Characterization and Detection of Vector-borne Diseases in Endemic Transmission Areas

By

Robin H. Miller

Dissertation submitted to the Faculty of the
Emerging Infectious Diseases (EID) Graduate Program
Uniformed Services University of the Health Sciences
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy 2016

FINAL EXAMINATION/PRIVATE DEFENSE FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE EMERGING INFECTIOUS DISEASES GRADUATE PROGRAM
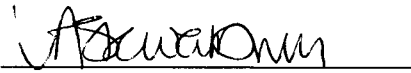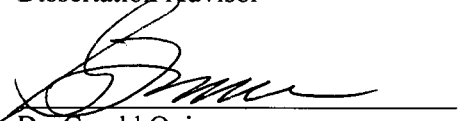
Name of Student: Robin Miller

Date of Examination: January 5, 2016
Time: 1:00 PM
Place: A2074

DECISION OF EXAMINATION COMMITTEE MEMBERS:

	PASS	FAIL
Dr. Stephen Davies DEPARTMENT OF MICROBIOLOGY & IMMUNOLOGY Committee Chairperson	✓	
Dr. V. Ann Stewart DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS Dissertation Advisor	✓	
Dr. Gerald Quinnan DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS Committee Member	✓	
Dr. Robert DeFraites DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS Committee Member	✓	

APPROVAL OF THE DOCTORAL DISSERTATION IN THE
EMERGING INFECTIOUS DISEASES GRADUATE PROGRAM

Title of Dissertation: "Characterization and Detection of Vector-Borne Diseases in Endemic Transmission Areas"

Name of Candidate: Robin Miller
Doctor of Philosophy Degree
January 5, 2016

DISSERTATION AND ABSTRACT APPROVED:

DATE:

1/18/2016

Dr. Stephen Davies
DEPARTMENT OF MICROBIOLOGY & IMMUNOLOGY
Committee Chairperson

5 Jan 2016

Dr. V. Ann Stewart
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Dissertation Advisor

01/05/16

Dr. Gerald Quinnan
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Committee Member

5 Jan 16

Dr. Robert DeFraites
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Committee Member

ACKNOWLEDGMENTS

DEDICATION

For my family,

R.H.M, L.E.E, J.R.M, & J.M.M

And always,

E.D.L

"Everything about malaria is so moulded by local conditions that it becomes a

thousand epidemiological puzzles."

-LW Hackett (1937)

COPYRIGHT STATEMENT

The author hereby certifies that the use of any copyrighted material in the dissertation [or thesis] manuscript entitled: "Characterization and Detection of Vector-borne Diseases in Endemic Transmission Areas" is appropriately acknowledged and, beyond brief excerpts, is with the permission of the copyright owner.

Robin H. Miller

ABSTRACT

Characterization And Detection Of Vector-Borne Diseases In Endemic Transmission Areas

Robin H. Miller, Doctor of Philosophy, 2016

Thesis directed by: V. Ann Stewart, DVM, PhD, Professor, Department of Preventive Medicine and Biostatistics, Division of Tropical Public Health

Vector-borne diseases contribute significantly to the global burden of infectious diseases and remain a major public health challenge worldwide. Detection and surveillance of the pathogen and vector are critical for the control of vector-borne diseases. Japanese encephalitis virus (JEV) and malaria cause a significant portion of disease and mortality due to vector-borne diseases globally. The research described in this dissertation aims to improve detection methods for both the vector and pathogens that facilitate JEV and malaria transmission in an effort to advance vector-borne disease control and potential elimination. First, we developed an ecological niche model to estimate the distribution of the JEV vector, *Culex tritaeniorhynchus*, based on environmental variables and known vector locations in endemic regions. We analyzed the overlap between Japanese encephalitis (JE) cases and predicted prevalence of *Cx. tritaeniorhynchus* distribution as well as the prevalence of predicted vector habitat within rice fields. Our novel ecological niche model can be used to target regions for vector

control strategies and vaccine campaigns in order to reduce JE disease burden within the human population. Second, we developed a novel real-time PCR (qPCR) assay for the detection of the newly described *Plasmodium ovale* subspecies, *P. ovale curtisi* and *P. ovale wallikeri*. Previous assays for the detection of *P. ovale* in our laboratory and others inadvertently detected only one of the two *P. ovale* subspecies due to genetic polymorphisms in the primer binding regions, resulting in an overall underestimation of *P. ovale* prevalence in malaria endemic regions. Our newly described *P. ovale* assay can successfully detect both subspecies with high sensitivity and specificity. Additionally, we report the first evidence that both *P. ovale curtisi* and *P. ovale wallikeri* circulate in a malaria holoendemic region in Western Kenya. Our *P. ovale*-specific qPCR assay can be used for future epidemiological studies to improve detection of *P. ovale* parasites and increase our understanding of the contribution of this neglected malaria parasite to the global malaria disease burden. Finally, we utilized an amplicon-based deep sequencing approach to detect multiclonal *P. falciparum* infections and analyze the Complexity of Infection (COI) based on several demographic factors from samples collected in the Democratic Republic of Congo (DRC). We deep sequenced the *P. falciparum* apical membrane antigen 1 (*pfama1*) gene from 79 individual samples and detected 68 unique *pfama1* haplotypes. We found that the majority (64.5%) of individuals had multiclonal infections (COI>1) and also found no association between *P. falciparum* COI and age, sex, HIV status, or geographical location within the DRC. Overall, we report high *pfama1* genetic diversity from asymptomatic malaria infections in the DRC and highlight the utility of amplicon-based deep sequencing to detect low frequency strains (variants) and multiclonal malaria infections that have been shown to impact malaria clinical

disease and transmission dynamics. In conclusion, the research presented herein describes both sensitive molecular and geospatial tools that can be utilized to improve detection of malaria parasites and the JEV vector, respectively.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Introduction

VECTOR-BORNE DISEASES IN HUMANS

Vector-borne diseases are globally distributed, cause severe morbidity and mortality, disproportionally affect the global poor, and contribute to the cycle of poverty (14; 275). Vector-borne diseases cause more than 17% of all infectious diseases in humans and result in over 1 million deaths each year worldwide (17). Additionally, the World Health Organization (WHO) estimates that over 1 billion people in over 100 countries are infected with a vector-borne disease, and that over half of the world's population is at risk for contracting a vector-borne disease (14). A summary of selected vector-borne diseases that occur in humans, shown in Table 1, highlights the diversity of vectors, pathogens, and diseases that together constitute a major global health challenge.

Vector-borne diseases are significant contributors to emerging infectious diseases (EIDs) globally and present several challenges that impede control efforts (112). The ability of vector-borne diseases to evade control strategies is due to several factors. First, vector-borne disease mitigation depends on appropriate and sustained vector control strategies that take into account the particular vector habitat and life cycle (14; 112). These vector control strategies can be complicated by anthropomorphic land-use changes to the environment, such as urbanization, agriculture, and deforestation that impact vector prevalence and can facilitate the emergence of vector-borne diseases into novel environments (113; 243; 315). Additionally, climate change is postulated to result in changes to vector habitat and distribution that may contribute to the expansion of vector-borne diseases into naïve populations (113; 243; 247; 277). Second, another challenge of

1

controlling vector-borne diseases is that many vector-bone diseases have a sylvatic cycle and infect non-human animal reservoirs. Vector-borne diseases have a range of natural reservoir hosts, including birds, bats, pigs, primates, rodents and several other vertebrate species (39). Controlling vector-borne diseases therefore requires identification of natural reservoir hosts and subsequent detection of the pathogen in these hosts to determine strategies to reduce human vector-borne disease prevalence. Third, vaccines are not commercially available for most vector-borne diseases, further confounding efforts to control these pathogens (14). Vaccine development is particularly difficult for vector-borne protozoan pathogens, such as malaria, which cause a significant proportion of vector-borne disease infections and deaths worldwide (4; 296). Lastly, detection of vector-borne disease-causing pathogens and their associated vectors requires sensitive methods for pathogen detection and field resources for entomological surveys in order to determine the prevalence of vector-borne disease within an endemic region and to identify and control a vector-borne disease outbreak (160). As vector-borne diseases continue to spread into new geographical locations and human populations, sensitive pathogen detection and vector surveillance methods are important tools for monitoring the emergence of vector-borne diseases (160). Detection of vector-borne disease associated pathogens and vectors can be hindered due to lack of resources and training to perform current pathogen detection methods, poor sensitivity and/or specificity of pathogen detection tools, and lack of resources and training for entomological surveys (7; 14). Based on these factors and others, control of vector-borne diseases remains a difficult and important global challenge.

The epidemiological triad (Figure 1) is an infectious disease model that illustrates the required combination of host, pathogen, environment, and vector to cause disease (7; 78). Complexities in the epidemiological triad model arise, as different vector-borne diseases require varying factors in terms of host, vector, pathogen, and environment for disease to occur (78). In addition, the host, vector, pathogen and environmental conditions that contribute to disease can vary based on geographic location (78). Japanese encephalitis and malaria, both mosquito-borne pathogens with distinct epidemiological factors that cause disease in humans, are discussed in greater detail below.

The Epidemiological Triad: Japanese Encephalitis

Disease

The majority of Japanese encephalitis virus (JEV) infections are asymptomatic, and severe Japanese encephalitis (JE) cases are estimated to occur in 1 out of every 250 JEV infections (21). The WHO estimates approximately 68,000 clinical JE cases occur every year and JE remains a major cause of viral encephalitis, particularly in children, in endemic Asian countries (15). The case fatality rate for clinical JE is estimated to be about 30%, and neurological sequelae occur in 30-50% of recovered individuals (15). Clinical JE symptoms include fever, chills, headache, myalgia, and vomiting, which can progress to more severe symptoms such as changes in mental status, confusion, seizures, and coma (20; 21). Despite the lack of specific treatment for JE, vaccines are available and have been shown to reduce cases and deaths due to JEV in endemic regions (164; 255).

Host

JEV Life Cycle

Humans are dead-end hosts for JEV as levels of viremia in humans are too low for transmission to biting mosquitoes (reviewed in (315)). JEV is maintained in an enzootic cycle in which wading birds and pigs serve as amplifying hosts and reservoirs for the mosquito vector (158).

Pathogen

Japanese Encephalitis Virus

Japanese encephalitis virus (JEV) is a member of the *Flaviviridae* family that includes other arboviruses such as West Nile Virus (WNV) and Dengue virus (DENV). JEV is a zoonotic virus and is comprised of five genotypes (I-V), often referred to as strains, which are restricted to different geographical locations throughout the JE endemic region (108).

JEV Detection

As JE symptoms are indistinguishable from other causes of acute encephalitis, the WHO recommends laboratory diagnosis for individuals that meet the clinical case definition of Acute Encephalitis Syndrome (AES) (123). The standard diagnostic test for JEV infection is an enzyme-linked immunosorbent assay (ELISA) for the detection of JEV-specific IgM in the cerebrospinal fluid (CSF), preferably, or in serum if CSF is not available (123).

Vector

Mosquitoes in the genus *Culex*, predominately *Cx. tritaeniorhynchus*, transmit JEV (56). *Cx. tritaeniorhynchus* is found throughout Southeast Asia and in several countries in sub-Saharan Africa and the Middle East (1). Breeding habitats include shallow pools, rice paddies, swamps, streams, marshes, puddles, and other temporary water sources (1). *Cx. tritaeniorhynchus* females are highly zoophilic and prefer to feed on pigs, cattle, birds, and other animals instead of humans (229).

Geographic information systems (GIS) are powerful tools that allow for visualization and spatial modeling of disease vectors. GIS can utilize large data sets consisting of vector collection coordinates, remote-sensing images from satellite imagery, and spatial modeling programs to display and even predict vector distribution with high resolution (reviewed in (219) and (135)). GIS maps can be used to target regions for vector control based on vector prevalence and abundance. Accordingly, the WHO recommends using GIS in vector control programs, particularly in resource-limited settings where entomological surveys are not practical (171). GIS are also useful ways to produce visual ecological niche models that predict the distribution of vector species based on previously collected geospatial (occurrence) points and environmental habitat preference (313). Several studies have shown the utility of GIS and ecological niche modeling of vector distribution based on climatic factors and land use data to assess vector-borne pathogen risk for many vector-borne diseases in a range of locations worldwide (32; 70; 115; 156; 190; 205; 226). Overall, GIS are robust mapping tools that allow for the visualization, analysis, and statistical modeling of geospatial data as it relates to vector-borne disease prevalence and transmission.

Environment

Proximity to rice fields has been shown to increase the risk of JEV transmission to humans, as rice fields are the preferred habitat for *Cx. tritaeniorhynchus* breeding (141). The presence of amplifying hosts, such as pigs and large wading birds has also been shown to increase the risk of JEV transmission due to high-level JEV viremias in these reservoir species (158).

JEV is an emerging virus throughout much of Southeast Asia and the Western Pacific (173). JEV is thought to have originated in Indonesia and Malaysia and later emerged in Japan, China, and Korea in the late 1800s (92; 315). JEV spread further into Thailand and Vietnam and more recently into Nepal, India, and Pakistan (92; 315). Additional studies indicate JEV is now expanding east into New Guinea and south into northern Australia (118; 173). JEV has emerged dramatically across Southeast Asia and the Western Pacific, often times following the expansion of rice cultivation and pig farming, and is currently a major risk for viral encephalitis throughout the entire endemic region (311).

The Epidemiological Triad: Malaria

Disease

Uncomplicated clinical malaria is classically described as an initial prodromal period followed by alternating cycles of fever and chill paroxysms. Uncomplicated malaria can be caused by any of the malaria parasite species and consists of several non-specific symptoms such as shivering, nausea, vomiting, headache, diarrhea, and body aches (37). Severe malaria refers to clinical malaria with evidence of vital organ dysfunction such as convulsions, respiratory distress, shock, kidney failure, jaundice,

prostration, severe anemia, and coma (11). Although the majority of severe malaria cases are due to *Plasmodium falciparum*, *P*. *vivax* and *P*. *knowlesi* also can cause severe malaria and death (11).

Vector

Mosquitoes of the genus *Anopheles* transmit human malaria. Although the genus *Anopheles* is comprised of over 430 species, only 70 species can transmit malaria, of which approximately 40 species are considered major vectors for human malaria transmission globally (257). Dominant malaria vector species vary by geographical location. For instance, *An*. *arabiensis* dominates in parts of eastern Africa, while *An*. *darlingi* is a major malaria vector in South America (144). In addition to geographical variation, anopheline species have varying feeding behaviors, such as preference for feeding inside (endophagic) or outside (exophagic), and preference for human (anthropophagic) or animal (zoophagic) hosts (144). Anopheline species also exhibit widely varied preferences for larval habitats, on the basis of factors such as salinity levels, sunlight, vegetation cover, and natural versus artificial water sources (144). Based on the geographical and behavioral differences of anopheline mosquitoes, it is important to identify the dominant and secondary malaria vectors contributing to malaria transmission in an endemic region for control measures to be effective.

Environment

The environmental factors that contribute to malaria disease transmission depend largely on the specific characteristics of the malaria endemic region and the behaviors and ecological requirements of the anopheline vector species in that region. Several climatic factors, including rainfall, temperature, and humidity are associated with malaria

transmission as they influence vector breeding preferences (257). Agricultural practices have also been shown to influence anopheline vector prevalence in some malaria endemic regions. For example, the presence of rice fields has been shown to increase malaria transmission intensity in Southeast Asia (257), but had no effect on malaria transmission in Cote d'Ivoire (77). In addition to ecological and climatic factors, regions of social or political disturbance are also associated with increased malaria transmission (221). Thus, there is a complex relationship between climate, ecology, anthropomorphic factors, and anopheline vector behavior that together can influence malaria transmission.

Host

Malaria endemicity is classified into four levels: holoendemic, hyperendemic, mesoendemic, and hypoendemic (Table 2). Regions with holoendemic malaria transmission typically experience intense, stable malaria transmission year round, and thus individuals living in these regions are routinely infected with malaria parasites (121). Clinical malaria in holoendemic regions is typically restricted to children less than five years old and pregnant women (269; 272). Older children and adults develop naturally acquired immunity (NAI) that protects against clinical malaria, but as NAI to malaria is non-sterile, these clinically protected individuals still maintain asymptomatic malaria infections (82; 154). The majority of malaria infections in holoendemic regions are therefore asymptomatic, and these asymptomatically infected individuals are important reservoirs for malaria transmission (29; 51; 153). This is in contrast to regions with less malaria transmission (hyperendemic, mesoendemic, and hypoendemic, see table 2), in which individuals of all ages are at risk for developing clinical malaria and asymptomatic infections are less common (208).

8

Pathogen

Malaria Life Cycle

The complex malaria life cycle includes ten parasite morphological changes that occur in several different types of host tissue in both in the anopheline mosquito (definitive host) and the human (intermediate host) (174). When an infectious female anopheline mosquito takes a blood meal from a susceptible human host, motile malaria sporozoites are injected into the skin. During the exo-erythrocytic stage of the life cycle, sporozoites travel to the liver, invade hepatocytes, and develop first into liver trophozoites and then into liver schizonts. Although malaria parasite development during the exo-erythrocyctic stage typically takes 7-14 days, both *P. ovale* and *P. vivax* can develop into hypnozoites that remain dormant in the liver for months to years after the initial infection and cause relapsing malaria. In all human malaria species, liver schizonts eventually rupture the hepatocyte and release merozoites that enter the blood stage, initiating the erythrocytic stage of the life cycle. Merozoites quickly invade the host red blood cells, where they develop into trophozoites and then into schizonts. The malaria schizont then ruptures the red blood cell and merozoites are released, which invade new red blood cells and initiate a subsequent round of the erythorcytic cycle. Depending on the malaria parasite species, the erythrocytic cycle takes approximately 24-72 hours (Table 3). A subset of malaria parasites during the erythrocytic cycle will develop into both male and female gametocytes, the malaria parasite stage infectious for the female anopheline mosquito (261). The malaria life cycle stages that occur in the mosquito are referred to as the sporogonic cycle, which begins when an anopheline mosquito takes a blood meal containing malaria gametocytes. Once in the anopheline mosquito midgut, the

9

male and female malaria gametocytes fuse to form the zygote. The zygote undergoes

meiosis and genetic recombination while in the mosquito midgut. The zygote then

develops into the motile ookinete, which invades the mosquito midgut wall. The ookinete

transforms into an oocyst, which continues to grow and divide in the midgut wall until it

ruptures and releases sporozoites. Sporozoites invade the mosquito salivary glands,

allowing for transmission to the human host during the next blood meal (261).

Malaria Parasite Complexity

Five different malaria species infect humans: *P. falciparum, P. vivax, P. malariae,*

P. knowlesi, and the newly described *P. ovale* subspecies: *P. ovale curtisi* and *P. ovale*

wallikeri (Table 3) (82; 213). In malaria endemic regions, individuals can be infected

with multiple malaria parasite species at the same time (53; 54; 322). These multi-species

infections have been shown to influence both clinical outcomes and responses to

antimalarial treatment (177; 264). For example, case studies in non-immune travelers

returning from malaria endemic regions have documented patients initially diagnosed and

treated for *P. falciparum* experiencing relapsing malaria due to a non-*P. falciparum*

malaria (64; 186; 256). In contrast to treatment for *P. falciparum*, the relapsing malaria

parasite species *P. vivax, P. ovale curtisi,* and *P. ovale wallikeri* must be treated with

primaquine to ensure dormant hypnozoites are cleared from the liver. Proper detection

and diagnosis of these complex malaria infections is therefore critical to ensure proper

antimalarial treatment is provided.

In addition to infection of multiple malaria species, humans can also be infected

with multiple strains of a single malaria species (31; 178; 265). The number of distinct

malaria strains infecting a single individual is referred to as the Complexity of Infection

(COI) (44). Multiclonal (COI>1) malaria infections have been shown to influence clinical disease, impact malaria transmission dynamics, and may be a useful marker for malaria transmission intensity in endemic areas (26; 52; 67; 72; 97). Detection of multiclonal infections requires sensitive molecular techniques that can detect multiple genetically related strains that occur at varying frequencies, including low-level minor variants, within an individual (31; 104; 134; 266).

Malaria Detection Methods

A summary of malaria detection methods is shown in Table 4. Microscopic diagnosis of malaria remains the "gold standard" for malaria detection (320). Despite the widespread use and utility of microscopic diagnosis of malaria parasites, microscopy poses several challenges for malaria detection and diagnosis. For example, proper microscopic diagnosis requires highly trained and/or expert microscopists who must undergo periodic evaluation to ensure quality control (QC) and quality assurance (QA) of the malaria diagnostic program (214; 320). Several studies have shown that misdiagnosis of malaria species can occur using light microscopic detection, especially in the context of mixed species infections (211). Highly skilled microscopists can detect low-level malaria parasites, although the accuracy of microscopic detection depends on the expertise of the microscopist and training methods (209; 211; 214).

Malaria Rapid Diagnostic Tests (mRDTs) are increasingly used as a method for malaria detection and diagnosis (Reviewed in (43)). mRDTs detect malaria antigen from a small volume of patient blood based on a lateral flow immune-choromatographic test strip (16). Over 200 different mRDTs are now commercially available (16). Yet, the majority of mRDTs detect *P. falciparum* only, while a select few can detect *P.*

falciparum plus additional non-falciparum species (16). The use of mRDTs for malaria detection and diagnosis has several advantages over microscopy. Training and skill level required for mRDT administration is minimal and does not require laboratory space or electricity, highlighting the utility of mRDTs in resource-limited areas (43). Despite the uptake and usage of mRDTs in several countries as part of their malaria control programs, there also remain pitfalls of mRDT usage for malaria detection and diagnosis (43; 176). These include the inability of some mRDTs to detect low-level malaria infections and the failure to detect non-falciparum malaria species, resulting in false negative results that can delay malaria drug treatment (43; 111).

