigate biological and epidemiological mechanisms by which symptomaticity influences transmission as well as estimate the feasibility and efficacy of targeting asymptomatic infections as a means to reduce transmission in highly-endemic settings.

Table 6.1. Comparison of pairings of participant and mosquito infections by symptomatic status

	Asymptomatic infections (N=3012)	Symptomatic infections (N=715)	*P*-value
Participant-level covariates			
Parasite density (parasites/μL), Median (IQR)	11.08 (0.96-251)	3229.46 (505-6581)	<0.001[a]
Age, N (%)			<0.001[b]
<5 years	326 (3.72)	112 (15.66)	
5-15 years	1372 (45.55)	111 (15.52)	
>15 years	1314 (43.63)	492 (68.81)	
Mosquito abundance, N (%)			<0.001[b]
Low	1564 (51.93)	227 (31.75)	
High	1448 (48.07)	488 (68.25)	
Number of *pfcsp* haplotypes, Median (IQR)	3.00 (1.00-8.00)	2.00 (1.00-3.00)	<0.001[a]
Village, N (%)			<0.001[b]
Maruti	2267 (75.27)	411 (57.48)	
Kinesamo	616 (20.45)	208 (29.09)	
Sitabicha	129 (4.28)	96 (13.43)	
Participant-mosquito pair-level covariates			
Probability of transmission, Median (IQR)			
Across all variables	0.05 (0.00-0.16)	0.00 (0.00-0.11)	<0.001[a]
Time interval[#]	1.00 (1.00-1.00)	1.00 (1.00-1.00)	NE
Distance interval	0.67 (0.54-0.84)	0.70 (0.53-0.80)	1.000[a]
pfcsp haplotype sharing and prevalence*	0.09 (0.00-0.24)	0.00 (0.00-0.15)	<0.001[a]
For those that shared *pfcsp* haplotypes	0.20 (0.09-0.32)	0.19 (0.10-0.37)	1.000[a]
Number *pfcsp* haplotypes shared, Median (IQR)**	1.00 (0.00-2.00)	0.00 (0.00-1.00)	<0.001[a]
For those that shared *pfcsp* haplotypes	2.00 (1.00-3.00)	1.00 (1.00-1.00)	<0.001[a]

Abbreviations: IQR, interquartile range; NE, not evaluated

[#] The probability of transmission based on the time interval was set as 1 for all participant-mosquito pairings where the mosquito was collected within 7 days prior to or 14 days after the participant's infection. All pairings outside of that time interval had a probability of transmission of 0.

*The probability of transmission based on the *pfcsp* haplotype sharing and prevalence is shown for all pairings regardless on if they shared haplotypes or not.

**The number of *pfcsp* haplotypes shared is shown for all pairs regardless on if they shared haplotypes or not.

[a] Wilcoxon Rank Sum test with continuity correction and Bonferroni correction for repeated measures

[b] Pearson's χ^2 test with Bonferroni correction for repeated measures

Figure 6.1. *P. falciparum* infections observed across study participants and female *Anopheles* mosquitoes across 14 months.
Female Anopheline mosquitoes were captured weekly by vacuum aspiration and their abdomens were tested using real-time PCR for the presence (red) or absence (gray) of *P. falciparum* (Pf) parasites. Symptomatic malaria infections were captured by passive case detection with clinical symptoms and positive *P. falciparum* results by both RDT and real-time PCR (blue). The number of participants who had malaria-like symptoms and requested a symptomatic visit but did not have a confirmed symptomatic infection were also identified (grey). Asymptomatic malaria infections were captured by active case detection at monthly follow-up visits with participants and real-time PCR-positive for *P. falciparum*. These monthly visits were conducted in different weeks for each of the 3 villages, with additional re-visits if needed to sample enrolled participants who were absent for the initial visit. Monthly counts of asymptomatic malaria infections (yellow) and uninfected participants (grey) were reported.

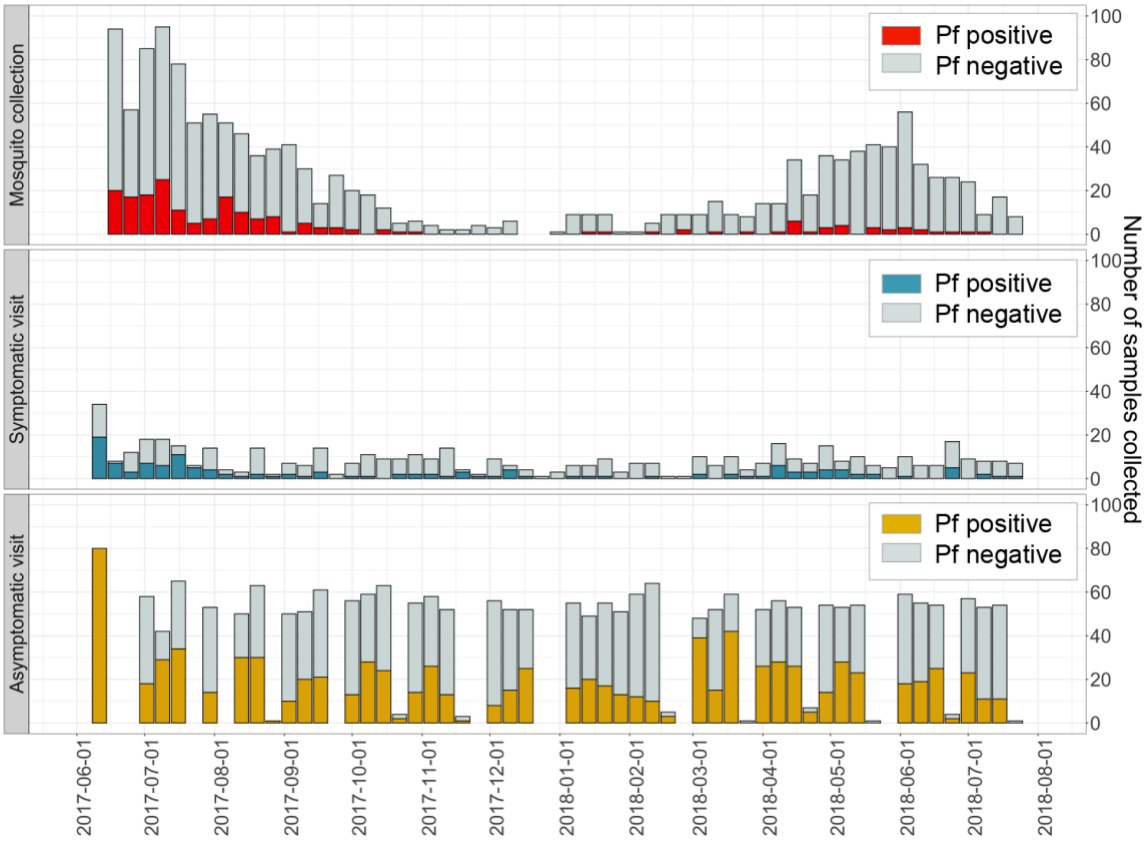


Figure 6.2. Distributions of *pfcsp* haplotypes across and within participants and mosquitoes.
6.2A: Distribution of the 75 most common *pfcsp* haplotypes) in mosquitoes (red), symptomatic infections (blue), and asymptomatic infections (yellow), ordered vertically by the number in the asymptomatic infections. A full plot of all 229 *pfcsp* haplotypes across sample types is in the supplement (**Figure S6.12**). **6.2B:** Multiplicity of infection (MOI) based on observed number unique *pfcsp* haplotypes in each mosquito abdomen (red), symptomatic infection (blue), and asymptomatic infection (yellow).

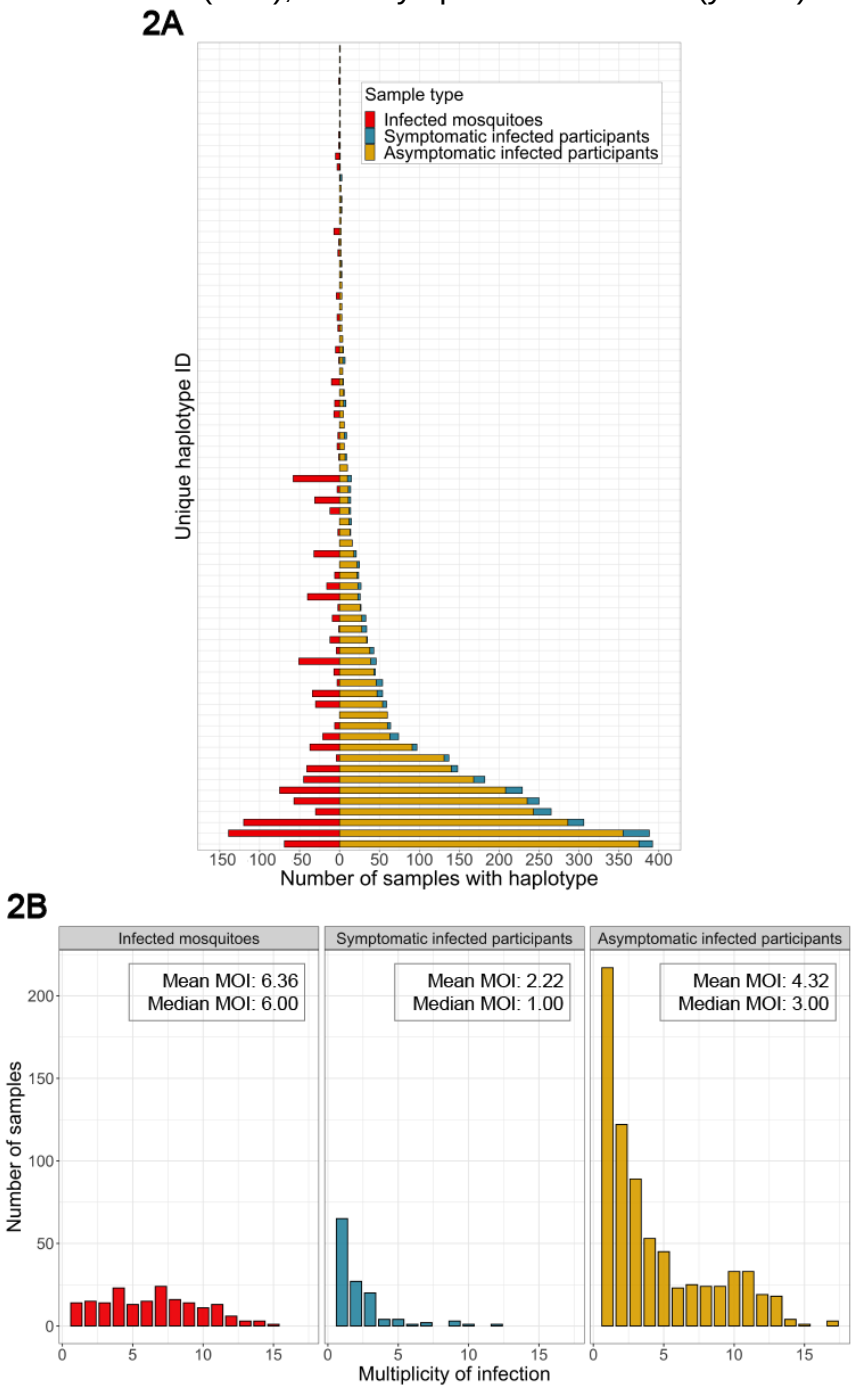


Figure 6.3. Comparison of the proportion of infected mosquitoes harboring a matching *pfcsp* haplotype for participants with both asymptomatic and symptomatic infections.
6.3A: Scatterplot of the proportion of pairings with a mosquito that shared a minimum of one haplotype for asymptomatic (y-axis) and symptomatic (x-axis) infections. Each dot is a participant who suffered at least one asymptomatic and symptomatic infection, and for participants with more than one of either type of infection, the plotted value is the median of proportions across infections within that type. Size of dots is relative to the total number of the participant's infections. **6.3B:** Odds ratios of the proportion of matched mosquitoes in a multi-level logistic regression model using the continuous coding of the proportion of participant-mosquito pairings that shared haplotypes for each infection. Dots are point estimates, and bars are 95% confidence intervals.

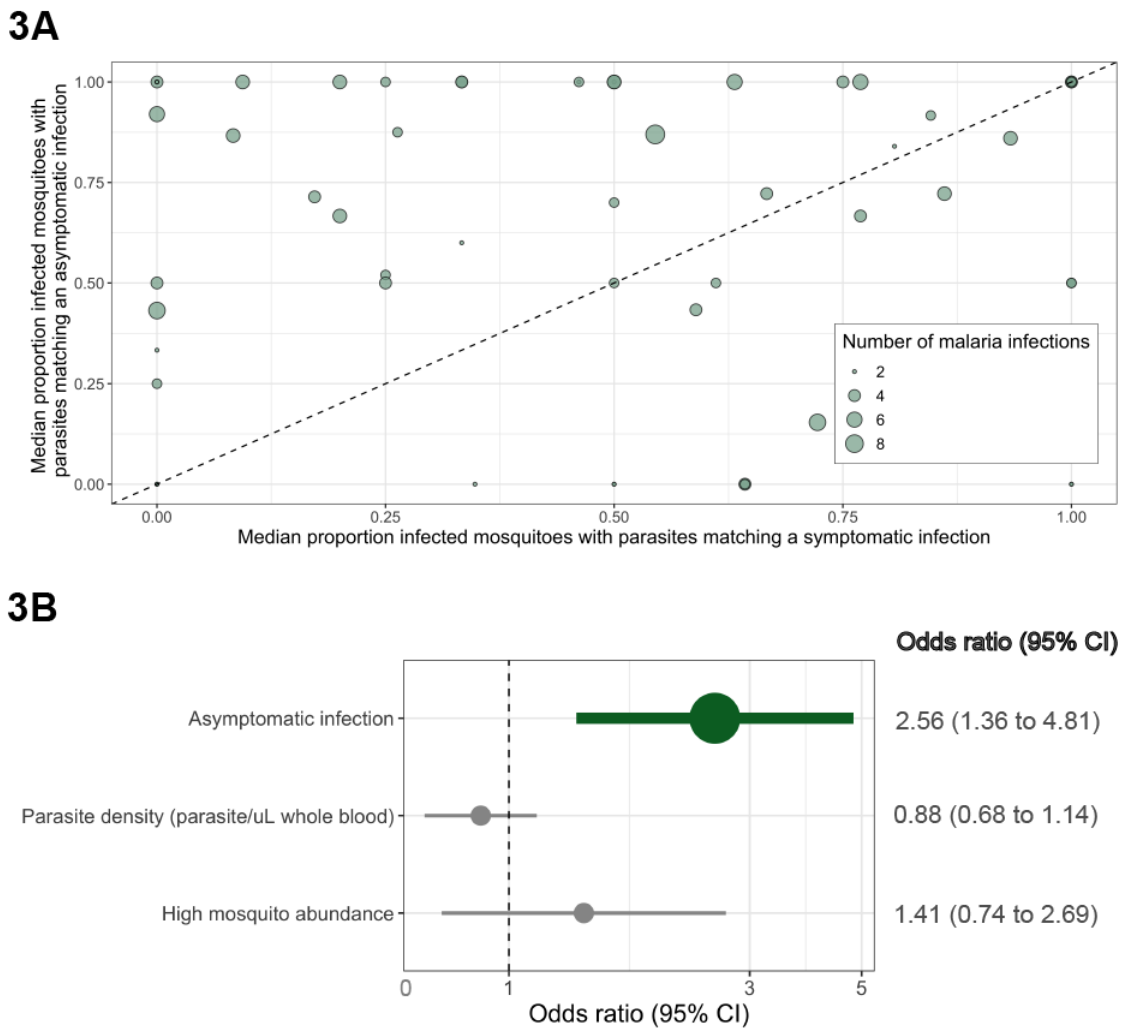


Figure 6.4. Modeling approach to estimate the probability of a *P. falciparum* transmission event to mosquitoes using the *pfcsp* gene target.
6.4A: Distribution of the interval in days between all possible pairings (N=159,285) of all infected participants and mosquito abdomens. Day 0 was set as the date of the mosquito infection, and therefore negative values indicate the mosquito was collected prior to the participant infection. The light green area indicates those pairings in which the mosquito was collected within 7 days prior to or 14 days after the participant's infection. Subsequent analysis was restricted to these pairings. **6.4B:** Distribution of the distance interval between all possible pairings of infected participants and mosquito abdomens. The light green area indicates those pairings within the same village and at a maximum distance of 3 kilometers, to which subsequent analysis was restricted. Across these pairings, a probability function was applied (**S6.6 Fig**) to upweight pairings with shorter distance intervals. The peaks result from differences in distance across the three villages. **6.4C:** Distribution of the estimated probabilities of transmission as a function of the number of *pfcsp* haplotypes shared within the participant-mosquito pair. These probabilities were estimated by upweighting pairings which shared more haplotypes and which shared haplotypes that were rare across the entire study population. **6.4D:** Distributions of final estimated probabilities of transmission events stratified by symptomatic status of the participant infection. Final probabilities were computed as the product of the individual probabilities based upon the time interval, distance interval, and *pfcsp* haplotypes of each pairing.

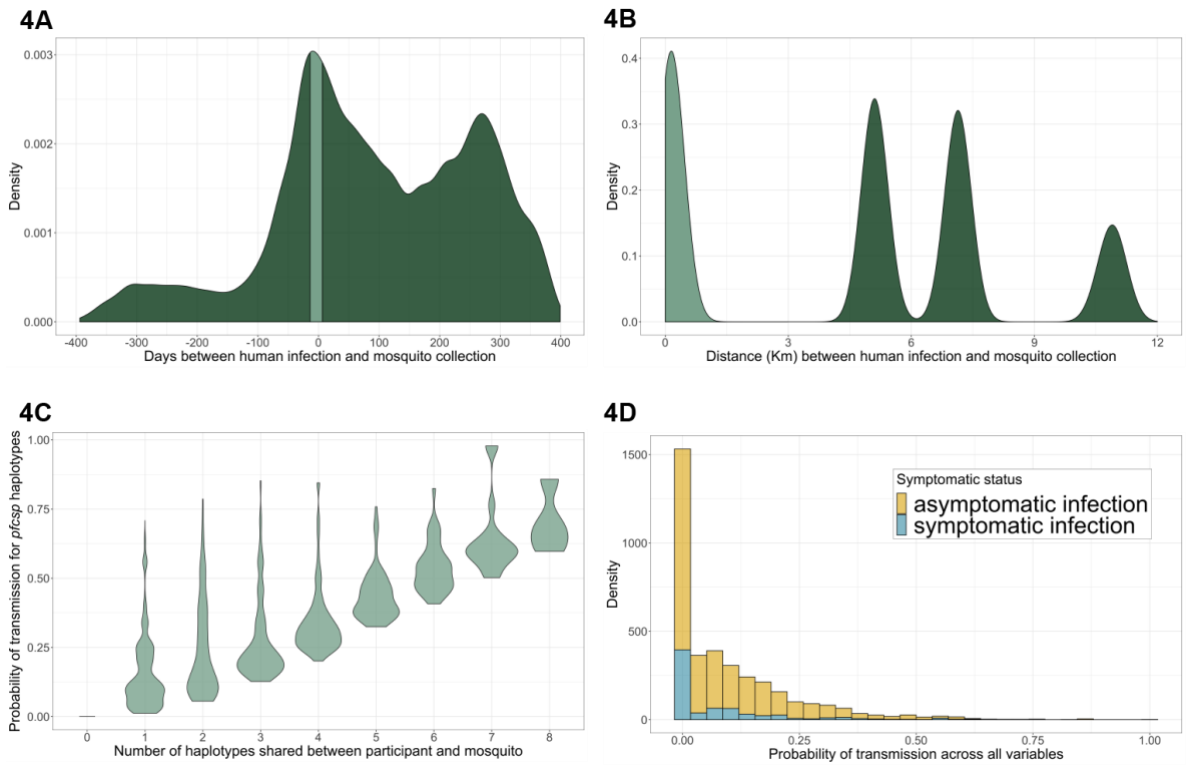
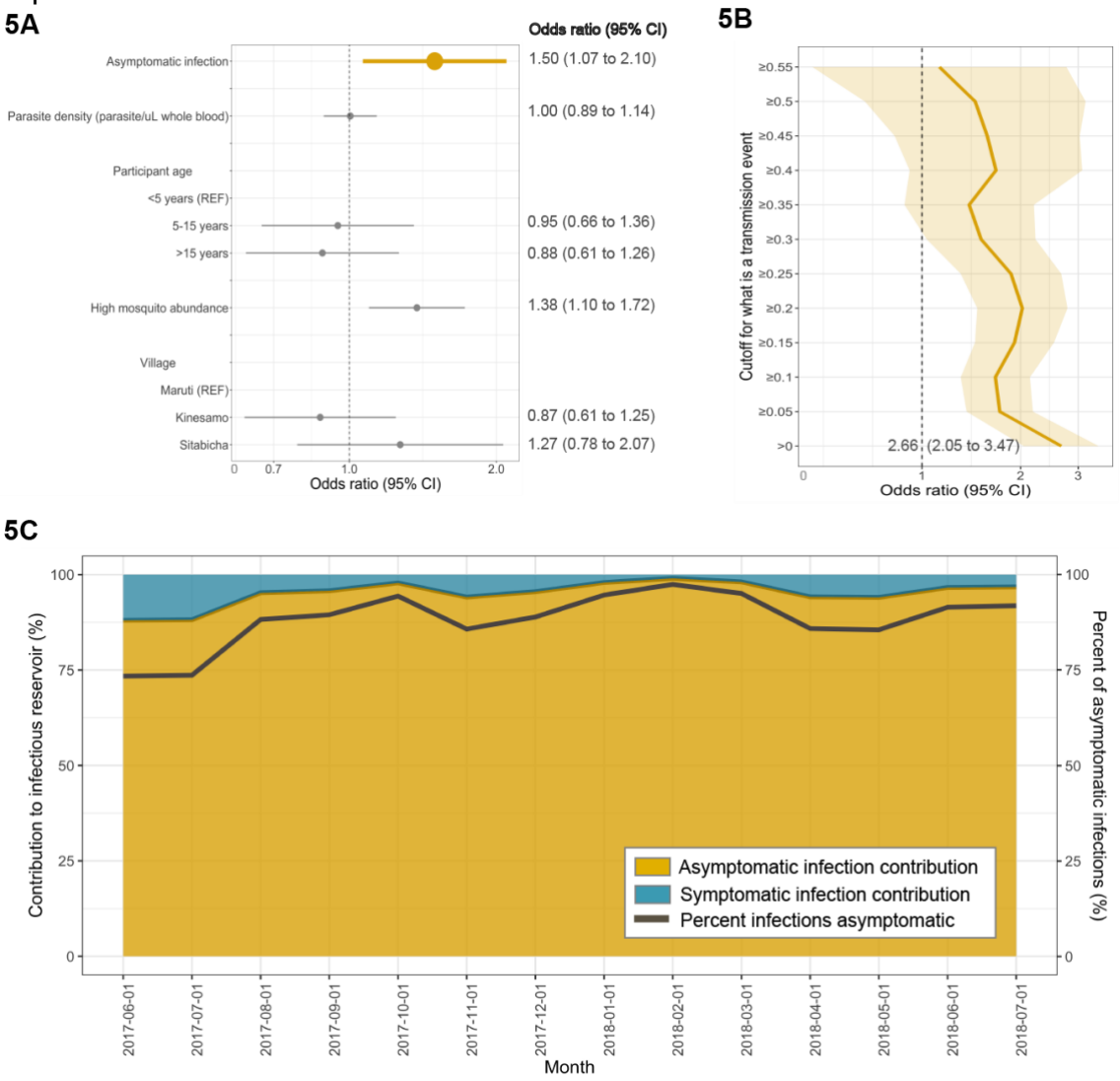


Figure 6.5. Multi-level logistic regression results for the odds of a participant-to-mosquito malaria transmission event from participants with asymptomatic compared to symptomatic infections using the *pfcsp* gene target.
6.5A: Odds ratios (ORs) of the probability of malaria transmission events from infected participants to mosquitoes. ORs were computed using a multi-level logistic regression model with the probability of transmission outcome coded continuously. Values above 1 indicate a factor that is associated with a greater likelihood of transmission of parasites to a mosquito, while values below 1 indicate a lesser probability. **6.5B:** ORs of the probability of transmission from infected participants to mosquitoes were re-estimated using multi-level logistic regression models with the outcome coded dichotomously. Models were computed iteratively by dichotomizing the probability of transmission at increasing values from 0.00 to 0.55, thereby increasing the stringency of the definition of a transmission event. The dark yellow line indicates the OR at each dichotomized level of the probability outcome, and the shaded area indicates the 95% confidence interval around each OR. **6.5C:** The contribution to the infectious reservoir was calculated using the odds of transmission to mosquitoes from participants with asymptomatic compared to symptomatic infections each month.



CHAPTER VII: CONCLUSIONS

Summary of Findings

This dissertation study was motivated by the sustained high *P. falciparum* malaria burden in many parts of sub-Saharan Africa in spite of increased transmission reduction efforts. We aimed to better understand the natural history of asymptomatic malaria and its potential role as a reservoir for sustained transmission; to do this, we investigated asymptomatic malaria's influence on attainment of future symptomatic infections at the individual and genomic levels as well as its transmission to mosquitoes. We hypothesized that prior asymptomatic infections would increase the short-term hazard of symptomatic infection, and similarly, new infections would increase the odds of developing symptomaticity. At the population level, we anticipated asymptomatic malaria would be a larger source of transmission to mosquitoes compared to symptomatic illness. These genomic, individual, and population-level effects were studied in a region in Western Kenya with high, perennial malaria transmission. Asymptomatic infections were common in our study site, highlighting the need to learn more about its influence on symptomaticity risk and its role as a reservoir.

The study used a longitudinal cohort of participants followed from June 2017 to November 2019 in Webuye, Kenya. Monthly, dried blood spots were collected from participants, and, weekly, mosquitoes collected from their households. We combined amplicon deep sequencing, probabilistic and multi-level modelling, and time-to-event

analysis methods to investigate the individual and genomic-level effects of asymptomatic infections acquired over time on symptomaticity as well as the population-level effects of asymptomatic malaria's transmission to mosquitoes.

At the individual level in aim 1A, we found, compared to being uninfected at monthly visits, asymptomatic infections greatly increased the short-term, 1-month hazard of symptomatic malaria, having a hazard 2.6 times that when uninfected; this association was similar when follow-up was expanded to 3 and 6 months but greatly weakened when following participants for 12 months or more.

At the genomic level in aim 1B, we observed that, compared to infections with haplotypes a person had been infected with before, incident infections with only new haplotypes increased odds of symptomatic malaria over 3-fold; however, this increased risk of symptomatic disease with new parasite haplotypes was attenuated when new haplotypes were mixed with recurrent ones or when people with persistent infections acquired new or recurrent parasite haplotypes.

At the population level in aim 2, we identified asymptomatic infections as the major source of onward malaria transmission, with asymptomatic infections having 2.6 times the odds of probable transmission to mosquitoes compared to symptomatic infections. Additionally, asymptomatic infections were the likely source of almost 95% of mosquito infections in the study site.

The dissertation findings suggest that, in high transmission areas, asymptomatic infections highly contribute to both future short-term symptomatic malaria risk as well as sustained malaria transmission. Taken together, this research provides a rationale for targeting asymptomatic infections for both medical and public health motives, as

reducing asymptomatic infections is expected to lessen the risk of symptomatic illness for individuals and create outsized transmission reduction for the population.

Strengths and Limitations

Specific strengths and limitations of each aim are described below across five main topics inherent in epidemiology training: confounding, measurement, missingness, selection, and generalizability.

Confounding

The minimally sufficient adjustment set of confounding covariates for each aim was chosen using a DAG analysis, with no unmeasured confounders identified. For causal inference analyses, there is a large assumption that there is no unmeasured confounding, and, thus, conditional exchangeability, comparing the exposure categories. In aim 1A, by allowing participant exposure to vary over time, we assumed exchangeability between the exposed and unexposed groups; approximately 94% of the study population changed exposure status at least once during follow-up, so lack of exchangeability between groups in our study population was not a large concern. In aim 1B, all haplotype categories were observed across both asymptomatic and symptomatic infections, suggesting exchangeability across haplotype categories.

Measurement

In aim 1A, we used a qPCR assay with high sensitivity,[129] but it is possible some infections were not detected. Pre-symptomatic infections could have also been misclassified as asymptomatic infections in aims 1A and 1B; sensitivity analyses were performed to detect how these pre-symptomatic infections influenced study results. In aim 1B, misclassification could have also occurred between recurrent and persistent

haplotypes; to minimize this risk, we excluded infections where participants had a symptomatic infection, were prescribed antimalarials, and had another infection with persistent haplotypes within 30 days of the initial infection.

In aim 2, we had no direct measurement of gametocytes due to the method of sample collection and gDNA extraction; however, we adjusted models for asexual parasite density, which has been suggested as a proxy for gametocyte density.[158] For aims 1B and 2, many infections in participants and mosquitoes had low parasite densities, which increases the risk of haplotype false discovery.[124] To mitigate this risk, we enforced stringent haplotype censoring based on read quality and haplotype abundance consistent with prior studies.[70,124,188]

Missingness

Across all three aims, asymptomatic infections were only captured at monthly follow-up visits, missing asymptomatic infections that arose and cleared between visits. We did not observe malaria infections participants acquired prior to the study but approximated previous malaria exposure by including participant age in the models in aims 1A, 1B, and 2. Additionally, in regards to aim 1B, misclassifying haplotypes as new when they had been acquired prior to the study would have biased results towards the null.

In regards to aims 1B and 2, while amplicon deep sequencing was a sensitive method for identifying different malaria infections,[110] it might not have captured all genetically distinct infections that occurred during the study. To account for this, we compared results across two unlinked parasite gene targets, *pfama1* and *pfcsp*. Missing data bias due to sequencing failure was also investigated for aims 1B and 2, with there

being a statistically significant correlation between lower parasite density and sequencing failure; however, a DAG representing missingness due to sequencing failure indicated that restricting the data set to samples that passed sequencing was unlikely to produce missing data bias (**Figure S6.4**). As a precaution to account for lower parasite density samples potentially being biased towards sequencing failure and a form of missing at random bias, we included a covariate in our models for parasite density.

In aim 2, we only measured human-to-mosquito malaria transmission directly within households, and could not capture events occurring in other settings. This limitation is mitigated by the known nocturnal feeding preference of the female *Anopheles* mosquitoes in this area.

Selection

Across all aims, participants were chosen using radial sampling of households across three villages with similar high malaria transmission. In aim 2, we may have under-detected asymptomatic infections and therefore over-represented symptomatic infections, owing either to the sparse monthly sampling for asymptomatic infections or the inability to capture transmission from symptomatic infections during their asymptomatic or pre-symptomatic phase. We expect that this would serve mainly to bias our analyses towards the null by providing relatively more opportunities for symptomatic infections to match to mosquitoes. Similarly, mosquito sampling was sparser than human sampling, which caused us to miss some transmission events and allowed for only relative estimations of onward transmission.

Generalizability

This dissertation study was based in a high malaria transmission region and included participants of all ages, and findings may be relevant to other areas of sub-Saharan Africa with high *P. falciparum* prevalence. It is important to note, however, in aim 1A, that the Hernán *et al.* method has limited transportability to other study sites with different infection dynamics.[109] Additionally, in aim 2, symptomatic infections were quickly diagnosed and treated with antimalarials under our protocol which likely reduced the duration of these infections and therefore limited their transmission potential. The access to diagnosis and treatment in our study is higher than is generally available across sub-Saharan Africa,[145] though recent reports indicate gradual improvement in quality clinical management.[187] Overall, this study demonstrates enhanced methodological approaches to study asymptomatic malaria at the genomic, individual, and population levels that could be applied to other study settings.

Public Health Implications and Future Directions

Broadly, our results highlight the major role asymptomatic *P. falciparum* plays in shaping malaria transmission dynamics in regions with high malaria prevalence. We found that asymptomatic malaria exposure over time increased the short-term hazard of symptomatic illness. We also observed an association between acquisition of new haplotypes in incident infections and increased odds of symptomatic disease, indicating the critical influence of parasite genetic diversity on symptomaticity. Additionally, these asymptomatic infections were large contributors to mosquito infection and onward malaria transmission.

Given the detrimental individual- and population-level health effects of

asymptomatic *P. falciparum* in our study site, expanded interventions to reduce the asymptomatic reservoir in high transmission settings could be beneficial. Examples include increasing access to insecticide-treated bed nets, ramping up administration of endectocides like ivermectin,[189] and expanding the use of targeted test and treat strategies for identifying and treating asymptomatic infections in addition to symptomatic ones.[34] Treating asymptomatic infections could have detrimental effects, such as increased drug resistance and risk of symptomatic malaria post-treatment,[18,190–194] possibly by limiting exposure to prior diverse infections. Reduced exposure would increase the likelihood that incident infections are composed of new haplotypes and therefore likely to manifest symptoms. However, if reduced transmission is accompanied by reductions in parasite genetic diversity as has been reported in several settings,[155,156] even with fewer prior infections the likelihood that a parasite will harbor recurrent haplotypes that attenuate symptoms would remain high. Precautions would still be needed, however, to limit drug resistance.

Future research could expand upon our work in several ways. We urge more studies to incorporate frequent longitudinal sampling of asymptomatic infections in people of all ages to enhance knowledge of time-to-symptomatic malaria in adults. Future work could also explore at the immunological level how infection with new haplotypes changes the risk of symptomatic malaria compared to infection with recurrent or persistent haplotypes; we were the first study to differentiate between these two types of previously seen haplotypes and were unable to directly assess immunological markers. Finally, high malaria transmission settings like ours rely on usage of insecticide-treated bed nets, rapid diagnostics for malaria, and treatment with

artemisinin-based combination therapies for malaria control.[34] The asymptomatic malaria burden has remained high in our study site despite adoption of all these methods, suggesting the need for enhanced interventions. Future studies could assess new ways to identify, treat, and reduce the asymptomatic malaria reservoir.

Using innovative methodological approaches, we learned more about the natural history of asymptomatic malaria at the genomic, individual, and population levels. In our high malaria transmission study site, we found that asymptomatic *P. falciparum* infections increased the short-term hazard of symptomatic illness and served as the primary source of infections in mosquito vectors. Results provide a clear rationale to target asymptomatic malaria in order to reduce the disease burden in highly-endemic settings.

APPENDIX

Supplementary Information for Chapter IV

Alternative Models Comparing Asymptomatic Malaria Exposure and Resulting Symptomatic Illness

As an alternative approach to assess how asymptomatic infection versus no infection affected the likelihood of having a symptomatic malaria infection, a multi-level logistic regression model was ran controlling for age (categorized: <5 years, 5-15 years, >15 years), sex (categorized: male, female), and regular bed net usage (categorized: averages > 5 nights a week sleeping under a bed net – yes, no) and including a random intercept at the participant level. Results suggested that the 30-day odds of symptomatic malaria in participants with asymptomatic infections during monthly visits was 2.70 (95% confidence interval (CI): 2.09 to 3.49) times the odds of those that were uninfected at those visits.

For an additional approach, we also reran the frailty Cox proportional hazards model described in **Equation 4.1** in the main text using a robust error estimator instead of random intercept to account for clustering at the participant-level. Results were similar to the original frailty model as well as the logistic regression, suggesting that people with asymptomatic infections had a hazard of symptomatic malaria 2.51 (95% CI: 1.89 to 3.34) times the hazard of people that were uninfected.

Symptomatic Malaria Primary Case Definition Effect Measure Modification Results

All effect measure modification results by age and sex are summarized in **Table S4.4**. Effect measure modification by age or sex was investigated for the 1-month effect of asymptomatic malaria exposure on the hazard of symptomatic illness. In the main

model, this relationship was not modified by age (*p*-value = 0.447 by log-likelihood ratio test): < 5 years [HR: 3.77, 95% CI: 2.02 to 7.04], 5-15 years [HR: 2.45, 95% CI: 1.79 to 3.35], and > 15 years [HR: 2.55, 95% CI: 1.57 to 4.15]. In contrast, the relationship was modified by sex (*p*-value = 0.006 by log-likelihood ratio test), with females having a stronger short-term hazard of symptomatic malaria [HR: 3.71, 95% CI: 2.62 to 5.24] after asymptomatic malaria exposure compared to males [HR: 1.76, 95% CI: 1.24 to 2.50] using the primary case definition for symptomatic malaria but not using the secondary stringent or permissive case definitions (**Figure S4.2**). In the pre-symptomatic malaria, this relationship was not modified by age (*p*-value = 0.507 by log-likelihood ratio test): < 5 years [HR: 2.85, 95% CI: 1.19 to 6.79], 5-15 years [HR: 1.61, 95% CI: 1.05 to 2.46], and > 15 years [HR: 1.90, 95% CI: 0.93 to 3.86] or sex (*p*-value = 0.094 by log-likelihood ratio test): males [HR: 1.24, 95% CI: 0.75 to 2.05] and females [HR: 2.34, 95% CI: 1.47 to 3.71]. In a post-treatment analysis assessing infections at least 14 days post-antimalarial treatment, the relationship between asymptomatic malaria exposure and the 1-month hazard of symptomatic illness was not modified by participant age (*p*-value = 0.864 by log-likelihood ratio test), with similar hazard ratios across age categories: < 5 years [HR: 2.06, 95% CI: 0.87 to 4.85], 5-15 years [HR: 2.61, 95% CI: 1.68 to 4.06], and > 15 years [HR: 2.96, 95% CI: 1.09 to 8.04]. No effect measure modification was observed by participant sex (*p*-value = 0.416 by log-likelihood test): males [HR: 2.11, 95% CI: 1.27 to 3.51] and females [HR: 3.00, 95% CI: 1.77 to 5.08].

For the 3-month hazard of symptomatic malaria, effect measure modification was not observed by participant age (*p*-value = 0.128 by log-likelihood ratio test). Adjusted 3-

month hazard ratios were similar across participant age categories: < 5 years [HR: 2.47, 95% CI: 1.59 to 3.84], 5-15 years [HR: 1.49, 95 CI: 1.21 to 1.85], and > 15 years [HR: 1.69, 95% CI: 1.23 to 2.32]. The 3-month hazard of symptomatic malaria was modified by sex (*p*-value = 0.009 by log-likelihood ratio test) with females [HR: 2.03, 95% CI: 1.62 to 2.55] having a stronger relationship than males [HR: 1.29, 95% CI: 1.01 to 1.64]. In the pre-symptomatic analysis, effect measure modification by age was not observed (*p*-value = 0.164 by log-likelihood ratio test): < 5 years [HR: 2.00, 95% CI: 1.20 to 3.34], 5-15 years [HR: 1.16, 95% CI: 0.90 to 1.48], > 15 years [HR: 1.35, 95% CI: 0.94 to 1.93]. Modification by sex was not observed (p-value = 0.064 by log-likelihood ratio test): females [HR: 1.52, 95% CI: 1.18 to 1.97] and males [HR: 1.05, 95% CI: 0.79 to 1.39].

For the 6-month hazard of symptomatic malaria, there was no effect measure modification by age (*p*-value = 0.197 by log-likelihood ratio test): < 5 years [HR: 1.94, 95% CI: 1.34 to 2.80], 5-15 years [HR: 1.32, 95% CI: 1.11 to 1.57], > 15 years [95% CI: 1.31, 95% CI: 1.01 to 1.70]. Modification was present by sex (*p*-value = 0.013 by log-likelihood ratio test) with females [HR: 1.62, 95% CI: 1.35 to 1.94] having a stronger association than males [HR: 1.13, 95% CI: 0.93 to 1.39]. In the pre-symptomatic analysis, age did not appear to modify the relationship between asymptomatic malaria exposure and symptomatic illness (*p*-value = 0.210 by log-likelihood ratio test): < 5 years [HR: 1.63, 95% CI: 1.08 to 2.46], 5-15 years [HR: 1.11, 95% CI: 0.92 to 1.34], and > 15 years [HR: 1.10, 95% CI: 0.83 to 1.46]. Slight modification was observed by sex (*p*-value = 0.050 by log-likelihood ratio test): females [HR: 1.33, 95% CI: 1.09 to 1.62] and males [HR: 0.98, 95% CI: 0.78 to 1.22].

For the 12-month hazard of symptomatic malaria, no effect measure modification by age was observed (*p*-value = 0.264 by log-likelihood ratio test). Age-stratified models could not be calculated due to data sparsity. Effect measure modification by sex was not observed (*p*-value = 0.122 by log-likelihood ratio test): females [HR: 1.21, 95% CI: 1.05 to 1.41] and males [HR: 1.10, 95% CI: 0.86 to 1.19]. In the pre-symptomatic analysis, age did not modify the relationship between asymptomatic malaria exposure and the hazard of symptomatic malaria (*p*-value = 0.201 by log-likelihood ratio test): < 5 years [HR: 1.24, 95% CI: 0.88 to 1.74], 5-15 years [HR: 1.00, 95% CI: 0.86 to 1.17], > 15 years [HR: 0.85, 95% CI: 0.68 to 1.07]. No modification was seen by sex (*p*-value = 0.364 by log-likelihood ratio test) with similar hazard ratios across females [HR: 1.04, 95% CI: 0.88 to 1.22] and males [HR: 0.91, 95% CI: 0.77 to 1.09].

The effect of asymptomatic malaria infection versus no infection on the 29-month hazard of symptomatic malaria was modified by participant age (*p*-value < 0.001 by log-likelihood ratio test) with the strongest association in children < 5 years [HR: 1.38, 95% CI: 1.05 to 1.81], second-strongest in children 5-15 years [HR: 1.16, 95% CI: 1.02 to 1.32], and weakest in adults > 15 years [HR: 0.96, 95% CI: 0.81 to 1.13] (**Figure S4.3**). No effect measure modification was observed by sex (*p*-value = 0.378 by log-likelihood ratio test) with males [HR: 1.08, 95% CI: 0.94 to 1.24] and females [HR: 1.14, 95% CI: 1.01 to 1.30] having similar hazard ratios for symptomatic malaria. In contrast to the main model results, the pre-symptomatic analysis model found no relationship between exposure to asymptomatic malaria compared to no infection and the hazard of symptomatic malaria [HR: 1.02, 95% CI: 0.92 to 1.12]. Effect measure modification by age was observed (*p*-value <0.001 by log-likelihood ratio test): < 5 years [HR: 1.23,

95% CI: 0.92 to 1.64], 5-15 years [HR: 1.06, 95% CI: 0.93 to 1.21], and >15 years [HR: 0.88, 95% CI: 0.74 to 1.05. Effect measure modification by sex could not be assessed due to data sparsity.

Pre-symptomatic Sensitivity Analysis Across Longer Follow-up Periods

A pre-symptomatic analysis was conducted to assess potential bias caused by misclassifying pre-symptomatic infections as asymptomatic at monthly follow-up visits. Similar to the 1-month analysis, for the 3-month follow-up the pre-symptomatic analysis found an increased hazard of symptomatic malaria within 3 months when a participant had an asymptomatic infection compared to being uninfected [HR: 1.28, 95% CI: 1.06 to 1.55]. In contrast to the main model results, there was no association between asymptomatic infections and the hazard of symptomatic malaria when expanding follow-up to 6 [HR: 1.16, 95% CI: 1.00 to 1.34], 12 [HR: 0.98, 95% CI: 0.87 to 1.10], and 29 [HR: 1.02, 95% CI: 0.92 to 1.12] months. Effect measure modification by age and sex in the subset analyses was similar to the full analyses and recorded in **Table S4.4**.

Symptomatic Malaria Secondary Permissive Case Definition Results

As a sensitivity analysis for defining symptomatic malaria, the 1-month and 29-month analyses were repeated using a secondary (permissive) case definition for symptomatic malaria. The secondary permissive case definition defined a symptomatic infection as one where at a participant had at least one symptom consistent with malaria during a sick visit and was *P. falciparum* positive by real-time PCR (qPCR). Under this case definition, there was a total of 5380 monthly follow-up visits with 1837 (34.2%) of visits indicating asymptomatic malaria exposure. Using the secondary permissive case definition, a total of 409 symptomatic infections occurred. Participants had a median of 1

(IQR: 0, 2) symptomatic infections during follow-up. Median time to symptomatic malaria when exposed to asymptomatic infections (137, IQR: 41, 308) was lower than when unexposed (190, IQR: 80, 333) (**Table S4.2**).

For the 1-month effect of asymptomatic malaria exposure, a multivariate frailty Cox proportional hazards model observed that asymptomatic infection exposure increased the short-term hazard of symptomatic malaria [HR: 1.97, 95% CI: 1.63 to 2.40] (**Figure S4.2**). This relationship was not modified by participant age (*p*-value = 0.482 by log-likelihood ratio test): < 5 years [HR: 2.27, 95% CI: 1.37 to 3.74], 5-15 years [HR: 2.09, 95% CI: 1.58 to 2.77], and > 15 years [HR: 1.77, 95% CI: 1.29 to 2.44]. The relationship was also not modified by sex (*p*-value = 0.293 by log-likelihood ratio test) with males [1.73, 95% CI: 1.30 to 2.32] and females [HR: 2.18, 95% CI: 1.68 to 2.83] having similar hazard ratios for symptomatic malaria (**Figure S4.2**).

Testing the 29-month effect of asymptomatic malaria exposure, a frailty Cox proportional hazards model controlling for age, sex, bed net usage, and village found that exposure to asymptomatic infections over time had no relationship with the long-term hazard of symptomatic malaria compared to having no malaria infections [Hazard ratio (HR): 1.20, 95% CI: 1.11 to 1.31] (**Figure S4.3**). Effect measure modification of this relationship was not observed by participant age (*p*-value = 0.494 by log-likelihood ratio test) with the long-term hazard of symptomatic malaria similar among children < 5 years [HR: 1.14, 95% CI: 0.90 to 1.44], children 5-15 years [HR: 1.23, 95% CI: 1.09 to 1.38], and adults > 15 years [HR: 1.19, 95% CI: 1.04 to 1.36] (**Figure S4.3**). Effect measure modification was not observed by sex (*p*-value = 0.159 by log-likelihood ratio test)

comparing hazard ratios across males [HR: 1.12, 95% CI: 0.99 to 1.26] and females [HR: 1.26, 95% CI: 1.13 to 1.41] (**Figure S4.3**).

Symptomatic Malaria Secondary Stringent Case Definition Results

As an additional sensitivity analysis for defining symptomatic malaria, the 1-month and 29-month analyses were repeated using a secondary (stringent) case definition for symptomatic malaria. The secondary stringent case definition defined a symptomatic infection as one where at a participant had a self-reported fever during a sick visit and was *P. falciparum* positive by both RDT and qPCR. Under this case definition, there was a total of 5374 monthly follow-up visits with 1848 (34.3%) of visits indicating asymptomatic malaria exposure. Using the secondary stringent case definition, a total of 215 symptomatic infections occurred. Participants had a median of 0 (IQR: 0, 1) symptomatic infections during follow-up. Median time to symptomatic malaria when exposed to asymptomatic infections (203, IQR: 52, 429) was lower than when unexposed (232, IQR: 104, 403) (**Table S4.3**). Median follow-up time to symptomatic malaria was also shorter when participants were living in the village Maruti (**Table S4.3**).

For the 1-month of asymptomatic malaria exposure, a multivariate frailty Cox proportional hazards model observed that asymptomatic infection exposure increased the short-term hazard of symptomatic malaria [HR: 2.76, 95% CI: 2.11 to 3.62] (**Figure S4.2**). This relationship was not modified by participant age (*p*-value = 0.438 by log-likelihood ratio test): < 5 years [HR: 3.94, 95% CI: 2.09 to 7.43], 5-15 years [HR: 2.64, 95% CI: 1.88 to 3.72], > 15 years [HR: 2.47, 95% CI: 1.36 to 4.49]. Sex did not modify this relationship either (*p*-value = 0.061 by log-likelihood ratio test): males [HR: 2.05, 95% CI: 1.39 to 3.02] and females [HR: 3.60, 95% CI: 2.46 to 5.28] (**Figure S4.2**).

Testing the 29-month effect of asymptomatic malaria exposure, a frailty Cox proportional hazards model controlling for age, sex, bed net usage, and village found that exposure to asymptomatic infections over time did not affect the long-term hazard of symptomatic malaria compared to having no malaria infections [Hazard ratio (HR): 1.02, 95% CI: 0.92 to 1.13] (**Figure S4.3**). This relationship was modified by participant age (*p*-value < 0.001 by log-likelihood ratio test) with the strongest association in children < 5 years [HR: 1.38, 95% CI: 1.05 to 1.81], second-strongest in children 5-15 years [HR: 1.10, 95% CI: 0.96 to 1.25], and weakest in adults > 15 years [HR: 0.73, 95% CI: 0.59 to 0.90] (**Figure S4.3**). Effect measure modification could not be assessed by sex due to data sparsity.

Table S4.1 Comparison of time-varying exposure coding approaches

Exposure coding method	Description	Rationale for including/excluding
Intention-to-treat	Take exposure status at baseline and apply it over the full follow-up period	Can misclassify person-time if exposure frequently changes over time, as happens with the exposure in our study; this was commonly done in previous time to symptomatic malaria studies[14–19,21–23]
Allow participants to change exposure group over follow-up	Exposure is reassessed for participants over time and summarized as the number of months exposed	Can have issues with left truncation bias for exposures that began before the study, as occurred in our study where participants could have been infected with asymptomatic malaria at baseline
Ever-never approach	Classify participant as exposed if were ever exposed during follow-up period	Many issues with misclassification and "look-back" bias; Buchwald *et al.*[20] did a modified version of this where participants were classified as unexposed until an asymptomatic infection occurred then classified as exposed for the remaining period afterward
Hernán *et al.* multiple month method[109]	Modified version of intention-to-treat where each month was treated as a baseline for follow-up; The exposure status of each monthly visit was applied to the subsequent follow-up period	Allows exposure to change over time with more precision than the typical intention-to-treat approach; produces effect estimate that is predictive of future risk regardless of prior exposure so not prone to left truncation bias; some misclassification bias still possible but less than alternative methods

Table S4.2. Covariate distribution across symptomatic events: secondary permissive case definition

	Total person-months* (N, %)	**Person-months ending in symptomatic infections**** (N, %)	**Median time to symptoms** (days, IQR)	***P*-value**
Main exposure				<0.001[a]
No infection	3537 (65.8)	2122 (67.0)	190 (80, 333)	
Asymptomatic infection	1837 (34.2)	1044 (33.0)	137 (41, 308)	
Age				1.000[b]
< 5 years	806 (15.0)	419 (13.2)	182 (57, 345)	
5-15 years	2280 (42.4)	1556 (49.1)	174 (64, 337)	
> 15 years	2288 (42.6)	1191 (37.6)	169 (64, 310)	
Sex				0.133[a]
Male	2374 (44.2)	1468 (46.4)	186 (67, 349)	
Female	3000 (55.8)	1698 (53.6)	163 (61, 308)	
Regular bed net usage[#]				1.000[a]
No	1445 (26.9)	876 (27.7)	179 (67, 342)	
Yes	3929 (73.1)	2290 (72.3)	171 (63, 320)	
Village				1.000[b]
Kinesamo	1846 (34.4)	1073 (33.9)	181 (65, 319)	
Maruti	1669 (31.1)	1013 (32.0)	159 (60, 314)	
Sitabicha	1859 (34.6)	1080 (34.1)	182 (68, 348)	

Abbreviations: IQR, interquartile range

[#]Regular bed net usage was defined as a person averaging > 5 nights a week sleeping under a bed net.

*Total person-months indicates the total number of monthly follow-up visits ending in a symptomatic infection or censoring.

**Symptomatic infections were defined using the secondary permissive case definition where a participant was *P. falciparum*-positive by qPCR as well as had at least one symptom consistent with malaria during a sick visit.

[a] Wilcoxon Rank Sum test with continuity correction and Bonferroni correction for repeated measures.

[b] Kruskal-Wallis test with Bonferroni correction for repeated measures.

Table S4.3. Covariate distribution across symptomatic events: secondary stringent case definition

	Total person-months visits* (N, %)	**Person-months ending in symptomatic infections**** (N, %)	**Median time to symptoms** (days, IQR)	***P*-value**
Main exposure				0.018[a]
No infection	3532 (65.7)	1311 (65.6)	232 (104, 403)	
Asymptomatic infection	1848 (34.3)	687 (34.4)	203 (52, 429)	
Age				0.106[b]
< 5 years	812 (15.1)	329 (16.5)	226 (82, 435)	
5-15 years	2286 (42.5)	1161 (58.1)	209 (78, 389)	
> 15 years	2282 (42.4)	508 (25.4)	254 (103, 459)	
Sex				1.000[a]
Male	2355 (43.8)	983 (49.2)	236 (88, 436)	
Female	3025 (56.2)	1015 (50.8)	210 (82, 398)	
Regular bed net usage[#]				1.000[a]
No	1427 (26.5)	645 (32.3)	210 (84, 380)	
Yes	3953 (73.5)	1353 (67.7)	233 (86, 432)	
Village				0.032[b]
Kinesamo	1853 (34.4)	695 (34.8)	241 (94, 440)	
Maruti	1680 (31.2)	643 (32.2)	186 (66, 376)	
Sitabicha	1847 (34.3)	660 (33.0)	238 (91, 427)	

Abbreviations: IQR, interquartile range
[#]Regular bed net usage was defined as a person averaging > 5 nights a week sleeping under a bed net.
*Total person-months indicates the total number of monthly follow-up visits ending in a symptomatic infection or censoring.
**Symptomatic infections were defined using the secondary stringent case definition where a participant was *P. falciparum*-positive by both RDT and qPCR as well as had a self-reported fever during a sick visit.
[a] Wilcoxon Rank Sum test with continuity correction and Bonferroni correction for repeated measures.
[b] Kruskal-Wallis test with Bonferroni correction for repeated measures.

Table S4.4. Age- and sex-stratified hazard ratios of time to symptomatic malaria

Comparison	Age HR (95% CI)			Sex HR (95% CI)	
	<5 years	5-15 years	>15 years	Male	Female
1-month main	3.77 (2.02,7.04)	2.45 (1.79,3.35)	2.55 (1.57,4.15)	**1.76 (1.24,2.50)**	**3.71 (2.62,5.24)**
1-month pre-symptomatic	2.85 (1.19,6.79)	1.61 (1.05,2.46)	1.90 (0.93,3.86)	1.24 (0.75,2.05)	2.34 (1.47,3.71)
1-month post-treatment	2.06 (0.87,4.85)	2.61 (1.68,4.06)	2.96 (1.09,8.04)	2.11 (1.27,3.51)	3.00 (1.77,5.08)
3-month main	2.47 (1.59,3.84)	1.49 (1.21,1.85)	1.69 (1.23,2.32)	**1.29 (1.01,1.64)**	**2.03 (1.62,2.55)**
3-month pre-symptomatic	2.00 (1.20,3.34)	1.16 (0.90,1.48)	1.35 (0.94,1.93)	1.05 (0.79,1.39)	1.52 (1.18,1.97)
6-month main	1.94 (1.34,2.80)	1.32 (1.11,1.57)	1.31 (1.01,1.70)	**1.13 (0.93,1.39)**	**1.62 (1.35,1.94)**
6-month pre-symptomatic	1.63 (1.08,2.46)	1.11 (0.92,1.34)	1.10 (0.83,1.46)	0.98 (0.78,1.22)	1.33 (1.09,1.62)
12-month main	Not calculated*	Not calculated*	Not calculated*	1.10 (0.86,1.19)	1.21 (1.05,1.41)
12-month pre-symptomatic	1.24 (0.88,1.74)	1.00 (0.86,1.17)	0.85 (0.68,1.07)	0.91 (0.77,1.09)	1.04 (0.88,1.22)
29-month main	**1.38 (1.05,1.81)**	**1.16 (1.02,1.32)**	**0.96 (0.81,1.13)**	1.08 (0.94,1.24)	1.14 (1.01,1.30)
29-month pre-symptomatic	**1.23 (0.92,1.64)**	**1.06 (0.93,1.21)**	**0.88 (0.74,1.05)**	Not calculated*	Not calculated*

Abbreviations: CI, confidence interval; HR, adjusted hazard ratio
*Not calculated due to data sparsity
Statistically significant effect measure modification by the log-likelihood ratio test is bolded.