The application of Polymerase Chain Reaction (PCR) for the detection of malaria parasites has increased the ability of researchers and clinicians to detect low-level and submicroscopic malaria parasitemias (248; 319). PCR detection of low-level malaria parasitemias is a critical tool for both understanding malaria epidemiology and facilitating malaria control. As more malaria endemic regions enter control and elimination phases, PCR detection of low-level malaria parasitemias is a sensitive surveillance tool to monitor malaria transmission (119; 128). Furthermore, PCR is often utilized for early detection of low-level drug resistant malaria parasites that persist after antimalarial treatment, which is critical for monitoring the spread and emergence of malaria drug resistance (8). In addition to increased sensitivity for the detection of malaria parasites, PCR assays that target species-specific nucleic acid sequences can detect and differentiate mixed malaria species infections that may be missed or undetected based on light microscopy and mRDT (266). The development of real-time PCR (qPCR) assays have further expanded the utility of these molecular based assays for

malaria detection by increasing specificity through the use of probe based technology (248).

Next-generation deep sequencing technologies are also highly sensitive methods for malaria detection (192; 310) and have been shown to successfully detect low frequency parasites (134; 310). Deep sequencing, also referred to as massively parallel sequencing, refers to nucleic acid sequencing methods that generate multiple sequence reads, resulting in high coverage of target sequences. Several next-generation sequencing platforms are currently available, each with distinct template/library preparation protocols, sequencing chemistries, detection methods, run times, output read lengths, number of reads generated per run, and cost (55; 59). However, the overall theory behind next-generation sequencing methods is similar. First, the target nucleic acid is sheared and further processed for library preparation (59). The library amplification step typically includes adapter ligation, immobilization of the fragment (usually on a bead or array), and PCR amplification. Next, the amplified libraries are used as template for deep sequencing, during which the NGS platform detects the addition of a nucleotide, usually by fluorescence, pyrophosphate detection, or proton release (55; 59). Finally, the last step consists of appropriately trimming, clustering, aligning, and analyzing the resulting sequence reads using the manufacturer or institutionally provided software or through a custom bioinformatics analysis pipeline (59). A summary of selected next-generation sequencing technologies is provided in Table 5.

There are several advantages of using next-generation deep sequencing for malaria genomics studies (129; 310), including detection of drug resistant loci, characterization of low frequency malaria parasitemias in multiclonal malaria infections,

13

malaria whole genome sequencing, population genetic studies, and malaria transmission surveillance (71; 72; 104; 129; 134; 169; 192; 310). The ongoing significant decreases in cost associated with next-generation sequencing further highlights their potential to enhance malaria detection and aid malaria control efforts (129).

Molecular based detection methods, such as PCR, qPCR, and next-generation sequencing technologies are highly sensitive for the detection of low-level parasitemias, mixed malaria species infections, and multiclonal malaria parasite infections compared to traditional methods, such as microscopy and mRDTs. As detection of the malaria parasite is critical for malaria control, utilizing sensitive molecular tools will likely improve current malaria detection and surveillance methods. Further, GIS can be utilized to map the geographical locations of malaria parasites based on the results of malaria detection assays or using mathematical models that estimate the distribution of malaria parasite species and strains (47; 126; 180; 258; 291).

APPLYING THE EPIDEMIOLOGICAL TRIAD MODEL FOR JEV AND MALARIA CONTROL

Control of vector-borne pathogens requires an understanding of the host, pathogen, environment, and vector that cause disease. In terms of public health nomenclature, pathogen control refers specifically to the reduction of disease morbidity and mortality, prevalence, and incidence within a geographical region (83). Pathogen elimination requires reducing disease incidence to zero over a sustained period of time within a geographical region (83). Both pathogen control and elimination require sustained intervention strategies to prevent pathogen reintroduction and disease reemergence (83). Pathogen eradication refers to the global reduction of disease incidence, thus continued intervention strategies are not required (83).

14

Vector-borne diseases have successfully been targeted for control and elimination in several countries around the world, and typically require a multifaceted approach that includes detection and surveillance of the pathogen in vectors, hosts (human and reservoir) and the environment. For example, Mexico recently validated the elimination of river blindness (onchocerciasis) using a PCR assay designed to detect *Onchocerca volvulus* in the black fly vector (244). To successfully eliminate lymphatic filariasis, several countries, including China and the Republic of Korea, utilized geospatial surveillance to identify regions for targeted mass drug administration (MDA) and monitor for reemergence of the pathogen in vector and human populations after MDA (130). Additionally, the WHO recommends that countries in the malaria elimination stage conduct disease surveillance, which includes detection of the malaria parasite in the human host and also monitoring the mosquito vector habitat and breeding sites (10).

Malaria and JEV are two important pathogens that together cause a substantial proportion of morbidity and mortality attributed to vector-borne diseases worldwide. The epidemiological triad model for disease control demonstrates that detection of both the pathogen and vector are critical for mitigating disease caused by JEV and malaria. Improvement in pathogen detection, through the use of highly sensitive molecular assays, and vector surveillance, through geospatial modeling, will help to better understand how these factors contribute to disease and provide a framework for the control and possible elimination of JEV and malaria.

RESEARCH GOAL, OBJECTIVES, AND RATIONALE

Research Goal

The overall goal of the research in this dissertation is to improve the detection of pathogens and vectors that cause vector-borne diseases in humans.

Research Objectives

The objectives of this dissertation research are to 1) utilize geospatial tools to estimate JEV vector prevalence and 2) develop highly sensitive molecular tools to detect malaria parasite species and strains in complex infections.

Rationale

Importance of Vector Detection for Disease Control

Understanding the prevalence and distribution of the vector is critically important for controlling vector-borne pathogen disease and transmission. Vector surveillance can be used to detect the emergence of a new vector into an environment and also for monitoring the effectiveness of vector control strategies, such as insecticides (24; 116; 142; 155; 237). Vector detection is conducted using a variety of methods, including field surveys using adult stage traps, immature stage (larval and pupal) surveys, and host landing and biting counts (22; 196; 257). Additionally, GIS can map vector geospatial sampling data and estimate disease risk within an endemic region based on ecological niche models (219; 313).

Importance of Pathogen Detection for Disease Control

As shown by the epidemiological triad, the pathogen is also a critical factor in disease causation. Detection of the pathogen within the host, vector, and/or environment

is key for understanding the prevalence and distribution of a vector-borne pathogen. Detection and surveillance of vector-borne pathogens provides important epidemiological information that can then be utilized for public health efforts to mitigate disease and prevent transmission (6; 8; 10). Vector-borne pathogen surveillance requires detection tools that can identify the pathogen, even at low levels (14). Use of highly sensitive nucleic acid based detection methods, such as qPCR and deep sequencing, have shown increased capabilities in detecting low-level pathogen infections and therefore provide valuable information about pathogen prevalence that can facilitate control of vector-borne diseases (134; 136; 169; 178; 267).

AIM 1

Develop an ecological niche model to estimate the JEV vector *Cx. tritaeniorhynchus* prevalence in Japanese encephalitis endemic regions.

Culex tritaeniorhynchus geospatial collection data and environmental variables will be utilized to develop an ecological niche model for Japanese encephalitis virus (JEV) vector prevalence in the Japanese encephalitis (JE) endemic region. The ecological niche model will be compared to locations of documented JE clinical cases and rice fields using GIS. Analysis of environmental and climatic factors that contributed significantly to the ecological niche model will be performed in order to investigate abiotic factors that influence JEV vector distribution. The ecological niche model for *Cx. tritaeniorhynchus* will identify regions with an increased risk of the dominant JEV vector and will also provide a framework for targeting vector control programs and vaccine administration to control JE disease in human populations.

17

AIM **2**

Improve detection of *P. ovale* in malaria endemic regions by targeting a conserved genetic region between *P. ovale curtisi* and *P. ovale wallikeri* subspecies using a real-time PCR (qPCR) platform.

P. ovale is a neglected malaria parasite species that was recently discovered to exist as two genetically distinct subspecies: *P. ovale curtisi* and *P. ovale wallikeri*. Previous molecular based methods for *P. ovale* detection inadvertently targeted only one of the two subspecies, thereby failing to detect all *P. ovale* infections. An inclusive real-time PCR (qPCR) assay will be developed based on a conserved genetic region between *P. ovale curtisi* and *P. ovale wallikeri* in order to detect both subspecies. Several qPCR validation steps will be performed in order to evaluate both sensitivity and specificity of the assay. Additionally, the presence of both *P. ovale curtisi* and *P. ovale wallikeri* in a malaria holoendemic region in Kenya will be investigated using a multilocus genotyping approach. The *P. ovale*-specific qPCR assay will allow for detection of both *P. ovale* subspecies, thereby improving detection of a neglected but clinically relevant malaria parasite species.

AIM **3**

Utilize a deep sequencing approach to detect multiclonal *P. falciparum* infections.

Multiclonal *P. falciparum* infections can impact malaria clinical outcomes, influence within-host transmission dynamics, and indicate malaria transmission intensity.

Detection of multiclonal *P. falciparum* remains a challenge due to lack of sensitive tools to detect low frequency strains (variants) within a multiclonal infection. An amplicon-based deep sequencing approach targeting the *P. falciparum* apical membrane antigen 1 (*pfama1)* gene will be utilized to explore multiclonal *P. falciparum* infections in the Democratic Republic of Congo (DRC). Amplicon-based deep sequencing allows for sensitive detection of minor variant strains within a multiclonal *P. falciparum* infection due to the extensive coverage of the target *pfama1* sequence. Several epidemiological factors associated with multiclonal infections (COI>1) will be analyzed, including age, sex, HIV status, and geographical location in the DRC. Detection of complex *P. falciparum* infections can be utilized to better understand the relationship of multiclonal *P. falciparum* infections and clinical disease and also potentially as a marker for transmission dynamics in malaria control settings.

Table 1. Summary of Selected Vector-Borne Diseases in Humans.
Table adapted from the WHO: A Global Brief on Vector-Borne Diseases (14).

Vector	Pathogen	Disease
Mosquitoes	*Plasmodium* spp., dengue virus, yellow fever virus, chikungunya virus, West Nile Virus, *Wuchereria bancrofti, Brugia malayi, B. timori*, Japanese encephalitis virus	Malaria, dengue, yellow fever, chikungunya, West Nile virus, lymphatic filariasis, Japanese encephalitis
Sandflies	*Leishmania* spp.	Leishmaniasis
Triatomine bugs	*Trypanosoma cruzi*	Chagas Disease
Ticks	*Borrelia burgdorferi*, Crimean-Congo hemorrhagic fever virus, *Rickettsia rickettsii*	Lyme disease, Crimean-Congo hemorrhagic fever, Rocky Mountain spotted fever
Fleas	*Yersinia pestis, Rickettsia typhi*	Plague, murine typhus
Flies	*Trypanosoma brucei gambiense, T. b rhodesiense, Onchocerca volvulus*	Human African trypanosomiasis, onchocerciasis (river blindness)
Freshwater snails	*Schistosoma mansoni, S. japonicum, S. haematobium*	Schistosomiasis/bilharzia

Table 2. Malaria endemicity levels.
        Modified from (268).

	Holoendemic	Hyperendemic	Mesoendemic	Hypoendemic
Parasite prevalence (2-9 year olds)	>75%*	>50%	11-50%	0-10%

* In infants (0-11 months)

Table 3. Defining characteristics of the five human malaria parasites (82).

	P. falciparun	*P. vivax*	*P. malariae*	*P. ovale*	*P. knowlesi*
Geographic Range	Pan-tropical	Pan-tropical, temperate	Pan-tropical	Africa, Southeast Asia	Southeast Asia
Prevalence	High	High	Low	Low	Low
Pre-erythrocytic stage (day)	5-7	6-8	14-16	9	8-9
Erythrocyctic cycle (hr)	48	48	72	~48	24
Severe malaria	+++	+++	+	+	+++
Drug Resistance	Yes	Yes	No	No	No
Relapse	No	Yes	No	Yes	No

Table 4. Malaria detection methods.
Table adapted from (285).

	Detection Limit (parasites/µl)	Expertise	Time (min)	Cost
Light microscopy	Expert: 5-10 Non-expert: >50	Medium to high	<60	Low
Malaria Rapid Diagnostic Tests (mRDTs)	50-100	Low	<15	Moderate
PCR/qPCR	Range: 0.002 to <1	High	<45 to > 360	High

Table 5. Summary of next-generation sequencing technologies.
Adapted from (55; 59).

Manufacturer	Platform	Detection	Run Time	Max read length (bp)	Usable reads per run (millions)
Life Technologies/ Applied Biosystems	Ion Torrent PGM	Proton release	2 h	400	4
	5500 SOLiD	Fluorescence detection	8 d	~75	>700
Illumina	HiSeq2000	Fluorescence detection	8.5 d	~100	3000
	MiSeq	Fluorescence detection	27 h	~150	7
Roche/454 Life Sciences	454 FLX	Pyrophosphate detection	10 h	600	1
	454 FLX+	Pyrophosphate detection	23 h	1000	1
Pacific Biosciences	RS II	Fluorescence detection	0.5-2 h	50% reads >10 kb	0.8
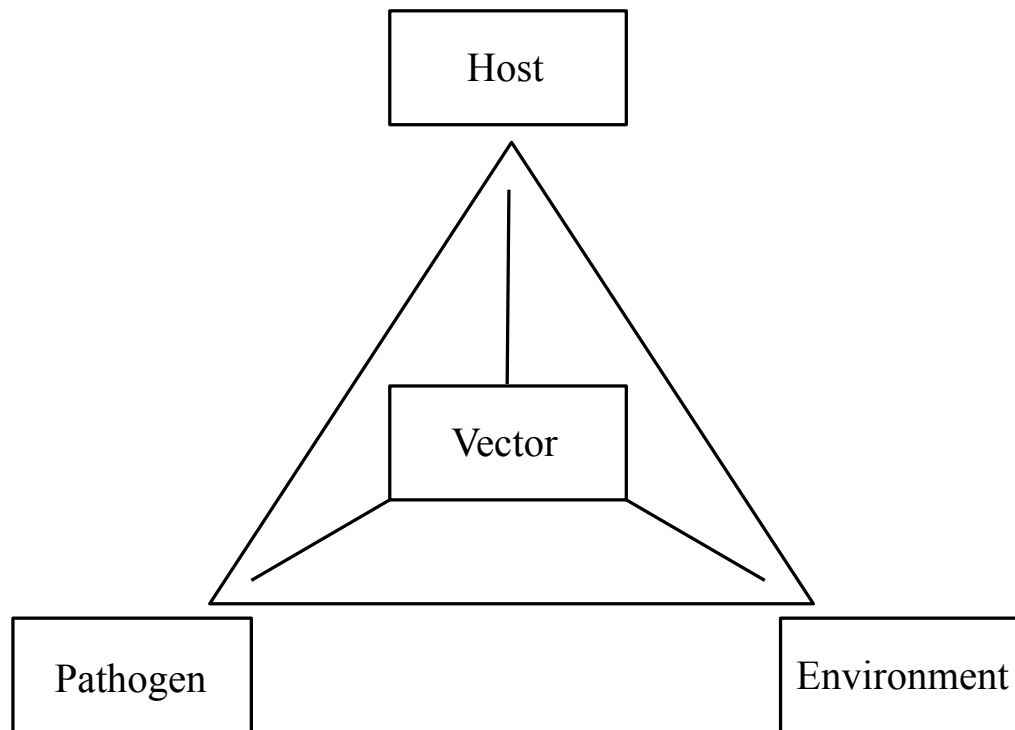
Figure 1. Epidemiological triad. Adapted from (78).

CHAPTER 2: Ecological Niche Modeling To Estimate The Distribution Of Japanese Encephalitis Virus In Asia

Robin H. Miller[1], Penny Masuoka[1], Terry A. Klein[2], Heung-Chul Kim[3], Todd Somer[4], John Grieco[1]

[1] Department of Preventive Medicine and Biometrics, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

[2] 65th Medical Brigade/USAMEDDACKorea, Unit 15281, Force Health Protection and Preventive Medicine, Yongsan Army Garrison, Seoul, Republic of Korea

[3] 168th Multifunctional Medical Battalion, 65th Medical Brigade, Unit 15247, 5th Medical Detachment, Yongsan Army Garrison, Seoul, Republic of Korea

[4] NST, National Geospatial-Intelligence Agency, Fort Detrick, MD, USA

Table, figure, and reference numbers have been changed to follow the USU thesis format guidelines.

ABSTRACT

Background

Culex tritaeniorhynchus is the primary vector of Japanese encephalitis virus
(JEV), a leading cause of encephalitis in Asia. JEV is transmitted in an enzootic cycle
involving large wading birds as the reservoirs and swine as amplifying hosts. The
development of a JEV vaccine reduced the number of JE cases in regions with
comprehensive childhood vaccination programs, such as in Japan and the Republic of
Korea. However, the lack of vaccine programs or insufficient coverage of populations in
other endemic countries leaves many people susceptible to JEV. The aim of this study
was to predict the distribution of *Culex tritaeniorhynchus* using ecological niche
modeling.

Methods/Principal Findings

An ecological niche model was constructed using the Maxent program to map the
areas with suitable environmental conditions for the *Cx. tritaeniorhynchus* vector.
Program input consisted of environmental data (temperature, elevation, rainfall) and
known locations of vector presence resulting from an extensive literature search and
records from MosquitoMap. The statistically significant Maxent model of the estimated
probability of *Cx. tritaeniorhynchus* presence showed that the mean temperatures of the
wettest quarter had the greatest impact on the model. Further, the majority of human
Japanese encephalitis (JE) cases were located in regions with higher estimated probability
of *Cx. tritaeniorhynchus* presence.

25

Conclusions/Significance

Our ecological niche model of the estimated probability of *Cx. tritaeniorhynchus* presence provides a framework for better allocation of vector control resources, particularly in locations where JEV vaccinations are unavailable. Furthermore, this model provides estimates of vector probability that could improve vector surveillance programs and JE control efforts.

AUTHOR SUMMARY

Japanese encephalitis virus (JEV) is transmitted predominately by the mosquito, *Culex tritaeniorhynchus*. The primary reservoirs of the virus are wading birds, with swine serving as amplifying hosts. Despite the development of a JEV vaccine, people remain unvaccinated in endemic countries and are susceptible to JEV infection. The distribution of the JEV vector(s) provides essential information for preventive measures. This study used an ecological niche modeling program to predict the distribution of *Cx. tritaeniorhynchus* based on collection records and environmental maps (climate, land cover, and elevation). The model showed that the mean temperatures of the wettest quarter had the greatest impact on the model. Of the 25 countries endemic for Japanese encephalitis (JE) endemic countries, seven possessed greater than 50% land area with an estimated high probability of *Cx. tritaeniorhynchus* presence. Our model provides a useful tool for JEV surveillance programs that focus on vector control strategies.

INTRODUCTION

Japanese encephalitis virus (JEV), the causative agent of Japanese encephalitis (JE), is an arbovirus that belongs to the family *Flaviviridae* and is endemic to Southeast and Northeast Asia, the Pacific Islands, and northern Australia (Figure 2) (124). The

primary vector of JEV is *Culex tritaeniorhynchus* Giles, but other *Culex* species (e.g.,

Culex annulirostris, *Culex vishnui* Theobald, *Culex bitaeniorhynchus* Giles, and *Culex*

pipiens Linnaeus) have also been implicated as important viral transmitters (35; 48; 56;

241). The larval habitat of *Cx. tritaeniorhynchus* is primarily low lying flooded areas

containing grasses and flooded rice paddies, but this species can also be found in urban

environments in close proximity to human populations (239). Within the past 40 years,

rice agriculture in JEV endemic countries has increased by 20%, thereby expanding *Cx.*

tritaeniorhynchus habitat and increasing human risk of exposure to vector populations

(141).

Swine, including domestic and feral pigs, serve as amplifying hosts of JEV in

endemic areas. The proximity of human populations to pig farms, sties or feral pig

populations increases the risks of JEV exposure (143; 253). Ardeid birds (large wading

birds) are an important JEV reservoir and can spread JEV to new regions through their

northern migration to breeding and feeding grounds in the spring and southern return in

the fall (56). Additional animals have been identified as host species for JEV, including

domesticated animals (chickens, goats, cows, and dogs), as well as bats, flying foxes,

ducks, snakes and frogs. However, these are considered dead-end hosts as they

infrequently develop sufficient viremias to infect mosquito vectors (228; 299; 307; 312).

Despite the introduction of an effective vaccine to the public in the mid-1900s,

JEV remains the leading cause of viral encephalitis globally (303). Comprehensive

vaccination programs in Japan, Republic of Korea (ROK), Brunei, Australia, and

Malaysia have significantly reduced the number of human cases (222). Rare occurrences

of neurological complications associated with the mouse-brain derived JEV vaccine

interrupted vaccination programs in some regions, initiating concerns of the reemergence

of JEV in an unvaccinated and non-immune population (150; 234). The prevalence of JE

is higher in countries with lower socioeconomic status, when compared to more affluent

neighboring countries, indicating the importance of economic and social stability as

additional risk factors that impact the transmission and prevalence of JE in non-immune

populations (222).

Recent developments in the field of ecological niche modeling and the

development of global environmental data sets have resulted in the ability to predict the

distribution of vector populations that directly relate to transmission of viruses, parasites,

and fungal pathogens and impact on animal and human health. Modeling to estimate the

distribution of disease vectors provides useful information in disease-endemic areas, in

addition to predicting how anthropogenic changes to the environment will affect disease

presence (109; 175; 181; 182).

In the current study, the Maxent ecological niche modeling program was utilized

to model the distribution of the primary vector of JEV, *Cx. tritaeniorhynchus* (230). The

resulting vector habitat suitability map was compared to the reported locations of JE

human cases and the current status of established JE vaccination programs by country.

Our ecological niche model can be used by public health officials and government

agencies in endemic regions to guide implementation of comprehensive vaccination

programs, vector control strategies, and public health awareness campaigns.

METHODS

Culex tritaeniorhynchus **Data Collection**

Geographical coordinates of known *Cx. tritaeniorhynchus* records were identified by performing a literature search in PubMed for all previous field collection studies. When exact geographical coordinates were not provided, locations were approximated by searching for the given city, town, or village using Google Earth software. Further geographical data points for the distribution of *Cx. tritaeniorhynchus* were obtained through MosquitoMap (http://www.mosquitomap.org/), a database of spatial data points of mosquitoes that is maintained by the Walter Reed Biosystematics Unit, Smithsonian Support Center, Silver Hill, MD. Additional *Cx. tritaeniorhynchus* collection data were obtained from Force Health Protection and Preventive Medicine, 65th Medical Brigade, Yongsan Army Garrison, ROK.