Table S4.5. Covariate distribution across symptomatic events comparing main analysis and post-treatment analysis datasets

	Main analysis Person-months ending in symptomatic infections* (N, %)	**Post-treatment analysis** Person-months ending in symptomatic infections* (N, %)	***P*-value comparing datasets[a]**
Main exposure			1.000
No infection	1580 (65.7)	570 (68.3)	-
Asymptomatic infection	826 (34.3)	264 (31.7)	-
Age			<0.001
< 5 years	329 (13.7)	63 (7.6)	-
5-15 years	1319 (54.8)	617 (74.0)	-
> 15 years	758 (31.5)	154 (18.5)	-
Sex			0.004
Male	1190 (49.5)	348 (41.7)	-
Female	1216 (50.5)	486 (58.3)	-
Regular bed net usage[#]			1.000
No	730 (30.3)	257 (30.8)	-
Yes	1676 (69.7)	577 (69.2)	-
Village			0.350
Kinesamo	876 (36.4)	259 (31.1)	-
Maruti	745 (31.0)	294 (35.3)	-
Sitabicha	785 (32.6)	281 (33.7)	-

[#]Regular bed net usage was a person averaging > 5 nights a week sleeping under a bed net.
*Symptomatic infections were defined using the primary case definition where a participant was *P. falciparum*-positive by both RDT and qPCR as well as had at least one symptom consistent with malaria during a sick visit.
[a] Pearson's χ^2 test with Bonferroni correction for 29 infections.

Figure S4.1. DAG illustrating covariate relationships for the association between exposure to asymptomatic malaria versus no infection and time to symptomatic malaria infection.

The shaded grey boxes represent the main exposure and main outcome. The boxes outlined with dotted lines represent potential effect measure modifiers (sex and age).

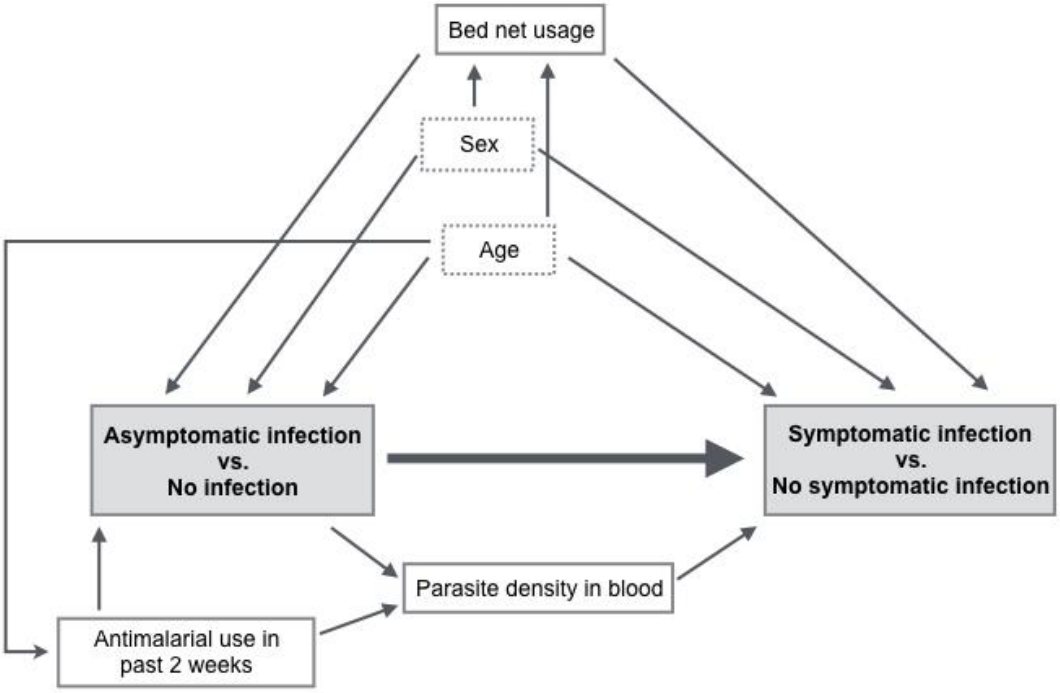


Figure S4.2. Frailty Cox proportional hazards model results comparing exposure to asymptomatic malaria infections versus no infection over time and <u>1-month</u> hazard of symptomatic malaria across the three case definitions for symptomatic malaria: primary, secondary permissive, and secondary stringent.
The main model controls for covariates participant age, sex, bed net usage, and village. The age-stratified model controls for covariates sex, bed net usage, and village. The sex-stratified model controls for participant age, bed net usage, and village. Statistically significant hazard ratios are bolded.

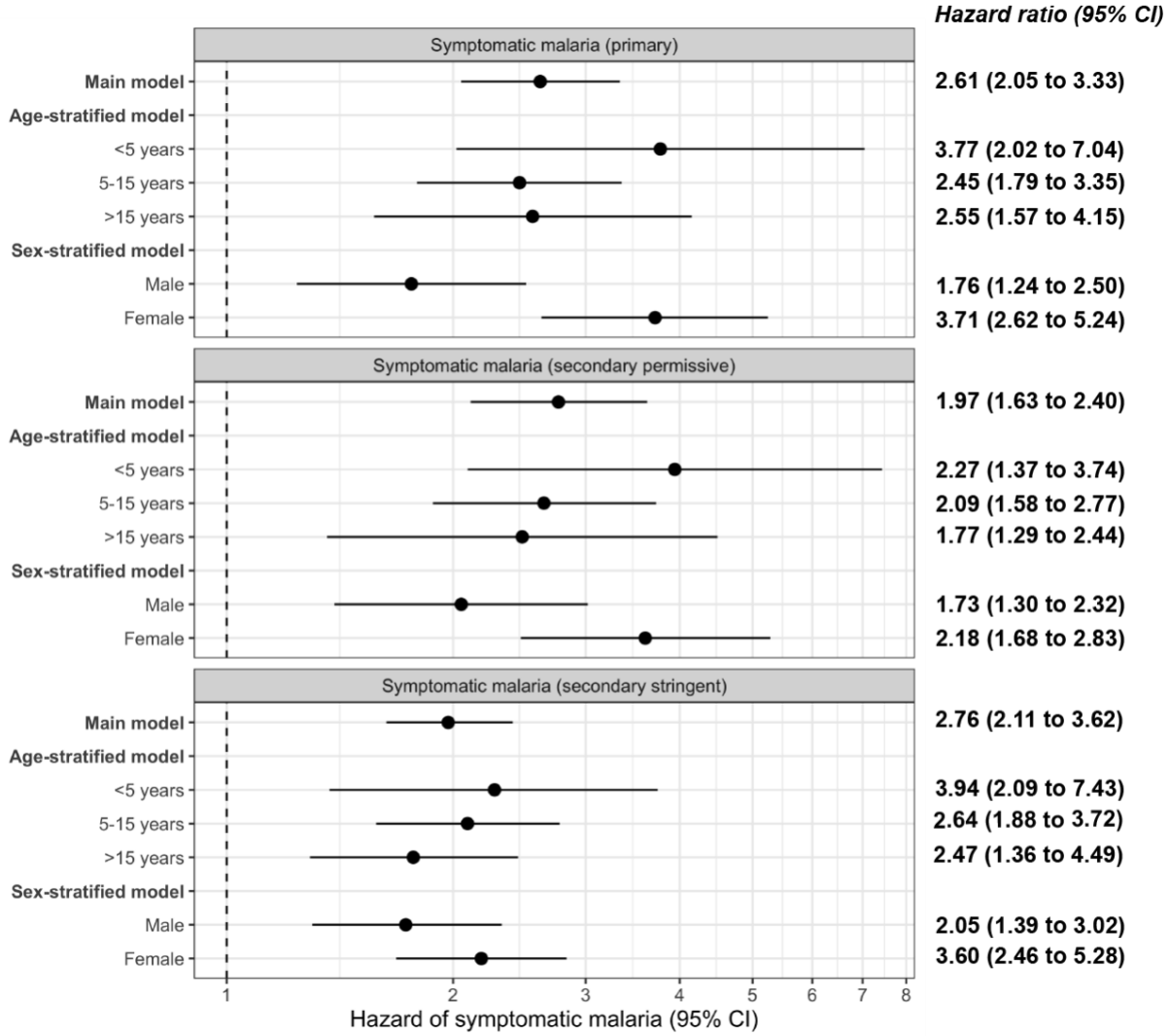
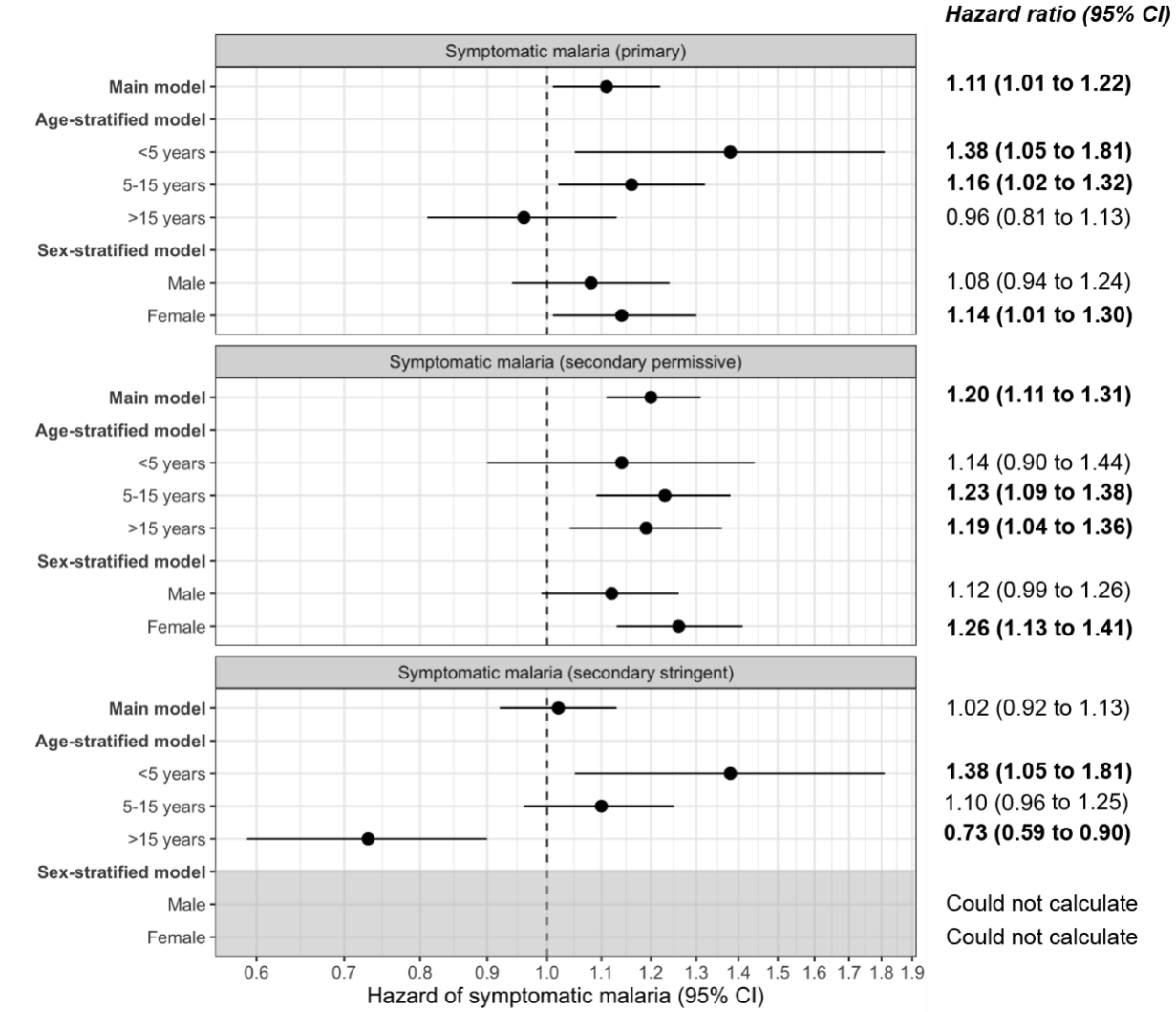


Figure S4.3. Frailty Cox proportional hazards model results comparing exposure to asymptomatic malaria infections versus no infection over time and 29-month hazard of symptomatic malaria across the three case definitions for symptomatic malaria: primary, secondary permissive, and secondary stringent.
The main model controls for covariates participant age, sex, bed net usage, and village. The age-stratified model controls for covariates sex, bed net usage, and village. The sex-stratified model controls for participant age, bed net usage, and village. Statistically significant hazard ratios are bolded.



Supplementary Information for Chapter V

Description of pfama1 Haplotype Results

For *pfama1,* we identified 193 haplotypes across 611 asymptomatic and 113 symptomatic infections in 204 participants. After censoring participant's initial infections and infections occurring within 14 days of antimalarial treatment, we observed 154 haplotypes in 430 asymptomatic and 72 symptomatic infections across 156 participants. Across all 502 infections, only new haplotypes were found in 194 infections, new and recurrent haplotypes observed in 91 infections, and only recurrent haplotypes identified in 60 infections (**Figure S5.2**). Persistent haplotypes were observed in 157 infections.

Assessing the potential for pre-symptomatic infections, persistent *pfama1* haplotypes (N=31) were not as commonly found in symptomatic infections as new or recurrent haplotypes (N=41); however, symptomatic infections with persistent *pfama1* haplotypes had a statistically significant shorter time between their most recent infection and symptomatic malaria (*p*-value <0.001 by Kruskal-Wallis χ^2 test), with most infections occurring within 12 days (range: 2 to 63) when persistent haplotypes were present as opposed to 50 (range: 7 to 327) when they were not (**Figure S5.3**).

We ran a multi-level logistic regression model on data restricted to only infections with new or recurrent *pfama1* haplotypes and controlling for within-individual random effects, the number of previous infections, transmission season, and age. Compared to infections with only recurrent haplotypes, model results found that infections with only new haplotypes [odds ratio (OR): 2.20, 95% confidence interval (CI): 0.66 to 7.40] or both new and recurrent haplotypes [OR: 1.21, 95% CI: 0.26 to 5.65] had higher odds of symptomatic malaria (**Figure S5.4**). Results were not statistically significant. Comparing

multi-level logistic regression models with and without interaction term for age indicated that the model accounting for effect measure modification by age was not a statistically significant better fit (*p*-value = 0.410 by log-likelihood ratio test). Due to data sparsity, age-stratified models were not able to be compared to assess effect measure modification by age using *pfama1* haplotypes.

Persistent *pfama1* haplotypes were identified in 157 infections categorized into: (i) only persistent (N=44); (ii) new and persistent (N=60); (iii) recurrent and persistent (N=21); and (iv) new, recurrent, and persistent (N=32). Across all four categories, the number of days since the previous infection ranged from 2 to 65, with most infections occurring within 28 days (**Figure S5.5**). No statistically significant differences were observed in the time since previous infection across haplotype categories (*p*-value = 0.432 by Kruskal-Wallis χ^2 test).

To test how presence of the persistent haplotypes affected the odds of developing a symptomatic compared to asymptomatic malaria infection, we ran a multi-level logistic regression model adjusting for within-individual random effects, age, the number of prior malaria infections, and transmission season. After restricting the model to consecutive infections with persistent haplotypes occurring within 30 days, the model assessed 109 infections with new or persistent *pfama1* haplotypes across 81 asymptomatic and 28 symptomatic infections. Compared to infections with only persistent *pfama1* haplotypes, infections with mixed types of haplotypes [OR: 0.24, 95% CI: 0.06 to 1.00] had lower odds of symptomatic malaria (**Figure S5.6**). Due to small sample sizes, age-stratified models and the log-likelihood ratio test were unable to be performed to assess effect measure modification.

Table S5.1. Distribution of symptomatic status across covariates for infections with *pfama1* new and recurrent haplotypes

	Asymptomatic infections (N=304)	Symptomatic infections (N=41)	*P*-value
Haplotype category, N (%)			0.230[a]
Only new	163 (53.6)	31 (75.6)	
New and recurrent	85 (28.0)	6 (14.6)	
Only recurrent	56 (18.4)	4 (9.8)	
Age, N (%)			0.097[a]
≤ 15 years	187 (60.5)	34 (82.9)	
> 15 years	117 (38.5)	7 (17.1)	
Number of prior malaria infections[#], N (%)			0.468[a]
Low	221 (72.7)	36 (87.8)	
High	83 (27.3)	5 (12.2)	
Transmission season*, N (%)			0.302[a]
Low	196 (64.5)	19 (46.3)	
High	108 (35.5)	22 (53.7)	
Multiplicity of infection**, N (%)			1.000[a]
Low	180 (59.2)	29 (70.7)	
High	124 (40.8)	12 (29.3)	

Abbreviations: IQR, interquartile range; NE, not evaluated

[#] The number of prior malaria infections a participant had was categorized into low versus high. Low indicated 3 infections or fewer during the study period. High represented more than 3 infections during the study period.

*Transmission season was categorized into low versus high based on the mosquito abundance across the study site and malaria seasonality. The low transmission season was when ≤ 50 mosquitoes were collected in the two weeks prior. The high transmission season was when > 50 mosquitoes were collected in the two weeks prior to the person's infection.

**Multiplicity of infection (MOI) was categorized into low versus high based on a functional form assessment. Low MOI was ≤ 2 haplotypes, whereas high MOI was > 2 haplotypes.

[a] Pearson's χ^2 test with Bonferroni correction for repeated measures for 8 infections

Table S5.2. Distribution of symptomatic status across covariates for consecutive infections with *pfama1* persistent and new or recurrent haplotypes

	Asymptomatic infections (N=81)	Symptomatic infections (N=28)	*P*-value
Haplotype category, N (%)			<0.001[a]
Mixed types of haplotypes	67 (82.7)	12 (42.9)	
Only persistent haplotypes	14 (17.3)	16 (57.1)	
Age, N (%)			1.000[a]
≤ 15 years	58 (71.6)	21 (75.0)	
> 15 years	23 (28.4)	7 (25.0)	
Number of prior malaria infections[#], N (%)			0.249[a]
Low	44 (54.3)	22 (78.6)	
High	37 (45.7)	6 (21.4)	
Transmission season*, N (%)			0.919[a]
Low	52 (64.2)	13 (46.4)	
High	29 (35.8)	15 (53.6)	
Multiplicity of infection**, N (%)			0.006[a]
Low	27 (33.3)	20 (71.4)	
High	54 (66.7)	8 (28.6)	

Abbreviations: IQR, interquartile range; NE, not evaluated

[#] The number of prior malaria infections a participant had was categorized into low versus high. Low indicated 3 infections or fewer during the study period. High represented more than 3 infections during the study period.

*Transmission season was categorized into low versus high based on the mosquito abundance across the study site and malaria seasonality. The low transmission season was when ≤ 50 mosquitoes were collected in the two weeks prior. The high transmission season was when > 50 mosquitoes were collected in the two weeks prior to the person's infection.

**Multiplicity of infection (MOI) was categorized into low versus high based on a functional form assessment. Low MOI was ≤ 2 haplotypes, whereas high MOI was > 2 haplotypes.

[a] Pearson's χ^2 test with Bonferroni correction for repeated measures for 6 infections

Figure S5.1. DAG of relationship between malaria haplotype categories and a participant's symptomatic status.
The haplotype categories differed depending on if new haplotypes were being compared to recurrent or persistent haplotypes. The DAG indicated that participant age, number of prior malaria infections, malaria transmission season, and multiplicity of infection were confounders.

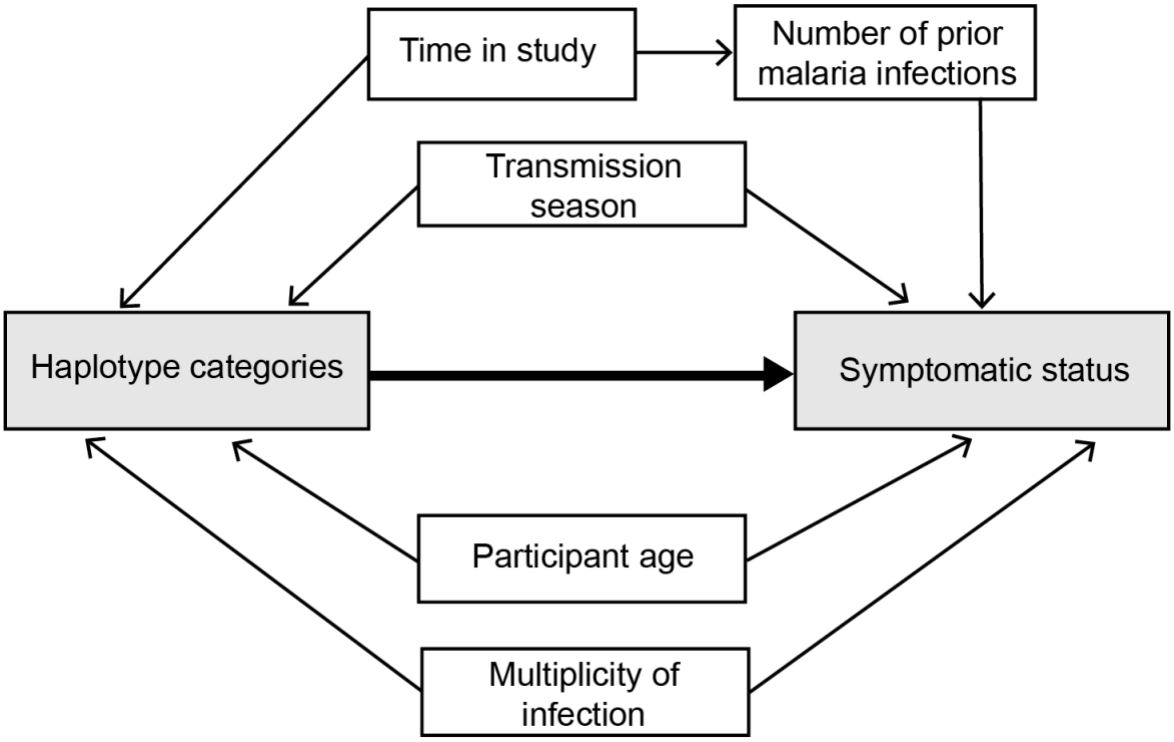


Figure S5.2. Venn diagram of *pfama1* haplotype categories across all malaria infections.
Each number indicates the number of infections that had haplotypes within each category: new, recurrent or persistent.

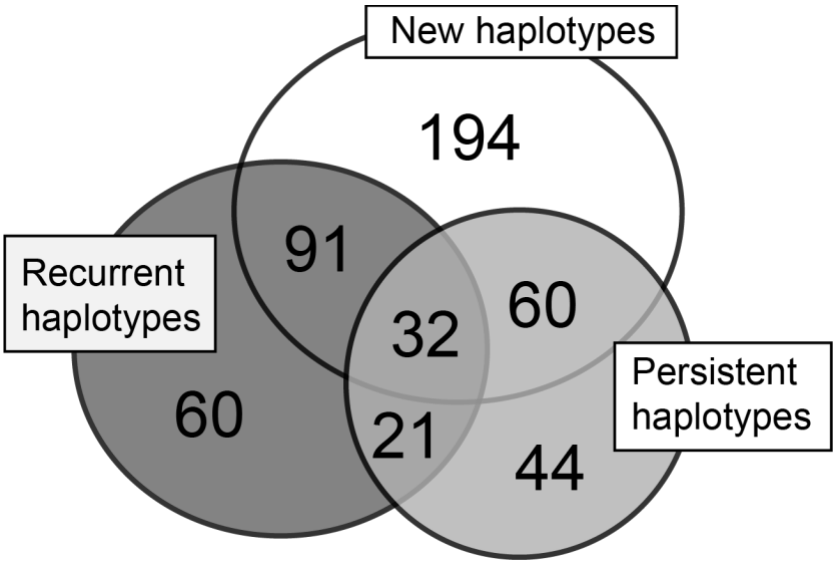


Figure S5.3. Comparison of time since previous infection for infection with or without persistent *pfama1* haplotypes.
Distribution of the number of days since previous infection for symptomatic malaria infection was stratified by whether or not the infection had persistent *pfama1* haplotypes or not.

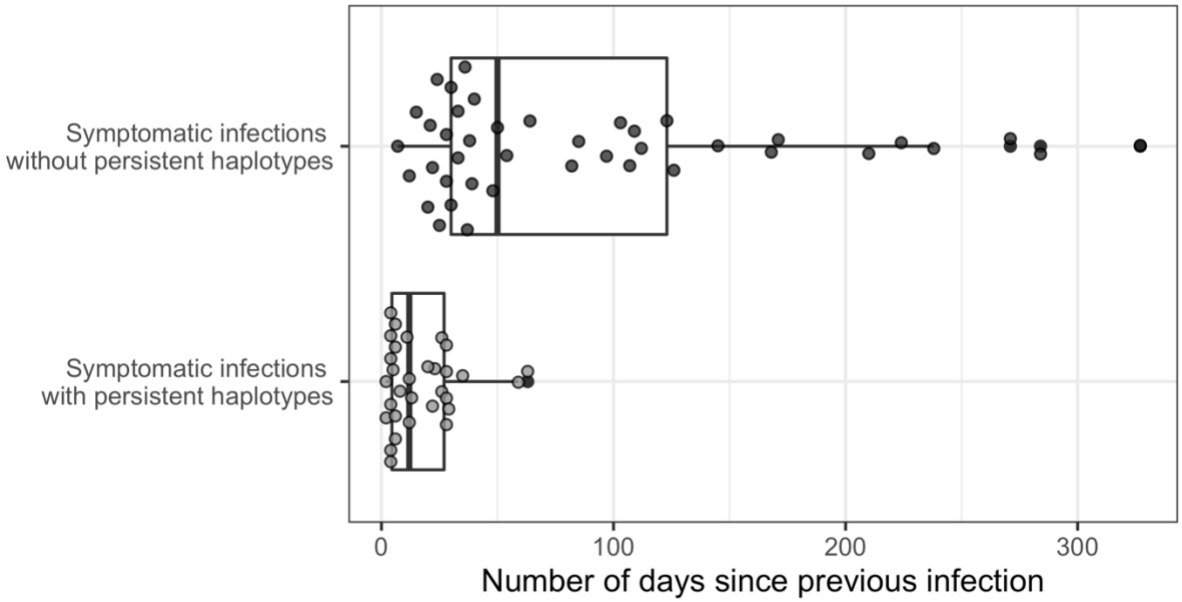


Figure S5.4. Odds of symptomatic malaria comparing infections with new versus recurrent *pfama1* haplotypes.
Multi-level logistic regression results for the odds of symptomatic malaria comparing (i) only new versus only recurrent (dark red) and (ii) new and recurrent versus only recurrent *pfama1* haplotypes (light red). Odds ratios are represented by the dots with the lines indicating the surrounding 95% confidence intervals.