In previous modeling work in the ROK (181), we found that a large number of collection records in a limited geographical area biased the model. As a result of the large number of collection sites for the ROK (96 unique locations), we reduced the number of records to 23 by deleting all but one randomly selected record per administrative district.

Identification of Japanese Encephalitis (JE) Human Cases

Approximate locations of known human JE cases were determined using locations provided in ProMED mail reports (www.promedmail.org) from 1994 through 2010 (Figure 3). Additional locations of confirmed JE cases were also determined through a PubMed literature search. Exact geographical coordinates were not reported for most documented human cases and were therefore extrapolated using the Google Earth

software to obtain the latitude/longitude coordinates of the reported city, town, or village in which JE was documented.

Vaccination Programs in JEV Endemic Countries

JEV vaccination program information was obtained from "Japanese Encephalitis Morbidity, Mortality and Disability: Reduction and Control by 2015" published in 2009 by the Program for Appropriate Technology in Health (PATH), Armed Forces Research Institute of Medical Sciences, and BIKEN (222). Additionally, JEV vaccination programs information was also obtained from the WHO/IVB database (216). Countries lacking a JEV vaccination program were identified using information in the above mentioned publications and confirmed with additional literature searches. A summary of these data are listed in Table 6.

Environmental Data

One kilometer resolution climate and elevation data were obtained from WorldClim (http://www.worldclim.org/bioclim). The WorldClim organization has processed 50 years of ground-based weather measurements to produce mean monthly minimum and maximum temperatures and precipitation in a grid format at several different resolutions. The data were further processed to produce bioclimatic variables (e.g., mean temperatures of the wettest quarters). For this project, the highest resolution data available from WorldClim (approximately 1 km) were downloaded. In addition to bioclimatic variables, global elevation data obtained from WorldClim was re-sampled to 1-km resolution from NASA's Shuttle Radar Topography Mission (SRTM). Descriptions of the bioclimatic and elevation variables used for this study are listed in Table 7. To better understand the effect of each environmental variable on *Cx. tritaeniorhynchus*

distribution, the values of each environmental layer at each site were extracted using ArcGIS (ESRI, Redlands, California, www.esri.com). This allowed for a comparison to the known environmental and distribution limitations for *Cx. tritaeniorhynchus* in the literature.

A map of rice growing areas was created by processing GeoCover-LC (Land Cover) data from MDA Information Systems, Inc. (http://www.mdafederal.com/geocover/geoco verlc). GeoCover was created by processing Landsat Thematic Mapper images to create land cover maps for most areas of the world. Each pixel within the GeoCover-LC represents 30 by 30 meters. To convert the image to a resolution that could be used in the Maxent model, ArcGIS was used to count the number of rice pixels within each square kilometer (33 by 33 pixels). Then, a rice percentage was calculated for each square kilometer (number of rice pixels divided by total number of pixels in 1 km) and stored in a final output image.

Ecological Niche Model

The Maxent 3.2.1 modeling program (http://www.cs.princeton.edu/~schapire/ma xent/) was utilized to model the distribution of *Cx. tritaeniorhynchus* based on previously obtained geographical locations. Maxent utilizes a maximum entropy algorithm to analyze values of environmental layers, such as temperature, precipitation, and elevation, at known locations of species occurrence (collection records) to estimate the probable range of the species over a geographic region (230; 231). This model is based on presence-only data instead of presence/absence data due to the lack of available absence data. Although absence data can be informative for modeling, ecological niche models based on presence-only data are useful in regions with limited collection data (230).

Without absence data, the true probability of presence cannot be modeled. In Maxent, which uses presence only data, the species distribution is output as an estimated probability map (163).

The Maxent program calculates the importance of environmental variables in developing predictive species distribution models by using the jackknife test of variable importance. The jackknife test runs the model 1) once with all variables, 2) dropping out each variable in turn, and 3) with a single variable at a time. Variables are considered import if they produce high training gains when used alone in a model. A variable is also important if the training gain is low when the variable is removed from the model (230).

Maxent utilizes two approaches to validate the accuracy of the model. The first method randomly selects occurrence points to be withheld from the model building to use as testing points. Using multiple definitions, a set of thresholds split the continuous probability values of the model into 'predicted presence or absence' categories. Maxent then calculates the p-value based on the null hypothesis that testing points will be predicted as "present" no better than by a random model. The second method calculates the Area Under the Curve (AUC) of the receiver operator characteristic (ROC), a graphical depiction of the sensitivity versus one (1) minus the specificity of the model often used to validate ecological niche models (230; 278). The AUC indicates whether the model predicts species location better than a random distribution. AUC values of ≤0.5 indicates a random distribution and AUC values >0.9 indicates high reliability of the model (230). To determine the best combination of environmental data for modeling, the model was run four times using different sets of input layers each time: 1) bioclimatic

layers, elevation and rice crop data, 2) bioclimatic layers and elevation data, 3)
bioclimatic layers only, and 4) elevation data only.

RESULTS

Ecological Niche Model of *Cx. tritaeniorhynchus*

A total of 139 unique sites of documented *Cx. tritaeniorhynchus* geographical
locations were utilized to construct the ecological niche model (Figure 4). Of the 139
total points, 105 (76%) were randomly designated as training points in order to build the
model and 34 (24%) points were used to test the model. The model was run four times
using different combinations of environmental layers (Table 8). Statistical results indicate
that the most accurate model included bioclimatic layers and elevation (Table 8), and
therefore this model was used in all subsequent analyses. Statistical evaluation showed
the model to have a high accuracy, with the AUC>0.9 and low p-values. The model is
available to view or download from www.vectormap.org.

In order to evaluate the contribution of each environmental variable to the model,
Maxent utilizes a jackknife test, which indicated that the annual precipitation (bio12)
environmental layer is the environmental variable with the highest gain when used in the
model by itself. The Maxent program also calculates a percent contribution for each
variable in the model. The annual precipitation variable contributed 16.2% of the
information used by the model, another indication that it is an important environmental
factor for estimating the distribution of *Cx. tritaeniorhynchus* (Table 7). The mean
temperature of the wettest quarter variable (bio08) contributed the highest percentage
(21.7%) of the information to the model. Elevation was also an important variable,
contributing 9.6% to the model. From the jackknife test, if elevation data were removed

33

from the model, the overall training gain would decrease the most, indicating the

elevation variable contained the most unique information of the variables in the *Cx.*

tritaeniorhynchus distribution model.

The values of each environmental variable at each recorded location of

occurrence were extracted using ArcGIS (Table 7). For example, the known locations for

Cx. tritaeniorhynchus used in the model fell within 0 and 838 meters of elevation. This is

consistent with the published reports that *Cx. tritaeniorhynchus* is rarely collected above

1,000 meters (220; 225).

Human JE Cases and Vector Presence Estimation

Ninety-six reported JE case locations were identified in endemic regions (Figure

3). ArcGIS analysis categorized human JE cases based on the estimated probability of

Cx. tritaeniorhynchus presence (Figure 5). Human JE cases were identified at locations

with a range of estimated probability of vector presence, including regions with 25% or

less estimated probability. However, the majority (>75%) of human JE cases were

reported from regions with greater than 25% estimated probability of *Cx.*

tritaeniorhynchus presence. Limited availability of location data of human JE cases

greatly impacts any associations between areas of high estimated vector probability and

disease. For instance, the lack of human JE cases in other regions of estimated high

probability of vector presence could be due to lack of reporting, improper diagnosis, or

due to successful prevention strategies.

***Cx. tritaeniorhynchus* Presence Estimation Per Country**

ArcGIS analysis determined the approximate percentage of each country with

>25% probability of *Cx. tritaeniorhynchus* presence based on the Maxent model (Table

6). Of the 25 endemic countries, seven possessed >50% of their land area with a higher probability of *Cx. tritaeniorhynchus* presence. Three countries (Bhutan, Pakistan, and Russia) possessed <1% of their total country area with a 25% probability of *Cx. tritaeniorhynchus* presence.

DISCUSSION

In this study, a statistically significant ecological niche model for *Cx. tritaeniorhynchus* was developed using mosquito presence records, climate, and elevation variables. Locations of human cases of JE generally fell within the higher probability areas of *Cx. tritaeniorhynchus* (Figure 5). Regions of estimated high probability of *Cx. tritaeniorhynchus* presence (Figure 4) are representative of preferred environments, based on temperature, precipitation and elevation where *Cx. tritaeniorhynchus* habitats occur. This model serves as a tool to fill in knowledge gaps regarding *Cx. tritaeniorhynchus* and can be utilized by health care professionals and policy officials in endemic regions to help guide the development and implementation of disease mitigating strategies in endemic regions.

The Maxent program identifies important environmental variables that are major contributors to the vector distribution model. Based on the jackknife test of variable importance, the annual precipitation (bio12) is an important contributor to the model. Additionally, the mean precipitation of the wettest quarters (bio08) and elevation also contributed greatly to the model for distribution of *Cx. tritaeniorhynchus* (Table 7). Previous studies that aimed to identify favorable ecological conditions of mosquitoes found that the optimal temperature of JEV vectors is between 22.8 and 34.5°C (200). The importance of temperature during the wet season in the model is attributed to

35

temperatures and flooded habitats that are optimal for larval development and adult survival. Locations in which the temperatures do not fall into the optimal range during the rainy season may therefore experience fewer mosquitoes, despite harboring the appropriate habitat. Temperature also plays a role in disease transmission rates, as higher temperatures increase the rates of virus replication and dissemination, while decreasing the time from mosquito infection to transmission of the virus to animal and human hosts (280).

Sampling bias is an issue that affects the accuracy of the model as the model was developed using existing data from the literature and VectorMap. Therefore, some regions have not been sampled in the study area and some have been oversampled. *Cx. tritaeniorhynchus* data for China were very limited (Figure 3), which may mean that some potential environmental conditions of *Cx. tritaeniorhynchus* were not represented in the model, in particular, the cooler Northeast region of China. Because modeling was limited to Cx. *tritaeniorhynchus*, there is a potential that for some regions, other primary or secondary vectors, i.e., *Cx. annulirostris*, *Cx. bitaeniorhynchus*, and *Cx. vishnui*, may predominate and maintain transmission of JEV in these areas. Collection records of *Cx. tritaeniorhynchus* were obtained spanning many decades and at different times during the year, furthering the impact of sampling bias on our model. In addition, the density of *Cx. tritaeniorhynchus* was not collected in this study and is an important limitation as vector abundance plays a crucial role in disease transmission. Further collection studies are therefore needed to determine the abundance of vector species in addition to presence in endemic regions.

Low-lying flooded areas containing grasses, including rice paddies, are the primary larval habitats for *Cx. tritaeniorhynchus*. An increase in the amount of flooded rice field habitat has shown to be positively correlated with increases in adult populations of *Cx. tritaeniorhynchus* in the ROK (240). Although the rice map derived from the GeoCover Land Cover map (Figure 6) does generally match the predicted occurrence of *Cx. tritaeniorhynchus*, there are some areas where the model predicts the presence of the mosquito, yet no rice crops were mapped. For some areas, rice may not have been identified correctly on the satellite images, since agricultural areas were limited or were adjacent to other predominant habitats. For example, rice is produced in Nepal (133), but no rice fields were identified by GeoCover in Nepal, since the identification of small rice fields in mountainous areas on satellite images can be difficult. Alternatively, this shows that environments other than rice fields are suitable habitat for *Cx. tritaeniorhynchus*.

The predicted probability of *Cx. tritaeniorhynchus* presence values were used to determine the percentage of a country at high risk (greater than 25% probability) for vector presence (Table 6). Many Asian countries have high percentages of their total land area with a >25% probability for the presence of *Cx. tritaeniorhynchus*. Cambodia, the ROK, Sri Lanka, and Thailand have over 75% of their land area with a >25% probability of *Cx. tritaeniorhynchus* presence. Countries demonstrating >50% of their total land area and with *Cx. tritaeniorhynchus* occurrence >25% probability includes: Bangladesh, East Timor, and Vietnam. However, some countries may have small areas of vector habitat close to large populations that can result in outbreaks despite low percentage estimated probability in the region overall. Further, this analysis does not take into account country size or vector abundance, which would also impact disease transmission. Although

additional factors contribute to JE disease risks, the distribution of the vector populations within a country is a valuable data set when considering the necessity of vaccination and other health risk reduction programs.

Human JE cases were categorized based on the estimated probability of vector presence at the reported location (Figure 5). Interestingly, a portion of human cases were reported from regions with 25% or less estimated probability of *Cx. tritaeniorhynchus* presence. JE cases not falling within high probability pixels could have been acquired in nearby locations. Even for precisely located case data, the high resolution of the model (one kilometer pixels) increases the likelihood that the predicted *Cx. tritaeniorhynchus* location does not match disease acquisition location as many people travel more than a kilometer in the course of a typical day. Alternatively, additional factors other than the presence of *Cx. tritaeniorhynchus* may be important when determining the risk of disease. For instance, other JEV vectors may dominant in these regions estimated with low *Cx. tritaeniorhynchus* presence. JE is one of many febrile illnesses that affect human populations in Asia. Difficulties arise in diagnosis of JE in patients based on symptoms alone that range from mild to very severe, with laboratory tests required for confirmation. Obtaining geographical data of where human cases were acquired is made difficult due to lack of confident diagnoses, patient travel history, and spatial data. The lack of precision of the reported case locations may also contribute to lower numbers of JE cases falling within high probability *Cx. tritaeniorhynchus* pixels. Identification of human JE cases in this study is extremely limited and does not represent all human cases in JE endemic regions. In many cases, only a village or city name was given for reported cases. A previous study to model the distribution of *Cx. tritaeniorhynchus* to predict JE in the

Republic of Korea found human cases to occur in areas of high estimated probability of vector presence (181). This study, however, utilized intensive vector collection methods and JE case data were obtained from the Korea Centers for Disease Control resulting in an overall more extensive and accurate model. This illustrates the need for increased surveillance of vector and human JE cases in order to generate more accurate risk models for JE. In order to evaluate the impact of vector presence on the risk of JE in humans, comprehensive efforts to identify specific locations of both symptomatic and asymptomatic JE cases across endemic regions are needed.

Ecological niche modeling inherently possesses limitations in that it makes predictions based solely on environmental variables that impact larval development and adult survival. Other important factors that influence vector distributions include: vector control strategies, public health campaigns, socioeconomic status, human population densities, anthropogenic changes to land (creation of vector habitat), vector species competition, and predator influences on their potential distribution and population densities. Further, the use of WorldClim data may underestimate or ignore environmental variables that occur during a short time period or transient habitat suitable for the vector to survive. Incorporation of these variables will undoubtedly increase the validity of the model. These factors are also important to take into consideration when implementing mosquito control initiatives and vaccination campaigns.

The reemergence of JEV remains possible due to multiple factors. Increases in the pig farming industry, modification and expansion of arable lands for wetland rice farming, and a fraction of the population unvaccinated/non-immune, in combination with optimal climatic conditions, contribute to the potential for periodic outbreaks of JE as the

one observed in the ROK (143). Genotype analyses of circulating JEV strains identified the reemergence of genotype V, which was unseen in Asia for over 50 years (162; 283). The identification of emerging/reemerging JEV strains is important for vaccine development and the implementation of effective vaccination programs. Increased surveillance in areas with known vector populations and additional risk factors, such as reservoir and amplifying hosts, will aid in the identification of circulating JEV strains as well as strains that are emerging in novel human and vector populations. Understanding the vector distribution is a key step to effectively understanding JEV risks and also to preventing additional outbreaks of JE in endemic countries.

Table 6. Summary of JEV vaccination programs in endemic countries and predicted percentage of land with greater than 25% estimated probability of *Cx. tritaeniorhynchus* presence based on the ecological niche model.

JEV Endemic Countries	JE vaccination program status[1]	Percentage of area >25% estimated probability of *Cx. tritaeniorhynchus* presence
Australia	Administered in endemic areas	4.6
Bangladesh	No Current Immunization Program	56.8
Bhutan	No Current Immunization Program	0
Brunei	No Current Immunization Program	16.3
Cambodia	No Current Immunization Program	79.4
China	National Vaccination Program 2010	3.6
East Timor	No Current Immunization Program	68.4
India	Vaccine administered in high risk areas, not integrated into routine immunization program	19.2
Indonesia	No Current Immunization Program	14
Japan	National Vaccination Program 2010	42.4
Laos	No Current Immunization Program	19.7
Malaysia	Regional vaccination	8.4
Myanmar	No Current Immunization Program	20.5
Nepal	Vaccine introduced in 2006, not widely implemented	2.8
Democratic People's Republic of Korea	DPRK originated vaccine provided in high risk areas	21.1
Pakistan	No Current Immunization Program	0.06
Papua New Guinea	No Current Immunization Program	11.1
Philippines	No wide scale vaccination program in place, vaccine trial in progress	46
Republic of Korea	Government mandated mass immunization began in 1971	78.8
Russia	No data	0.01
Singapore	No data	10.7
Sri Lanka	18 of 26 districts receive vaccine annually, plans to extend to all districts	85.2
Taiwan	National Vaccination Program	34.2

Thailand	National Vaccination Program 2010	80.9
Viet Nam	Vaccine distributed in high risk districts	61
Western Pacific (Guam, Saipan)	No data	No data

[1] Obtained from PATH: Japanese Encephalitis Morbidity, Mortality, and Disability: Reduction and Control by 2015 (222) and WHO/IVB database, 193 WHO Member States. Data as of September 2011 (216).

Table 7. Minimum, maximum, mean values and percent contribution of environmental
data layers for the *Cx. tritaeniorhynchus* model.

Variable	Description	Min	Max	Mean	Percent Contribution
Alt	Altitude (elevation above sea level), m	0	838	153.4	9.6
Bio01	Annual mean temperature, ^{o}C	8.2	28.9	23.3	4.4
Bio02	Mean diurnal range (Mean of monthly (max temp-min temp)),^{o}C	4.9	15	9.4	4.7
Bio03	Isothermality [(Bio2/Bio7)*100], ^{o}C	2.2	9	5.1	1.8
Bio04	Temperature Seasonality (standard deviation * 100), ^{o}C	30.1	1031.3	366	3.3
Bio05	Max temperature of the warmest month, ^{o}C	25.8	42.5	33.4	1.2
Bio06	Min temperature of the coldest month, ^{o}C	-12.5	24.4	12.4	3.3
Bio07	Temperature annual range (Bio5-Bio6), ^{o}C	7.2	40.7	21	3.2
Bio08	Mean temperature of the wettest quarter[1], ^{o}C	16.9	32.8	26.2	21.7
Bio09	Mean temperature of the driest quarter, ^{o}C	-4.9	28.7	19.5	0.3
Bio10	Mean temperature of the warmest quarter, ^{o}C	20.1	34.3	27.7	0.6
Bio11	Mean temperature of the coldest quarter, ^{o}C	-4.9	27.2	18.3	1
Bio12	Annual precipitation, mm	152	4005	1610.3	16.2
Bio13	Precipitation of the wettest month, mm	41	1011	319.4	5.9
Bio14	Precipitation of the driest month, mm	0	233	30.8	6

Bio15	Precipitation seasonality (coefficient of variation), mm	18	138	74.5	5.8
Bio16	Precipitation of the wettest quarter, mm	95	2455	797.4	1
Bio17	Precipitation of the driest quarter, mm	0	786	114.9	0.7
Bio18	Precipitation of the warmest quarter, mm	62	1015	467.2	0.7
Bio19	Precipitation of the coldest quarter, mm	11	1812	260.7	8.5

[1]A quarter is a period of three months

Table 8. Maxent model accuracy analysis using different sets of environmental data inputs.

Environmental Data Input	AUC Training Points	AUC Test Points	P-value Minimum Training Presence
Bioclimatic and Elevation data	0.971	0.932	<0.0001
All layers (including rice crop)	0.968	0.919	<0.0001
Bioclimatic data only	0.968	0.929	<0.0001
Elevation data only	0.822	0.849	<0.0001
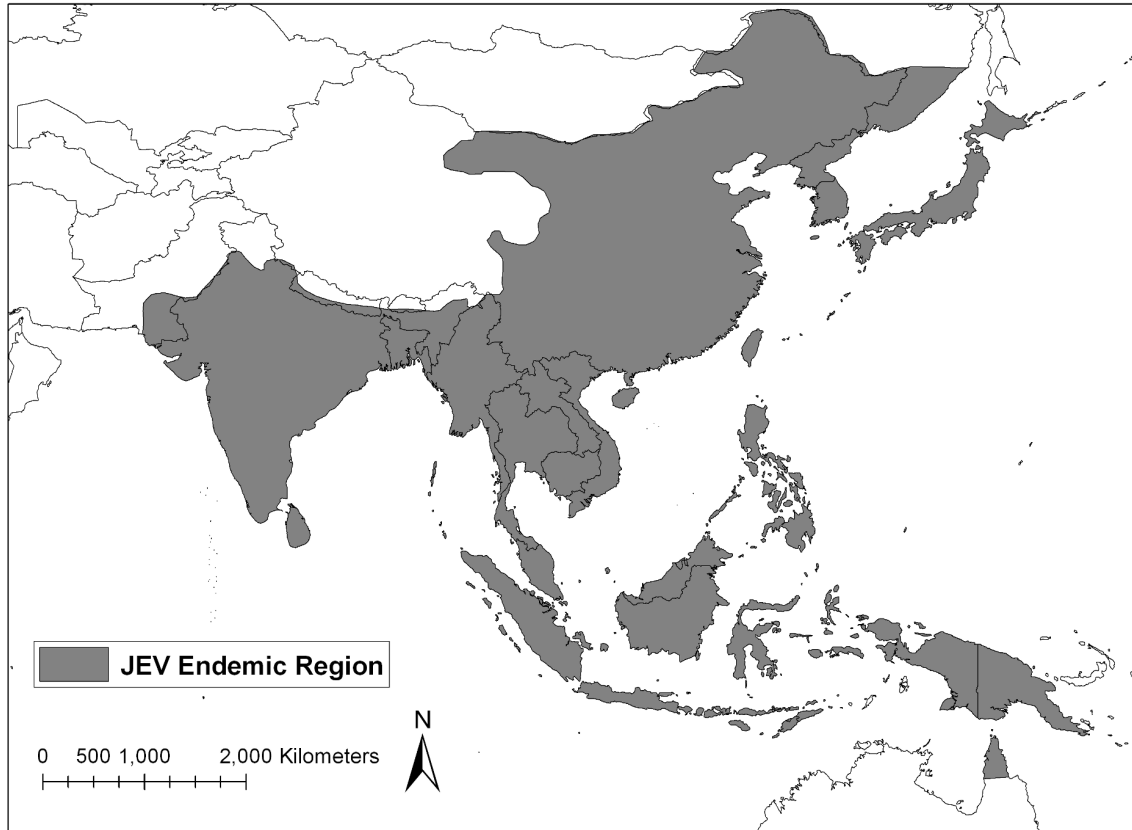
Figure 2. Japanese encephalitis virus endemic area.
Map adapted from CDC.
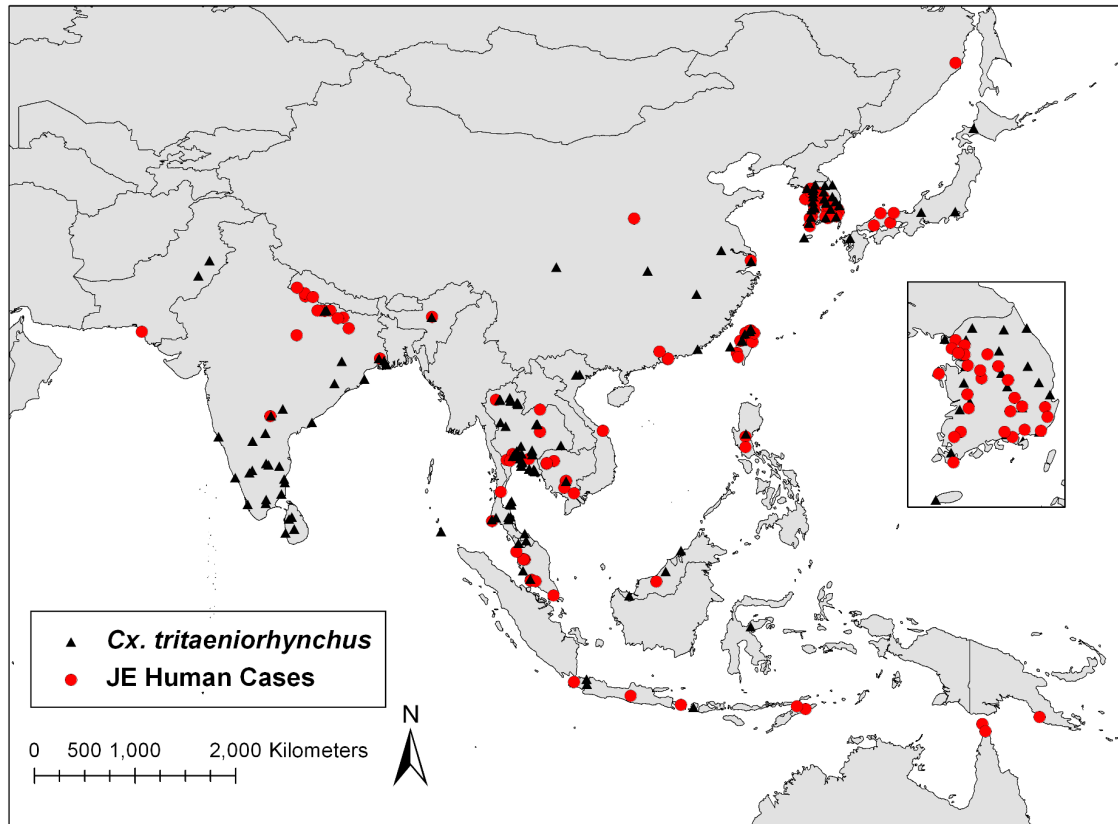
Figure 3. Distribution of known *Cx. tritaeniorhynchus* locations and documented human cases of JE within endemic region.
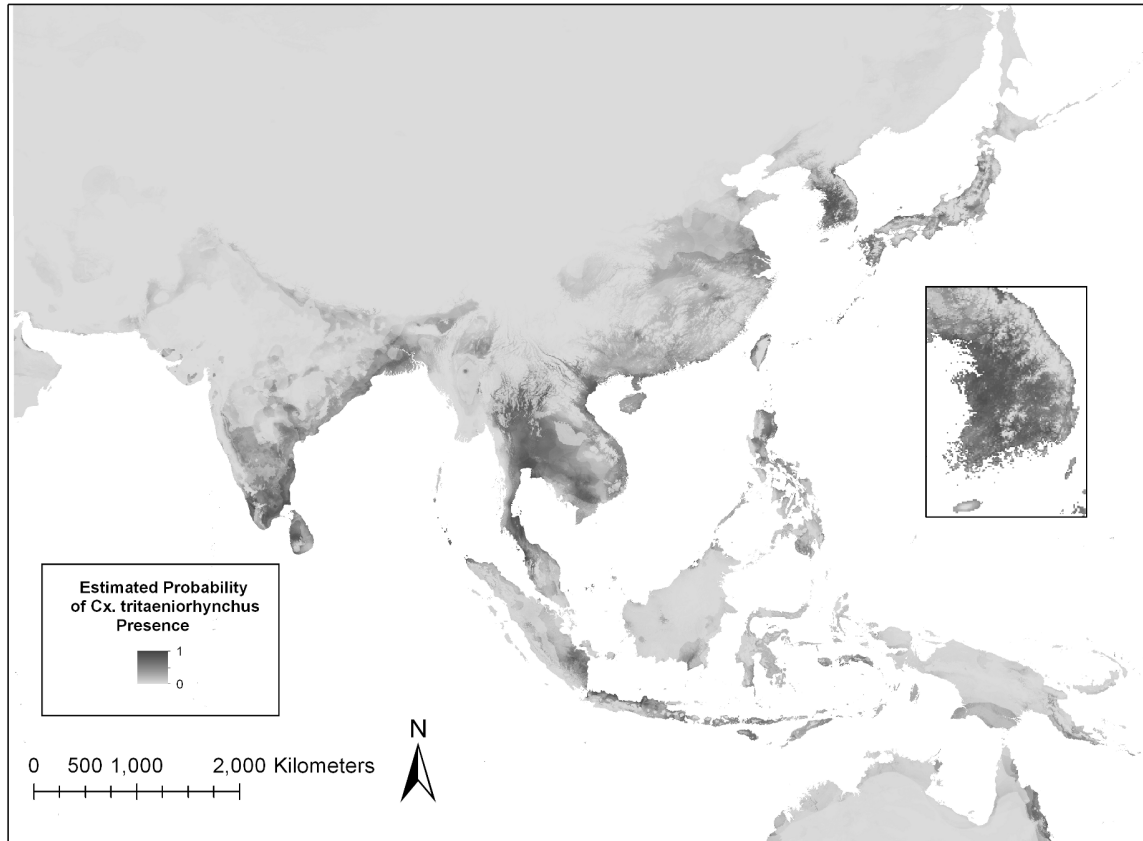
Figure 4. Maxent model estimation of the probability of *Cx. tritaeniorhynchus* distribution in the JE endemic region.

   Darker areas indicate areas that are likely to have suitable habitat for this vector species while lighter areas indicate areas of that are less suitable for the vector.
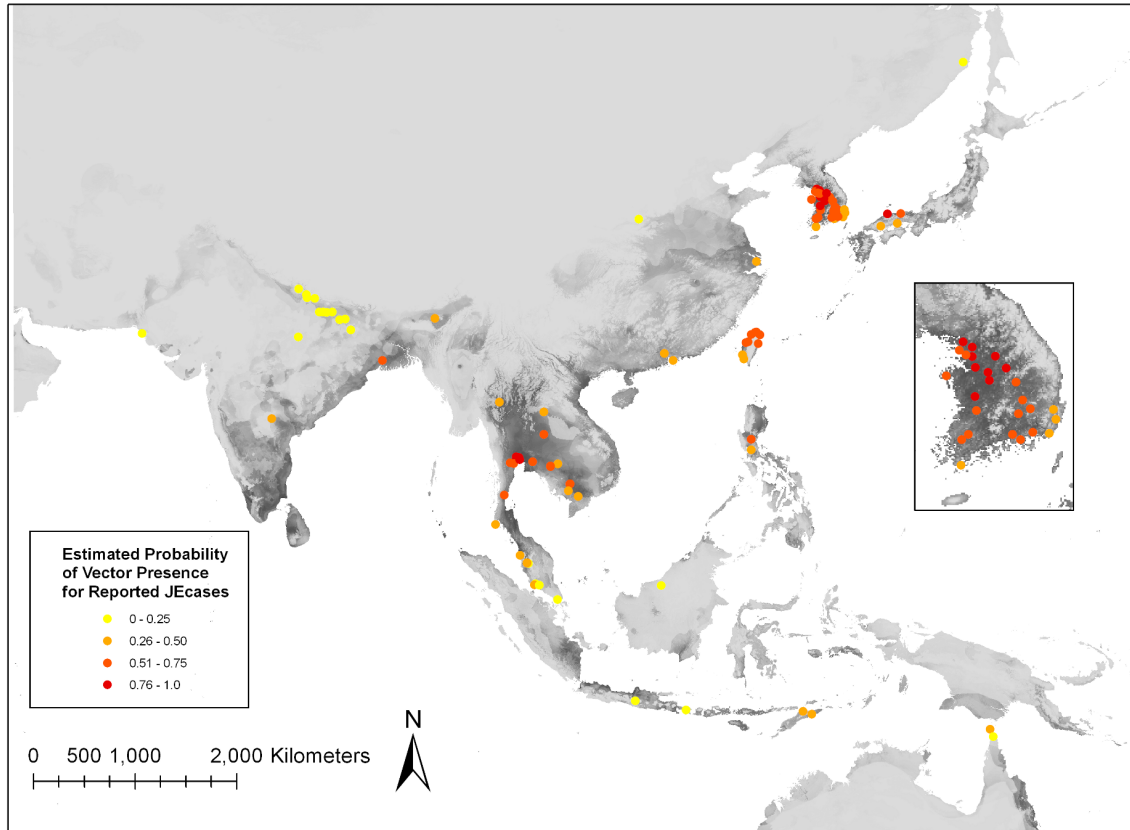
Figure 5. Human JE cases categorized by color based on the estimated probability of *Cx. tritaeniorhynchus* presence.
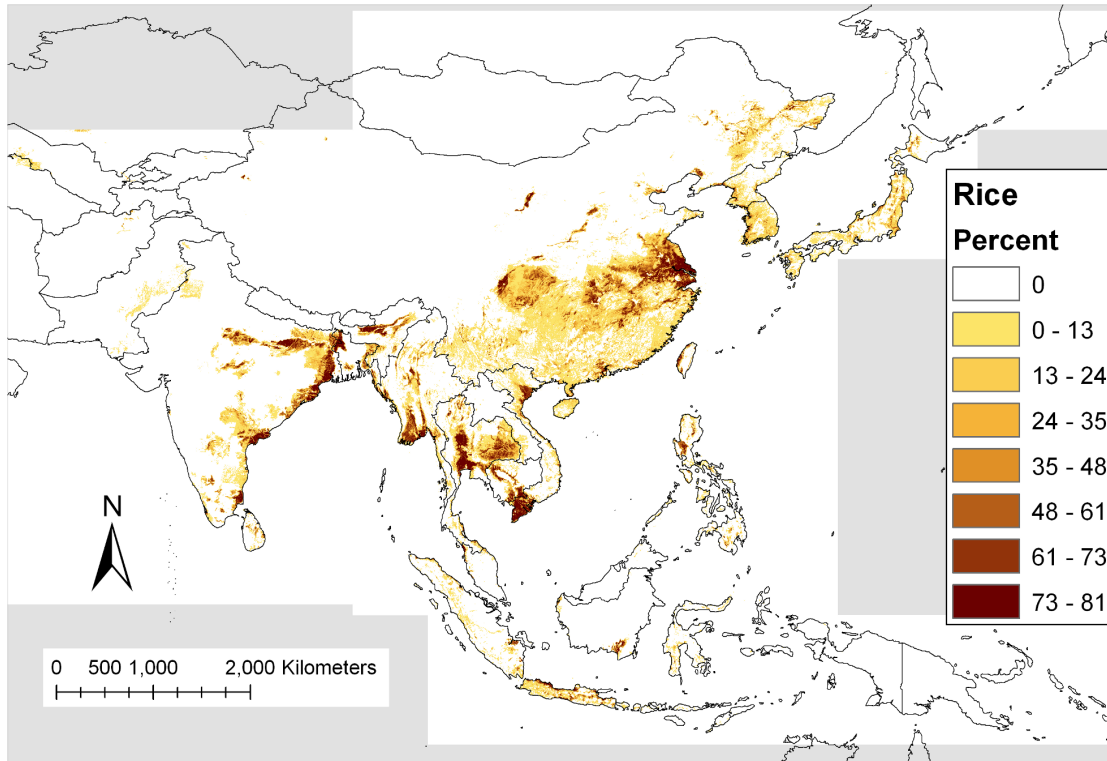
Figure 6. Percent of 30 meter pixels classified as rice land cover within 1 one square kilometer derived from the GeoCover Land Cover product.
    Gray areas indicate no data.

CHAPTER 3: Characterization of *Plasmodium ovale curtisi* and *P. ovale wallikeri* in Western Kenya Utilizing a Novel Species-specific Real-time PCR Assay

Robin H. Miller[1], Clifford O. Obuya[2], Elizabeth W. Wanja[2], Bernhards Ogutu[3], John Waitumbi[3], Shirley Luckhart[4], V. Ann Stewart[1]

[1]Preventive Medicine and Biometrics, Uniformed Services University, Bethesda, MD, USA

[2]Kondele Laboratory, U.S. Army Medical Research Unit-Kenya, Kisumu, Kenya

[3]Walter Reed Project, Kenya Medical Research Institute, Kisumu, Kenya

[4]Department of Medical Microbiology and Immunology, University of California Davis School of Medicine, Davis, CA, USA

Table, figure, and reference numbers have been changed to follow the USU thesis format guidelines.

ABSTRACT

Background

   Plasmodium ovale is comprised of two genetically distinct subspecies, *P. ovale* curtisi and *P. ovale wallikeri*. Although *P. ovale* subspecies are similar based on morphology and geographical distribution, allelic differences indicate that *P. ovale curtisi* and *P. ovale wallikeri* are genetically divergent. Additionally, potential clinical and latency duration differences between *P. ovale curtisi* and *P. ovale wallikeri* demonstrate the need for investigation into the contribution of this neglected malaria parasite to the global malaria burden.

Methods

   In order to detect all *P. ovale* subspecies simultaneously, we developed an inclusive *P. ovale*-specific real-time PCR assay based on conserved regions between *P. ovale curtisi* and *P. ovale wallikeri* in the reticulocyte binding protein 2 (*rbp2*) gene. Additionally, we characterized the *P. ovale* subspecies prevalence from 22 asymptomatic malaria infections using multilocus genotyping to discriminate *P. ovale curtisi* and *P. ovale wallikeri*.

Results

   Our *P. ovale rbp2* qPCR assay validation experiments demonstrated a linear dynamic range from 6.25 *rbp2* plasmid copies/microliter to 100,000 *rbp2* plasmid copies/microliter and a limit of detection of 1.5 *rbp2* plasmid copies/microliter. Specificity experiments showed the ability of the *rbp2* qPCR assay to detect low-levels of P. ovale in the presence of additional malaria parasite species, including *P. falciparum*, *P. vivax*, and *P. malariae*. We identified *P. ovale curtisi* and *P. ovale wallikeri* in Western

51

Kenya by DNA sequencing of the tryptophan-rich antigen gene, the small subunit ribosomal RNA gene, and the *rbp2* gene.

Conclusions

Our novel P. ovale *rbp2* qPCR assay detects *P. ovale curtisi* and *P. ovale wallikeri* simultaneously and can be utilized to characterize the prevalence, distribution, and burden of *P. ovale* in malaria endemic regions. Using multilocus genotyping, we also provided the first description of the prevalence of *P. ovale curtisi* and *P. ovale wallikeri* in Western Kenya, a region holoendemic for malaria transmission.

AUTHOR SUMMARY

Humans can be infected with five malaria parasite species: *Plasmodium falciparum*, *P. vivax*, *P. malariae*, *P. knowlesi*, and *P. ovale*. Although the vast majority of malaria morbidity and mortality worldwide can be attributed to *P. falciparum*, non-falciparum malaria parasites can also cause clinical disease. Researchers use nucleic acid based detection methods, such a polymerase chain reaction (PCR), to detect low-density malaria parasitemias that can evade microscopic detection. *P. ovale* was recently identified to exist as two subspecies, *P. ovale curtisi* and *P. ovale wallikeri*, that look identical but differ genetically. In this study, we developed a novel real-time PCR (qPCR) assay to detect all *P. ovale* parasites, based on a conserved gene between *P. ovale curtisi* and *P. ovale wallikeri*. We also used DNA sequencing to differentiate between *P. ovale curtisi and P. ovale wallikeri* from a small sample of *P. ovale* asymptomatic infections in Western Kenya. Through the use of our novel *rbp2* qPCR assay, we aim to characterize the prevalence of *P. ovale* in future epidemiological studies in order to better understand this neglected malaria parasite species.

INTRODUCTION

Plasmodium ovale, the causative agent of benign tertian malaria, was identified as

a distinct malaria parasite species in 1922 based on its characteristic oval morphology in

infected erythrocytes (273). *P. ovale* rarely causes severe disease in humans living in

malaria endemic regions, but can cause serious clinical disease in naive travelers (64; 74;

157; 165; 185; 249; 256; 274). The actual prevalence and clinical relevance of *P. ovale* is

likely underestimated for the following reasons. First, *P. ovale* is often found as a mixed

infection with other malaria parasite species (46; 79; 198). This can confound

microscopic identification of *P. ovale* due to difficulties in differentiating *P. ovale* from

other morphologically similar malaria parasites, such as *P. vivax.* Second, the

characteristic low-level parasitemia of *P. ovale* infection further complicates microscopic

detection due to the difficulty in finding and identifying low numbers of *P. ovale*

parasites (65). Finally, malaria Rapid Diagnostic Tests (RDTs) show a reduced ability to

detect *P. ovale* compared to other human malaria parasites, resulting in false negative

cases (49; 73; 111). However, the use of extremely sensitive molecular detection

methods, such as polymerase chain reaction (PCR), have revealed a higher prevalence of

P. ovale and expanded the geographical distribution of this malaria parasite compared to

what was previously identified based on microscopy (27; 80; 139; 198; 262).

Recent findings demonstrated that *P. ovale* exists as two genetically distinct

sympatric subspecies, *P. ovale curtisi* and *P. ovale wallikeri* (102; 213; 276; 318).

Morphological differences between the two *P. ovale* subspecies have not been identified,

thereby limiting the use of microscopy to differentiate *P. ovale curtisi* and *P. ovale*

wallikeri. As recent studies suggest potential clinical and latency duration differences

53

between the two *P. ovale* subspecies, (207; 245), a discriminatory assay to differentiate

P. ovale curtisi and *P. ovale wallikeri* is clinically relevant. Additionally, initial *P. ovale*-

specific assays developed by our group and others were unknowingly designed based on

gene sequences specific to only one subspecies, thereby failing to detect the other *P.*

ovale subspecies. PCR assays that target conserved genetic regions between the two

subspecies are, therefore, necessary to determine the true *P. ovale* prevalence and

distribution (38; 57; 101; 286).

Small-subunit ribosomal RNA (ssrRNA) genes are common targets for malaria

parasite species-specific assays based on nucleotide polymorphisms that facilitate

specific detection of the species of interest (57; 101; 250). Although rRNA based PCR

assays have proven useful for the detection of low-level parasitemias of a single malaria

parasite species, Demas et al. demonstrated that alternative gene targets may be more

sensitive for species-specific detection in the context of mixed species infections (76). A

quality control program to determine the ability of 10 different laboratories to detect

malaria parasite species based on rRNA PCR revealed detection of *P. ovale* to be the

most difficult, with a detection rate of 70% (290). Additionally, allelic diversity within

the *P. ovale* ssrRNA alleles may further limit the ability of rRNA specific PCR assays to

detect *P. ovale* infections (161). Due to these difficulties in the detection of *P. ovale*, we

designed a novel *P. ovale*-specific assay based on a gene found only in *P. ovale curtisi*

and *P. ovale wallikeri* and not present in other human malaria parasite species. This

approach reduces aberrant amplification of non-target malaria species and allows for the

detection of low-level *P. ovale* infections in the presence of high parasitemias of other

malaria parasite species, such as *P. falciparum*.

Several epidemiology surveys of extant malaria species have established the endemicity of *P. ovale* in Western Kenya based on microscopic identification, entomological studies, and nucleic acid detection methods (42; 65; 172; 211; 300). Clinical cases due to *P. ovale* relapse in non-immune individuals after traveling to Western Kenya have also been reported, including a single case of a returned traveller with *P. ovale curtisi* infection (207; 224). However, the lack of data on the prevalence and distribution of *P. ovale curtisi* and *P. ovale wallikeri* in Western Kenya represents a critical gap in our understanding of the true malaria epidemiology in this region that could impact both patient treatment and malaria control strategies.

In this study, we developed a novel, highly specific, real-time PCR (qPCR) assay to detect all *P. ovale* subspecies simultaneously based on a conserved region of the *P. ovale*-specific reticulocyte binding protein 2 (*rbp2)* gene. This inclusive *P. ovale rbp2* qPCR assay was characterized and validated to determine the sensitivity, limit of detection, limit of quantification, specificity, repeatability, and reproducibility. In addition, the occurrence of both *P. ovale* subspecies (*P. ovale curtisi* and *P. ovale wallikeri)* was documented in Western Kenya using multilocus genotyping. Our *P. ovale* species-specific assay can be utilized to better characterize the presence, parasitemia, geographical distribution, and the contribution of this malaria parasite species to mixed species infections and to clinical disease in malaria endemic regions.