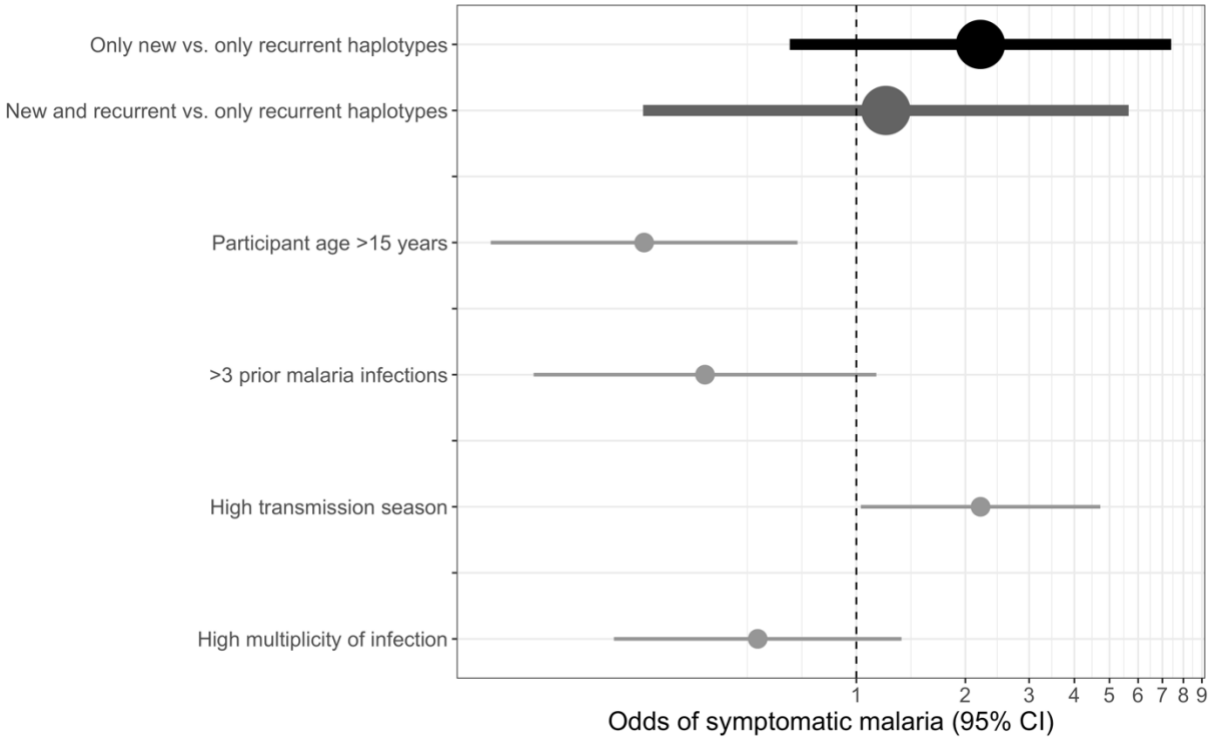


Figure S5.5. Distribution of days since previous infection for infections with persistent *pfama1* haplotypes.
Distribution of the number of days since previous infection for malaria infections with persistent *pfama1* haplotypes. Infections were categorized into: (i) only persistent; (ii) new and persistent; (iii) recurrent and persistent; and (iv) new, recurrent, and persistent. Asymptomatic infections were represented by circles and symptomatic ones by triangles.

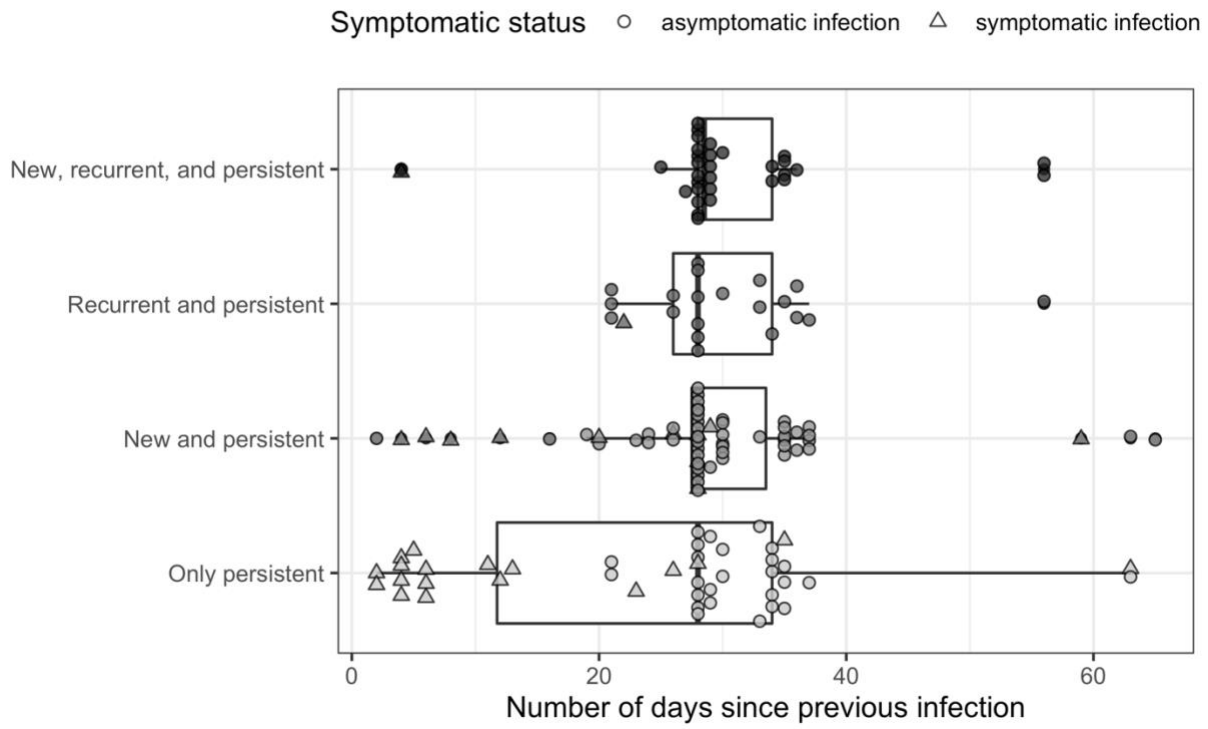
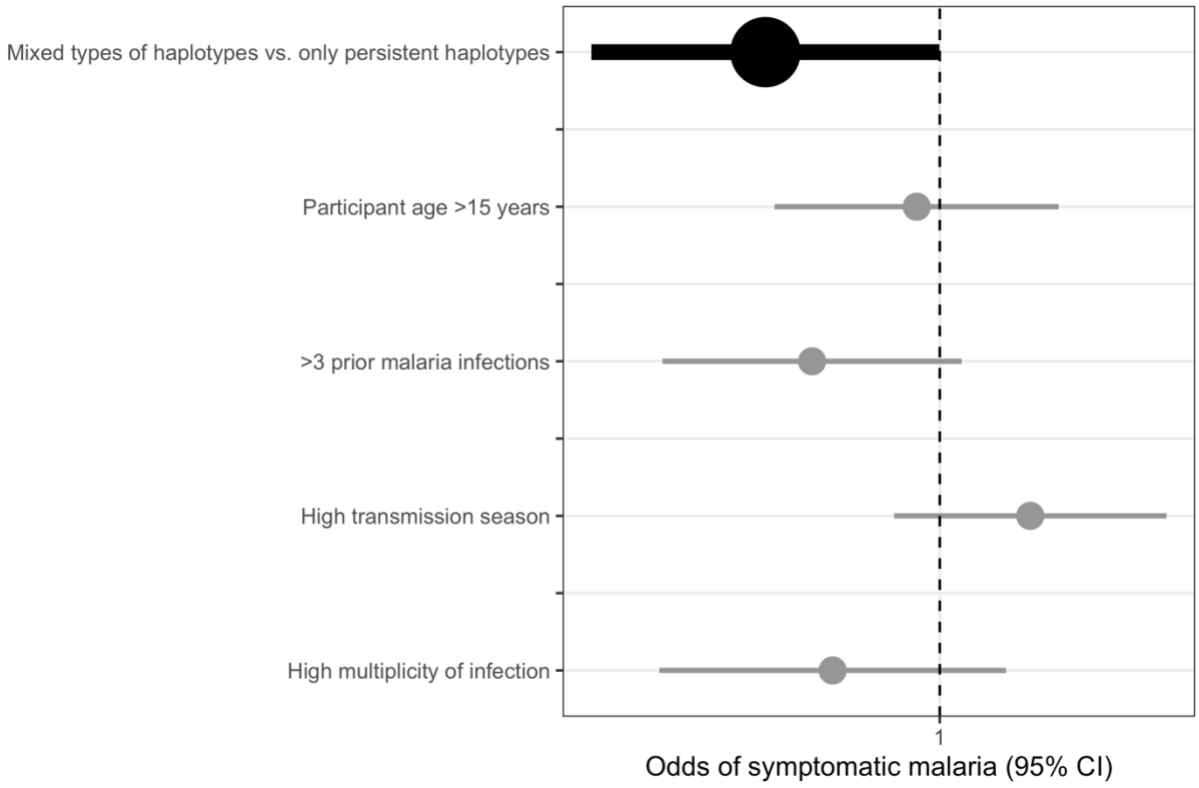


Figure S5.6. Odds of symptomatic malaria comparing persistent infections with mixed types of *pfama1* haplotypes versus only persistent haplotypes.
Adjusted multi-level logistic regression results for the odds of symptomatic malaria comparing consecutive infections with mixed types of haplotypes versus only persistent haplotypes (dark blue). Odds ratios are represented by the dots with the lines indicating the surrounding 95% confidence intervals.



Supplementary Information for Chapter VI

Additional Sample Processing Methodology

Participant DBS samples and mosquito abdomens were shipped to Duke University in Durham, North Carolina, where they were processed to determine *P. falciparum* infection status and haplotypes. Mosquito parts were individually ground in 1% Saponin using a micro tube homogenizer system fitted with a pestle, and the homogenate was transferred to unique wells of a deep 96-well plate. Single 6mm punches from the DBS were likewise distributed in deep well plates and genomic DNA was extracted from mosquito and DBS samples using a Chelex-100 protocol.[128] As described in Taylor *et al.,*[129] each sample was tested in duplicate for *P. falciparum* parasites using a duplex TaqMan real-time PCR assay targeting the *P. falciparum* pfr364 motif and the human β-tubulin gene.

P. falciparum positive DBS gDNA was prepared for genotyping based on qPCR Ct-values. Samples with Ct 25 to 30 were applied to Genomic DNA Clean & Concentrator-10 columns, and for samples with Ct >30, gDNA from a second punch of each identical DBS was added to the initial sample and the total applied to RNA Clean & Concentrator-5 columns. *P. falciparum* positive mosquito gDNA samples were applied to DNeasy PowerClean Pro Cleanup columns and the eluate concentrated by EtOH precipitation.

Library preparation for sequencing followed methods described in Nelson *et al.*[148] but with the following exceptions. PCR1 reactions contained 300 nM of each primer and 2 μL of template gDNA when DBS sample Ct was < 25, 5 μL when Ct 25 to 30, 9 μL when Ct > 30, and 7 μL for mosquito gDNA. PCR2 reactions contained 2 μL template

when DBS sample Ct < 25, 9 μL when Ct ≥ 25, and 3 μL for mosquito template. Dual-indexed libraries were prepared for the polymorphic *P. falciparum* parasite gene targets encoding apical membrane antigen-1 (*pfama1*) and circumsporozoite protein (*pfcsp*), then pooled and sequenced on an Illumina MiSeq platform.[117]

Additional Haplotype Calling Information for Samples for pfama1 and pfcsp

pfama1 and *pfcsp* haplotypes were called using the amplicon deep sequencing reads. As in Nelson *et al.*,[148] CutAdapt, Trimmomatic, and BBmap were used to trim *pfama1* and *pfcsp* primers and adapters, quality filter reads with an average Phred Quality Score < 15 over a sliding window of 4 nucleotides, remove reads less than 80 nucleotides long, and map sample reads to the 3D7 reference sequences for *pfama1* and *pfcsp* to differentiate between the two gene targets.[118–120,148] Quality-filtered reads were input into the R (version 3.6.1) package DADA2 (version 1.8) to join paired-end reads, perform an additional quality filter based on modeled error frequency, call haplotypes, and remove chimeras.[123,126] This process outputted haplotypes (distinct sequences of the *pfama1* or *pfcsp* gene target) to be used as a measure of parasite genetic diversity. Because sequencing low parasite densities has been associated with an increased risk of haplotype false discovery,[124] haplotypes were further filtered in order to mitigate the risk of false discovery by removing haplotypes from a sample that met any of the following criteria: (i) supported by < 250 reads within the sample; (ii) supported by < 3% of the sample's total read depth; (iii) deviation from the expected nucleotide length of 300 for *pfama1* or 288 for *pfcsp*; or (iv) a minority haplotype distinguished by a one single-nucleotide polymorphism (SNP) difference from another haplotype within the sample that had a read depth > 8 times the read depth of the

minority haplotype.[124] Finally, we removed haplotypes from the overall population if each haplotype was defined by a single variant position that was only variable within that haplotype.

We defined censoring criteria empirically by analyzing sequences of *pfama1* and *pfcsp* obtained from controlled mixtures of *P. falciparum* strains 3D7, V1/S, 7g8, Dd2, and FCR3. The figure and table show results for the *pfcsp* region sequenced (**Figure S6.1**). *P. falciparum* V1/S and Dd2 strains were identical within the *pfcsp* region sequenced, so results are presented with the reads combined. To develop haplotype censoring criteria, the controls were sequenced in differing proportions (control mixtures C1-C6). After quality-filtering reads and applying the haplotype censoring criteria, the final percentage of reads of each strain was similar to what was expected from the control mixtures, as indicated in the figure and table. Because the censoring criteria filtered out reads that were present in < 3% of the sample's total reads, the 3D7 and 7g8 controls were filtered out in control mixture 6. Similar results were produced for *pfama1*.

Comparison of Target Variant Positions with Prior Studies

Across all samples, we compared the variant positions that we identified in the sequenced fragments of *pfcsp* and *pfama1* with those identified in prior studies. To do so, we compiled variant positions in these fragments from PlasmoDB (accessed August 1, 2019),[163] the Pf3k database (accessed July 30, 2019),[164] as well as an external data set (Neafsey *et al.*).[162] For the latter, we downloaded raw sequencing reads and processed these with the haplotype inference criteria described above (**Figure S6.2**).[162] Through these searches, the number of variant positions in our sequenced fragment of *pfcsp* was 30 in PlasmoDB, 44 in Pf3k, and 39 in Neafsey *et al.*[162–164] Overall, these

databases yielded a total of 57 variant positions, and 37 of these were among the 72 nucleotide positions that we identified in our sequences.

Haplotype Distributions Between Sample Types

Because low parasite density samples were sequenced and strict filtering criteria were used, some samples failed sequencing and were not genotyped for *pfcsp* or *pfama1*. A total of 1242 samples were sequenced across 902 asymptomatic participant infections, 137 symptomatic participant infections, and 203 mosquito abdomens. After censoring criteria was applied, we identified *pfcsp* haplotypes in 185 mosquito abdomens, 733 asymptomatic infections, and 128 symptomatic infections (**Figure S6.3**). For *pfama1,* we identified haplotypes in 177 mosquito abdomens, 611 asymptomatic participants, and 113 symptomatic participants. Based on these numbers, *pfcsp* had a sequencing failure rate of 196/1242 (15.78%) and *pfama1* had a sequencing failure rate of 341/1242 (27.46%). Using the Wilcoxon Rank Sum test with continuity correction, there was a statistically significant correlation between parasite density and sequencing failure for *pfcsp* (p-value < 0.001) and *pfama1* (p-value < 0.001), with more sequencing failures for *pfama1* than *pfcsp*. While there were statistically significant differences between parasite density and the likelihood of sequencing failure, a DAG representing missingness due to sequencing failure indicated that restricting the data set to samples that passed sequencing was unlikely to produce missing data bias (**Figure S6.4**); however, as a precaution to account for lower parasite density samples potentially being biased towards sequencing failure and a form of missing at random bias, we included a covariate in our models for parasite density.

Inferred pfama1 Haplotypes Across Samples

For *pfama1*, 348 unique haplotypes were identified across 177 mosquito abdomens, 611 asymptomatic participants, and 113 symptomatic participants. Haplotypes produced from *pfama1* had a median MOI of 7 for mosquito abdomens, 1 for symptomatically-infected participants, and 2 for asymptomatically-infected participants.

Functional Form Assessment for Continuous Variables

A functional form assessment was conducted for continuous variables included in the models: parasite density in the participant samples, participant age at study enrollment, and mosquito abundance. The functional form assessment indicated that the optimal coding for parasite density was linear and rescaled to have a mean value of 0.0 due to its interpretability and similar functional form (**Table S6.3**). For participant age, the categorical coding (categorized: <5 years, 5-15 years, >15 years) was the best choice, because it had the lowest Akaike information criteria (AIC) value, fit the functional form, and was a commonly used coding of age in malaria literature (**Table S6.4**). For mosquito abundance, a binary coding was chosen (expressed as the total number of female *Anopheles* mosquitoes collected within the week following the participant infection stratified at <75 mosquitoes or ≥75 mosquitoes), because that functional form had the lowest AIC, was easily interpretable, and had a similar functional form to the variable (**Table S6.5**).

Within-participant Modeling of Transmissibility for pfama1

Using the *pfama1* haplotypes shared as a proxy for transmission, we selected 56 participants who suffered at least one asymptomatic and one symptomatic infection that

passed genotyping for *pfama1*. The participants had multiple infections matched with mosquitoes consisting of 1197 participant-mosquito pairs. Mosquitoes were collected between 7 days before and 14 days after the participant infection and were within 3 kilometers of the participant's household. Asymptomatic infections (Median: 0.34) had a higher median proportion of pairings that shared at least one *pfama1* haplotype with a mosquito compared to symptomatic infections (Median: 0.25) across the participants (**Figure S6.13**). In a multi-level logistic regression model controlling for parasite density and mosquito abundance, compared to symptomatic infections, asymptomatic infections had higher odds of sharing parasite haplotypes with infected mosquitoes [OR: 1.30, 95% CI: 0.63 to 2.69] (**Figure S6.14**).

Probabilistic Modeling of Transmission Across All Participants for pfama1

For a more comprehensive analysis of all participants, we conducted an additional analysis of transmissibility using a probabilistic modelling framework. After applying time and distance constraints to participant-mosquito pairings, the final *pfama1* analysis data set consisted of 3160 observations of participant-mosquito pairs found across 178 participants, 172 mosquitoes, and 36 households. 2537 pairs had a participant with an asymptomatic infection and 623 pairs had a participant with a symptomatic infection. The overall probability of transmission outcome measure, $P(TE_{all})$, ranged from 0.00 to 0.99 with a median of 0.00. Using the continuous coding of $P(TE_{all})$ and controlling for confounding covariates: parasite density in participant samples in parasites/μL, participant age, mosquito abundance, and village, we found that over 14 months participants with asymptomatic infections had an odds of participant-to-mosquito malaria transmission that was 1.22 (95% CI: 0.82 to 1.82) times

the odds of transmission for participants with symptomatic infections (**Figure S6.15**).

Table S6.1. Comparison of participant-mosquito pairs among 65 participants included in within-participant modeling to full data set of all participants

	Analysis data set 65 participants (1565 pairings)	**Full data set 198 participants (3727 pairings)**	**P-value**
Participant-level covariates			
Parasite density (parasites/μL), Median (IQR)	290.55 (3654.96)	43.49 (731.76)	<0.001[a]
Age, N (%)			<0.001[b]
<5 years	179 (11.44)	438 (11.75)	
5-15 years	1105 (70.61)	1806 (48.46)	
>15 years	281 (17.96)	1483 (39.79)	
Number of *pfcsp* haplotypes, Median (IQR)	1.00 (2.00)	3.00 (6.00)	0.211[a]
Number of infections per participant, Median (IQR)	3.00 (2.00)	2.00 (3.00)	<0.001[a]

Abbreviations: IQR, interquartile range
[a] Wilcoxon Rank Sum test with continuity correction and Bonferroni correction for repeated measures
[b] Pearson's χ^2 test with Bonferroni correction for repeated measures

Table S6.2. Differences between participant-mosquito pairs that were excluded from the analysis due to time and distance constraints across model covariates

	Asymptomatic infections (N=132,593)	Symptomatic infections (N=22,965)
Participant-level covariates		
Parasite density (parasites/μL), Median (IQR)	6.73 (166.45)	1,545.74 (6,370.95)
Age, N (%)		
<5 years	12,994 (9.80)	3,033 (13.21)
5-15 years	68,061 (51.33)	14,493 (63.11)
>15 years	51,538 (38.87)	5,439 (23.68)
Mosquito abundance, N (%)		
Low	115,911 (87.42)	15,128 (65.87)
High	16,682 (12.58)	7,837 (34.13)
Number of *pfcsp* haplotypes, Median (IQR)	3.00 (5.00)	1.00 (2.00)
Village, N (%)		
Maruti	45,463 (34.29)	7,544 (32.85)
Kinesamo	37,124 (28.00)	6,267 (27.29)
Sitabicha	50,006 (37.71)	9,154 (39.86)
Participant-mosquito pair-level covariates		
Probability of transmission, Median (IQR)		
Across all variables[#]	0.00 (0.00)	0.00 (0.00)
Time interval	0.00 (0.00)	0.00 (0.00)
Distance interval	0.00 (0.51)	0.00 (0.47)
pfcsp haplotype sharing and prevalence*	0.09 (0.24)	0.00 (0.17)
For those that shared *pfcsp* haplotypes	0.20 (0.19)	0.20 (0.23)
Number *pfcsp* haplotypes shared, Median (IQR)**	1.00 (2.00)	0.00 (1.00)
For those that shared *pfcsp* haplotypes	2.00 (2.00)	1.00 (0.00)

Abbreviations: IQR, interquartile range

[#] The probability of transmission across all variables was 0.00 because the participant-mosquito pairs were not within the distance and time restraints to be a likely participant-to-mosquito transmission.

*The probability of transmission based on the *pfcsp* haplotype sharing and prevalence is shown for all pairings regardless on if they shared haplotypes or not.

**The number of *pfcsp* haplotypes shared is shown for all pairs regardless on if they shared haplotypes or not.

Table S6.3. Results of multi-level logistic regression models of probability of a transmission event using different functional forms of the *P. falciparum* parasite density in humans

Coding Choice / Term	Coefficient	SE	Log Likelihood	AIC
Linear			-601.0	1210.0
Parasite density	1.13	1.07		
Quadratic			-599.3	1208.7
Parasite density	0.70	1.34		
Parasite density squared	1.08	1.04		
Cubic			-599.3	1210.6
Parasite density	0.75	1.48		
Parasite density squared	1.03	1.21		
Parasite density cubed	1.01	1.02		
Binary			0598.3	1204.7
<100 p/μL (under cRDT detection)	Ref	Ref		
≥100 p/μL (over cRDT detection)	0.60	1.20		
Categorical			-596.0	1203.9
< 1.93 p/μL	Ref	Ref		
≥ 1.93 and < 51.64 p/μL	1.34	1.25		
≥ 51.64 and 773.53 p/μL	0.57	1.31		
≥ 773.53 p/μL	0.75	1.29		
Natural Log			-600.1	1208.2
Parasite density ln	0.95	1.03		

Abbreviations: SE, standard error; AIC, Akaike information criteria; cRDT, conventional rapid diagnostic test

Table S6.4. Results of multi-level logistic regression models of probability of a transmission event using different functional forms of participant age

Coding Choice / Term	Coefficient	SE	Log Likelihood	AIC
Linear			-602.2	1212.5
Age	0.94	1.09		
Quadratic			-602.2	1214.5
Age	0.92	1.16		
Age squared	1.01	1.08		
Cubic			-599.5	1211.1
Age	0.86	1.17		
Age squared	0.69	1.20		
Age cubed	1.14	1.06		
Categorical			-600.2	1210.5
<5 years	Ref	Ref		
5-15 years	1.59	1.37		
>15 years	1.14	1.39		
Natural Log			-602.5	1212.9
Age ln	1.01	1.10		

Abbreviations: SE, standard error; AIC, Akaike information criteria

Table S6.5. Results of multi-level logistic regression models of probability of a transmission event using different functional forms of the total number of female *Anopheles* mosquitoes collected within one week following participant infection

Coding Choice / Term	Coefficient	SE	Log Likelihood	AIC
Linear			-602.0	1211.9
Mosquito abundance	1.09	1.09		
Quadratic			-601.9	1213.8
Mosquito abundance	1.09	1.09		
Mosquito abundance squared	1.04	1.09		
Cubic			-600.8	1213.6
Mosquito abundance	1.30	1.16		
Mosquito abundance squared	1.08	1.10		
Mosquito abundance cubed	0.92	1.06		
Binary				
<75 mosquitoes	Ref	Ref	-601.9	1211.7
75-147 mosquitoes	1.21	1.19		
Natural Log			-602.2	1212.5
Mosquito abundance ln	1.09	1.14		

Abbreviations: SE, standard error; AIC, Akaike information criteria

Figure S6.1. Expected and observed *pfcsp* haplotype frequencies in control mixtures of genomic DNA from *P. falciparum* reference lines.
Expected strain mixtures were based on the input amounts of genomic DNA of each reference parasite strain. Strains V1/S and Dd2 share identical *pfcsp* haplotypes and therefore could not be resolved. Haplotypes in "C6" that mapped to 3D7 and 7g8 were censored because they were present in ≤ 3% in the overall read yield for that template. NA: not applicable.