METHODS

Sample Collection

Anonymized human whole blood samples were collected with signed informed consent under approved protocols (Walter Reed Army Institute of Research Human Use

and Review Committee Protocols #1720 and 1306, Kenya Medical Research Institute

(KEMRI) SSC#2008 and 1111). Clinically healthy (asymptomatic) adult individuals in

Nyanza Province, Kenya were screened (active detection) with the Parascreen Pan/Pf [®]

malaria Rapid Diagnostic Test (Zephyr Biomedicals, Verna, Goa, India) for the

presence/absence of malaria parasites from March through September of 2008. Thin and

thick smears were examined subsequently by up to 5 expert microscopists in the Malaria

Diagnostic Centre (MDC), Kisumu, Kenya for malaria species designation and estimation

of quantitative parasitemia (214). Samples identified as positive for *P*. *ovale* (n=22) via

microscopy, in which all were mixed infections with other malaria species, were targeted

for DNA extraction and PCR based analysis. DNA was extracted from 200 microliters of

whole blood using the QIAamp DNA Minikit (Qiagen, Venlo, Netherlands) following the

manufacturer's protocol. DNA was eluted in 200 microliters of Buffer EB and samples

were stored at -20°C until time of use. A human-specific RNaseP based qPCR assay was

performed for each sample in duplicate to confirm successful nucleic acid extraction

(136).

Characterization of *P*. *ovale* subspecies in Western Kenya

Tryptophan-rich antigen (tra) gene

The *P*. *ovale*-specific tryptophan-rich antigen (*tra*) gene was recently identified as

a target to discriminate between *P*. *ovale* subspecies based on DNA sequence length and

single nucleotide polymorphisms (SNPs) (213; 276; 286). We utilized the PoTRA fwd3

and PoTRA rev3 primers reported in Oguike et al. 2011 for PCR analysis (213). Primers

(Table 9) were synthesized by Integrated DNA Technologies (IDT, Coralville, IA, USA)

and purified by standard desalting methods. Each PCR assay consisted of 1X Sigma

JumpStart REDTaq ReadyMix (20 mM Tris-HCl, 100 mM KCl, 4 mM MgCl$_2$, 0.4 mM

of each dNTP, 0.03 unit/µl of Taq DNA polymerase, Sigma, Balcatta, WA, USA), 8.75

picomoles of each primer, and one microliter of template with a final volume of 25

microliters. PCR cycling conditions were: initial denaturation for 2 minutes at 95°C

followed by 45 cycles of 95°C for 30 seconds, 58°C for 45 seconds, 72°C for 1 min and a

final extension at 75°C for 5 minutes. All conventional PCRs were performed on a DNA

Engine PTC-200 Thermal Cycler (MJ Research, Waltham, MA, USA).

Reticulocyte binding protein 2 (rbp2) gene

The reticulocyte binding protein 2 (*rbp2)* gene was utilized by Oguike et al. 2011

to differentiate between *P. ovale* subspecies using qPCR melt curve profiles based on six

SNPs present within a 120 base pair fragment. We designed a novel set of primers (Table

9, IDT) using Primer Express software (Life Technologies, version 3.0; Frederick, MD,

USA) to amplify a smaller, 74 base pair region of the *rbp2* gene for assay development.

Our primers (PoRBP2f and PoRBP2r) are located within conserved DNA sequences of

the *P. ovale* subspecies to ensure detection and amplification of both *P. ovale* subspecies.

The amplicon also contains a single SNP to distinguish *P. ovale* subspecies by DNA

sequencing. Figure 7 shows the single SNP in the *rbp2* amplicon at position 431, in

which *P. ovale curtisi* contains an adenine and *P. ovale wallikeri* contains a thymine.

Primer BLAST was utilized to ensure our primers were specific for *P. ovale* and would

not amplify non-*P. ovale* malaria parasite DNA or human DNA. PCRs consisted of 1X

Sigma JumpStart REDTaq ReadyMix Reaction Mix, 25 picomoles of each primer, and

one microliter of template, with a final volume of 25 microliters. PCR cycling conditions

were as follows: initial denaturation at 95°C for 2 minutes followed by 40 cycles of 95°C

for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds, and a final extension at 72°C for 10 minutes.

Small subunit ribosomal RNA (ssrRNA) gene

We utilized *P. ovale*-specific primers (Table 9, IDT) reported by Fuehrer et al. 2012 (rOVA1WC and rOVA2WC) to further characterize *P. ovale* positive samples based on differences within the small subunit ribosomal RNA (ssrRNA) gene (101). PCRs consisted of 1X Sigma JumpStart REDTaq ReadyMix Reaction Mix, 25 picomoles of each primer, one microliter of template, and a final volume of 25 microliters. PCR cycling conditions were as follows: initial denaturation at 95°C for 4 minutes followed by 35 cycles of 94°C for 1 minute, 58°C for 2 minutes, 72°C for 2 minutes, and a final extension at 72°C for 5 minutes.

DNA sequencing

PCR products were visualized on 0.7% agarose gels stained with ethidium bromide. PCR products were cloned into the pCR 2.1-TOPO TA vector (Life Technologies) based on manufacturer's guidelines. Plasmid purification was performed using the QIAprep Spin Miniprep kit (Qiagen) and used as template for sequencing reactions. PCR products were sequenced using the M13 Forward (-20) Primer (Life Technologies) at the Biomedical Instrumentation Center at the Uniformed Services University or GENEWIZ Inc (Germantown, MD, USA) using the ABI 3500XL Genetic Analyzer and the ABI 3730XL DNA Analyzer, respectively. Sequencing facility was chosen based on temporal availability. DNA sequences were aligned and analyzed with previously published sequences using SeqMan software (DNAStar Lasergene Version

8.1.5, Madison, WI, USA). Reference sequences utilized for DNA alignments are shown in Table 10.

Real-time PCR assay to detect *P. ovale*

Primer Express software (Life Technologies, version 3.0) was utilized to design a hydrolysis probe (Table 9) for use with our *rbp2* primers on the ABI 7500 fast real-time PCR (qPCR) platform (Life Technologies). An alignment of the *P. ovale rbp2* DNA sequences was constructed using the Clustal Omega Program provided by the European Molecular Biology Laboratory – European Bioinformatics Institute (EMBL-EBI) (110; 259). We utilized the Jalview output tool to visualize the DNA sequence alignment (Figure 7)(314). Primers and probe were designed in order to amplify a conserved region within the *rbp2* gene to ensure detection of both *P. ovale* subspecies by our qPCR assay at the same time. *In silico* analyses were performed to ensure primers and probe were specific to *P. ovale* and would not amplify genes of other malaria parasites or human DNA. Each qPCR reaction consisted of the following: 1X TaqMan Fast Universal PCR Master Mix, No AmpErase UNG (Life Technologies, Cat No. 4364103), 5 picomoles of each primer and probe, and one microliter of template in a final volume of 20 microliters. Real-time PCR was performed utilizing fast thermal cycling conditions (95°C for 20 seconds, followed by 40-60 cycles of 95°C for 3 seconds and 60°C for 30 seconds). Analysis of qPCR results was performed using ABI 7500 Fast Real-Time PCR Systems Software (Life Technologies, Version 2.0.5). Basic statistical analyses (means, standard deviations, coefficient of variation), generation of standard curve graphs, calculation of slopes, and coefficient of correlation were performed in Microsoft Excel or GraphPad Prism (GraphPad Prism Software Version 6, La Jolla, CA, USA).

Plasmid standard curve

We cloned the 74 base pair *rbp2* amplicon into the pCR 2.1-TOPO TA vector (Life Technologies) following manufacturer's guidelines and eluted the *rbp2* plasmid in PCR grade water. The approximate *rbp2* amplicon copy number per microliter was determined based on spectrophotometer (Nanodrop 2000c) concentration in nanograms per microliter. Plasmids with the *rbp2* amplicon (*rbp2* plasmid) were diluted in water to generate a ten-fold serial dilution from 100,000 *rbp2* copies per microliter to 0.1 *rbp2* copies per microliter. The resulting non-linearized ten-fold serial dilution series was utilized as a standard curve in subsequent validation experiments including determination of the linear dynamic range, specificity, reproducibility, repeatability, and limit of detection. The effect of the conformation of the *rbp2* plasmid on standard curve linearity was analyzed by linearizing the *rbp2* plasmid using the NotI restriction enzyme (New England BioLabs Inc, Ipswich, MA, USA) according to the manufacturer's protocol. *Rbp2* plasmid linearization was confirmed by gel electrophoresis on a 0.7% agarose gel stained with ethidium bromide. Linearized *rbp2* plasmid was purified using the Qiagen PCR Purification Kit following the manufacturer's protocol. The approximate *rbp2* copy number per microliter of the linearized *rbp2* plasmid was determined and diluted in water to generate a ten-fold serial dilution (100,000 to 0.1 copies per microliter). The *rbp2* standard curve PCR efficiency and coefficient of correlation (R^2) were determined and the Pearson product-moment correlation was used to compare the linearized and non-linearized *rbp2* plasmid standard curves (GraphPad Prism).

Validation experiments

Real-time PCR efficiency was determined using a standard curve of 10-fold serial dilutions of the non-linearized *rbp2* plasmid. Efficiency (E) was calculated using the following formula: $E = 10^{(-1/slope)} - 1$. *Rbp2* plasmid standard curve samples were run at least in duplicate and the mean quantification cycle (Cq) value was utilized to generate the standard curve. The limit of detection was defined as the concentration of *rbp2* plasmid in copies per microliter that gave a positive signal in at least one replicate well in two separate qPCR experiments. Limit of quantification was defined as the range of *rbp2* plasmid concentrations that maintained linearity and therefore could be used to quantify *P. ovale* concentration from test samples.

Specificity was analyzed using DNA template from non-*P. ovale* malaria parasite species and uninfected human DNA. Genomic DNAs from *P. falciparum* strains 3D7 (WRAIR), FCR3CSA (ATCC/BEI Resources, MR4, Manassas, Virginia), Dd2 (ATCC/BEI Resources, MR4), and NF54 (ATCC/BEI Resources, MR4) were utilized as template to assess specificity. *P. vivax* genomic DNA was extracted from frozen whole parasites (kind gift of Dr. J. Prachumsri, Mahidol University, Bangkok, Thailand). Since pure *P. malariae* positive samples were unavailable, we utilized three samples collected as part of the blood collection protocol in Kenya that were positive for *P. malariae* as well as *P. falciparum* by microscopy and PCR, but were negative for *P. ovale*. The *P. malariae* parasitemias ranged from approximately 30 to 2400 parasites per microliter. Additionally, genomic DNAs from *P. knowlesi*, *P. simiovale*, *P. fragile*, and *P. cynomolgi* (ATCC/BEI Resources, MR4), were also utilized as templates. Specificity was further analyzed by performing spiking experiments in which a known concentration of

61

rbp2 plasmid was added to template containing *P. falciparum* 3D7 DNA (10,000 parasites per microliter) or *P. vivax* DNA (517 parasites per microliter). One-way analysis of variance (ANOVA) was used to determine differences in Cq values for spiking experiments (GraphPad Prism).

Within-run repeatability was defined as the variation of Cq values within a single run and was analyzed by calculating the percent coefficient of variation (%CV) of Cq values in replicate wells. Between-run repeatability was defined as the variation of Cq values in separate qPCR runs and was determined by calculating the percent coefficient of variation (%CV) of mean Cq values based on six separate qPCR experiments. Reproducibility was evaluated by comparing the assay performance by a technician at the USAMRU-K laboratory in Kisumu, Kenya and the Uniformed Services University in Bethesda, Maryland, USA.

Quantification comparison: microscopy versus rbp2 qPCR

Parasitemias were determined for *P. ovale* positive blood films using standard microscopic methods at the Malaria Diagnostic Centre, affiliated with both USAMRU-K and KEMRI, in Kisumu, Kenya. DNA was extracted from microscopy-positive *P. ovale* samples and tested using the *P. ovale*-specific *rbp2* qPCR assay. Approximate *rbp2* copy number per microliter was determined based on the *rbp2* plasmid standard curve. Parasitemias as determined by expert microscopy (parasites per microliter) were compared to *rbp2* copy number per microliter as determined by the *P. ovale*-specific qPCR in order to examine potential correlation between *rbp2* plasmid copy number and microscopic parasitemias.

P. ovale subspecies characterization

Human-specific RNaseP qPCR

A previously described qPCR assay based on the human-specific RNaseP gene was performed to confirm the presence of nucleic after DNA extraction (136). The human RNaseP gene was detected from all 22 samples (Average Cq=29.12, Cq Range=28.2-32.87, standard deviation=1.02), indicating extraction methods yielded DNA suitable for subsequent PCR experiments.

Tryptophan-rich antigen (tra) gene

Alignments of *tra* gene sequences revealed nine samples (40.9%) positive for *P. ovale curtisi* type 1, two samples (9.1%) positive for *P. ovale curtisi* type 2, six samples (27.3%) positive for *P. ovale wallikeri* type 1, and three samples (13.6%) positive for *P. ovale wallikeri* type 2 (Table 11). Previously published GenBank accession numbers were utilized as reference sequences for alignment and are shown in Table 10. Representative *P. ovale curtisi* type 1, *P. ovale curtisi* type 2, *P. ovale wallikeri* type 1, and *P. ovale wallikeri* type 2 *tra* DNA sequences were deposited under GenBank accession numbers KM494978-KM494981, respectively, and are identical to the reference sequences. As shown in Table 12, unique polymorphisms within the *tra* gene were also detected and confirmed by at least two separate sequencing reactions for 5 samples: Po05, Po12, Po20, Po06, and Po07 (Accession numbers KM494982-KM494986, respectively). Samples Po12 and Po20 contained an 18 base pair insertion between nucleotide positions 171 and 172 (based on *P. ovale wallikeri* type 1 HM594180 reference sequence), which represents a short sequence repeated throughout the *tra* gene.

Two samples, Po9 and Po18, failed to amplify with the *tra* primers despite multiple PCR attempts.

Reticulocyte binding protein 2 (rbp2) gene

DNA sequences of the *rbp2* gene were obtained for all 22 *P. ovale* samples (Table 11). Table 14 contains the 74 pair *rbp2* amplicon for both *P. ovale curtisi* and *P. ovale wallikeri*. These sequences were not eligible for submission as the minimum length requirement for GenBank is 200 nucleotides. *P. ovale* subspecies results based on *rbp2* gene sequences agreed with subspecies results based on the *tra* gene sequences. Thirteen (59%) of the *P. ovale* samples were positive for *P. ovale curtisi* and 9 (41%) were positive for *P. ovale wallikeri*. None of our samples failed to amplify with the *rbp2* primers.

Small subunit rRNA (ssrRNA) gene

Nineteen of the 22 *P. ovale* positive samples were detected by the ssrRNA gene assay (Table 11). *P. ovale curtisi* and *P. ovale wallikeri* ssrRNA sequences were approximately 99% identical to previously published sequences at this locus. Representative *P. ovale curtisi* and *P. ovale wallikeri* ssrRNA sequences were deposited in GenBank as KM494987 and KM494988, respectively. *P. ovale* subspecies results based on ssrRNA gene sequences agreed with subspecies results based on *tra* and *rbp2* gene sequences. Three samples, Po9, Po11, and Po18, failed to amplify using the ssrRNA primers despite a second attempt using an additional microliter of template DNA.

Real-time PCR to detect *P. ovale*

Plasmid standard curve analysis of rbp2 qPCR assay

Since all 22 *P. ovale* microscopy positive samples were successfully amplified and sequenced using the *rbp2* primers, we developed an *rbp2* based qPCR assay to detect all *P. ovale* subspecies simultaneously in a single assay. Efficiency of the *rbp2* qPCR assay was analyzed using the non-linearized *rbp2* plasmid 10-fold serial dilution standard curve. Efficiency ranged from 90%-99% for six consecutive qPCR experiments with a coefficient of correlation (R^2) greater than 0.99. A representative qPCR amplification plot and standard curve are shown in Figures 8 and 9, respectively. All 22 *P. ovale* samples identified as *P. ovale* positive by expert microscopy were detected using our *rbp2* qPCR assay. There was no difference in PCR efficiency or R^2 value based on the conformation (linearized vs. non-linearized) of the *rbp2* plasmid standard curve (Pearson product-moment correlation=0.998, P<0.001).

Limit of quantification and limit of detection

The linear dynamic range of the *rbp2* qPCR assay was determined to be between 6.25 copies per microliter and 100,000 copies per microliter based on serial dilutions of the *rbp2* plasmid. Two-fold serial dilutions of known concentrations of the *rbp2* plasmid were performed in at least duplicate to determine the limit of detection. Dilutions containing 1.5 copies per microliter of the *rbp2* plasmid were detected by at least one replicate well in two separate qPCR experiments.

Specificity

In order to test the specificity of our *rbp2* assay for *P. ovale,* we performed qPCR using DNA isolated from cultured *P. falciparum* 3D7 (10,000 parasites per microliter)

65

and *P. vivax* DNA (517 parasites per microliter). Based on a series of ten separate qPCR

experiments, DNA from *P. falciparum* and *P. vivax* were uniformly negative. To ensure

no background from other *P. falciparum* strains, we tested genomic DNAs from strains

Dd2, NF54, and FCR3CSA, which were also not detected by our assay. We tested DNA

from *P. knowlesi, P. fragile,* and *P. cynomolgi* and found DNA from these malaria

parasite species were undetectable by our *rbp2* qPCR assay. As we were unable to obtain

pure *P. malariae* samples, we examined DNA samples isolated from the blood of

individuals co-infected with both *P. malariae* and *P. falciparum*. These *P. falciparum* and

P. malariae co-infected samples were also negative, indicating that our *rbp2* qPCR assay

does not detect *P. malariae* DNA. Two different control DNA samples from malaria

uninfected human blood were also uniformly negative. All specificity experiments were

carried out to 60 cycles in an attempt to capture non-specific amplification, which was

never seen, although the standard curve and the *P. ovale*-containing field samples

amplified appropriately.

Spiking experiments, in which *P. falciparum* DNA or *P. vivax* DNA was added to

the *rbp2* plasmid standard curve samples and subsequently utilized as template for the

rbp2 qPCR did not significantly alter the Cq values compared to when the standard curve

plasmid samples were run alone (ANOVA, P = 0.9993, Figure 10).

Interestingly, our *rbp2* qPCR assay detected *P. simiovale* genomic DNA isolated

from filter paper. DNA sequencing utilizing the *rbp2* primers revealed that the 74 base

pair *rbp2* region in *P. simiovale* is identical to that in *P. ovale curtisi*. Subsequent

attempts using additional primers to sequence the full-length *rbp2* gene of *P. simiovale*

were not successful. As these additional primers successfully amplified *P. ovale* positive

samples, the inability to amplify the full-length *P. simiovale rbp2* gene is likely due to sequence polymorphisms between *P. ovale* and *P. simiovale* in the primer binding regions.

Repeatability

Within-run repeatability of the *rbp2* plasmid standard curve Cq values was high, with the percent coefficient of variation (%CV) of dilution series replicates between 0.00-2.23% (Table 13). Results were also repeatable between runs, with the percent coefficient of variation (%CV) between 1.17-3.43% (Table 13). Repeatability was determined using results from six separate consecutive qPCR experiments.

Reproducibility

Analysis of the efficiency of the *rbp2* assay was performed independently at the USAMRU-K laboratory. A known concentration of non-linearized *rbp2* plasmid was diluted in PCR grade water to generate a 10-fold dilution standard curve for PCR efficiency analysis. The assay was performed with the same *P. ovale*-specific primers and probe utilized in validation experiments in a final volume of 50 microliters of Life Technologies TaqMan Fast Master Mix for the USAMRU-K ABI 7500. Despite slight variations in qPCR set up and cycling conditions, the Kenya laboratory obtained a PCR efficiency of 93.6% with an R^2 >0.99 for the standard curve analysis. These results are identical to the PCR efficiencies and R^2 values obtained at USU. The USAMRU-K laboratory also performed specificity experiments and demonstrated no amplification from *P. falciparum* DNA, DNA from uninfected human blood, or from negative template controls.

Quantification comparison: microscopy versus rbp2 qPCR

Quantitative parasitemia determined by expert microscopy (parasites per microliter) was compared to the *rbp2* copy number per microliter based on the *rbp2* plasmid standard curve (Figure 11). A modest correlation was determined (R^2=0.6595). This lack of a strong correlation is not surprising, as all *P. ovale* parasitemias were low, ranging from 16-3800 parasites/µl, and such low-level parasitemias are notoriously difficult to quantify accurately by microscopy (84; 209; 210; 214). Additionally, the samples utilized for comparison were mixed malaria species infections, mainly with *P. falciparum*. Mixed species infections create further difficulties for the accurate quantification of *P. ovale*-specific parasitemia based on light microscopy, but single-species *P. ovale* infected samples were not available.

DISCUSSION

Based on multilocus genotyping using the *rbp2*, ssrRNA, and *tra* genes, we detected both *P. ovale curtisi* and *P. ovale wallikeri* in approximately equal frequencies in a small sample set from Western Kenya, a region in which *P. ovale* subspecies characterization had not been previously performed. The presence of both *P. ovale* subspecies in Western Kenya is in agreement with other studies in sub-Saharan Africa and *P. ovale* endemic regions that describe the sympatric distribution of *P. ovale curtisi* and *P. ovale wallikeri* (38; 101; 213). We also identified additional allelic diversity within the *tra* gene in *P. ovale* samples from Kenya (Table 12) compared to what was previously identified in *P. ovale* samples from other malaria endemic regions (213). This allelic diversity at the *P. ovale tra* gene is consistent with reports of other *tra* variants

identified by DNA sequencing, however our *tra* sequences are unique from previously published *tra* gene sequences (286).