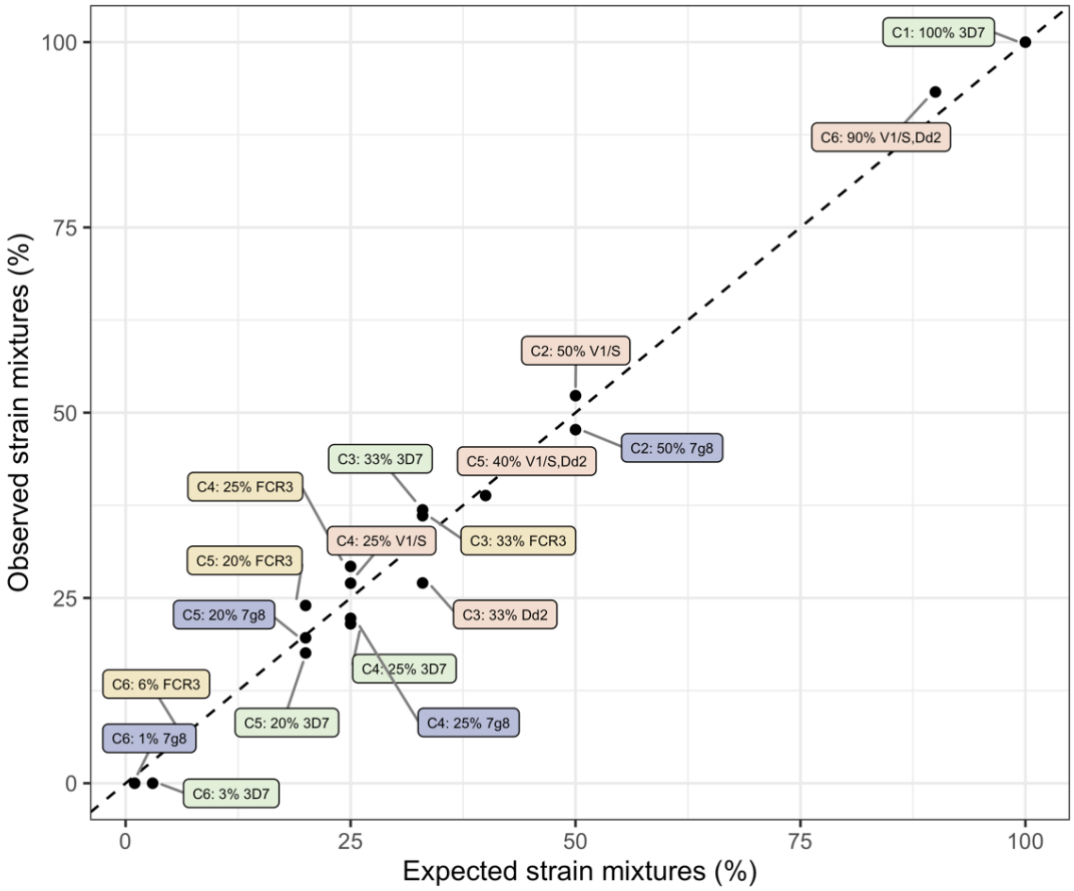


Figure S6.2. Comparison of overlap in the variant nucleotide positions within the sequenced *pfcsp* fragment identified in our study and in prior studies.
The total number of variant nucleotide positions for each set was: Neafsey *et al.* = 39, PlasmoDB = 30, Pf3k = 44, and this study = 72.[162–164]

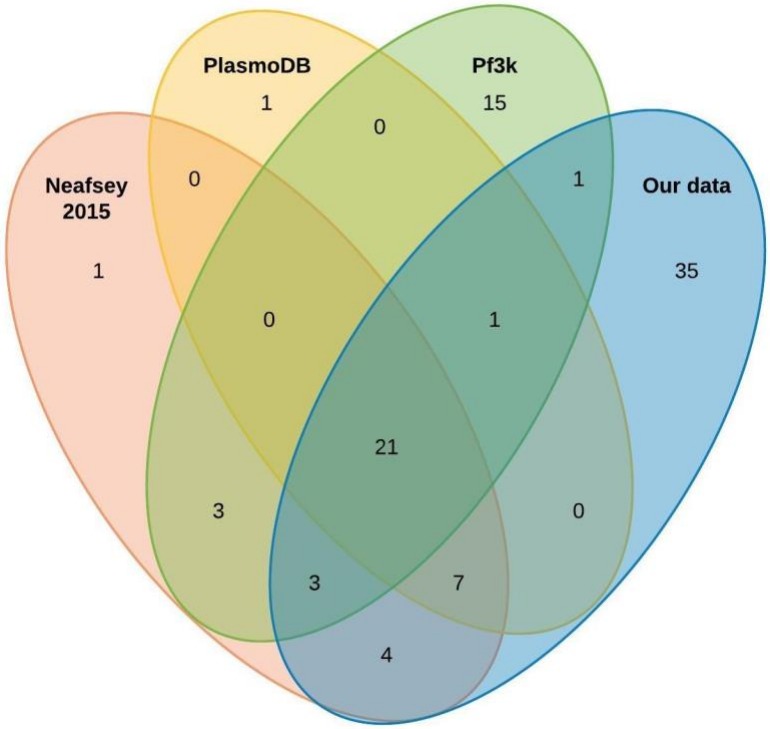


Figure S6.3. Sample processing flow-diagram from original samples to censored, high-quality haplotypes.
The number of samples and reads returned from each step of sample processing is shown for amplicon deep sequencing of *pfcsp*. The same process was done for *pfama1*.

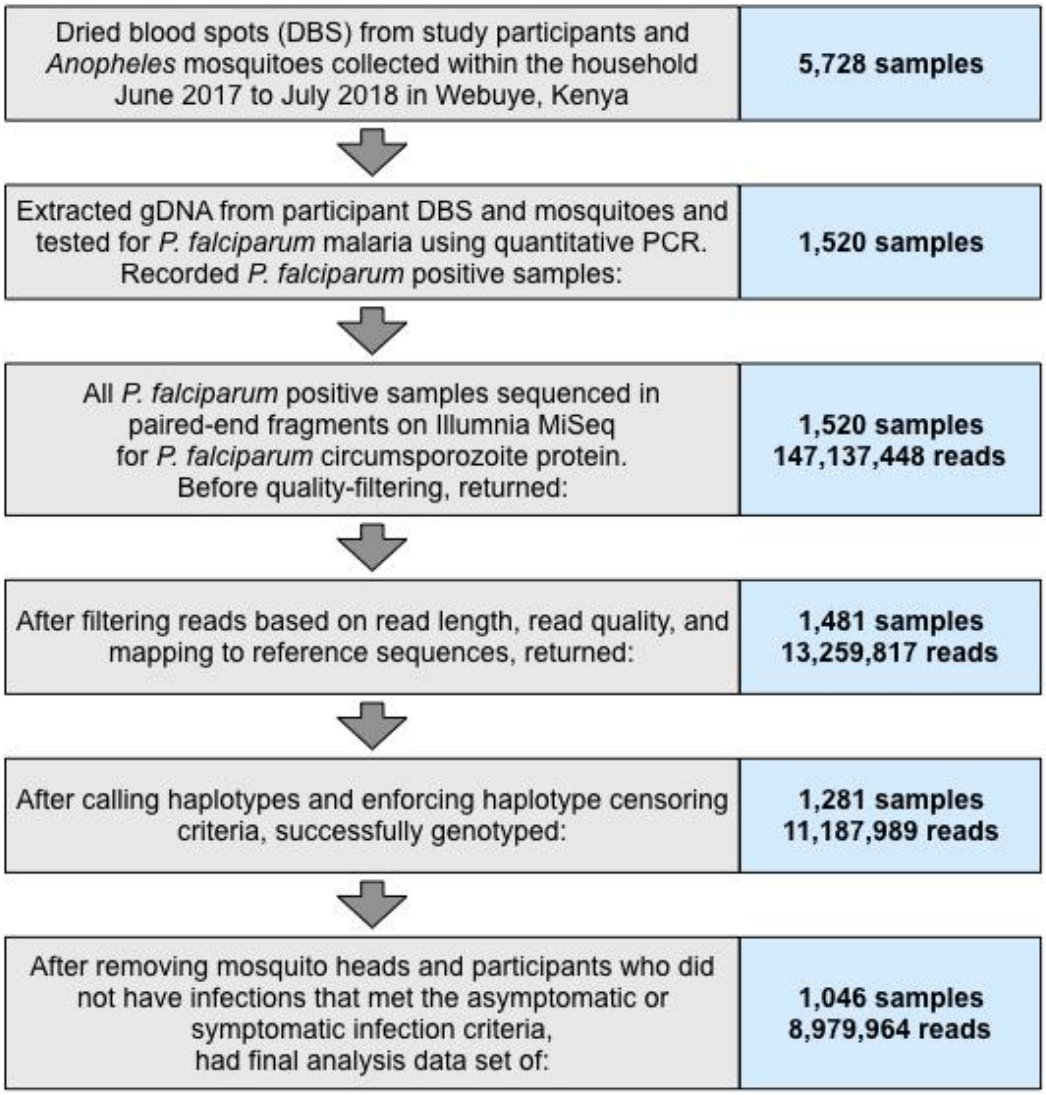


Figure S6.4. DAG investigating potential for missing data bias in samples that failed sequencing.
A DAG was used to assess potential bias caused by data missing at random based on sequencing failure.

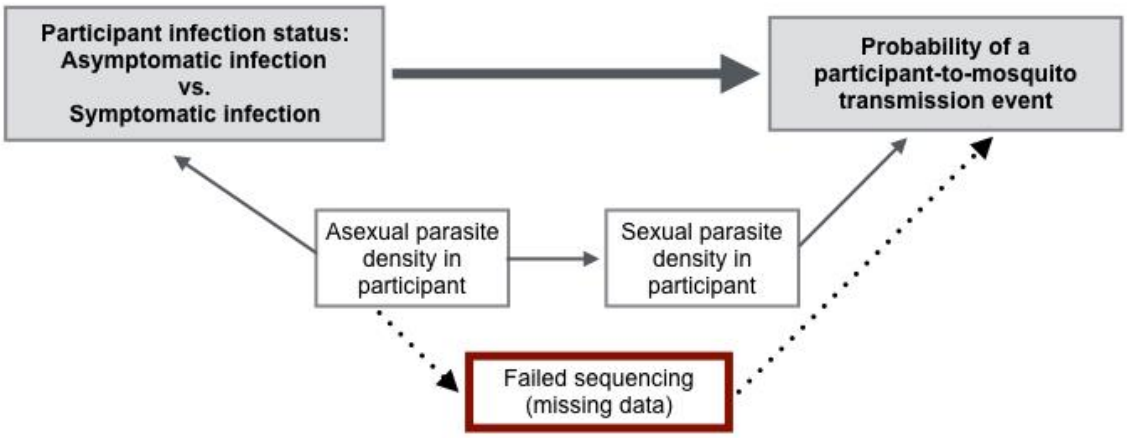


Figure S6.5. Probability of transmission over time.
The probability of transmission over time [$P(TE_t)$] distribution had a flat, high probability of transmission from -14 to 7 days to allow for each participant sample to have the same number of mosquito collections and the same probability of transmission within the time range. The distribution was restricted to only allow a transmission event to occur when a mosquito was collected within 14 days (i.e. -14 days) after the participant infection or 7 (i.e. +7 days) days prior to the participant infection. Any participant-mosquito pair within this time range, had $P(TE_t)$ = 1. Any time outside of the time range, had $P(TE_t)$ = 0.

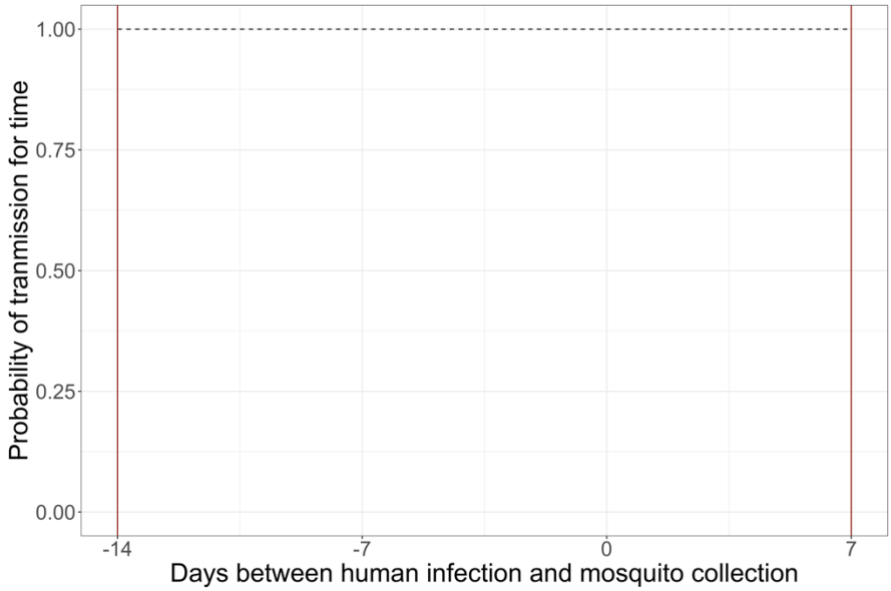


Figure S6.6. Sensitivity analysis for probability of transmission over time.
A sensitivity analysis was done to comparing different time windows for the probability of transmission over time [$P(TE_t)$] and the effect on the relationship observed. The distribution was restricted to only allow a transmission event to occur when a mosquito was collected within 30 to 14 days (i.e. -30 to -14 days) after the participant infection or 7 (i.e. +7 days) days prior to the participant infection. The multi-level logistic regression model was reran comparing the probability of transmission to mosquitoes across participants with asymptomatic compared to symptomatic infections using each time window. Each time window is shown on the y-axis and the associated odds ratio for transmission on the x-axis. The *pfcsp* haplotypes were used for this sensitivity analysis.

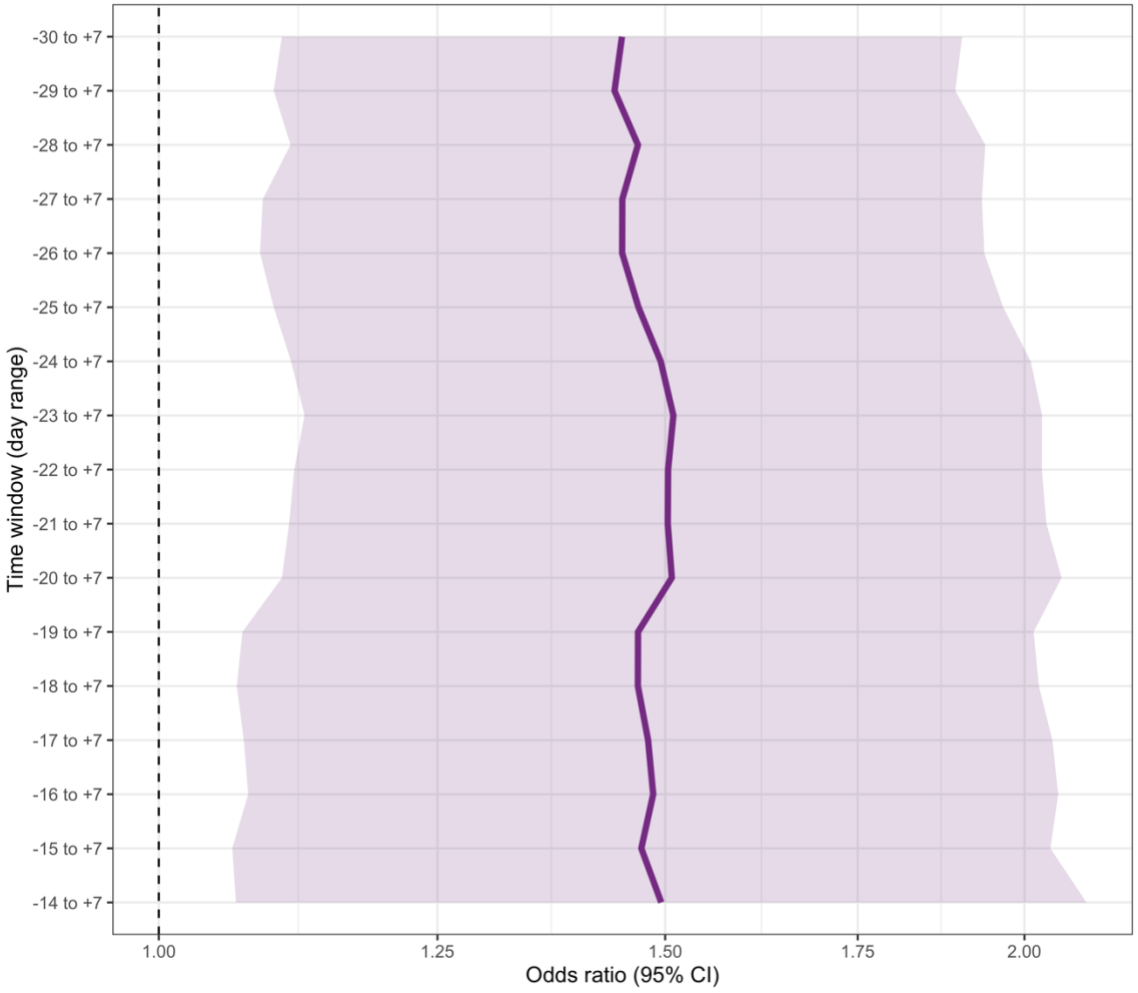


Figure S6.7. Probability of transmission over distance.
The distribution of the probability of transmission over distance [$P(TE_d)$] decreased to a low probability of transmission as the distance between the participant infection and mosquito collection increased. At any distance greater than 3 kilometers, estimated $P(TE_d)$ = 0, allowing transmission events to occur across households but not villages. The curve stops at 0.56 kilometers, because no participants and mosquitoes were collected within a village at a distance greater than 0.56 kilometers.

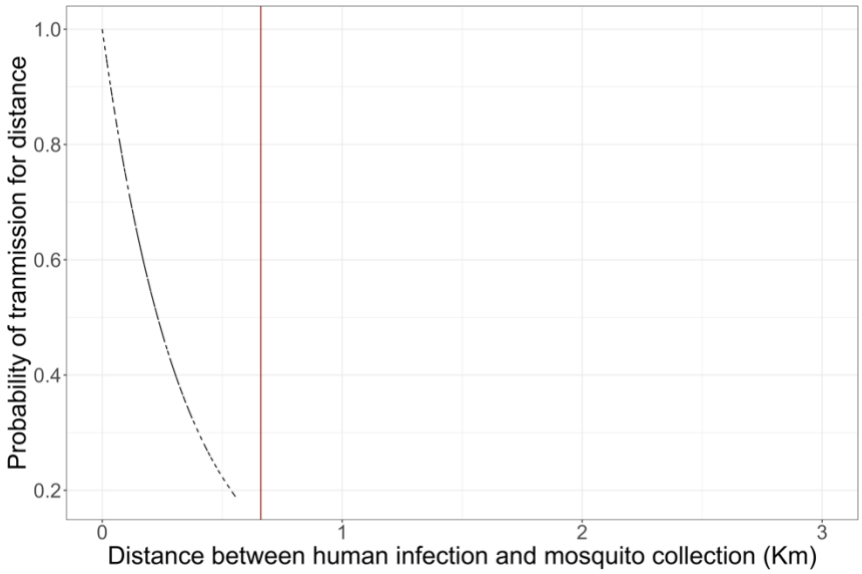


Figure S6.8. Sensitivity analysis for probability of transmission over distance. A sensitivity analysis was done to comparing different distance cutoffs for the probability of transmission and the effect on the relationship observed. The multi-level logistic regression model was reran comparing the probability of transmission to mosquitoes across participants with asymptomatic compared to symptomatic infections using each distance cutoff. Each maximum distance cutoff is shown on the y-axis and the associated odds ratio for transmission to mosquitoes on the x-axis. The *pfcsp* haplotypes were used for this sensitivity analysis.

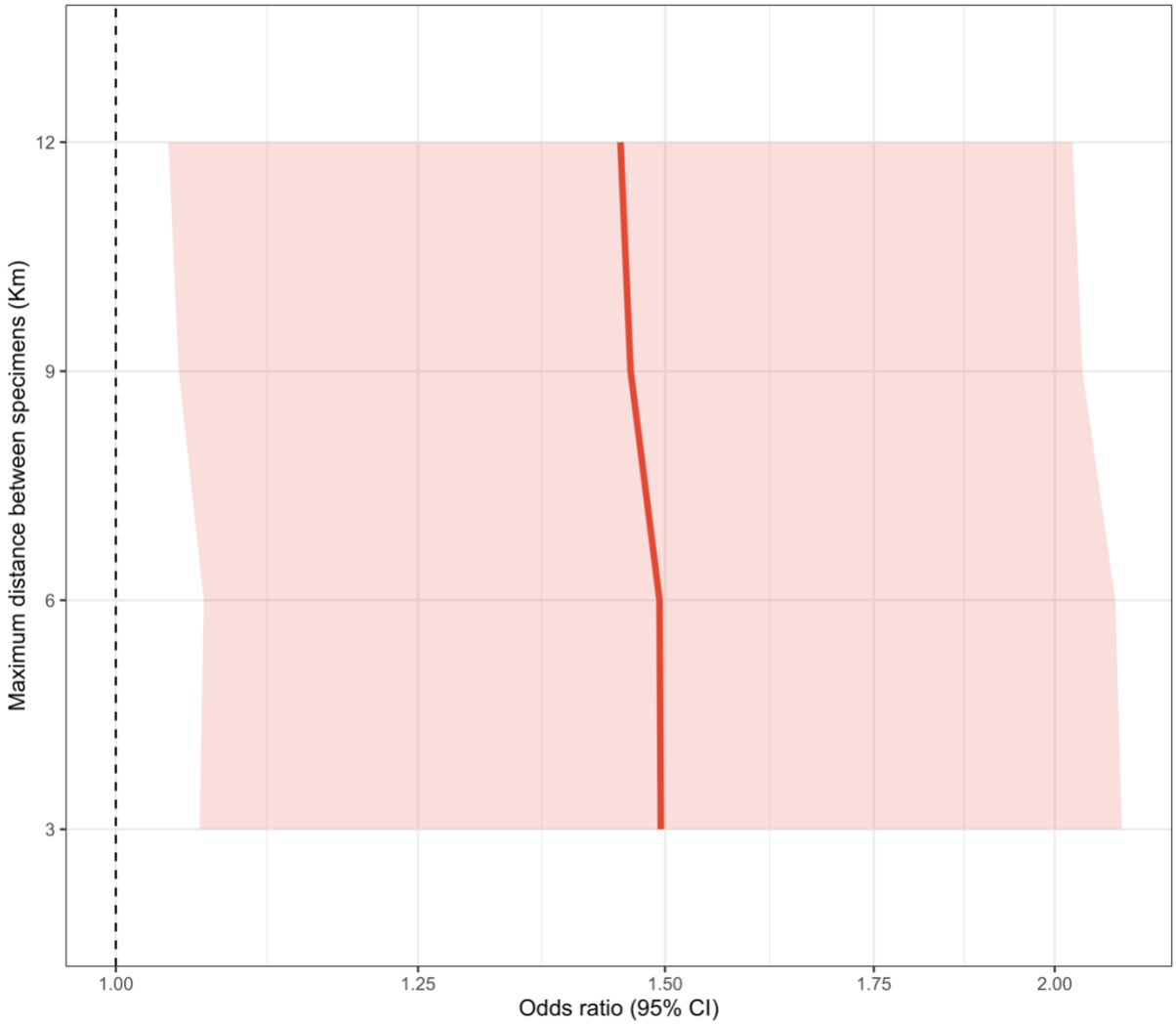


Figure S6.9. Number of shared *pfcsp* haplotypes between participants and mosquitoes at < 3 kilometers and ≥ 3 kilometers.
The number of pfcsp haplotype shared between specimens collected at a distance < 3 kilometers and ≥ 3 kilometers was compared.

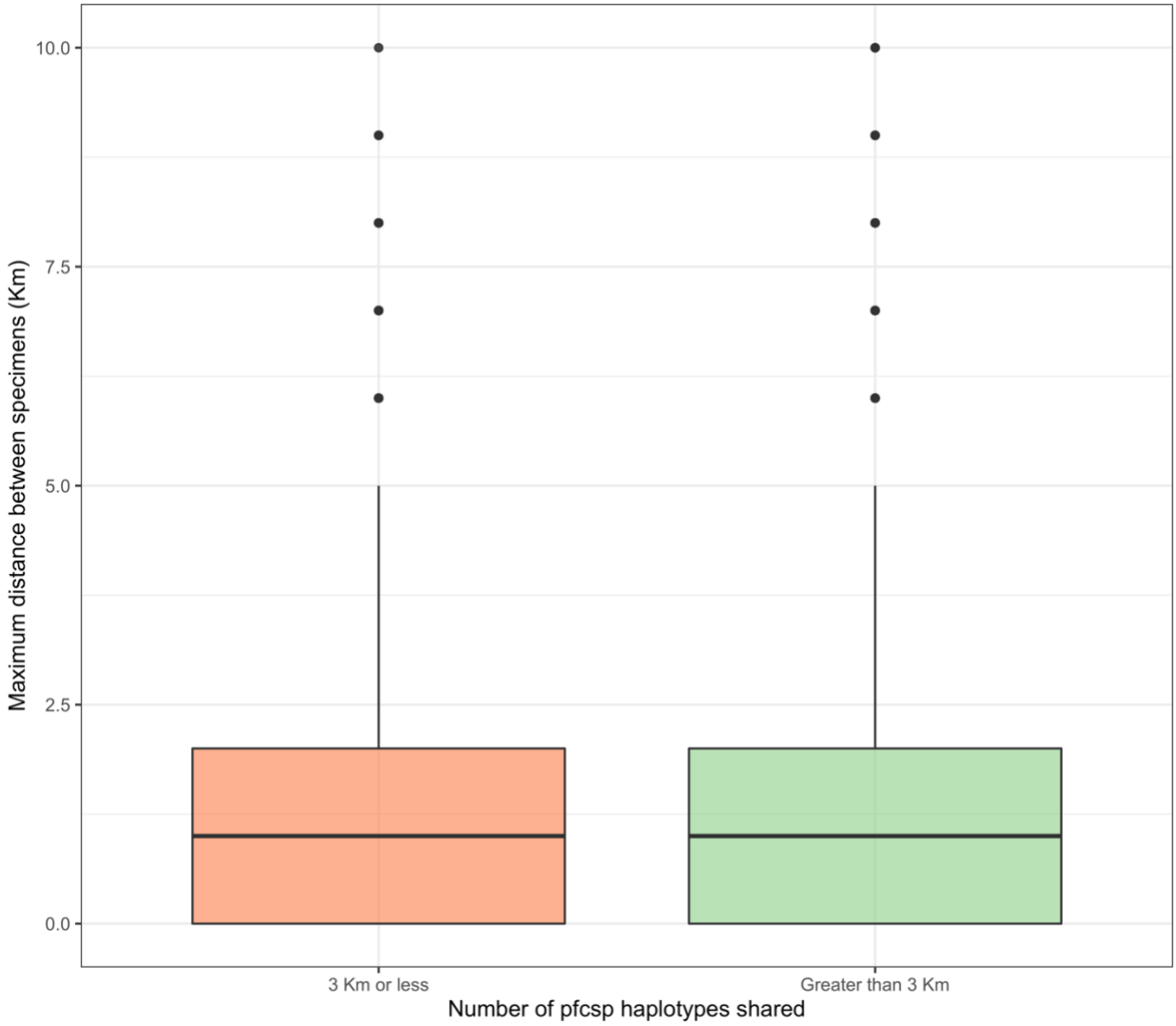


Figure S6.10. Sensitivity analysis for probability of transmission over haplotypes. A sensitivity analysis was done using a different coding for the $P(TE_h)$ term where it was no longer calculated separately for *pfama1* and *pfcsp* but instead calculated as a combined value using both *pfama1* and *pfcsp* haplotypes. The multi-level logistic regression model was reran comparing the probability of transmission to mosquitoes across participants with asymptomatic compared to symptomatic infections.

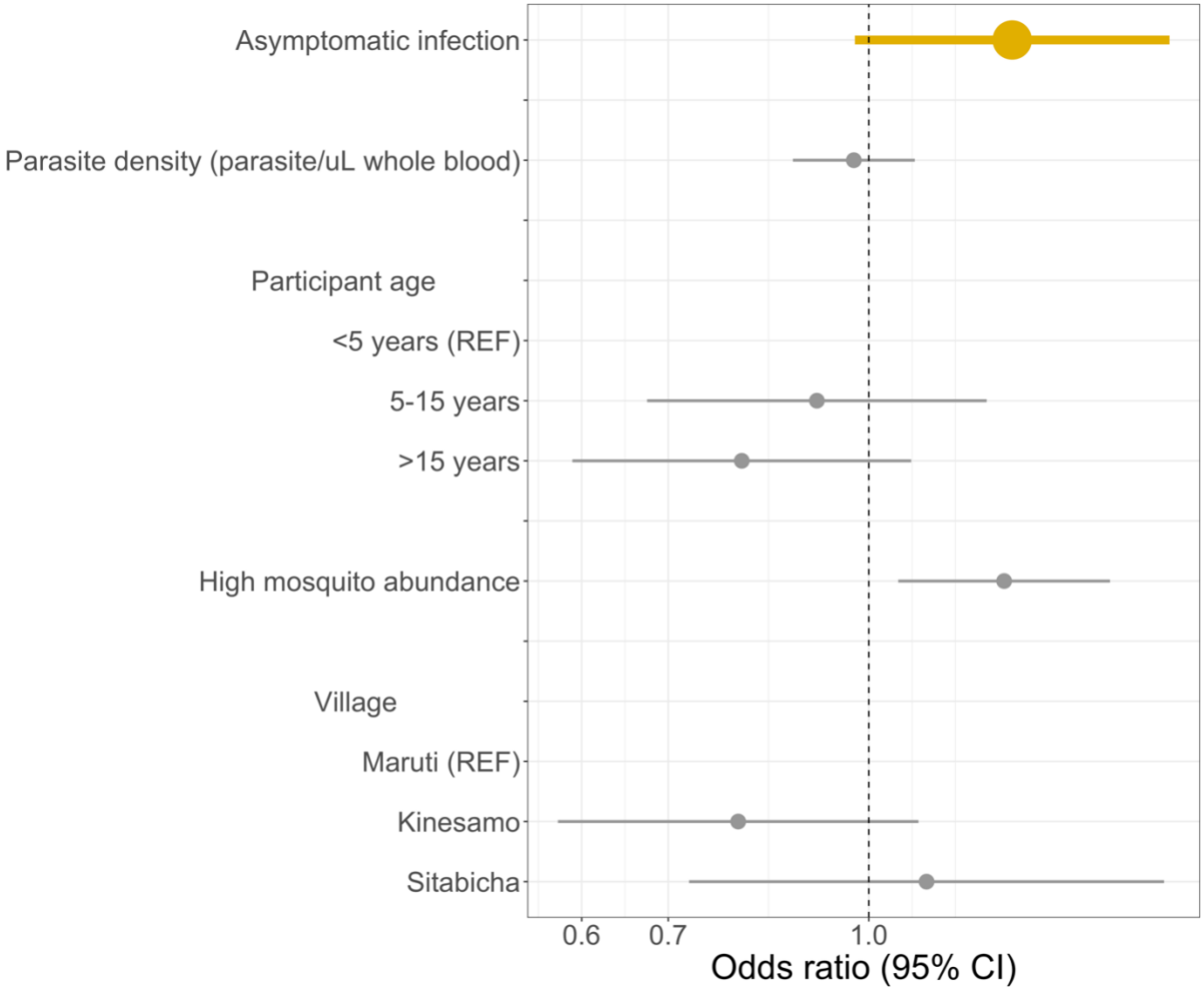


Figure S6.11. DAG of causal relationship between a participant's symptomatic status and probability of a participant-to-mosquito transmission event.
The DAG identified four confounding covariates that needed to be controlled for in assessing the effect of participants' malaria symptomatic status on the probability of a participant-to-mosquito transmission event: age, parasite density in the participant samples in parasites/μL, total number of female *Anopheles* mosquitoes collected within the week following the participant infection, and multiplicity of infection (MOI) in participants. MOI was controlled for in **Equation 6.3** for $P(TE_h)$.

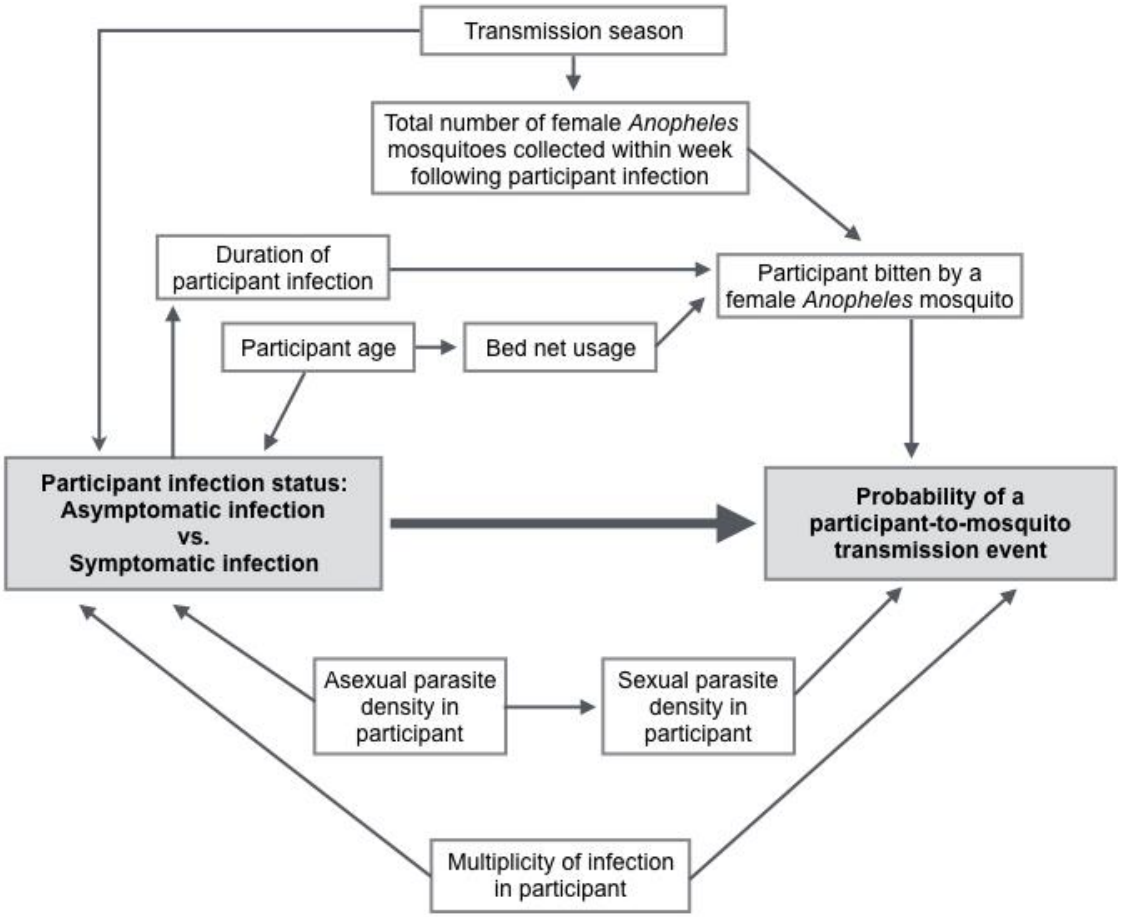


Figure S6.12. Haplotype distribution across sample types for *pfcsp.*
The full distribution of *pfcsp* haplotypes across mosquitoes, asymptomatic infections, and symptomatic infections is shown here. This plot shows all *pfcsp* haplotypes regardless of how many samples they were found in.

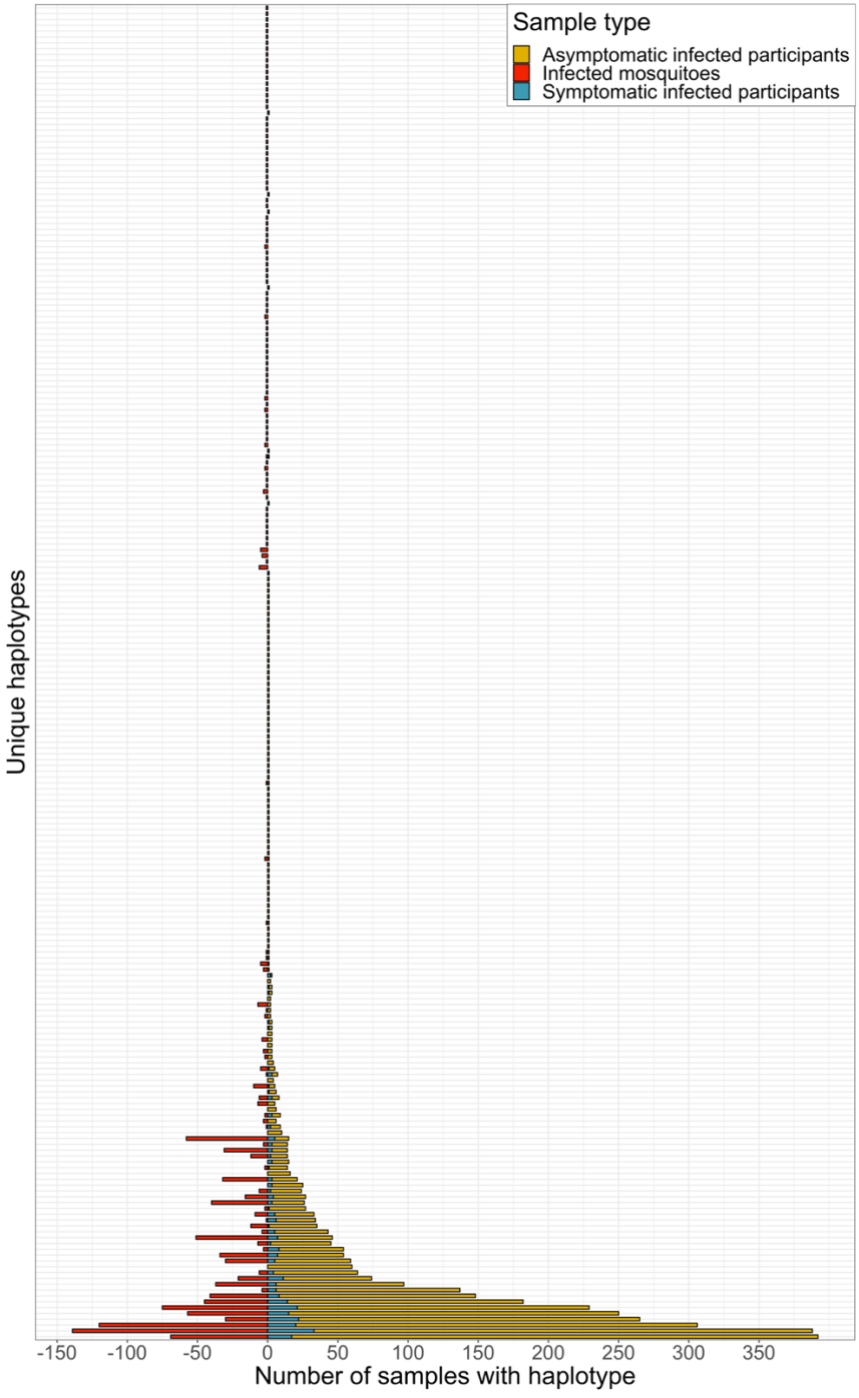


Figure S6.13. Comparison of likelihood of transmission to mosquitoes for participants with both asymptomatic and symptomatic infections using the *pfama1* gene target.

For each participant (N=56), the median proportion of pairings with a mosquito that shared a minimum of one haplotype was calculated for asymptomatic and symptomatic infections to represent the average likelihood of transmission to a mosquito. Using the *pfama1* gene target, asymptomatic infections had a higher median likelihood of transmission, as indicated by the higher number of dots to the left of the dotted diagonal line; however, this difference was not statistically significant as shown in **Figure S6.14.**

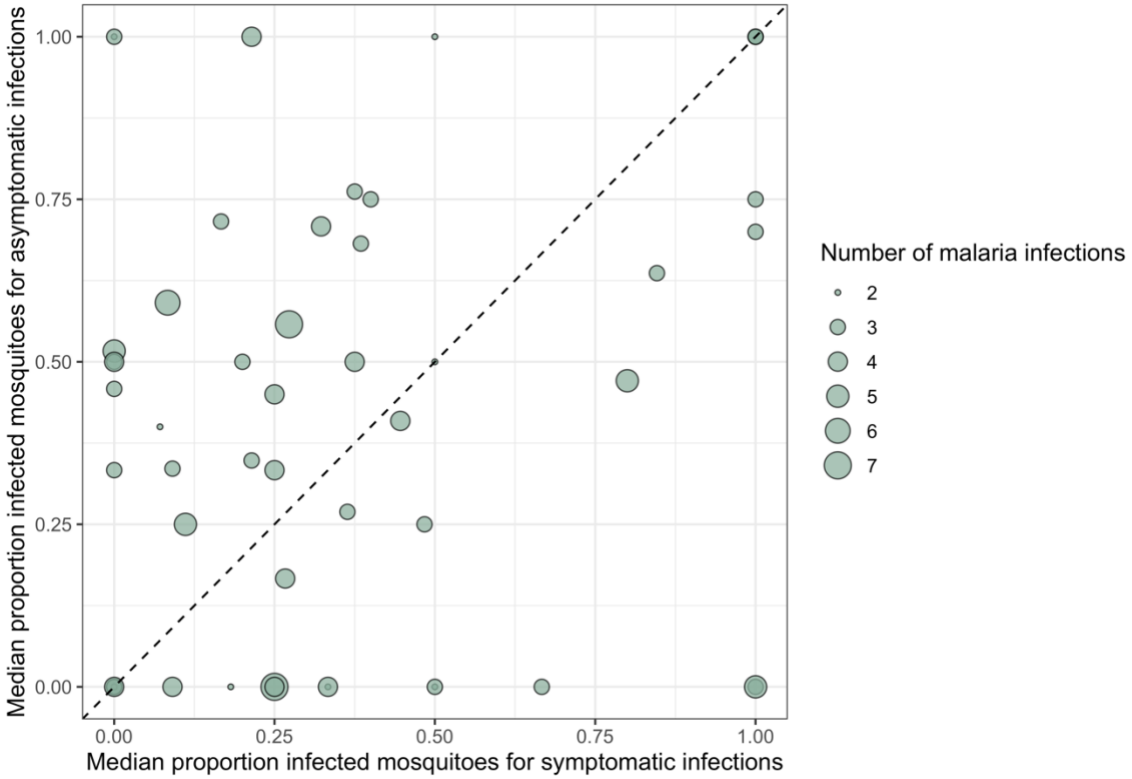


Figure S6.14. Comparison of likelihood of transmission to mosquitoes for participants with both asymptomatic and symptomatic infections using the *pfama1* gene target.

We ran a multi-level logistic regression using the continuous coding of the proportion of participant-mosquito pairings that shared at least one *pfama1* haplotype for each infection. The model controlled for covariates: parasite density in the participant samples in parasites/μL (linear) and the mosquito abundance (binary: <75 mosquitoes, ≥75 mosquitoes). Model results suggested higher odds of asymptomatic compared to symptomatic malaria transmission to mosquitoes, but results were not statistically significant.

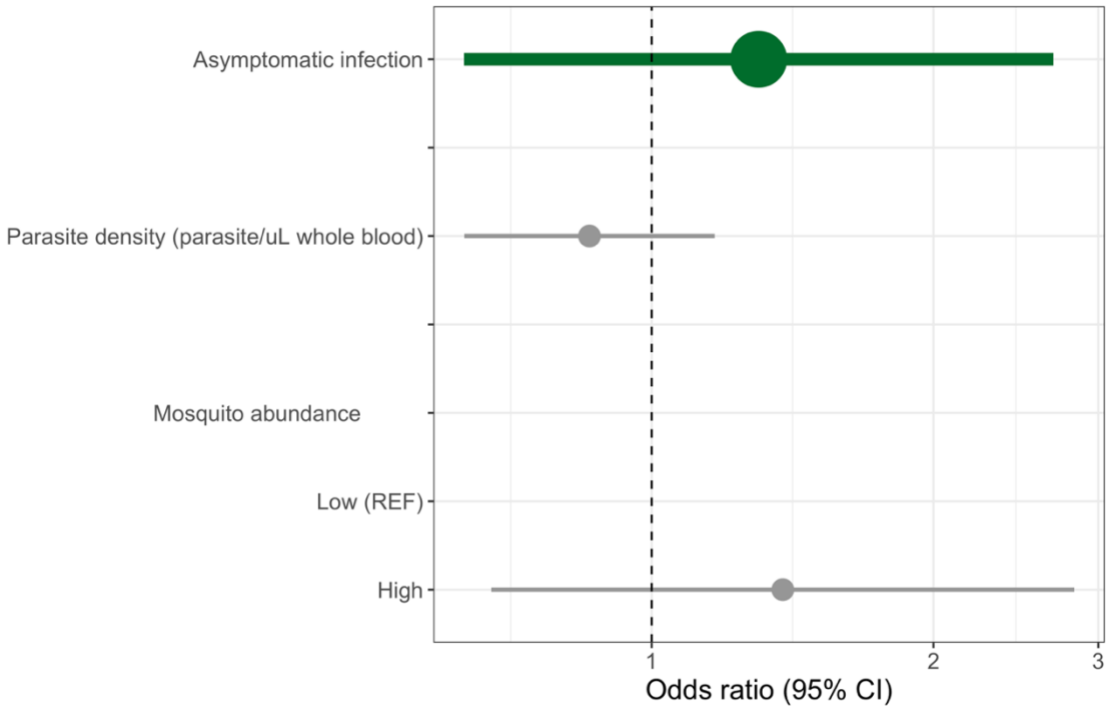
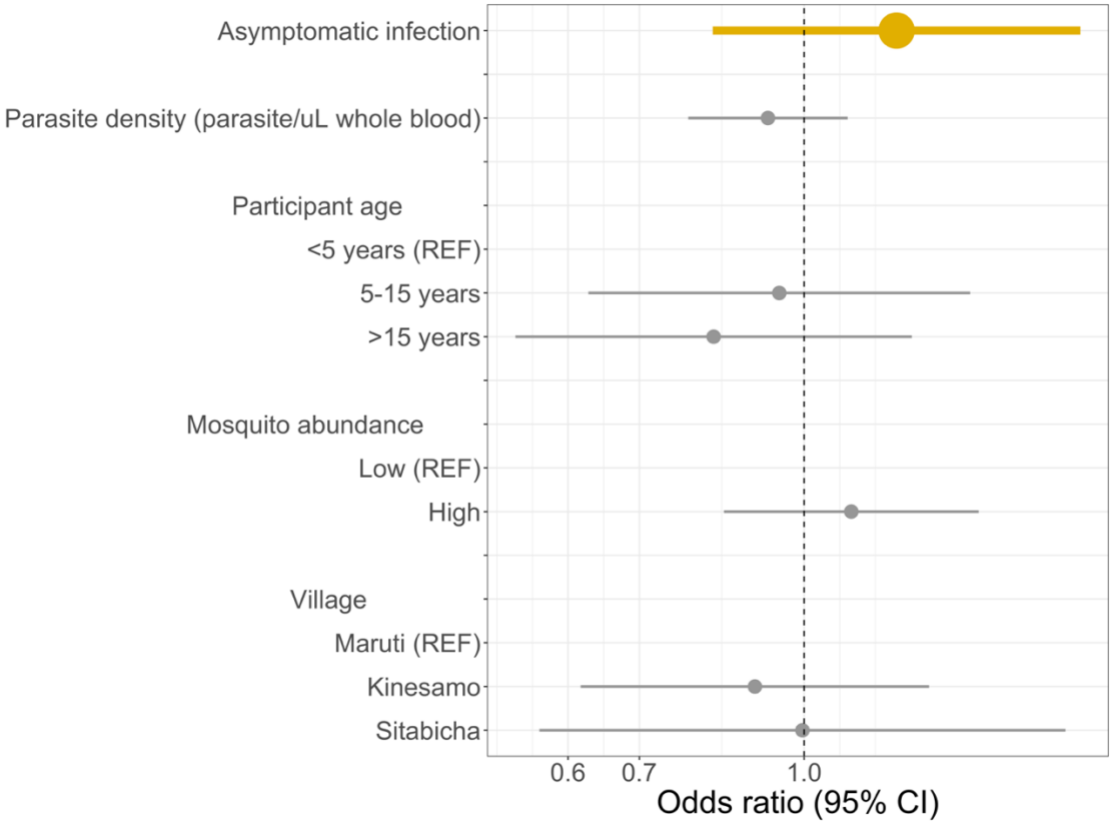


Figure S6.15. Multi-level logistic regression results for odds of a participant-to-mosquito malaria transmission from participants with asymptomatic compared to symptomatic infections using the *pfama1* malaria gene target.
We ran a multi-level logistic regression using the continuous coding of $P(TE_{all})$ and *pfama1* haplotypes. The model controlled for covariates: parasite density in the participant samples in parasites/μL (linear), age (categorized: <5 years, 5-15 years, >15 years), the mosquito abundance (binary: <75 mosquitoes, ≥75 mosquitoes), and village. Model results suggested higher odds of asymptomatic compared to symptomatic malaria transmission to mosquitoes, but results were not statistically significant.



REFERENCES

1. World Health Organization. Country Profiles for Malaria: Kenya, Africa Region. Published online 2017. http://www.who.int/malaria/publications/country-profiles/profile_ken_en.pdf?ua=1

2. World Health Organization. Malaria. Published 2018. Accessed October 10, 2018. http://www.who.int/ith/diseases/malaria/en/

3. World Health Organization. *Kenya Malaria Mortality and Burden of Disease*.; 2016. http://www.who.int/countries/ken/en/

4. Ernst KC, Hayden MH, Olsen H, Cavanaugh JL, Ruberto I, Agawo M, Munga S. Comparing ownership and use of bed nets at two sites with differential malaria transmission in western Kenya. *Malar J*. 2016;15(1):217. doi:10.1186/s12936-016-1262-1

5. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, Battle KE, Moyes CL, Henry A, Eckhoff PA, Wenger EA, Briët O, Penny MA, Smith TA, Bennett A, Yukich J, Eisele TP, Griffin JT, Fergus CA, Lynch M, Lindgren F, Cohen JM, Murray CLJ, Smith DL, Hay SI, Cibulskis RE, Gething PW. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature*. 2015;526(7572):207-211. doi:10.1038/nature15535

6. Zhou G, Afrane YA, Vardo-Zalik AM, Atieli H, Zhong D, Wamae P, Himeidan YE, Minakawa N, Githeko AK, Yan G. Changing patterns of malaria epidemiology between 2002 and 2010 in western kenya: The fall and rise of malaria. von Seidlein L, ed. *PLoS One*. 2011;6(5):e20318. doi:10.1371/journal.pone.0020318

7. Zhou G, Lee M-C, Githeko AK, Atieli HE, Yan G. Insecticide-Treated Net Campaign and Malaria Transmission in Western Kenya: 2003–2015. *Front Public Heal*. 2016;4:15. doi:10.3389/fpubh.2016.00153

8. Kapesa A, Kweka EJ, Atieli H, Afrane YA, Kamugisha E, Lee MC, Zhou G, Githeko AK, Yan G. The current malaria morbidity and mortality in different transmission settings in western Kenya. *PLoS One*. 2018;13(8). doi:10.1371/journal.pone.0202031

9. Nguyen TN, von Seidlein L, Nguyen TV, Truong PN, Hung S Do, Pham HT, Nguyen TU, Le TD, Dao VH, Mukaka M, Day NP, White NJ, Dondorp AM, Thwaites GE, Hien TT. The persistence and oscillations of submicroscopic Plasmodium falciparum and Plasmodium vivax infections over time in Vietnam: an open cohort study. *Lancet Infect Dis*. 2018;18(5):565-572. doi:10.1016/S1473-3099(18)30046-X

10. Rovira-Vallbona E, Contreras-Mancilla JJ, Ramirez R, Guzmán-Guzmán M, Carrasco-Escobar G, Llanos-Cuentas A, Vinetz JM, Gamboa D, Rosanas-Urgell

A. Predominance of asymptomatic and sub-microscopic infections characterizes the Plasmodium gametocyte reservoir in the Peruvian Amazon. *PLoS Negl Trop Dis*. 2017;11(7). doi:10.1371/journal.pntd.0005674

11. De Jong SE, Asscher VER, Wammes LJ, Wiria AE, Hamid F, Sartono E, Supali T, Smits HH, Luty AJF, Yazdanbakhsh M. Longitudinal study of changes in γδ T cells and CD4+T cells upon asymptomatic malaria infection in Indonesian children. *Sci Rep*. 2017;7(1):8844. doi:10.1038/s41598-017-09099-z

12. Daou M, Kouriba B, Ouédraogo N, Diarra I, Arama C, Keita Y, Sissoko S, Ouologuem B, Arama S, Bousema T, Doumbo OK, Sauerwein RW, Scholzen A. Protection of Malian children from clinical malaria is associated with recognition of multiple antigens. *Malar J*. 2015;14(1). doi:10.1186/s12936-015-0567-9

13. Lawaly R, Konate L, Marrama L, Dia I, Diallo D, Sarr FD, Schneider BS, Casademont I, Diallo M, Brey PT, Sakuntabhai A, Mecheri S, Paul R. Impact of mosquito bites on asexual parasite density and gametocyte prevalence in asymptomatic chronic plasmodium falciparum infections and correlation with IgE and IgG titers. *Infect Immun*. 2012;80(6):2240-2246. doi:10.1128/IAI.06414-11

14. Njama-Meya D, Kamya MR, Dorsey G. Asymptomatic parasitaemia as a risk factor for symptomatic malaria in a cohort of Ugandan children. *Trop Med Int Heal*. 2004;9(8):862-868. doi:10.1111/j.1365-3156.2004.01277.x

15. Le Port A, Cot M, Etard JF, Gaye O, Migot-Nabias F, Garcia A. Relation between Plasmodium falciparum asymptomatic infection and malaria attacks in a cohort of Senegalese children. *Malar J*. 2008;7. doi:10.1186/1475-2875-7-193

16. Nsobya SL, Parikh S, Kironde F, Lubega G, Kamya MR, Rosenthal PJ, Dorsey G. Molecular Evaluation of the Natural History of Asymptomatic Parasitemia in Ugandan Children. *J Infect Dis*. 2004;189(12):2220-2226. doi:10.1086/421281

17. Portugal S, Tran TM, Ongoiba A, Bathily A, Li S, Doumbo S, Skinner J, Doumtabe D, Kone Y, Sangala J, Jain A, Davies DH, Hung C, Liang L, Ricklefs S, Homann MV, Felgner PL, Porcella SF, Färnert A, Doumbo OK, Kayentao K, Greenwood BM, Traore B, Crompton PD. Treatment of chronic asymptomatic plasmodium falciparum infection does not increase the risk of clinical malaria upon reinfection. *Clin Infect Dis*. 2017;64(5):645-653. doi:10.1093/cid/ciw849

18. Liljander A, Bejon P, Mwacharo J, Kai O, Ogada E, Peshu N, Marsh K, Färnert A. Clearance of asymptomatic P. falciparum infections interacts with the number of clones to predict the risk of subsequent malaria in Kenyan children. *PLoS One*. 2011;6(2). doi:10.1371/journal.pone.0016940

19. Sonden K, Doumbo S, Hammar U, Vafa Homann M, Ongoiba A, Traord B, Bottai M, Crompton PD, Färnert A. Asymptomatic Multiclonal Plasmodium falciparum Infections Carried Through the Dry Season Predict Protection Against Subsequent Clinical Malaria. *J Infect Dis*. 2015;212(4):608-616.

doi:10.1093/infdis/jiv088

20. Buchwald AG, Sixpence A, Chimenya M, Damson M, Sorkin JD, Wilson ML, Seydel K, Hochman S, Mathanga DP, Taylor TE, Laufer MK. Clinical Implications of Asymptomatic Plasmodium falciparum Infections in Malawi. *Clin Infect Dis*. 2019;21201(1):106-112. doi:10.1093/cid/ciy427

21. Males S, Gaye O, Garcia A. Long-Term Asymptomatic Carriage of Plasmodium falciparum Protects from Malaria Attacks: a Prospective Study among Senegalese Children. *Clin Infect Dis*. 2008;46(4):516-522. doi:10.1086/526529

22. Henning L, Schellenberg D, Smith T, Henning D, Alonso P, Tanner M, Mshinda H, Beck HP, Felger I. A prospective study of Plasmodium falciparum multiplicity of infection and morbidity in Tanzanian children. *Trans R Soc Trop Med Hyg*. 2004;98(12):687-694. doi:10.1016/j.trstmh.2004.03.010

23. Wamae K, Wambua J, Nyangweso G, Mwambingu G, Osier F, Ndung'u F, Bejon P, Ochola-Oyier LI. Transmission and Age Impact the Risk of Developing Febrile Malaria in Children with Asymptomatic Plasmodium falciparum Parasitemia. *J Infect Dis*. Published online 2018. doi:10.1093/infdis/jiy591

24. Mangeni JN, Menya D, Obala A, Platt A, O'Meara WP. Development and validation of a rapid assessment tool for malaria prevention. *Malar J*. 2016;15(1):1-12. doi:10.1186/s12936-016-1575-0

25. Rusk A, Smith N, Menya D, Obala A, Simiyu C, Khwa-Otsyula B, Omeara W. Does anti-malarial drug knowledge predict anti-malarial dispensing practice in drug outlets? A survey of medicine retailers in western Kenya. *Malar J*. 2012;11:1. doi:10.1186/1475-2875-11-263

26. Rusk A, Goodman C, Naanyu V, Koech B, Obala A, O'Meara WP. Expanding Access to Malaria Diagnosis through Retail Shops in Western Kenya: What Do Shop Workers Think? *Malar Res Treat*. 2013;2013:1-9. doi:10.1155/2013/398143

27. Rusk A, Highfield L, Wilkerson JM, Harrell M, Obala A, Amick B. Geographically-weighted regression of knowledge and behaviour determinants to anti-malarial recommending and dispensing practice among medicine retailers in western Kenya: capacitating targeted interventions. *Malar J*. 2016;15(1):1-14. doi:10.1186/s12936-016-1599-5

28. Rusk A, Highfield L, Michael Wilkerson J, Harrell M, Obala A, Amick B. Spatial distribution and cluster analysis of retail drug shop characteristics and antimalarial behaviors as reported by private medicine retailers in western Kenya: Informing future interventions. *Int J Health Geogr*. 2016;15(1):9. doi:10.1186/s12942-016-0038-8

29. Rutto EK, Nyagol J, Oyugi J, Ndege S, Onyango N, Obala A, Simiyu CJ, Boor G, Cheriro WC, Otsyula B, Estambale B. Effects of HIV-1 infection on malaria

parasitemia in milo sub-location, western Kenya. *BMC Res Notes*. 2015;8(1):303. doi:10.1186/s13104-015-1270-1

30. Smith N, Obala A, Simiyu C, Menya D, Khwa-Otsyula B, O'Meara W. Accessibility, availability and affordability of anti-malarials in a rural district in Kenya after implementation of a national subsidy scheme. *Malar J*. 2011;10:316. doi:10.1186/1475-2875-10-316

31. O'Meara WP, Obala A, Thirumurthy H, Khwa-Otsyula B. The association between price, competition, and demand factors on private sector anti-malarial stocking and sales in western Kenya: Considerations for the AMFm subsidy. *Malar J*. 2013;12(1). doi:10.1186/1475-2875-12-186

32. Obala AA, Mangeni JN, Platt A, Aswa D, Abel L, Namae J, O'Meara WP. What is threatening the effectiveness of insecticide-treated bednets? A case-control study of environmental, behavioral, and physical factors associated with prevention failure. *PLoS One*. 2015;10(7). doi:10.1371/journal.pone.0132778

33. Marete IK, Mutugi M, Lagat Z, Obala A, Simba J, Mwangi A, Esamai F. Malaria parasitemia among febrile children infected with human immunodeficiency virus in the context of prophylactic cotrimoxazole as standard of care: a cross-sectional survey in Western Kenya. *East Afr Med J*. 2014;91(1):21-28. Accessed November 26, 2018. http://www.ncbi.nlm.nih.gov/pubmed/26862632

34. World Health Organization. *World Malaria Report 2020*. Vol WHO/HTM/GM.; 2020. doi:ISBN 978 92 4 1564403

35. Snow RW, Sartorius B, Kyalo D, Maina J, Amratia P, Mundia CW, Bejon P, Noor AM. The prevalence of Plasmodium falciparum in sub-Saharan Africa since 1900. *Nature*. 2017;550(7677):515-518. doi:10.1038/nature24059

36. Bousema T, Okell L, Felger I, Drakeley C. Asymptomatic malaria infections: Detectability, transmissibility and public health relevance. *Nat Rev Microbiol*. 2014;12(12):833-840. doi:10.1038/nrmicro3364

37. Centers for Disease Control and Prevention. Malaria: Biology. Published 2018. Accessed February 1, 2019. https://www.cdc.gov/malaria/about/biology/index.html

38. Warrell DA. Clinical features of malaria. In: Warrell DA, Gilles HM, eds. *Essential Malariology*. 4th ed. CRC Press; 2017:352.