Our new inclusive *P. ovale*-specific qPCR assay is based on *rbp2*, a gene that contains conserved regions between *P. ovale curtisi* and *P. ovale wallikeri* but that is absent from other human malaria parasite species. The *rbp2* qPCR assay described herein allows simultaneous detection of both *P. ovale* subspecies using a single set of primers and probe. All 22 samples were detected and sequenced using our *rbp2* primers, highlighting the utility of these primers for *P. ovale* identification. *P. ovale* subspecies differentiation by DNA sequencing of the 74 base pair *rbp2* sequence region was in absolute agreement with *tra* and ssrRNA DNA sequencing results. This again emphasizes the utility of the PoRBP2fwd1 and PoRBP2rev1 primers for *P. ovale* subspecies discrimination based on a single SNP at position 431 (Figure 7) located between these primers. In agreement with other previous studies, these data demonstrate perfect dimorphism between *P. ovale curtisi* and *P. ovale wallikeri*, providing further support for the separation of the two *P. ovale* subspecies (101; 102; 140; 213; 276; 279; 318). As we begin to understand potential clinical, pathological, and biological differences between the two *P. ovale* subspecies, molecular methods to distinguish *P. ovale curtisi* and *P. ovale wallikeri* will aid in these research efforts. Additionally, as genomic data and full genome sequences become available for *P. ovale curtisi* and *P. ovale wallikeri*, phylogenetic analyses to determine the evolutionary relatedness between these and other malaria species will likely further our understanding of these newly characterized but poorly understood human parasites.

Using the *rbp2* plasmid as a standard curve, the linear dynamic range of our assay was determined to be between 6.25 copies of *rbp2* per microliter to 100,000 copies of *rbp2* per microliter. The lower, non-linear but still clearly positive limit of detection of our assay was determined to be 1.5 copies of *rbp2* per microliter, confirming this assay's capacity to detect low-level parasitemias. *P. ovale* parasitemias are characteristically lower than other malaria species, so we limited the testing of our upper dynamic range to 100,000 *rbp2* copies per microliter, as higher copy numbers would likely be epidemiologically and clinically irrelevant. We used the *rbp2* plasmid to determine the linear dynamic range and limit of detection because of difficulties obtaining pure *P. ovale* infected samples from malaria endemic regions and the inability to culture *P. ovale* parasites. The paucity of published genomic information for *P. ovale* also hinders the determination of copy number of *P. ovale*-specific genes, such as the *rbp2, tra,* and ssrRNA genes, utilized in this study. Thus, we are further limited in our attempts to appropriately correlate *rbp2* copy number and *P. ovale* parasitemias. Despite these limitations, we demonstrate the utility of our *P. ovale*-specific assay to detect low-levels of the *rbp2* plasmid and to detect low *P. ovale* parasitemias (as low as 16 parasites per microliter) from human blood samples collected in Western Kenya. Our study was also limited by only testing samples collected in Western Kenya and additional validation is therefore needed to confirm the ability of the *rbp2* qPCR assay to detect total *P. ovale* from other malaria endemic regions. As the 22 samples included in this study were identified as *P. ovale* by microscopy, further studies are needed to test the *P. ovale rbp2* qPCR assay with submicroscopic and asymptomatic *P. ovale* infections with a range of parasitemias.

Repeatability and reproducibility of qPCR assays are important components of assay validation as they indicate the assay's capacity to provide consistent and reliable results in different environments. Different users under modified laboratory conditions performed this assay successfully, with high PCR efficiency and equivalent quantification.

Specificity experiments showed no cross reactivity of our assay with *P. falciparum*, *P. vivax*, *P. malariae*, *P. cynomolgi*, *P. knowlesi*, *P. fragile*, and DNA from uninfected human blood even when qPCR was performed for 60 cycles. The complete lack of background amplification from human and other malaria parasite DNA, verifies the exquisite specificity of the assay. Further, assay performance was unchanged in the presence of DNA from other malaria parasite species. This is of particular importance for *P. ovale*, as this malaria species is often found as a co-infection with other malaria species. Interestingly, our *rbp2* qPCR assay also detected DNA obtained from *P. simiovale*. As *P. simiovale rbp2* sequence information is not available, we attempted to amplify the full-length *P. simiovale rbp2* gene using additional primers based on the *P. ovale rbp2* gene. However, we were unable to amplify the full *P. simiovale rbp2* gene, suggesting the *P. ovale* and *P. simiovale rbp2* genes may be similar but not identical. These results warrant further investigation of the *P. simiovale rbp2* and additional specificity experiments of other *P.ovale* assays that may also unknowingly detect *P. simiovale*.

Of the 22 samples identified as *P. ovale* positive by expert microscopy, three samples (Po9, Po11, Po18) failed to amplify at two of the three loci tested despite multiple attempts (Table 11). However, the *rbp2* gene was successfully amplified for all

22 samples as was a human-specific RNaseP endogenous control. These data, along with

the parasitemia data from multiple expert microscopists, indicate that the 22 samples

were *P. ovale* positive and that DNA template quality was unlikely to be the cause of the

failed amplifications at the *tra* and ssrRNA loci. The inability to successfully amplify at

all three loci could be explained by several reasons including: sequence polymorphisms,

template degradation, low *P. ovale* density, and inter-laboratory variability due to

reagents, equipment, and personnel. Additional investigation into potential reasons for

the failure to amplify at all loci was prevented due to limited sample volume.

The limited correlation between microscopy and *rbp2* qPCR results (Figure 11) is

not surprising as parasitemia calculations for *P. ovale* human samples at low parasitemias

are notoriously difficult, particularly in co-infected samples (84). Our *P. ovale* positive

samples from Western Kenya are all co-infected with either *P. falciparum* or *P. malariae*,

thus likely complicating the microscopy quantitation further. Variation between

parasitemia and *rbp2* copy number could also be explained by the *P. ovale* parasite stage.

For example, a *P. ovale* ring stage counts as a single parasite by microscopy and DNA

extracted from a *P. ovale* ring stage parasite represents one genome. However, a *P. ovale*

schizont is counted as a single parasite by microscopy but DNA extracted from a *P. ovale*

schizont may contain up to 14 genomes. This is a limitation of our study, as any

relationship between *P. ovale* parasitemia and *rbp2* copy number based on qPCR would

depend on the parasite stages observed under the microscope and present in the blood

sample obtained for DNA extraction.

Utilizing a plasmid standard curve for qPCR assays provides an efficient method

for standardizing assays that does not require culturing organisms or using human

samples. However, recent studies have highlighted important concerns regarding the

plasmid template conformation that could lead to quantification bias of plasmid template

by qPCR (127; 168). After linearizing our template plasmid to compare with a non-

linearized plasmid standard curve, we found no difference in Cq value, R^2, slope, or PCR

efficiency with the *rbp2* qPCR assay. This is in agreement with another recent study,

which also found no difference in plasmid standard curve based on the plasmid

confirmation (linearized versus non-linearized) (215). These results indicate that the

effect of plasmid conformation on standard curve quantification may be assay specific. In

addition to plasmid conformation, several additional quality control factors were

optimized, including plasmid isolation methods, purification, storage, and developing

appropriate laboratory protocols to minimize freeze-thawing, handling, and

contamination.

Conventional PCR assays targeting the multi-copy small subunit ssrRNA genes

are sensitive methods to detect and differentiate malaria species (267). Initial *P. ovale*-

specific ssrRNA PCR protocols showed limited capability to detect both *P. ovale*

subspecies and have since been adapted to target conserved regions between the two

subspecies. (58; 103; 232). Although ssrRNA conventional PCR protocols have shown

high sensitivity and specificity for malaria detection, we aimed to develop a novel *P.*

ovale-specific assay based on a gene target that is found only in *P. ovale* and is absent

from other malaria species infecting humans. We believe this approach enhances the

specificity of our *P. ovale*-specific assay and eliminates the potential for nonspecific

amplification of non-*P. ovale* species. Additionally, allelic variation within the ssrRNA

genes of *P. ovale curtisi* and *P. ovale wallikeri* may limit the ability of ssrRNA based

assays to capture all *P. ovale* infections due to sequence polymorphisms (161). We found no allelic variation in the primer and probe-binding regions of the *rbp2* gene from 22 *P. ovale* positive samples, indicating the potential utility of *rbp2* for *P. ovale* subspecies detection.

While nested PCR is often utilized to enhance sensitivity for malaria PCR detection, we chose a single step qPCR protocol, as a nested PCR approach requires additional labor and cost to perform the second PCR. Nested PCR also increases the risk of laboratory contamination of PCR product and requires separate laboratory space to minimize the risk of contamination. Our *P. ovale*-specific qPCR assay maintains high sensitivity while also minimizing the additional cost, labor, designated laboratory space, and potential PCR product contamination that can be associated with nested PCR protocols.

Our *P. ovale*-specific qPCR assay provides several advantages for our future epidemiological studies of this neglected, and clinically relevant, malaria parasite species. First, fast qPCR conditions allow for a reaction to be completed in less than 1 hour. Second, the qPCR platform bypasses the need for gel electrophoresis, reducing the risk of amplicon contamination of the laboratory. Third, the use of a hydrolysis probe increases specificity compared to double stand DNA (dsDNA) based qPCR product detection. Our *P. ovale*-specific *rbp2* qPCR assay can be utilized to better characterize the presence, parasitemia, geographical distribution, and the contribution of *P. ovale* to mixed-species infections and to clinical disease in malaria endemic regions.

Table 9. Primer and probe sequences utilized for conventional PCR and qPCR experiments.

Target	Primer/Probe	Primer/Probe Sequence	Reference
Tryptophan-rich antigen (*tra*)	PoTRAfwd3	5'-GCACAAAAATGGTGCTAACC-3'	Oguike et.al 2011 (213)
	PoTRArev3	5'-ATCCATTTACCTTCCATTGC-3'	
Small subunit rRNA (ssrRNA)	rOVA1WC	5'-TGTAGTATTCAAACGCAGT-3'	Fuehrer et.al 2012 (103)
	rOVA2WC	5'-TATGTACTTGTTAAGCCTTT-3'	
Reticulocyte binding protein 2 (*rbp2*)	PoRBP2fwd1	5'-CCACAGATAAGAAGTCTCAAGTACGATATT-3'	
	PoRBP2rev1	5'-TTGGAGCACTTTTGTTTGCAA-3'	
	PoRBP2p	5'-6FAM-TGAATTGCTAAGCGATATC-MGB-3'	

Table 10. GenBank accession numbers used for DNA alignment of the *P. ovale curtisi* and *P. ovale wallikeri tra*, *rbp2*, and ssrRNA DNA sequences.

Target	Reference Sequence	GenBank Accession Number
Tryptophan-rich antigen (*tra*)	*P. ovale curtisi* type 1	HM594182
	P. ovale curtisi type 2	HM594183
	P. ovale wallikeri type 1	HM594180
	P. ovale wallikeri type 2	HM594181
Small subunit rRNA (ssrRNA)	*P. ovale curtisi*	JF894405
	P. ovale wallikeri	JF894406
Reticulocyte binding protein 2 (*rbp2*)	*P. ovale curtisi*	GU813971
	P. ovale wallikeri	GU813972

Table 11. *P. ovale* subspecies identification by DNA sequencing of the of the tryptophan-rich antigen *(tra)* gene, the reticulocyte binding protein 2 *(rbp2)* gene, and the small subunit ribosomal RNA (ssrRNA) gene.

Sample ID	Co-infecting malaria species (parasites/µl)	*P. ovale* microscopy (parasites/µl)	Tryptophan-rich antigen *(tra)*	Reticulocyte binding protein 2 *(rbp2)*	Small subunit rRNA *(ssrRNA)*
Po1	*P. falciparum* (7334) *P. malariae* (110.8)	57.6	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po2	*P. falciparum* (653.4) *P. malariae* (114)	156	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po3	*P. falciparum* (67121.1)	458	*P. ovale curtisi* type 2	*P. ovale curtisi*	*P. ovale curtisi*
Po4	*P. falciparum* (571.5) *P. malariae* (56)	42	*P. ovale wallikeri* type 2	*P. ovale wallikeri*	*P. ovale wallikeri*
Po5	*P. falciparum* (101.8)	121.78	*P. ovale wallikeri* type 2	*P. ovale wallikeri*	*P. ovale wallikeri*
Po6	*P. falciparum* (306.1) *P. malariae* (1320)	2321.78	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po7	*P. falciparum* (3284.2) *P. malariae* (1648.6)	69.33	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po8	*P. falciparum* (4568) *P. malariae* (320)	296.35	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po9	*P. falciparum* (515) *P. malariae* (255.3)	16	No data	*P. ovale curtisi*	No data
Po10	*P. falciparum* (1897.3)	456.89	*P. ovale curtisi* type 2	*P. ovale curtisi*	*P. ovale curtisi*
Po11	*P. falciparum* (412.7)	16	*P. ovale curtisi* type 1	*P. ovale curtisi*	No data

	P. malariae (583.3)				
Po12	*P. falciparum* (158.9) *P. malariae* (48)	331.26	*P. ovale wallikeri* type 1	*P. ovale wallikeri*	*P. ovale wallikeri*
Po13	*P. falciparum* (613.8) *P. malariae* (453.1)	365.54	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po14	*P. falciparum* (16703.3) *P. malariae* (32)	157.33	*P. ovale wallikeri* type 1	*P. ovale wallikeri*	*P. ovale wallikeri*
Po15	*P. falciparum* (28976)	3738.88	*P. ovale wallikeri* type 1	*P. ovale wallikeri*	*P. ovale wallikeri*
Po16	*P. falciparum* (3889.9) *P. malariae* (211.8)	32	*P. ovale wallikeri* type 2	*P. ovale wallikeri*	*P. ovale wallikeri*
Po17	*P. falciparum* (16) *P. malariae* (24)	1118.08	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po18	*P. falciparum* (382.1) *P. malariae* (52.4)	26.67	No data	*P. ovale curtisi*	No data
Po19	*P. falciparum* (8954.2) *P. malariae* (409.6)	58.67	*P. ovale curtisi* type 1	*P. ovale curtisi*	*P. ovale curtisi*
Po20	*P. falciparum* (197.3) *P. malariae* (304)	350.61	*P. ovale wallikeri* type 1	*P. ovale wallikeri*	*P. ovale wallikeri*
Po21	*P. falciparum* (4299.5) *P. malariae* (172)	84.36	*P. ovale wallikeri* type 1	*P. ovale wallikeri*	*P. ovale wallikeri*
KSI	*P. falciparum* (no data)	No data	*P. ovale wallikeri* type 1	*P. ovale wallikeri*	*P. ovale wallikeri*

Table 12. Five *P. ovale* positive samples contained unique *tra* gene polymorphisms identified by DNA sequencing.

Nucleotide Position (Genbank accession number)	71	99	171-172	523	595
P. ovale wallikeri type 1 (HM594180)[a]	T	G	-[b]	T	G
Po05 (KM494982)	C	G	-[b]	C	AA
Po12 (KM494983)	T	C	ATAAATGCTATAACCCCC	T	G
Po20 (KM494984)	T	C	ATAAATGCTATAACCCCC	T	G
Nucleotide Position	280	307	664		
P. ovale curtisi type 1 (HM594182)[c]	A	A	G		
Po06 (KM494985)	G	A	A		
Po07 (KM494986)	A	G	G		

[a] *P. ovale wallikeri* type 1 (HM594180, nucleotides 1-1171) was utilized as a reference for DNA sequence alignment of *P. ovale wallikeri* positive samples with unique polymorphisms (Po05, Po12, and Po20).
[b] Dashes (-) indicate lack of an insertion. Samples Po12 and Po20 contained an 18 base pair insertion between nucleotide position 171 and 172 based on the reference sequence.
[c] *P. ovale curtisi* type 1 (HM594182, nucleotides 1-1117) was utilized as a reference for DNA sequence alignment of *P. ovale curtisi* positive samples with unique polymorphisms (Po06 and Po07).

Table 13. Repeatability and reproducibility of the *rbp2* plasmid standard curves determined via Cq values from six separate qPCR experiments.

Rbp2 plasmid Copies/µl	100,000	10,000	1,000	100	10
Within-run Repeatability (%CV)	0.25-0.74	0.41-1.00	0.048-1.47	0.00-1.52	0.19-2.23
Between-run repeatability (%CV)	2.21	1.53	1.17	1.46	3.43

Table 14. *P. ovale curtisi* and *P. ovale wallikeri* reticulocyte binding protein 2 (*rbp2*) DNA sequences

P. ovale subspecies	*Rbp2* DNA sequences (5'-3')
P. ovale curtisi	CCACAGATAAGAAGTCTCAAGTACGATATTAATGAATTG CTAAGCGATATCAATTGCAAACAAAAGTGCTCCAA
P. ovale wallikeri	CCACAGATAAGAAGTCTCAAGTACGATATTAATGAATTG CTAAGCGATATCATTTGCAAACAAAAGTGCTCCAA
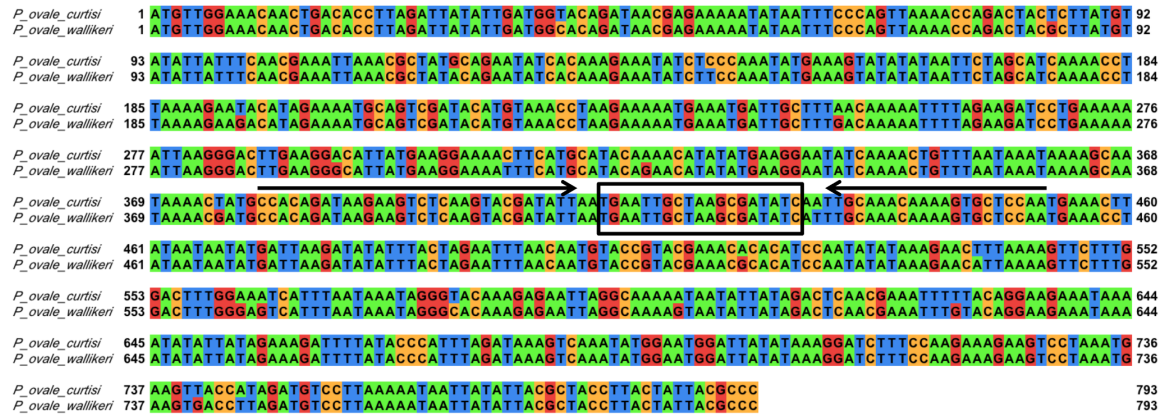
Figure 7. *P. ovale* reticulocyte binding protein 2 (*rbp2)* sequence alignment.
The *P. ovale curtisi* (GU813971) and *P. ovale wallikeri* (GU813972) *rbp2* sequences were aligned using EMBL-EBI Clustal Omega program and visualized in Jalview with the default Jalview nucleotide color scheme (green for adenine, orange for cytosine, red for guanine, and blue for thymine). Primers and probe were designed based on conserved regions between the two *P. ovale* subspecies. The forward (PoRBP2fwd1) and reverse (PoRBP2rev1) primers are indicated by arrows and the hydrolysis probe (PoRBP2p) binding site (boxed) is located in between the forward and reverse primer.