39. Ashley EA, White NJ. The duration of Plasmodium falciparum infections. *Malar J*. 2014;13(1):500. doi:10.1186/1475-2875-13-500

40. Sinden RE. The cell biology of malaria infection of mosquito: Advances and opportunities. *Cell Microbiol*. 2015;17(4):451-466. doi:10.1111/cmi.12413

41. Aly ASI, Vaughan AM, Kappe SHI. Malaria Parasite Development in the Mosquito

and Infection of the Mammalian Host. *Annu Rev Microbiol*. 2009;63(1):195-221. doi:10.1146/annurev.micro.091208.073403

42. Boreham PF, Chandler JA, Jolly J. The incidence of mosquitoes feeding on mothers and babies at Kisumu, Kenya. *J Trop Med Hyg*. 1978;81(4):63-67. Accessed March 29, 2019. http://www.ncbi.nlm.nih.gov/pubmed/25970

43. Bousema T, Griffin JT, Sauerwein RW, Smith DL, Churcher TS, Takken W, Ghani A, Drakeley C, Gosling R. Hitting hotspots: Spatial targeting of malaria for control and elimination. *PLoS Med*. 2012;9(1). doi:10.1371/journal.pmed.1001165

44. Churcher TS, Bousema T, Walker M, Drakeley C, Schneider P, Ouédraogo AL, Basáñez MG. Predicting mosquito infection from Plasmodium falciparum gametocyte density and estimating the reservoir of infection. *Elife*. 2013;2013(2). doi:10.7554/eLife.00626

45. Ouédraogo AL, Bousema T, Schneider P, de Vlas SJ, Ilboudo-Sanogo E, Cuzin-Ouattara N, Nábí I, Roeffen W, Verhave JP, Luty AJF, Sauerwein R. Substantial contribution of submicroscopical Plasmodium falciparum gametocyte carriage to the infectious reservoir in an area of seasonal transmission. *PLoS One*. 2009;4(12). doi:10.1371/journal.pone.0008410

46. Drakeley C, Sutherland C, Bousema JT, Sauerwein RW, Targett GAT. The epidemiology of Plasmodium falciparum gametocytes: weapons of mass dispersion. *Trends Parasitol*. 2006;22(9):424-430. doi:10.1016/j.pt.2006.07.001

47. Laishram DD, Sutton PL, Nanda N, Sharma VL, Sobti RC, Carlton JM, Joshi H. The complexities of malaria disease manifestations with a focus on asymptomatic malaria. *Malar J*. 2012;11(1):29. doi:10.1186/1475-2875-11-29

48. Bottius E, Guanzirolli A, Trape JF, Regier C, Konate L, Druilhe P. Malaria: Even more chronic in nature than previously thought; Evidence for subpatent parasitaemia detectable by the polymerase chain reaction. *Trans R Soc Trop Med Hyg*. 1996;90(1):15-19. doi:10.1016/S0035-9203(96)90463-0

49. Eke RA, Chigbu LN, Nwachukwu W. High Prevalence of Asymptomatic Plasmodium Infection in a Suburb of. *Ann Afr Med*. 2006;5(1):4245. Accessed February 19, 2019. http://www.bioline.org.br/pdf?am06010

50. Dal-Bianco MP, Köster KB, Kombila UD, Kun JFJ, Grobusch MP, Ngoma GM, Matsiegui PB, Supan C, Salazar CLO, Missinou MA, Issifou S, Lell B, Kremsner P. High prevalence of asymptomatic Plasmodium falciparum infection in Gabonese adults. *Am J Trop Med Hyg*. 2007;77(5):939-942. Accessed February 19, 2019. http://www.ncbi.nlm.nih.gov/pubmed/17984357

51. De Andrade ALSS, Martelli CMT, Oliveira RM, Arias JR, Zicker F, Pang L. High prevalence of asymptomatic malaria in gold mining areas in Brazil. *Clin Infect Dis*. 1995;20(2):475. doi:10.1093/clinids/20.2.475

52. Lindblade KA, Steinhardt L, Samuels A, Kachur SP, Slutsker L. The silent threat: Asymptomatic parasitemia and malaria transmission. *Expert Rev Anti Infect Ther*. 2013;11(6):623-639. doi:10.1201/b19294

53. Alves FP, Durlacher RUIR, Menezes MJ, Krieger H, Pereira LH, Camargo EP. High prevalence of asymptomatic Plasmodium vivax and Plasmodium falciparum infectiosn in native Amazonian populations. *Am J Trop Med Hyg*. 2002;66(6):641-648. Accessed October 29, 2018. https://www.ncbi.nlm.nih.gov/pubmed/12224567

54. Pinto J, Sousa CA, Gil V, Ferreira C, Gonçalves L, Lopes D, Petrarca V, Charlwood JD, Do Rosário VE. Malaria in Sao Tome and Principe: Parasite prevalences and vector densities. *Acta Trop*. 2000;76(2):185-193. doi:10.1016/S0001-706X(00)00100-5

55. Bousema T, Drakeley C. Epidemiology and infectivity of Plasmodium falciparum and Plasmodium vivax gametocytes in relation to malaria control and elimination. *Clin Microbiol Rev*. 2011;24(2):377-410. doi:10.1128/CMR.00051-10

56. Wampfler R, Timinao L, Beck HP, Soulama I, Tiono AB, Siba P, Mueller I, Felger I. Novel genotyping tools for investigating transmission dynamics of Plasmodium falciparum. *J Infect Dis*. 2014;210(8):1188-1197. doi:10.1093/infdis/jiu236

57. Nilsson SK, Childs LM, Buckee C, Marti M. Targeting Human Transmission Biology for Malaria Elimination. *PLoS Pathog*. 2015;11(6). doi:10.1371/journal.ppat.1004871

58. Stone W, Gonçalves BP, Bousema T, Drakeley C. Assessing the infectious reservoir of falciparum malaria: Past and future. *Trends Parasitol*. 2015;31(7):287-296. doi:10.1016/j.pt.2015.04.004

59. Alves FP, Gil LHS, Marrelli MT, Ribolla PEM, Camargo EP, Da Silva LHP. Asymptomatic Carriers of Plasmodium spp. as Infection Source for Malaria Vector Mosquitoes in the Brazilian Amazon. *J Med Entomol*. 2005;42(5):777-779. doi:10.1093/jmedent/42.5.777

60. Young MD, Hardman NF, Burgess RW, Frohne WC, Sabrosky CW. The infectivity of native malarias in South Carolina to Anopheles quadrimaculatus. *Am J Trop Med Hyg*. 1948;28(2):303-311. Accessed January 28, 2019. http://www.ajtmh.org/docserver/fulltext/14761645/s1-28/2/TMs1-280020303.pdf?expires=1548700062&id=id&accname=12474&checksum=69154703BB369C6BB2A888FA6690FD42

61. Gouagna LC, Ferguson HM, Okech BA, Killeen GF, Kabiru EW, Beier JC, Githure JI, Yan G. Plasmodium falciparum malaria disease manifestations in humans and transmission to Anopheles gambiae: A field study in Western Kenya. *Parasitology*. 2004;128(3):235-243. doi:10.1017/S003118200300444X

62. Tadesse FG, Slater HC, Chali W, Teelen K, Lanke K, Belachew M, Menberu T, Shumie G, Shitaye G, Okell LC, Graumans W, Van Gemert GJ, Kedir S, Tesfaye A, Belachew F, Abebe W, Mamo H, Sauerwein R, Balcha T, Aseffa A, Yewhalaw D, Gadisa E, Drakeley C, Bousema T. The Relative Contribution of Symptomatic and Asymptomatic Plasmodium vivax and Plasmodium falciparum Infections to the Infectious Reservoir in a Low-Endemic Setting in Ethiopia. *Clin Infect Dis*. 2018;66(12):1883-1891. doi:10.1093/cid/cix112

63. Vantaux A, Samreth R, Piv E, Khim N, Kim S, Berne L, Chy S, Lek D, Siv S, Taylor WR, Ménard D. Contribution to Malaria Transmission of Symptomatic and Asymptomatic Parasite Carriers in Cambodia. *J Infect Dis*. 2018;217(10):1561-1568. doi:10.1093/infdis/jiy060

64. Coalson JE, Cohee LM, Buchwald AG, Nyambalo A, Kubale J, Seydel KB, Mathanga D, Taylor TE, Laufer MK, Wilson ML. Simulation models predict that school-age children are responsible for most human-to-mosquito Plasmodium falciparum transmission in southern Malawi. *Malar J*. 2018;17:147. doi:10.1186/s12936-018-2295-4

65. Okell LC, Bousema T, Griffin JT, Ouédraogo AL, Ghani AC, Drakeley CJ. Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. *Nat Commun*. 2012;3:1237. doi:10.1038/ncomms2241

66. Boudin C, Olivier M, Molez JF, Chiron JP, Ambroise-Thomas P. High human malarial infectivity to laboratory-bred Anopheles gambiae in a village in Burkina Faso. *Am J Trop Med Hyg*. 1993;48(5):700-706. doi:10.4269/ajtmh.1993.48.700

67. Schneider P, Teun Bousema J, Gouagna LC, Otieno S, Van De Vegte-Bolmer M, Omar SA, Sauerwein RW. Submicroscopic Plasmodium falciparum gametocyte densities frequently result in mosquito infection. *Am J Trop Med Hyg*. 2007;76(3):470-474. Accessed January 28, 2019. http://www.ajtmh.org/docserver/fulltext/14761645/76/3/0760470.pdf?expires=1548698803&id=id&accname=guest&checksum=9D103B9ABF42527C0EDA2A5F602B7ED0

68. Gonçalves BP, Kapulu MC, Sawa P, Guelbéogo WM, Tiono AB, Grignard L, Stone W, Hellewell J, Lanke K, Bastiaens GJH, Bradley J, Nébié I, Ngoi JM, Oriango R, Mkabili D, Nyaurah M, Midega J, Wirth DF, Marsh K, Churcher TS, Bejon P, Sirima SB, Drakeley C, Bousema T. Examining the human infectious reservoir for Plasmodium falciparum malaria in areas of differing transmission intensity. *Nat Commun*. 2017;8(1):1133. doi:10.1038/s41467-017-01270-4

69. Ouédraogo AL, Gonçalves BP, Gnémé A, Wenger EA, Guelbeogo MW, Ouédraogo A, Gerardin J, Bever CA, Lyons H, Pitroipa X, Verhave JP, Eckhoff PA, Drakeley C, Sauerwein R, Luty AJF, Kouyaté B, Bousema T. Dynamics of the Human Infectious Reservoir for Malaria Determined by Mosquito Feeding Assays and Ultrasensitive Malaria Diagnosis in Burkina Faso. *J Infect Dis*.

2016;213(1):90-99. doi:10.1093/infdis/jiv370

70. Grignard L, Gonçalves BP, Early AM, Daniels RF, Tiono AB, Guelbéogo WM, Ouédraogo A, van Veen EM, Lanke K, Diarra A, Nebie I, Sirima SB, Targett GA, Volkman SK, Neafsey DE, Wirth DF, Bousema T, Drakeley C. Transmission of molecularly undetectable circulating parasite clones leads to high infection complexity in mosquitoes post feeding. *Int J Parasitol*. 2018;48(8):671-677. doi:10.1016/j.ijpara.2018.02.005

71. Muirhead-Thomson RC. The Malarial Infectivity of an African Village Population to Mosquitoes (Anopheles Gambiae). *Am J Trop Med Hyg*. 1957;6(6):971-979. doi:10.4269/ajtmh.1957.6.971

72. Githeko AK, Brandling-Bennett AD, Beier M, Atieli F, Owaga M, Collins FH. The reservoir of Plasmodium falciparum malaria in a holoendemic area of western Kenya. *Trans R Soc Trop Med Hyg*. 1992;86(4):355-358. doi:10.1016/0035-9203(92)90216-Y

73. Bonnet S, Gouagna LC, Paul RE, Safeukui I, Meunier JY, Boudin C. Estimation of malaria transmission from humans to mosquitoes in two neighbouring villages in south Cameroon: Evaluation and comparison of several indices. *Trans R Soc Trop Med Hyg*. 2003;97(1):53-59. doi:10.1016/S0035-9203(03)90022-8

74. Graves PM, Burkot TR, Carter R, Cattani JA, Lagog M, Parker J, Brabin BJ, Gibson FD, Bradley DJ, Alpers MP. Measurement of malarial infectivity of human populations to mosquitoes in the Madang area, Papua New Guinea. *Parasitology*. 1988;96(02):251. doi:10.1017/S003118200005825X

75. Drakeley CJ, Akim NI, Sauerwein RW, Greenwood BM, Targett GA. Estimates of the infectious reservoir of Plasmodium falciparum malaria in The Gambia and in Tanzania. *Trans R Soc Trop Med Hyg*. 2000;94(5):472-476. Accessed March 12, 2019. http://www.ncbi.nlm.nih.gov/pubmed/11132369

76. Gaye A, Bousema T, Libasse G, Ndiath MO, Konaté L, Jawara M, Faye O, Sokhna C. Infectiousness of the human population to Anopheles arabiensis by direct skin feeding in an area hypoendemic for malaria in Senegal. *Am J Trop Med Hyg*. 2015;92(3):648-652. doi:10.4269/ajtmh.14-0402

77. Lin JT, Ubalee R, Lon C, Balasubramanian S, Kuntawunginn W, Rahman R, Saingam P, Heng TK, Vy D, San S, Nuom S, Burkly H, Chanarat N, Ponsa C, Levitz L, Parobek C, Chuor CM, Somethy S, Spring M, Lanteri C, Gosi P, Meshnick SR, Saunders DL. Microscopic Plasmodium falciparum Gametocytemia and Infectivity to Mosquitoes in Cambodia. *J Infect Dis*. 2016;213(9):1491-1494. doi:10.1093/infdis/jiv599

78. Newell K, Kiggundu V, Ouma J, Baghendage E, Kiwanuka N, Gray R, Serwadda D, Hobbs C V, Healy SA, Quinn TC, Reynolds SJ. Longitudinal household surveillance for malaria in Rakai, Uganda. *Malar J*. 2016;15(1):77.

doi:10.1186/s12936-016-1128-6

79. Djimdé AA, Doumbo OK, Traore O, Guindo AB, Kayentao K, Diourte Y, Niare-Doumbo S, Coulibaly D, Kone AK, Cissoko Y, Tekete M, Fofana B, Dicko A, Diallo DA, Wellems TE, Kwiatkowski D, Plowe C V. Clearance of drug-resistant parasites as a model for protective immunity in Plasmodium falciparum malaria. *Am J Trop Med Hyg*. 2003;69(5):558-563. Accessed February 19, 2019. http://www.ajtmh.org/docserver/fulltext/14761645/69/5/0690558.pdf?expires=1550632573&id=id&accname=guest&checksum=07CB49B24F02ABBDFB31C81DF2F1D1C5

80. Snow RW, Omumbo JA, Lowe B, Molyneux CS, Obiero J-O, Palmer A, Weber MW, Pinder M, Nahlen B, Obonyo C, Newbold C, Gupta S, Marsh K. Relation between severe malaria morbidity in children and level of Plasmodium falciparum transmission in Africa. *Lancet*. 1997;349(9066):1650-1654. doi:10.1016/S0140-6736(97)02038-2

81. Tran TM, Li S, Doumbo S, Doumtabe D, Huang CY, Dia S, Bathily A, Sangala J, Kone Y, Traore A, Niangaly M, Dara C, Kayentao K, Ongoiba A, Doumbo OK, Traore B, Crompton PD. An intensive longitudinal cohort study of malian children and adults reveals no evidence of acquired immunity to plasmodium falciparum infection. *Clin Infect Dis*. 2013;57(1):40-47. doi:10.1093/cid/cit174

82. al-Yaman F, Genton B, Reeder JC, Anders RF, Smith T, Alpers MP. Reduced risk of clinical malaria in children infected with multiple clones of Plasmodium falciparum in a highly endemic area: a prospective community study. *Trans R Soc Trop Med Hyg*. 1997;91(5):602-605. doi:10.1016/S0035-9203(97)90046-8

83. Bereczky S, Liljander A, Rooth I, Faraja L, Granath F, Montgomery SM, Färnert A. Multiclonal asymptomatic Plasmodium falciparum infections predict a reduced risk of malaria disease in a Tanzanian population. *Microbes Infect*. 2007;9(1):103-110. doi:10.1016/j.micinf.2006.10.014

84. Färnert A, Williams TN, Mwangi TW, Ehlin A, Fegan G, Macharia A, Lowe BS, Montgomery SM, Marsh K. Transmission-Dependent Tolerance to Multiclonal Plasmodium falciparum Infection Europe PMC Funders Group. *J Infect Dis*. 2009;(7):1166-1175. doi:10.1086/605652

85. Mayor A, Saute F, Aponte JJ, Almeda J, Gómez-Olivé FX, Dgedge M, Alonso PL. Plasmodium falciparum multiple infections in Mozambique, its relation to other malariological indices and to prospective risk of malaria morbidity. *Trop Med Int Heal*. 2003;8(1):3-11. doi:10.1046/j.1365-3156.2003.00968.x

86. Roper C, Richardson W, Elhassan IM, Giha H, Hviid L, Satti GMH, Theander TG, Arnot DE. Seasonal changes in the Plasmodium falciparum population in individuals and their relationship to clinical malaria: A longitudinal study in a Sudanese village. *Parasitology*. 1998;116(6):501-510. doi:10.1017/S0031182098002650

87. Felger I, Maire M, Bretscher MT, Falk N, Tiaden A, Sama W, Beck HP, Owusu-Agyei S, Smith TA. The Dynamics of Natural Plasmodium falciparum Infections. *PLoS One*. 2012;7(9). doi:10.1371/journal.pone.0045542

88. Barry AE, Schultz L, Senn N, Nale J, Kiniboro B, Siba PM, Mueller I, Reeder JC. High levels of genetic diversity of Plasmodium falciparum populations in Papua New Guinea despite variable infection prevalence. *Am J Trop Med Hyg*. 2013;88(4):718-725. doi:10.4269/ajtmh.12-0056

89. Daubersies P, Sallenave-Sales S, Magne S, Trape JF, Contamin H, Fandeur T, Rogier C, Mercereau-Puijalon O, Druilhe P. Rapid turnover of Plasmodium falciparum populations in asymptomatic individuals living in a high transmission area. *Am J Trop Med Hyg*. 1996;54(1):18-26. doi:10.4269/ajtmh.1996.54.18

90. Färnert A, Rooth I, Svensson Å, Snounou G, Björkman A. Complexity of Plasmodium falciparum Infections Is Consistent over Time and Protects against Clinical Disease in Tanzanian Children. *J Infect Dis*. 1999;179(4):989-995. doi:10.1086/314652

91. Lines J, Armstrong JRM. For a few parasites more: Inoculum size, vector control and strain-specific immunity to malaria. *Parasitol Today*. 1992;8(11):381-383. doi:10.1016/0169-4758(92)90176-3

92. Marsh K, Howard RJ. Antigens induced on erythrocytes by P. falciparum: expression of diverse and conserved determinants. *Science*. 1986;231(4734):150-153. Accessed January 28, 2019. http://www.ncbi.nlm.nih.gov/pubmed/2417315

93. Giha HA, Theander TG, StaalsØ T, Roper C, Elhassan IM, Babiker H, Satti GMH, Arnot DE, Hviid L. Seasonal variation in agglutination of Plasmodium falciparum-infected erythrocytes. *Am J Trop Med Hyg*. 1998;58(4):399-405. doi:10.4269/ajtmh.1998.58.399

94. Gupta S, Day KP. A strain theory of malaria transmission. *Parasitol Today*. 1994;10(12):476-481. doi:10.1016/0169-4758(94)90160-0

95. Ntoumi F, Bakoua D, Fesser A, Kombo M, Vouvoungui J C, Koukouikila Koussounda F. Characterization of asymptomatic Plasmodium falciparum infection and its risk factors in pregnant women from the Republic of Congo. *Acta Trop*. 1995;153:111-115. doi:10.1016/j.actatropica.2015.10.009

96. Takala SL, Coulibaly D, Thera MA, Batchelor AH, Cummings MP, Escalante AA, Ouattara A, Traoré K, Niangaly A, Djimdé AA, Doumbo OK, Plowe C V. Extreme polymorphism in a vaccine antigen and risk of clinical malaria: Implications for vaccine development. *Sci Transl Med*. 2009;1(2):2-5. doi:10.1126/scitranslmed.3000257

97. Contamin H, Fandeur T, Rogier C, Bonnefoy S, Konate L, Trape JF, Mercereau-

Puijalon O. Different genetic characteristics of Plasmodium falciparum isolates collected during successive clinical malaria episodes in Senegalese children. *Am J Trop Med Hyg*. 1996;54(6):632-643. doi:10.4269/ajtmh.1996.54.632

98. Kun JFJ, Missinou MA, Lell B, Sovric M, Knoop H, Bojowald B, Dangelmaier O, Kremsner PG. *New Emerging Plasmodium Falciparum Genotypes in Children during the Tranisition Phase from Asymptomatic Parasitemia to Malaria*.; 2002. Accessed February 17, 2019. http://www.ajtmh.org/docserver/fulltext/14761645/66/6/12224569.pdf?expires=1550443674&id=id&accname=guest&checksum=4AAC187957B4AD64E4AFED9F79A16CE7

99. Babiker HA, Abdel-Muhsin A-MA, Ranford-Cartwright LC, Satti G, Walliker D. *Characteristics of Plasmodium Falciparum Parasites That Survive the Lengthy Dry Season in Eastern Sudan Where Malaria Transmission Is Markedly Seasonal*. Vol 59.; 1998. Accessed February 17, 2019. http://www.ajtmh.org/docserver/fulltext/14761645/59/4/9790434.pdf?expires=1550444163&id=id&accname=guest&checksum=16939BE4C1D869CDCF13D90006E9BB4C

100. Chen I, Clarke SE, Gosling R, Hamainza B, Killeen G, Magill A, O'Meara W, Price RN, Riley EM. "Asymptomatic" Malaria: A Chronic and Debilitating Infection That Should Be Treated. *PLOS Med*. 2016;13(1):e1001942. doi:10.1371/journal.pmed.1001942

101. Maketa V, Mavoko HM, da Luz RI, Zanga J, Lubiba J, Kalonji A, Lutumba P, Van geertruyden J-P. The relationship between Plasmodium infection, anaemia and nutritional status in asymptomatic children aged under five years living in stable transmission zones in Kinshasa, Democratic Republic of Congo. *Malar J*. 2015;14(1):83. doi:10.1186/s12936-015-0595-5

102. Sifft KC, Geus D, Mukampunga C, Mugisha JC, Habarugira F, Fraundorfer K, Bayingana C, Ndoli J, Umulisa I, Karema C, von Samson-Himmelstjerna G, Aebischer T, Martus P, Sendegeya A, Gahutu JB, Mockenhaupt FP. Asymptomatic only at first sight: malaria infection among schoolchildren in highland Rwanda. *Malar J*. 2016;15(1):553. doi:10.1186/s12936-016-1606-x

103. de Mast Q, Brouwers J, Syafruddin D, Bousema T, Baidjoe AY, de Groot PG, van der Ven AJ, Fijnheer R. Is asymptomatic malaria really asymptomatic? Hematological, vascular and inflammatory effects of asymptomatic malaria parasitemia. *J Infect*. 2015;71(5):587-596. doi:10.1016/j.jinf.2015.08.005

104. Cottrell G, Moussiliou A, Luty AJF, Cot M, Fievet N, Massougbodji A, Deloron P, Tuikue Ndam N. Submicroscopic Plasmodium falciparum Infections Are Associated With Maternal Anemia, Premature Births, and Low Birth Weight. *Clin Infect Dis*. 2015;60(10):1481-1488. doi:10.1093/cid/civ122

105. Matangila JR, Lufuluabo J, Ibalanky AL, Inocêncio da Luz RA, Lutumba P, Van

Geertruyden J-P. Asymptomatic Plasmodium falciparum infection is associated with anaemia in pregnancy and can be more cost-effectively detected by rapid diagnostic test than by microscopy in Kinshasa, Democratic Republic of the Congo. *Malar J*. 2014;13(1):132. doi:10.1186/1475-2875-13-132

106. Tagbor H, Bruce J, Browne E, Greenwood B, Chandramohan D. Malaria in pregnancy in an area of stable and intense transmission: is it asymptomatic? *Trop Med Int Heal*. 2008;13(8):1016-1021. doi:10.1111/j.1365-3156.2008.02111.x

107. Mabey DCW, Brown A, Greenwood BM. Plasmodium falciparum malaria and salmonella infections in gambian children. *J Infect Dis*. 1987;155(6):1319-1321. doi:10.1093/infdis/155.6.1319

108. Nankabirwa J, Wandera B, Kiwanuka N, Staedke SG, Kamya MR, Brooker SJ. Asymptomatic Plasmodium infection and cognition among primary schoolchildren in a high malaria transmission setting in Uganda. *Am J Trop Med Hyg*. 2013;88(6):1102-1108. doi:10.4269/ajtmh.12-0633

109. Hernán MA, Robins JM, García Rodríguez LA. Discussion on “statistical issues arising in the Women’s Health Initiative.” *Biometrics*. 2005;61(4):922-930. doi:10.1111/j.0006-341X.2005.454_6.x

110. Juliano JJ, Porter K, Mwapasa V, Sem R, Rogers WO, Ariey F, Wongsrichanalai C, Read A, Meshnick SR. Exposing malaria in-host diversity and estimating population diversity by capture-recapture using massively parallel pyrosequencing. *Proc Natl Acad Sci*. 2010;107(46):20138-20143. doi:10.1073/pnas.1007068107

111. Levitt B, Obala A, Langdon S, Corcoran D, O’Meara WP, Taylor SM. Overlap extension barcoding for the next generation sequencing and genotyping of plasmodium falciparum in individual patients in Western Kenya. *Sci Rep*. 2017;7(January):1-11. doi:10.1038/srep41108

112. Mburu MM, Mzilahowa T, Amoah B, Chifundo D, Phiri KS, van den Berg H, Takken W, McCann RS. Biting patterns of malaria vectors of the lower Shire valley, southern Malawi. *Acta Trop*. 2019;197(June):105059. doi:10.1016/j.actatropica.2019.105059

113. Murray L, Stewart LB, Tarr SJ, Ahouidi AD, Diakite M, Amambua-Ngwa A, Conway DJ. Multiplication rate variation in the human malaria parasite Plasmodium falciparum. *Sci Rep*. 2017;7(1):1-8. doi:10.1038/s41598-017-06295-9

114. World Bank. *Kenya County Fact Sheet*.; 2011. Accessed January 20, 2019. http://siteresources.worldbank.org/INTAFRICA/Resources/257994-1335471959878/Kenya_County_Fact_Sheets_Dec2011.pdf

115. AccessBio. CareStart Malaria. Published online 2019.

http://www.accessbio.net/eng/products/products01_02.asp

116. John W. Hock Company. Improved Prokopack Aspirator. Published online 2019. https://johnwhock.com/products/aspirators/improved-prokopack-aspirator/

117. Illumina. Illumnia Miseq. Published 2018. Accessed December 5, 2018. https://www.illumina.com/systems/sequencing-platforms/miseq.html

118. Bushnell B. BBMap short-read aligner, and other bioinformatics tools. Published online 2015.

119. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-2120. doi:10.1093/bioinformatics/btu170

120. Martin M. *Embnet.News : European Molecular Biology Network Newsletter.* Vol 17. EMBnet, Administration Office; 2011. Accessed January 20, 2019. https://journal.embnet.org/index.php/embnetjournal/article/view/200/479

121. Gardner MJ, Hall N, Fung E, White O, Berriman M, Richard W, Carlton JM, Pain A, Nelson KE, Bowman S, Ian T, James K, Eisen JA, Rutherford K, Salzberg SL, Kyes S, Chan M, Nene V, Shallom SJ, Peterson J, Angiuoli S, Pertea M, Allen J, Haft D, Mather MW, Vaidya AB, Martin DMA, Alan H. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature*. 2002;419(6906):498-511. doi:10.1038/nature01097.Genome

122. Hall N, Pain A, Berriman M, Churcher B, Harris D, Harris K. Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13. *Nature*. 2002;419:527-531.

123. Callahan B, McMurdie P, Rosen M, Han A, AJA J, SP H. High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581-583.

124. Early AM, Daniels RF, Farrell TM, Grimsby J, Volkman SK, Wirth DF, MacInnis BL, Neafsey DE. Detection of low-density Plasmodium falciparum infections using amplicon deep sequencing. *Malar J*. 2019;18(1):219. doi:10.1186/s12936-019-2856-1

125. R Core Team. R: A language and environment for statistical computing. Published online 2020. https://www.r-project.org/

126. R Core Team. R: A language and environment for statistical computing. Published online 2019. https://www.r-project.org/

127. O'Meara WP, Simmons R, Bullins P, Freedman B, Abel L, Mangeni J, Taylor SM, Obala AA. Mosquito exposure and malaria morbidity; a micro-level analysis of household mosquito populations and malaria in a population-based longitudinal cohort in western Kenya. *J Infect Dis*. 2019;27708:1-9. doi:10.1093/infdis/jiz561

128. Plowe C V., Djimde A, Bouare M, Doumbo O, Wellems TE. Pyrimethamine and proguanil resistance-conferring mutations in Plasmodium falciparum dihydrofolate reductase: Polymerase chain reaction methods for surveillance in Africa. *Am J Trop Med Hyg*. 1995;52(6):565-568. doi:10.4269/ajtmh.1995.52.565

129. Taylor SM, Sumner KM, Freedman B, Mangeni JN, Obala AA, O'Meara WP. Direct estimation of sensitivity of Plasmodium falciparum rapid diagnostic test for active case detection in a high-transmission community setting. *Am J Trop Med Hyg*. 2019;101(6):1416-1423. doi:10.4269/ajtmh.19-0558

130. Wickham H, Averick M, Bryan J, Chang W, D'Agostino McGowan L, Francois R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Muller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughn D, Wilke C, Woo K, Yutani H. Welcome to the tidyverse. *J Open Source Softw*. 2019;4(43):1686. doi:10.21105/joss.01686

131. Kassambara A, Kosinski M, Biecek P. survminer: Drawing survival curves using "ggplot2". R package version 0.4.8. Published online 2020.

132. Therneau TM, Grambsch PM. A package for survival analysis in R. R package version 3.2.3. Published online 2000.

133. Therneau TM. coxme: Mixed effects cox models. R package version 2.2-16. Published online 2020.

134. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1):1-48. doi:10.18637/jss.v067.i01

135. Brunson JC. ggalluvial: Layered grammar for alluvial plots. *J Open Source Softw*. 2020;5(49):2017. doi:doi.org/10.21105/joss.02017

136. Slater HC, Ross A, Felger I, Hofmann NE, Robinson L, Cook J, Gonçalves BP, Björkman A, Ouedraogo AL, Morris U, Msellem M, Koepfli C, Mueller I, Tadesse F, Gadisa E, Das S, Domingo G, Kapulu M, Midega J, Owusu-Agyei S, Nabet C, Piarroux R, Doumbo O, Doumbo SN, Koram K, Lucchi N, Udhayakumar V, Mosha J, Tiono A, Chandramohan D, Gosling R, Mwingira F, Sauerwein R, Riley EM, White NJ, Nosten F, Imwong M, Bousema T, Drakeley C, Okell LC. The temporal dynamics and infectiousness of subpatent Plasmodium falciparum infections in relation to parasite density. *Nat Commun*. 2019;10(1). doi:10.1038/s41467-019-09441-1

137. Houngbedji CA, Dri PBN, Hürlimann E, Yapi RB, Silué KD, Soro G, Koudou BG, Acka CA, Assi S, Vounatsou P, Goran EKN, Fantodji A, Utzinger J, Raso G. Disparities of Plasmodium falciparum infection,among school-aged children In Cote D'Ivoire. *Malar J*. 2015;14(7).

138. Mulu A, Legesse M, Erko B, Belyhun Y, Nugussie D, Shimelis T, Kassu A, Elias D, Moges B. Epidemiological and clinical correlates of malaria-helminth co-

infections in southern Ethiopia. *Malar J*. 2013;12(1):1-7. doi:10.1186/1475-2875-12-227

139. Briggs J, Teyssier N, Nankabirwa JI, Rek J, Jagannathan P, Arinatiwe E, Bousema T, Drakeley C, Murray M, Crawford E, Hathaway N, Staedke SG, Smith D, Rosenthal PJ, Kamya M, Dorsey G, Rodriguez-Barraquer I, Greenhouse B. Sex-based differences in clearance of chronic Plasmodium falciparum infection. *Elife*. 2020;9:e59872. doi:10.1101/2020.06.10.20127720

140. Hernán MA. The hazards of hazard ratios. *Epidemiology*. 2010;21(1):13-15. doi:10.1097/EDE.0b013e3181c1ea43.The

141. Portugal S, Carret C, Recker M, Armitage A, Gonçalves L, Epiphanio S, Sullivan D, Roy C, Newbold C, Drakesmith H, Mora M. Host mediated regulation of superinfection in malaria Silvia. 2011;17(6):732-737. doi:10.1038/nm.2368.Host

142. Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, Manson JE, Robins JM. Observational studies analyzed like randomized experiments: An application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology*. 2008;19(6):766-779. doi:10.1097/EDE.0b013e3181875e61

143. Danaei G, Rodríguez LAG, Cantero OF, Logan R, Hernán MA. Observational data for comparative effectiveness research: An emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res*. 2013;22(1):70-96. doi:10.1177/0962280211403603

144. Secora AM, Shin JI, Qiao Y, Alexander GC, Chang AR, Inker LA, Coresh J, Grams ME. Hyperkalemia and Acute Kidney Injury with Spironolactone Use Among Patients with Heart Failure. *Mayo Clin Proc*. 2020;95(11):2408-2419. doi:10.1016/j.mayocp.2020.03.035

145. World Health Organisation (WHO). *World Malaria Report 2019*.; 2019.

146. Adomako-Ankomah Y, Chenoweth MS, Durfee K, Doumbia S, Konate D, Doumbouya M, Keita AS, Nikolaeva D, Tullo GS, Anderson JM, Fairhurst RM, Daniels R, Volkman SK, Diakite M, Miura K, Long CA. High Plasmodium falciparum longitudinal prevalence is associated with high multiclonality and reduced clinical malaria risk in a seasonal transmission area of Mali. *PLoS One*. 2017;12(2). doi:10.1371/journal.pone.0170948

147. Smith T, Felger I, Tanner M, Beck HP. The epidemiology of multiple Plasmodium falciparum infections 11. Premunition in Plasmodium falciparum infection: Insights from the epidemiology of multiple infections. *Trans R Soc Trop Med Hyg*. 1999;93(SUPPL. 1):59-64. doi:10.1016/S0035-9203(99)90329-2

148. Nelson CS, Sumner KM, Freedman E, Saelens JW, Obala AA, Mangeni JN, Taylor SM, O'Meara WP. High-resolution micro-epidemiology of parasite spatial and temporal dynamics in a high malaria transmission setting in Kenya. *Nat*

Commun. 2019;10(1). doi:10.1038/s41467-019-13578-4

149. Sumner KM, Freedman E, Abel L, Obala A, Pence BW, Wesolowski A, Meshnick SR, O'Meara WP, Taylor SM. Genotyping cognate Plasmodium falciparum in humans and mosquitoes to estimate onward transmission of asymptomatic infections. *Nat Commun*. 2021;12(909):1-12. doi:10.1038/s41467-021-21269-2

150. Boutlis CS, Yeo TW, Anstey NM. Europe PMC Funders Group Malaria tolerance – for whom the cell tolls ? 2006;22(8):371-377. doi:10.1016/j.pt.2006.06.002.Malaria

151. Read AF, Taylor LH. The ecology of genetically diverse infections. *Science (80-)*. 2001;292(5519):1099-1102. doi:10.1126/science.1059410

152. Sergent ED, Parrot L. L'immunité, la prémunition et la résistance innée. *Arch l'Institut Pasteur d'Algérie*. 1935;13:279.

153. Chaorattanakawee S, Nuchnoi P, Hananantachai H, Tumkosit U, Saunders D, Naka I, Ohashi J, Patarapotikul J. Sequence variation in Plasmodium falciparum merozoite surface protein-2 is associated with virulence causing severe and cerebral malaria. *PLoS One*. 2018;13(1):1-18. doi:10.1371/journal.pone.0190418

154. Snow RW, Marsh K. The consequences of reducing transmission of Plasmodium falciparum in Africa. *Adv Parasitol*. 2002;52:235-264. doi:10.1016/S0065-308X(02)52013-3

155. Daniels RF, Schaffner SF, Wenger EA, Proctor JL, Chang HH, Wong W, Baro N, Ndiaye D, Fall FB, Ndiop M, Ba M, Milner DA, Taylor TE, Neafsey DE, Volkman SK, Eckhoff PA, Hartl DL, Wirth DF. Modeling malaria genomics reveals transmission decline and rebound in Senegal. *Proc Natl Acad Sci U S A*. 2015;112(22):7067-7072. doi:10.1073/pnas.1505691112

156. Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, Phyo AP, Nosten F, Anderson TJC. Population genetic correlates of declining transmission in a human pathogen. *Mol Ecol*. 2013;22(2):273-285. doi:10.1111/mec.12099

157. Dantzler KW, Ma S, Ngotho P, Stone WJR, Tao D, Rijpma S, De Niz M, Nilsson Bark SK, Jore MM, Raaijmakers TK, Early AM, Ubaida-Mohien C, Lemgruber L, Campo JJ, Teng AA, Le TQ, Walker CL, Hermand P, Deterre P, Davies DH, Felgner P, Morlais I, Wirth DF, Neafsey DE, Dinglasan RR, Laufer M, Huttenhower C, Seydel K, Taylor T, Bousema T, Marti M. Naturally acquired immunity against immature Plasmodium falciparum gametocytes. *Sci Transl Med*. 2019;11(495). doi:10.1126/scitranslmed.aav3963

158. Lin JT, Saunders DL, Meshnick SR. The role of submicroscopic parasitemia in malaria transmission: What is the evidence? *Trends Parasitol*. 2014;30(4):183-190. doi:10.1016/j.pt.2014.02.004

159. Talley AK, Healy SA, Finney OC, Murphy SC, Kublin J, Salas CJ, Lundebjerg S, Gilbert P, Van Voorhis WC, Whisler J, Wang R, Ockenhouse CF, Heppner DG, Kappe SH, Duffy PE. Safety and comparability of controlled human plasmodium falciparum infection by mosquito bite in malaria-naïve subjects at a new facility for sporozoite challenge. *PLoS One*. 2014;9(11):1-10. doi:10.1371/journal.pone.0109654

160. Thomas CJ, Cross DE, Bøgh C. Landscape Movements of Anopheles gambiae Malaria Vector Mosquitoes in Rural Gambia. *PLoS One*. 2013;8(7). doi:10.1371/journal.pone.0068679

161. Midega JT, Mbogo CM, Mwambi H, Wilson MD, Ojwang G, Mwangangi JM, Nzovu JG, Githure JI, Yan G, Beier JC. Estimating Dispersal and Survival of Anopheles gambiae and Anopheles funestus Along the Kenyan Coast by Using Mark–Release–Recapture Methods . *J Med Entomol*. 2007;44(6):923-929. doi:10.1093/jmedent/44.6.923

162. Neafsey DE, Juraska M, Bedford T, Benkeser D, Valim C, Griggs A, Lievens M, Abdulla S, Adjei S, Agbenyega T, Agnandji ST, Aide P, Anderson S, Ansong D, Aponte JJ, Asante KP, Bejon P, Birkett AJ, Bruls M, Connolly KM, D'Alessandro U, Dobaño C, Gesase S, Greenwood B, Grimsby J, Tinto H, Hamel MJ, Hoffman I, Kamthunzi P, Kariuki S, Kremsner PG, Leach A, Lell B, Lennon NJ, Lusingu J, Marsh K, Martinson F, Molel JT, Moss EL, Njuguna P, Ockenhouse CF, Ogutu BR, Otieno W, Otieno L, Otieno K, Owusu-Agyei S, Park DJ, Pellé K, Robbins D, Russ C, Ryan EM, Sacarlal J, Sogoloff B, Sorgho H, Tanner M, Theander T, Valea I, Volkman SK, Yu Q, Lapierre D, Birren BW, Gilbert PB, Wirth DF. Genetic Diversity and Protective Efficacy of the RTS,S/AS01 Malaria Vaccine. *N Engl J Med*. 2015;373(21):2025-2037. doi:10.1056/NEJMoa1505819

163. Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, Heiges M, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Miller JA, Nayak V, Pennington C, Pinney DF, Roos DS, Ross C, Stoeckert CJ, Treatman C, Wang H. PlasmoDB: A functional genomic database for malaria parasites. *Nucleic Acids Res*. 2009;37(SUPPL. 1):539-543. doi:10.1093/nar/gkn814

164. MalariaGEN Resource Centre. Pf3k: MalariaGen.

165. Kenea O, Balkew M, Tekie H, Gebre-Michael T, Deressa W, Loha E, Lindtjørn B, Overgaard HJ. Human-biting activities of Anopheles species in south-central Ethiopia. *Parasites and Vectors*. 2016;9(1):1-12. doi:10.1186/s13071-016-1813-x

166. Ossè RA, Tokponnon F, Padonou GG, Sidick A, Aïkpon R, Fassinou A, Koukpo CZ, Sèwadé W, Akinro B, Sovi A, Aïssi M, Akogbéto MC. Involvement of Anopheles nili in Plasmodium falciparum transmission in North Benin. *Malar J*. 2019;18(1):1-10. doi:10.1186/s12936-019-2792-0

167. Buckling A, Ranford-Cartwright LC, Miles A, Read AF. Chloroquine increases

Plasmodium falciparum gametocytogenesis in vitro. *Parasitology*. 1999;118(4):339-346. doi:10.1017/s0031182099003960

168. McCann RS, Messina JP, MacFarlane DW, Bayoh MN, Gimnig JE, Giorgi E, Walker ED. Explaining variation in adult Anopheles indoor resting abundance: The relative effects of larval habitat proximity and insecticide-treated bed net use. *Malar J*. 2017;16(1):61-63. doi:10.1186/s12936-017-1938-1

169. Do M, Babalola S, Awantang G, Toso M, Lewicky N, Tompsett A. Associations between malaria-related ideational factors and care-seeking behavior for fever among children under five in Mali, Nigeria, and Madagascar. *PLoS One*. 2018;13(1):1-15. doi:10.1371/journal.pone.0191079

170. Bradley J, Stone W, Da DF, Morlais I, Dicko A, Cohuet A, Guelbeogo WM, Mahamar A, Nsango S, Soumaré HM, Diawara H, Lanke K, Graumans W, Siebelink-Stoter R, van de Vegte-Bolmer M, Chen I, Tiono A, Gonçalves BP, Gosling R, Sauerwein RW, Drakeley C, Churcher TS, Bousema T. Predicting the likelihood and intensity of mosquito infection from sex specific plasmodium falciparum gametocyte density. *Elife*. 2018;7. doi:10.7554/eLife.34463

171. Bailey JA, Mvalo T, Aragam N, Weiser M, Congdon S, Kamwendo D, Martinson F, Hoffman I, Meshnick SR, Juliano JJ. Use of massively parallel pyrosequencing to evaluate the diversity of and selection on plasmodium falciparum csp T-cell epitopes in Lilongwe, Malawi. *J Infect Dis*. 2012;206(4):580-587. doi:10.1093/infdis/jis329

172. Pringle JC, Carpi G, Almagro-Garcia J, Zhu SJ, Kobayashi T, Mulenga M, Bobanga T, Chaponda M, Moss WJ, Norris DE. RTS,S/AS01 malaria vaccine mismatch observed among Plasmodium falciparum isolates from southern and central Africa and globally. *Sci Rep*. 2018;8(1). doi:10.1038/s41598-018-24585-8

173. Pringle JC, Wesolowski A, Berube S, Kobayashi T, Gebhardt ME, Mulenga M, Chaponda M, Bobanga T, Juliano JJ, Meshnick S, Moss WJ, Carpi G, Norris DE. High Plasmodium falciparum genetic diversity and temporal stability despite control efforts in high transmission settings along the international border between Zambia and the Democratic Republic of the Congo. *Malar J*. 2019;18(1):1-13. doi:10.1186/s12936-019-3023-4

174. Osier FHA, Weedall GD, Verra F, Murungi L, Tetteh KKA, Bull P, Faber BW, Remarque E, Thomas A, Marsh K, Conway DJ. Allelic diversity and naturally acquired allele-specific antibody responses to Plasmodium falciparum apical membrane antigen 1 in Kenya. *Infect Immun*. 2010;78(11):4625-4633. doi:10.1128/IAI.00576-10

175. Taylor SM, Parobek CM, De Conti DK, Kayentao K, Coulibaly SO, Greenwood BM, Tagbor H, Williams J, Bojang K, Njie F, Desai M, Kariuki S, Gutman J, Mathanga DP, Mårtensson A, Ngasala B, Conrad MD, Rosenthal PJ, Tshefu AK, Moormann AM, Vulule JM, Doumbo OK, Ter Kuile FO, Meshnick SR, Bailey JA,

Juliano JJ. Absence of putative artemisinin resistance mutations among Plasmodium falciparum in sub-Saharan Africa: A molecular epidemiologic study. *J Infect Dis*. 2015;211(5):680-688. doi:10.1093/infdis/jiu467

176. Rosenberg R. Malaria: some considerations regarding parasite productivity. *Trends Parasitol*. 2008;24(11):487-491. doi:10.1016/j.pt.2008.07.009

177. Gnémé A, Guelbéogo WM, Riehle MM, Sanou A, Traoré A, Zongo S, Eiglmeier K, Kabré GB, Sagnon N, Vernick KD. Equivalent susceptibility of Anopheles gambiae M and S molecular forms and Anopheles arabiensis to Plasmodium falciparum infection in Burkina Faso. *Malar J*. 2013;12(1). doi:10.1186/1475-2875-12-204

178. Balasubramanian S, Rahman RS, Lon C, Parobek C, Ubalee R, Hathaway N, Kuntawunginn W, My M, Vy D, Saxe J, Lanteri C, Lin FC, Spring M, Meshnick SR, Juliano JJ, Saunders DL, Lin JT. Efficient Transmission of Mixed Plasmodium falciparum/vivax Infections From Humans to Mosquitoes. *J Infect Dis*. 2020;221(3):428-437. doi:10.1093/infdis/jiz388

179. Scott TW, Takken W. Feeding strategies of anthropophilic mosquitoes result in increased risk of pathogen transmission. *Trends Parasitol*. 2012;28(3):114-121. doi:10.1016/j.pt.2012.01.001

180. Scott TW, Githeko AK, Fleisher A, Harrington LC, Yan G. DNA profiling of human blood in anophelines from lowland and highland sites in western Kenya. *Am J Trop Med Hyg*. 2006;75(2):231-237. doi:75/2/231 [pii]

181. Gillies MT. The recognition of age-groups within populations of Anopheles gambiae by the pre-gravid rate and the sporozoite rate. *Ann Trop Med Parasitol*. 1954;48:58-74.

182. Koella JC, Sorensen FL, Anderson RA. The malaria parasite, Plasmodium falciparum, increases the frequency of multiple feeding of its mosquito vector, Anopheles gambiae. *Proc R Soc B Biol Sci*. 1998;265(1398):763-768. doi:10.1098/rspb.1998.0358

183. Baliraine FN, Afrane YA, Amenya DA, Bonizzoni M, Menge DM, Zhou G, Zhong D, Vardo-Zalik AM, Githeko AK, Yan G. High Prevalence of Asymptomatic Plasmodium falciparum Infections in a Highland Area of Western Kenya: A Cohort Study. *J Infect Dis*. 2009;200(1):66-74. doi:10.1086/599317

184. Sturrock HJW, Hsiang MS, Cohen JM, Smith DL, Greenhouse B, Bousema T, Gosling RD. Targeting Asymptomatic Malaria Infections: Active Surveillance in Control and Elimination. *PLoS Med*. 2013;10(6):e1001467. doi:10.1371/journal.pmed.1001467

185. World Health Organization (WHO). Global technical strategy for malaria 2016-2030. Published online 2015.

186. Helb DA, Tetteh KKA, Felgner PL, Skinner J, Hubbard A, Arinaitwe E, Mayanja-Kizza H, Ssewanyana I, Kamya MR, Beeson JG, Tappero J, Smith DL, Crompton PD, Rosenthal PJ, Dorsey G, Drakeley CJ, Greenhouse B, Rayner JC. Novel serologic biomarkers provide accurate estimates of recent Plasmodium falciparum exposure for individuals and communities. *Proc Natl Acad Sci U S A*. 2015;112(32):E4438-E4447. doi:10.1073/pnas.1501705112

187. Cohen JL, Leslie HH, Saran I, Fink G. Quality of clinical management of children diagnosed with malaria: A cross-sectional assessment in 9 sub-Saharan African countries between 2007-2018. *PLoS Med*. 2020;17(9):e1003254. doi:10.1371/journal.pmed.1003254

188. Boyce RM, Hathaway N, Fulton T, Reyes R, Matte M, Ntaro M, Mulogo E, Waltmann A, Bailey JA, Siedner MJ, Juliano JJ. Reuse of malaria rapid diagnostic tests for amplicon deep sequencing to estimate Plasmodium falciparum transmission intensity in western Uganda. *Sci Rep*. 2018;8(1):1-10. doi:10.1038/s41598-018-28534-3

189. Billingsley P, Binka F, Chaccour C, Foy BD, Gold S, Gonzalez-Silva M, Jacobson J, Jagoe G, Jones C, Kachur P, Kobylinski K, Last A, Lavery J V., Mabey D, Mboera L, Mbogo C, Mendez-Lopez A, Rabinovich NR, Rees S, Richards F, Rist C, Rockwood J, Ruiz-Castillo P, Sattabongkot J, Saute F, Slater H, Steer A, Xia K, Zulliger R. A roadmap for the development of ivermectin as a complementary malaria vector control tool. *Am J Trop Med Hyg*. 2020;102(Suppl 2):3-24. doi:10.4269/ajtmh.19-0620

190. Liljander A, Chandramohan D, Kweku M, Olsson D, Montgomery SM, Greenwood B, Färnert A. Influences of intermittent preventive treatment and persistent multiclonal Plasmodium falciparum infections on clinical malaria risk. *PLoS One*. 2010;5(10). doi:10.1371/journal.pone.0013649

191. Ouédraogo A, Tiono AB, Diarra A, Nébié IO, Konaté AT, Sirima SB. The effects of a pre-season treatment with effective antimalarials on subsequent malaria morbidity in under five-year-old children living in high and seasonal malaria transmission area of Burkina Faso. *Trop Med Int Heal*. 2010;15(11):1315-1321. doi:10.1111/j.1365-3156.2010.02618.x

192. Tiono AB, Diarra A, Sanon S, Nébié I, Konaté AT, Pagnoni F, Sirima SB. Low specificity of a malaria rapid diagnostic test during an integrated community case management trial. *Infect Dis Ther*. 2013;2(1):27-36. doi:10.1007/s40121-013-0006-6

193. Larsen DA, Bennett A, Silumbe K, Hamainza B, Yukich JO, Keating J, Littrell M, Miller JM, Steketee RW, Eisele TP. Population-wide malaria testing and treatment with rapid diagnostic tests and artemether-lumefantrine in Southern Zambia: A community Randomized Step-Wedge Control Trial Design. *Am J Trop Med Hyg*. 2015;92(5):913-921. doi:10.4269/ajtmh.14-0347

194. Owusu-Agyei S, Binka F, Koram K, Anto F, Adjuik M, Nkrumah F, Smith T. Does radical cure of asymptomatic Plasmodium falciparum place adults in endemic areas at increased risk of recurrent symptomatic malaria? *Trop Med Int Heal*. 2002;7(7):599-603. doi:10.1046/j.1365-3156.2002.00902.x