81
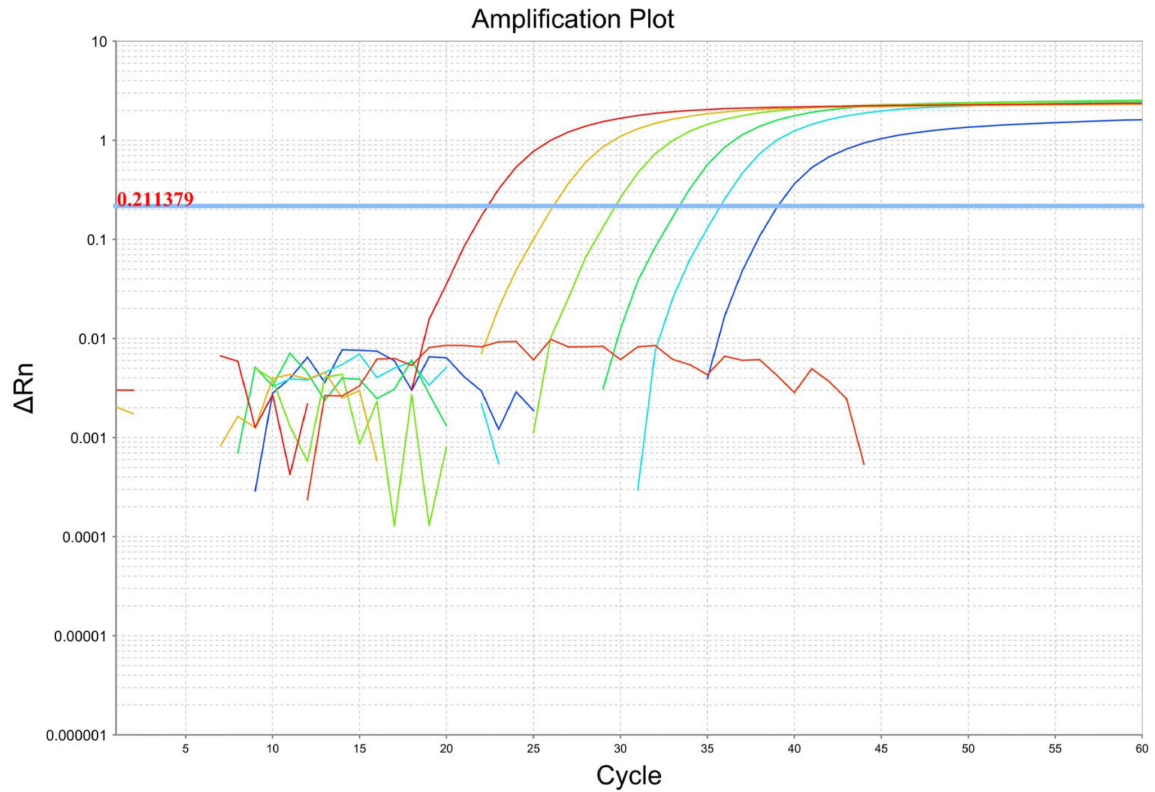
Figure 8. *P. ovale rbp2* qPCR dynamic range.
    A ten-fold serial dilution of *rbp2* plasmid (1 to 100,000 copies/μl) is shown in the amplification plot. The cycle threshold was determined automatically by the ABI 7500 fast system software program. The negative control sample (red line) shows no amplification over the cycle threshold for 60 cycles.
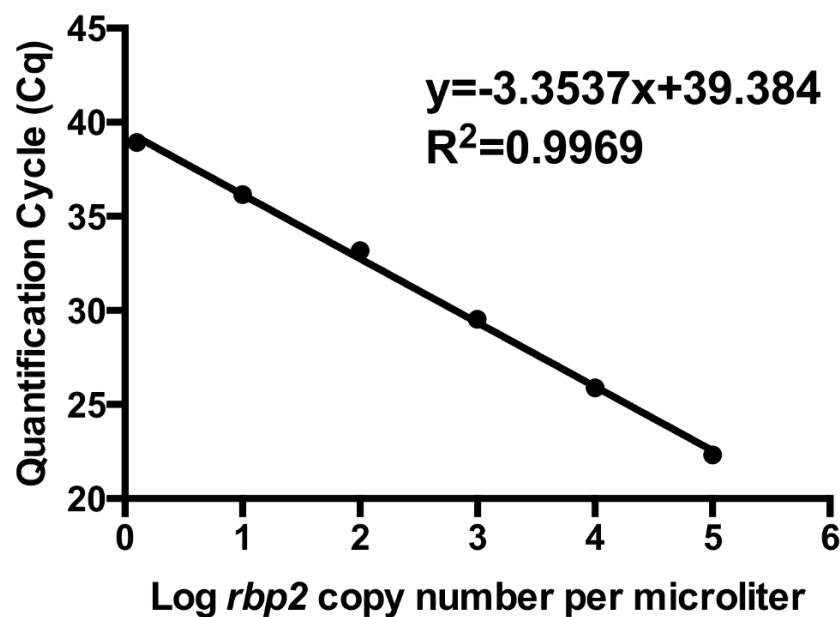
Figure 9. *P. ovale rbp2* plasmid standard curve.
  A representative standard curve demonstrates linearity based on 10-fold serial dilutions (1 to 100,000 copies/µl) of *rbp2* plasmid.
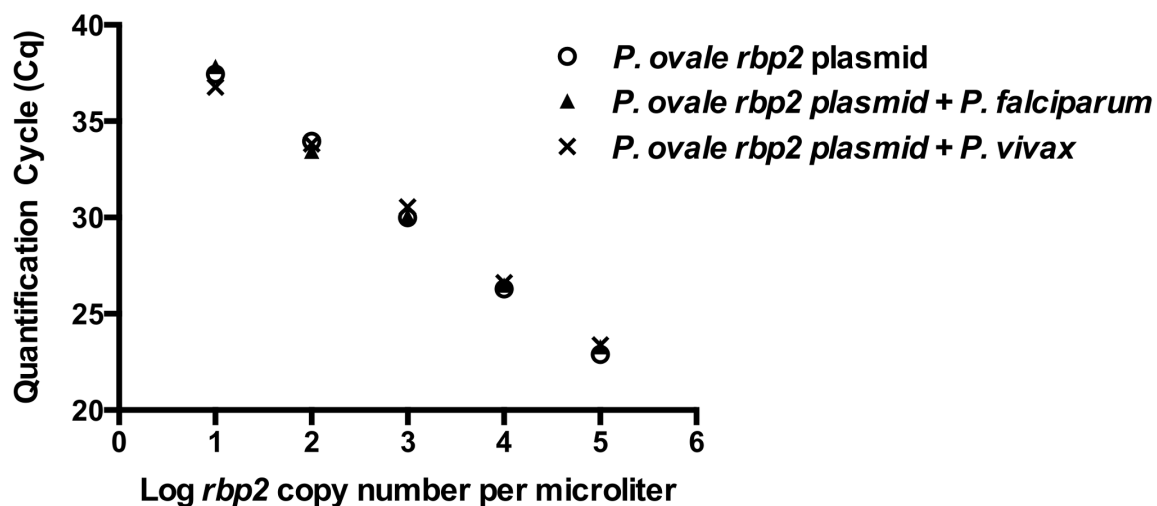
Figure 10. *P. ovale rbp2* qPCR specificity.
  Serial dilutions of *rbp2* plasmid were spiked with *P. falciparum* DNA (10,000 parasites/µl) or *P. vivax* DNA (517 parasites/µl). Cq values were unchanged in the presence of DNA from additional malaria parasite species (P=0.9993).

83
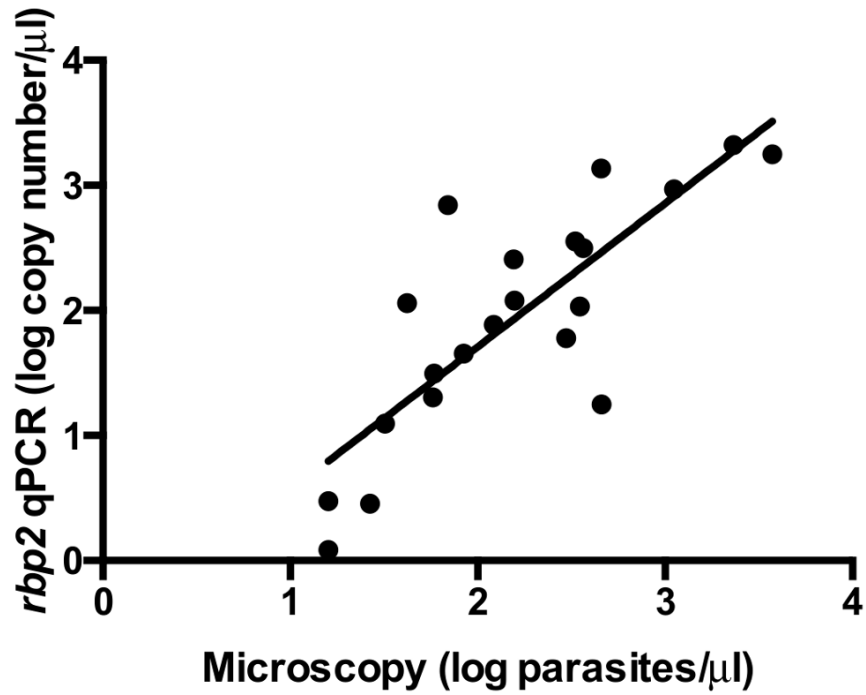
Figure 11. Comparison of microscopy and *P. ovale rbp2* qPCR results.
P. ovale parasitemias based on microscopy (log parasites/µl) were compared to *rbp2* qPCR results (log *rbp2* copy number/µl). A limited correlation was found between parasitemia and *rbp2* plasmid copy number (R^2=0.6595).

CHAPTER 4: A Deep Sequencing Approach to Estimate Malaria Complexity of Infection in the Democratic Republic of Congo

Robin H. Miller[1], Nicholas J. Hathaway[2], Oksana Kharabora[3], Kashamuka Mwandagalirwa[4], Antoinette Tshefu[5], Steven R. Meshnick[6], Steve M. Taylor[7], Jeffrey A. Bailey[2], Jonathan J. Juliano[3], V. Ann Stewart[1]

[1]Uniformed Services University, Bethesda, MD, USA

[2]University of Massachusetts School of Medicine, Worcester, MA, USA

[3]University of North Carolina School of Medicine, Chapel Hill, NC, USA

[4]Kinshasa General Hospital (HGK), Kinshasa-Gombe, Democratic Republic of Congo

[5]Ecole de Sante Publique, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[6]University of North Carolina Gillings School of Global Public Health, Chapel Hill, NC, USA

[7]Duke Global Health Institute, Duke University Medical Center, Durham, NC, USA

Chapter in preparation for publication

ABSTRACT

In regions of stable malaria transmission, human hosts can be infected with multiple strains of *Plasmodium falciparum*. The number of unique strains present within an individual is referred to as the Complexity of Infection (COI). Multiclonal (COI>1) *P. falciparum* infections can influence malaria clinical outcomes, response to drug treatment, and within-host parasite dynamics. In this study, we utilized an amplicon-based deep sequencing approach to detect minor frequency *P. falciparum* haplotypes and estimate the COI within individual samples collected as part of the 2007 Demographic and Health Survey in the Democratic Republic of Congo (DRC). We targeted malaria positive dried blood spot samples for PCR amplification and deep sequencing of the polymorphic *P. falciparum* apical membrane antigen 1 (*pfama1*) gene. Deep sequencing results were analyzed using the SeekDeep targeted amplicon analysis pipeline. We identified at total of 88 unique *pfama1* haplotypes and found 64.5% of individuals had multiclonal (COI>1) *P. falciparum* infections. We found no difference in the *P. falciparum* COI based on HIV status, age, or sex and no geospatial clustering of *pfama1* haplotypes within DRC provinces was identified. Overall, high sequence diversity within the *pfama1* gene was observed in *P. falciparum* parasites from the DRC.

INTRODUCTION

Complexity of Malaria Infections

Several molecular epidemiological studies have demonstrated that multiple *Plasmodium falciparum* strains circulate in malaria holoendemic regions (31; 52; 104; 138; 145; 178; 265). The number of genetically distinct *P. falciparum* strains within a single infected individual is referred to as the Complexity of Infection (COI) (44). The intensity of malaria transmission corresponds to the *P. falciparum* COI, as individuals living in malaria holoendemic regions typically have higher *P. falciparum* COIs compared to areas with seasonal or low malaria endemicity (310). Multiclonal (COI>1) infections are the result of a single mosquito inoculation with several *P. falciparum* strains or several mosquito inoculations with different *P. falciparum* strains (31; 86). A range of COIs have been described in the literature, with the experimental methodology, geographic location, and transmission intensity all factoring into the ability to harbor or maintain detectable multiclonal *P. falciparum* infections (31; 41; 44; 134; 138; 178; 254; 266). The majority of research focused on malaria COI is limited to *P. falciparum*, however, recent studies also have begun to elucidate the genetic complexity of multiclonal *P. vivax* infections (30; 75; 107; 169; 170).

Detection of multiclonal *P. falciparum* infections is important, as these complex infections have been shown to impact clinical outcomes, indicate malaria transmission intensity, and influence malaria parasite evolution (26; 52; 72; 174; 206; 310). Several studies have demonstrated that multiclonal *P. falciparum* infections are associated with increased risk of clinical malaria (52; 184; 212; 246), however, other studies have shown multiclonal infections protect against clinical malaria (26; 45; 96; 146; 199; 265). Differences in the association between *P. falciparum* COI and clinical malaria may be

influenced by malaria transmission intensity (97). Thus, the impact of *P. falciparum* COI on clinical malaria outcome remains unclear and further exploration is needed to understand the within-host dynamics of malaria parasites as it relates to disease in humans (104).

Malaria population genetics studies have recently described the utility of *P. falciparum* COI as a marker for malaria transmission intensity (310). As malaria transmission falls, due to seasonality or malaria intervention strategies, the *P. falciparum* COI in a population has been shown to decrease as well (72; 104; 306). Therefore, detection of multiclonal *P. falciparum* infections and surveillance of *P. falciparum* COI over time are potential useful tools for measuring malaria transmission before and after implementation of malaria control programs (206).

In multiclonal infections, *P. falciparum* undergoes genetic recombination during the obligate sexual cycle in the mosquito vector to produce genetically distinct offspring, facilitating the evolution of the malaria parasite (132; 197). Complex multiclonal malaria infections are therefore a potential source for the generation of novel parasite genotypes capable of evading host immune responses or antimalarial therapy (120; 174)

Factors that Influence *P. falciparum* COI

In regions with high malaria transmission, *P. falciparum* COI is typically low in young children (under 1 year) (44; 184; 265), increases as children mature, and then decreases when individuals reach adulthood (44; 184; 265). The relationship between COI and age in malaria endemic regions is likely due to the development of naturally acquired immunity to malaria and the ability of these semi-immune individuals to control malaria parasitemia (310). Malaria vaccines have also been shown to reduce malaria COI

in vaccinated individuals compared to unvaccinated controls, which further suggests that the development of malaria immunity aids in the control of multiclonal infections (40; 91).

Pregnant women in malaria endemic regions have higher *P. falciparum* COIs when compared to non-pregnant women, younger pregnant women have higher COIs compared to older pregnant women (254), and primigravid pregnant women have higher COIs compared to multigravid pregnant women (41; 159). Further, *P. falciparum* COI in pregnancy has been shown to impact the health of both the mother and child. For example, in primigravid women, multiclonal (COI>1) *P. falciparum* infections in the placenta were associated with low birth weight babies compared to monoclonal (COI=1) placental infections (114). A study in Mozambique found that pregnant women with higher COIs had increased prevalence of anemia compared to pregnant women with lower COIs (41).

Malaria and HIV Co-infections

Regions endemic for both malaria and HIV pose a significant challenge for the control of both diseases. Mathematical models estimate 3 million cases of malaria and 65,000 malaria deaths can be attributed to HIV in Africa annually (151). Furthermore, epidemiological modeling suggests that HIV and malaria co-infections fuel the transmission of both pathogens in endemic regions (23; 308). The impact of HIV on clinical malaria depends on age, level of immunosuppression, pregnancy, and malaria transmission intensity (reviewed in (99)). In regions holoendemic for malaria transmission, co-infection with HIV in adults is associated with increased prevalence of clinical malaria, increased malaria parasite density, and increased prevalence of malaria

parasitemia (100; 131; 137; 223; 287; 309; 316). The risks of clinical malaria and increased malaria parasitemias in HIV infected individuals are inversely correlated with CD4 T cell counts (316).

Evidence of the impact of HIV on *P. falciparum* COI is limited. A study in pregnant women in Malawi demonstrated that HIV positive women had significantly higher *P. falciparum* COIs compared to HIV negative women (152). Alternatively, a study in the Central African Republic (CAR) demonstrated that HIV seropositivity was significantly associated with lower *P. falciparum* COIs compared to HIV seronegative individuals during clinical malaria episodes (81). The impact of HIV on *P. falciparum* COI in malaria asymptomatic individuals, who make up the majority of malaria infections in endemic regions and are the reservoir for malaria transmission, remains unclear.

Methods for Determining COI

Numerous genetic tools and strategies have been employed to analyze malaria COI. The most widely utilized method for determining *P. falciparum* COI is PCR amplification of polymorphic genes, such as the merozoite surface protein genes (*msp1*, *msp2*) and *Glurp* (52; 98; 145; 149; 254; 266). These PCR based methods rely on DNA sequence length polymorphisms of the target gene, which are visualized via gel electrophoresis and the COI within an individual sample is calculated by counting the number of distinct bands. However, this approach limits the detection of distinct *P. falciparum* strains differing by only a few nucleotides in length or that contain single nucleotide polymorphisms (SNPs). Further, a multicenter comparison of related PCR genotyping methods for COI demonstrated high variability in the number of *P.*

falciparum strains detected between research laboratories due to differences in sample preparation and PCR conditions (95). Therefore, more sensitive and standardized methods for the detection of *P. falciparum* COI will improve our understanding of the impact of multiclonal infections on malarial disease and epidemiology.

Novel approaches based on next-generation sequencing (NGS) technologies have increased capabilities to detect low frequency, or minor variant, *P. falciparum* strains and therefore have enhanced capacity to characterize the COI within an individual or population (31; 72; 104; 105; 134; 242). These methods employ NGS technology to generate millions of sequence reads and utilize sophisticated bioinformatics tools to analyze large data sets. NGS has several advantages over conventional PCR and shotgun-sequencing for detecting low frequency *P. falciparum* infections (134). For example, NGS generates millions of sequencing reads that can be analyzed to detect multiple SNPs that occur at different frequencies within the parasite population (129; 310). NGS is also useful for estimating the frequency or proportion of different *P. falciparum* strains based on sequence read frequencies to identify major and minor strains present within a single infection or population (134). Finally, methods for barcoding samples and the decreasing costs of NGS reagents and equipment have significantly reduced the overall price of NGS technology for large malaria population genomic and epidemiological studies. In this study, we utilized a NGS deep sequencing approach to detect multiclonal *P. falciparum* infections and analyze *P. falciparum* COI.

Apical Membrane Antigen 1 (AMA1)

The *P. falciparum* apical membrane antigen 1 (*pfama1*) gene encodes the 83 kDa type I integral membrane AMA1 microneme protein (69; 125). AMA1 is expressed in the last

91

four hours of erythrocytic development in the dividing schizont, where it is secreted onto the merozoite surface before the parasite ruptures the host erythrocyte (202). The AMA1 prodomain is cleaved on the surface of the merozoite, resulting in a 66 kDa membrane bound AMA1 protein (202). AMA1 is comprised of a signal sequence, three ectodomains (N-terminal domain I, central domain II, and C-terminal domain III) linked by eight disulfide bonds followed by a transmembrane domain (Figure 12) (125; 167; 201). AMA1 is postulated to be involved in host cell invasion, as AMA1 specific antibodies block the ability of the malaria parasite to infect host erythrocytes (148; 297). However, studies to understand the function of *P. falciparum* AMA1 in erythrocyte invasion are extremely difficult, as *pfama1* is an essential gene and therefore cannot be targeted for knockout experiments (36; 122; 252; 302). Several studies using a related apicomplexan parasite, *Toxoplasma gondii,* indicate that AMA1 is important for host cell attachment and formation of the moving junction (MJ), a complex of proteins involved in host cell invasion (28; 122; 194; 195). AMA1 has been shown to interact with other erythrocyte invasion proteins including the rhoptry neck 2 (RON2) protein, which is also involved in MJ formation (271). A recent study demonstrated that conditional expression of *P. falciparum* AMA1 to 20% compared to wild type levels reduced the ability of *P. falciparum* to form a MJ during erythrocyte invasion and reseal the erythrocyte post invasion (321). Interestingly, AMA1 is also expressed on the surface of the sporozoite and may facilitate hepatocyte invasion (260).

Individuals routinely exposed to malaria develop AMA1 specific antibodies that have been shown to increase with age and predominately target domains I and II (61; 68; 236; 238; 298). AMA1 specific antibodies protect against malaria infection and clinical

disease in endemic areas (25; 236; 301). In addition to the magnitude of AMA1 specific antibodies, the breadth of AMA1 specific antibodies against several *P. falciparum* strains has also been identified to mitigate the risk of developing malaria clinical disease (33; 217). Natural exposure to malaria parasites also induces an AMA1 T-cell specific response, although this response has been shown to be short-lived compared to AMA1 specific antibody levels (304).

AMA1 remains a viable malaria vaccine antigen candidate, despite its high levels of heterogeneity. Several AMA1 based vaccine challenge studies have been conducted including different adjuvants, animal models, and challenge doses; resulting in generally positive clinical outcomes (Reviewed in (238)). A phase II clinical trial recently conducted in Mali utilizing an AMA1 based vaccine showed no significant reduction in overall clinical malaria, however, the authors reported a significant reduction in clinical malaria caused by the vaccine strain (270; 295). These results demonstrate that protection against clinical malaria based on AMA1 is highly strain-specific, and therefore circulating *P. falciparum* strain diversity must be considered in order to develop an effective malaria vaccine based on the AMA1 antigen. To increase the ability of AMA1 based vaccines to target heterologous *P. falciparum* strains, several efforts have been underway to develop an AMA1 based vaccine that incorporates AMA1 antigen from multiple strains to increase protection against the multitude of *P. falciparum* strains circulating in malaria endemic regions (33; 85; 88; 294).

The 1.8 kilobase (kb) single copy *pfama1* gene is located on chromosome 11 in the *P. falciparum* genome. The *pfama1* gene is highly polymorphic, containing several single nucleotide polymorphisms (SNPs), the majority of which are located in domain I

(69; 87; 90; 93; 106; 147; 179; 227; 235; 281). The diversity within the *pfama1* gene is maintained by balancing selection, which is likely due to immune selection within the host (30; 218; 235; 281). Studies in malaria endemic regions have described over 60 polymorphic sites within the *pfama1* gene and over 200 distinct *pfama1* haplotypes were sequenced from a malaria endemic region in Mali (89; 281). Based on the highly polymorphic nature of the *pfama1* gene and the relevance of its diversity to the development of an AMA1 based malaria vaccine, we chose to target *pfama1* for deep sequencing.

Malaria in the Democratic Republic of Congo

Malaria is a leading cause of morbidity and mortality in the Democratic Republic of Congo (DRC) with 100% of a population of over 79 million people considered to be at risk for malaria (12; 18; 19). The DRC Ministry of Health and the 2014 President's Malaria Initiative (PMI) Malaria Operation Plan for DRC reports that 97% of individuals living in the DRC reside in regions that experience malaria transmission from 8-12 months each year with over 95% of malaria infections estimated to be cause by *P. falciparum* (13). The primary malaria vectors in the DRC include *Anopheles gambiae sensu stricto* and *An. funestus* (50; 62; 63).

Demographic and Health Surveys (DHS) are utilized to collect data to identify health and social needs, implement policies, and evaluate and monitor programs. The DHS Program (www.dhsprogram.com) is funded by USAID along with several partnering institutions and managed by ICF International. In 2007, a collaboratively funded DHS was conducted in the DRC (EDS-RDC) by the DRC Ministry of Health and Ministry of Planning in order to collect cross-sectional data on several health and social

indicators. The survey was the first of its kind in the DRC and included questions concerning family planning, nutrition, mortality, domestic violence, fertility, maternal and child health, demographic information, attitudes towards HIV/AIDS, and use of insecticide-treated bednets (ITNs) to prevent malaria (2).

The EDS-RDC results indicated that ITN usage was low in 2007, with only nine percent of households owning at least one ITN and six percent of children under five slept under an ITN the night before the survey was performed (2). Approximately seven percent of pregnant women reported sleeping under an ITN the night before the survey and twelve percent of pregnant women received a single dose of intermittent preventive treatment (IPT) during pregnancy (2). Only five percent of pregnant women reported receiving the recommended two doses of IPT during pregnancy (2).

Blood samples collected for HIV testing during the 2007 EDS-RDC were also utilized to study malaria prevalence and epidemiology in the DRC (291). Using a panel of real-time PCR (qPCR) assays, malaria prevalence was reported as 33.5% in adults (age 15-59 years) with the vast majority (>90%) of infections were *P. falciparum* either as a monoinfection, or as a co-infection with *P. ovale* or *P. malariae* (291). Multivariate analysis based on data collected as part of the 2007 EDS-RDC identified variables that were significantly related to malaria prevalence and showed males to be 24% more likely to have malaria parasitemia. Additionally, individuals living further from urban areas had higher malaria prevalence and wealthier individuals had a decreased risk of malaria prevalence (191). Malaria positive samples collected as part of the 2007 EDS-RDC will be utilized in this study for deep sequencing of the *pfama1* gene and analysis of *P. falciparum* COI based on demographic factors and geographical locations in the DRC.

The aim of this study is to utilize a deep sequencing approach to detect multiclonal *P. falciparum* infections from blood samples collected as part of the 2007 EDS-RDC. To address this aim, we will 1) utilize a sensitive, amplicon-based deep sequencing approach targeting the polymorphic *pfama1* gene to detect low frequency and multiclonal *P. falciparum* infections and 2) analyze *P. falciparum* Complexity of Infection (COI) based on demographic factors including HIV status, geographic location, age, and sex in the DRC.

METHODS

EDS-RDC Methods and Sample Collection

The 2007 EDS-RDC was performed using a stratified sampling method in which each of the 11 provinces was separated into three strata: major cities, towns, and rural areas (187). Census data were utilized to identify neighborhoods in urban areas and villages in rural areas (187). 300 villages and neighborhoods, referred to as clusters, were identified randomly to represent urban (41% of the population) and rural areas according to census data (187). The 300 clusters contained 9,000 households, of which 99.3% were successfully interviewed for the 2007 EDS-RDC (2). Data collection for the EDS-RDC was conducted from January to March and also from May to August 2007 (2). The survey included 9,995 women (age 15-49) and 4,757 men (age 15-59) from each of the 11 DRC provinces, representing both urban and rural areas (2). To protect the privacy of the participants, geographical coordinates were randomly displaced by 5 kilometers in rural areas and 2 kilometers in urban areas (187).

Blood samples were obtained via filter paper for voluntary HIV testing from all the men and half of the women included in the survey (2; 187). The HIV testing protocol included an initial ELISA test (Vironostika) and subsequent retesting of all positive ELISA samples and 10% of negative ELISA samples with a second ELISA test (Enzygnost) (3, Purvis, 2015 #3152). Samples with discordant results from the first and second ELISAs were retested via Western blot (3). The 2007 EDS-RDC reports HIV prevalence as 1.3% in individuals aged 15-59 years (2). Both men and women living in urban areas showed a higher HIV prevalence compared to rural areas (2). In women, HIV prevalence was highest in the most educated and most wealthy compared the least educated and least wealthy (2).

DNA Extraction Methods

Remaining dried blood spots collected for HIV testing were utilized for detection of malaria parasitemia (291). Sample collection, storage, and extraction methods are published elsewhere (291). Briefly, genomic DNA was extracted from blood spots and utilized as template for a *Plasmodium* genus-specific 18S ribosomal RNA based qPCR assay and additional species-specific 18S ribosomal RNA based qPCR assays to detect *P. falciparum*, *P. ovale*, and *P. malariae* (289; 291).

Individual Sample Identification

For this study, we identified HIV positive samples (n=23) that were also positive for *P. falciparum* but negative for all other malaria parasite species based on the 18S species-specific qPCR. We then matched the 23 double positive (HIV/*P. falciparum*) samples to *P. falciparum* positive, HIV negative samples based on age and sex. To account for the low number of double positive samples, matching was done on a 4:1 ratio

of HIV negative (n=92) to HIV positive samples (n=23) for a total of 115 individual samples. A sample size of 115 samples has 81% power to detect a mean COI difference of 1.1 between HIV positive and HIV negative individuals, based on a two-sided 95% confidence interval.

Geographical Clusters Identification

To analyze geospatial differences of *P. falciparum* haplotypes, we selected 88 clusters from the original 300 clusters identified in the 2007 EDS-RDC that were the same or geographically close to the clusters from which the individual samples were identified. We combined DNA samples to generate a pooled "geographical cluster sample" that contained the DNA from all samples identified in that particular cluster (n=88). The number of samples that were pooled for each geographical cluster sample ranged from 1-25 samples. We utilized these geographical cluster samples in order to compare the *pfama1* haplotypes found in individual samples to *pfama1* haplotypes found in the same or surrounding geographical clusters. Additionally, pooling samples from a particular geographical cluster reduces the cost and labor of sequencing each sample individually.

***P. falciparum* lactate dehydrogenase qPCR**

To confirm the presence of *P. falciparum,* all individual samples (n=115) were screened for the *P. falciparum* lactate dehydrogenase (*pfldh*) gene by qPCR on the ABI Viia7 platform (Life Technologies). Primer/probe sequences and PCR cycling conditions are described elsewhere (233). Approximate *pfldh* concentrations were calculated for individual samples using a standard curve generated from known concentrations of *P. falciparum* DNA.

P. falciparum apical membrane antigen 1 (*pfama1*) PCR

Heminested primers were designed using Primer3 (305) to amplify a region of the *pfama1* gene (GenBank Reference XM_001347979.1). A heminested approach was used to increase sensitivity for detection of low-level concentrations of *P. falciparum* DNA. The Ama1OF and Ama1R primers amplify a 266 bp region and the Ama1F and Ama1R primers amplify a 236 base pair (bp) region (Table 15 and Figure 12). A 10 nucleotide Multiplex Identifier (MID) barcode sequence was added to the 5' end of the Ama1F primer to allow for pooling of individual sample PCR products during library preparation (9). Twenty-two unique MID sequences were utilized (Table 16).

Following a heminested PCR approach, an initial PCR (round 1) was performed and the PCR product utilized as template for a second PCR (round 2). For round 1 PCR, Ama1OF and Ama1R primers were used. For round 2 PCR, Ama1F and Ama1R primers were used. Details for the *pfama1* nested PCR and cycling conditions are shown in Table 15. PCR products were visualized on 1% agarose gels stained with ethidium bromide. Positive PCR products were purified using the Invitrogen PureLink Pro 96 PCR Purification Kit and purified PCR product concentration was calculated in duplicate using the Invitrogen Quant-iT PicoGreen dsDNA Assay Kit according to manufacturer's instructions.

Pfama1 Amplicon-based Deep Sequencing

Purified PCR products from individual samples and geographical cluster samples were organized into 18 indexes such that each index contained PCR products with unique MIDs. We added ten nanograms of each PCR product to the appropriate index. The DNA concentration of each of the 18 indexes was analyzed using the Agilent High Sensitivity

D1000 ScreenTape Assay on the 2200 TapeStation (Agilent Technologies) according to manufacturer's instruction to ensure DNA concentration between indexes was similar.

Ion Torrent library preparation for amplicon-based deep sequencing was performed following the "Preparing Short Amplicon (<350) Libraries Using the Ion Plus Fragment Library Kit" manual (Life Technologies, MAN0006846, revision 3.0). After library preparation, we determined the DNA concentration of each of the 18 libraries using the Agilent High Sensitivity D1000 ScreenTape Assay according to manufacturer's protocol. Equal concentrations of each library were subsequently pooled together and spilt across two Ion 318 Chips (Life Technologies) utilizing 400bp chemistry on the Ion Torrent PGM platform (Life Technologies) at the University of North Carolina Chapel Hill Microbiome Core Facility.

SeekDeep Bioinformatics Pipeline

We utilized the SeekDeep targeted amplicon bioinformatics pipeline for sequence data extraction, processing, and analysis. SeekDeep (http://baileylab.umassmed.edu/SeekDeep) was developed by Nicholas Hathaway and Dr. Jeffrey Bailey at the University of Massachusetts School of Medicine (169). The pipeline consists of three steps: extractor, qluster, and processClusters. An overview of the SeekDeep process is shown in Figure 13. The first step is extractor, in which the sequence data is "de-multiplexed" such that samples are identified and separated based on MID barcodes. The extractor step also applies several quality control steps, including removing short sequence reads and reads with poor quality scores. The second step, qluster, filters and compiles the raw sequence reads to separate unique haplotypes through an iterative process. The last step, processClusters, compares haplotypes found in

replicate samples, performs additional filtering of poor quality and chimeric reads, and generates the final population of haplotypes. Output files were analyzed in the popClusteringViewer, which is generated in the processClusters step and creates a local HTTP server based interactive site to view sequencing results.

Pfama1 Haplotype Analysis

DnaSP (v5.10.1) was used for analysis of nucleotide polymorphisms and haplotype diversity (166; 251). Population pairwise (F_{ST}) comparisons were determined between DRC provinces using Analysis of Molecular Variance (AMOVA) tool in the Arlequin (v3.5.2.2) population genetics data analysis program (94). We used the program Network (version 4.613) along with the DNA Alignment (v1.3.3.2) and Network Publisher (v2.0.01) add-ons (www.fluxus-engineering.com) to generate a median-joining (MJ) network diagram for visualization of phylogenetic relationships between *pfama1* haplotypes (34). EstimateS (v9.1.0) was utilized for developing rarefaction curves (66). Phylogenetic analysis was performed using MEGA version 6 (284).

Statistical Analyses and Data Visualization

Microsoft Excel, GraphPad Prism (v6), and SPSS (v22) were used for statistical analyses. Maps were created in ArcGIS ArcMap (ESRI, v.10.2.2). The DRC province boundary map was obtained from the Spatial Data Repository, part of the DHS Program (5).

RESULTS

P. falciparum lactate dehydrogenase qPCR

All individual samples (n=115) were tested for *P. falciparum* lactate dehydrogenase (*pfldh*) by qPCR. Of the 115 individual samples, 99 were positive for *pfldh* (Table 17). The concentration of *pfldh* ranged from less than 0.1 ng/ml to over 1,000 ng/ml (Table 17) based on standard curve analysis. Approximately 56% (n=55) of samples that were positive for *pfldh* were found at concentrations of less than 0.1 ng/ml.

Pfama1 Conventional PCR

P. falciparum ama1 (*pfama1*) specific conventional PCR was performed on all individual samples (n=115) and geographical cluster samples (n=88).

Pfama1 Results for Individual Samples

For the individual samples, 12 out of the 16 (75%) samples that were negative for the *pfldh* qPCR were also negative with the *pfama1* PCR (Table 17), suggesting that *P. falciparum* nucleic acid was absent or below the limit of detection for these assays. Four samples that were negative for *pfldh* were positive for *pfama1* (Table 17). Twenty-one samples were positive for *pfldh* but negative for *pfama1* and 16 of these *pfama1* negative samples had *pfldh* concentrations <0.1 ng/ml. Therefore, low *P. falciparum* nucleic acid concentrations may limit the detection ability of the *pfama1* PCR assay. Overall, 81 of the 115 (70.4%) individual samples were positive for *pfama1* and subsequently utilized for downstream deep sequencing reactions.

Pfama1 Results for Geographical Cluster Samples

Of the 88 pooled geographical cluster samples, 82 were successfully amplified using the *pfama1* PCR assay and utilized for downstream deep sequencing reactions.

Pfama1 **Amplicon-based Deep Sequencing**

We successfully sequenced 79 (68.7%) of the individual samples, 76 (86.4%) of the pooled geographical cluster samples, and six sequencing controls (n=161). We employed a conservative 2.5% haplotype cutoff, meaning we included *pfama1* haplotypes generated from sequence reads that occurred at a frequency of ≥2.5% within a given sample. A 2.5% cutoff was chosen in order to limit the inclusion of nonspecific and chimeric sequences in the dataset. Using the stringent 2.5% cutoff, we obtained 3,739,195 sequencing reads with an average sequence length of approximately 195 bp. We estimated the average number of reads per individual sample/geographical cluster sample to be 23,225 reads (3,739,195 total reads /161 samples = 23,225 reads). Therefore, based on a 2.5% cutoff, approximately 581 sequencing reads were used to construct the sequence of minor frequency haplotypes (23,225 x 2.5% = 581 reads). Rarefaction curve analysis using the 2.5% cut-off indicates that the inclusion of more samples would likely result in the detection of additional haplotypes, as we have not yet reached the asymptote (Figure 14). We identified a total of 88 unique *pfama1* haplotypes from both the individual samples and the pooled geographical cluster samples.

Several additional quality control methods were utilized to reduce background and limit inclusion of sequencing artifacts. First, all individual samples and geographical cluster samples were PCR amplified and sequenced in duplicate. Sequencing results for each duplicate were then compared to ensure the haplotypes matched in sequence and

fraction between duplicates. Second, a high fidelity PCR polymerase was utilized to reduce PCR error and generation of SNPs during PCR cycles. Finally, in order to reduce sequencing bias during library preparation (i.e preferential amplification of dominant haplotypes within pooled samples), equal concentrations of PCR products from individual and geographical cluster samples were calculated and added during library preparation steps.

A sequencing quality control sample was generated by mixing known concentrations of DNA from several *P. falciparum* strains and subsequently used as template in duplicate for three separate *pfama1* PCRs followed by library preparation and deep sequencing. The sequencing control contained *P. falciparum* DNA from V1s, Ro33, dd2, 7g8, and K1 strains at 5, 10, 15, 30, and 40 percent, respectively. As shown in Figure 15, the six sequencing controls show the expected haplotypes at approximately the same frequencies.

Individual Samples

For the individual samples (n=79), a total of 68 haplotypes were identified. Demographic characteristics for the 79 individual samples are shown in Table 18. The majority of individual samples showed multiclonal *P. falciparum* infections (64.5%) defined as a COI>1. The mean COI for individual samples was 2.43 and ranged from 1 to 9 haplotypes (Table 9). Figure 16 shows the geographical location of the 79 individual samples within the DRC. The size of the pie chart reflects the COI of the individual while the colors within the pie chart represent the particular haplotype fractions within a single individual. No clustering of particular *pfama1* haplotypes was observed based on geographical location in Figure 16.

No differences in *P. falciparum* COI estimated by *pfama1* haplotypes were observed based on age (Figure 17), sex (Figure 18), and HIV status (Figure 19). We utilized the median (COI=2) as the cut off to identify individual samples with a high COI (COI>2) versus a low COI (COI≤2). Two by two contingency analysis based on HIV status and a high COI versus low COI showed no difference (two-tailed t-test P=0.7955). A logistic regression analysis to estimate the odds ratio showed no significant impact of age, sex, or HIV status on high COI or low COI (Figure 20). We also used linear regression to analyze COI as a continuous variable and also found no impact of age, sex, or HIV status based on COI. As elevation has been shown to influence malaria transmission based on suitability for maintaining mosquito vector habitat (183; 293), we extrapolated elevation data based on geographical locations of individual samples. No relationship was observed between COI and elevation (Figure 21).

To analyze COI differences between the 11 DRC provinces, we averaged individual COIs within each province (Figure 22A, Table 19). We also compared average COI within a particular province with the percent of samples possessing a monoclonal (COI=1) versus percent polyclonal (COI>1) *P. falciparum* infection (Figure 22B). Kasaï-Oriental province had the highest average COI (COI=3.5), while Bandundu, Nord-Kivu, and Orientale had COIs>2.5 (Figure 22A, Table 19). As several studies have shown that higher COIs coincide with regions of high malaria prevalence (72; 310), we compared individual COI with malaria prevalence based on data from the 2007 EDS-RDC samples reported in a separate study (291). However, we did not find a significant correlation between COI and malaria prevalence (Figure 23).

Pooled Geographical Cluster Samples

Sixty-two haplotypes were identified in the pooled geographical cluster samples (n=76) and 82.9% were polyclonal (COI>1). Comparison of haplotypes found in the same geographical cluster or nearby geographical cluster between the individual samples and the pooled geographical cluster samples revealed that unique haplotypes were found infecting individuals that were not found in the same or nearby pooled geographical cluster sample. Further, we identified 27 unique *pfama1* haplotypes in individual samples that were not identified in the pooled geographical cluster samples and similarly, we identified 20 unique *pfama1* haplotypes in the geographical cluster samples that we did not find in individual samples. These data suggest that there is no geospatial clustering of particular *pfama1* haplotypes within our particular sample set from the DRC.

Population Genetics Analyses

Several population genetics analysis methods were utilized to investigate the *pfama1* haplotype diversity within individual samples in the DRC (Table 19). There were a total of 38 polymorphic sites (S) within the targeted region of *pfama1* across all 79 individual samples. We analyzed the nucleotide diversity (π), which is a measure of polymorphism based on the average nucleotide differences per site between randomly chosen sequences (204). Nucleotide diversity was similar between HIV negative and HIV positive individuals, and across different DRC provinces. We also determined haplotype diversity (Hd), another measure of population diversity, that calculates the probability that two randomly selected haplotypes in a population are different (117; 203). For all individual samples, the Hd was 0.998 (Table 19), indicating high haplotype diversity. We

106

compared Hd between haplotypes identified based on HIV status and DRC province and found high Hd across all groups (Table 19).

To investigate geospatial haplotype diversity, we compared the population fixation index (F_{ST}) between DRC provinces from individual samples. F_{ST} is a measure of the interpopulation heterogeneity based on allele frequencies and ranges between 0 and 1, indicating random mating (panmixis) and isolated populations, respectively (Table 20). The average F_{ST} value is 0.009, ranging from 0 to 0.04312. The low F_{ST} values suggest that *pfama1* haplotypes are not restricted based on DRC province.

To further explore whether certain *pfama1* haplotypes or related groups of haplotypes dominated in geographical regions within the DRC, we constructed a Median-Joining Network Diagram using the 68 haplotypes found in individual samples (Figure 24). Our Network Diagram analysis showed no clear clustering of related *pfama1* sequences based on DRC province, but instead illustrates the diversity of *pfama1* haplotypes found throughout the DRC. We also constructed a phylogenetic tree from the 68 *pfama1* haplotypes from individual samples to further explore the relatedness between haplotypes (Figure 25).

DISCUSSION

This study shows the utility of amplicon-based deep sequencing and the SeekDeep analysis pipeline to detect minor variants and analyze multiclonal *P. falciparum* infections. Using sensitive deep sequencing technology, we showed widespread sequence polymorphism of *pfama1* in *P. falciparum* parasites collected from asymptomatic individuals in the DRC. We also found that the majority of *P. falciparum* infections in the DRC are multiclonal (COI>1). We report no association between several

107

demographic factors, including age, sex, elevation, and HIV status, on the *P. falciparum* COI within individuals.

We also explored the geospatial distribution of *pfama1* haplotypes within the DRC and found no spatial clustering of haplotypes. Our findings are in agreement with another recent study that describes the population genetics of *P. falciparum* within the DRC as a "complex and fragmented landscape"(60). These results suggest that *P. falciparum* parasites are not restricted based on geography within the DRC, but instead are likely moving along with their human hosts between provinces and neighboring countries (60). Other possible explanations for the geospatial diversity of *pfama1* haplotypes in the DRC include movement of infected mosquito vectors between provinces and immune selection that maintains extensive malaria antigenic diversity within the DRC (282).

Selection of amplicon-based deep sequencing of *pfama1* for detection of multiclonal *P. falciparum* infections and COI analyses was based on several factors. Amplicon-deep sequencing is a cost effective method for targeted sequencing that results in very high coverage compared to whole genome sequencing, which can be more labor intensive and expensive. Additionally, using our amplicon-based deep sequencing approach we were able to pool several dozen samples, which also reduced overall cost and allowed us to analyze a larger sample size. *Pfama1* was chosen as a deep sequencing target for several reasons including 1) prior success of deep sequencing *pfama1* from dried blood spot samples, 2) *pfama1* is a single copy gene, 3) fewer repetitive sequences are present in *pfama1* compared to other targets for COI analysis that can cause sequencing errors, and 4) relevance of *pfama1* haplotype diversity to malaria AMA1

based vaccine development. We chose the Ion Torrent PGM (Life Technologies) as our deep sequencing platform because of the availability at the University of North Carolina Chapel Hill, prior experience of laboratory personnel, and reduced cost compared to other deep sequencing platforms.

Our study has several limitations. First, despite the sensitivity of deep sequencing for detecting minor variants within a population, we likely missed *P. falciparum* haplotypes due to one or more of the following: 1) degradation of *P. falciparum* nucleic acid in dried blood spots 2) sequence polymorphisms in the *pfama1* primer binding regions, 3) haplotype frequency below the limit of detection of *pfama1* PCR, and 4) haplotype sequence frequency less than the 2.5% cutoff for SeekDeep analysis. Second, we analyzed a subset of the total *P. falciparum* positive samples collected from the 2007 EDS-RDC and therefore may have limited our ability to detect significant differences in COI based on age, sex, and HIV status. Third, our analyses of COI and *P. falciparum* population genetics are restricted to asymptomatic malaria infections in the DRC. The inclusion of symptomatic malaria infections might have provided additional information on the *pfama1* haplotypes that cause clinical disease or the relationship between COI and disease in the DRC. Finally, we used an amplicon-based deep sequencing approach based on *pfama1* as a marker to identify different *P. falciparum* strains. Although similar approaches have been used previously to determine multiclonal/multistrain malaria infections (72; 134; 169), this method assumes that genetic polymorphisms within the single copy *pfama1* reflect genetically distinct *P. falciparum* strains. As the *P. falciparum* genome encompasses over 5,000 protein-coding genes, the use of *pfama1* as a

marker of genetic diversity likely underestimates the true genomic diversity of *P*. *falciparum* strains circulating in the DRC.

Somewhat unexpectedly, we found no significant correlation between *P*. *falciparum* prevalence and COI. However, our inability to detect a positive correlation between *P falciparum* COI and malaria prevalence may be due to the small sample size included in this study or due to the exclusion of minor frequency haplotypes below the 2.5% cutoff. Since we did not include all malaria positive samples from the 2007 EDS-RDC in this current study, inclusion of additional samples for COI analysis could potentially improve the correlation between *P*. *falciparum* prevalence and COI. Although several studies have recently described the significant relationship between COI and malaria prevalence (72; 104; 206; 306; 310), our data indicate that additional research is needed to confirm this relationship in the DRC.

Despite high malaria prevalence, relatively few studies have been conducted to understand malaria epidemiology and population genetics in the DRC (60; 187-189; 191; 288; 291; 292; 317). Our results have several implications for future malaria epidemiology studies in the DRC. First, AMA1 is a highly studied malaria vaccine candidate and is currently undergoing clinical trials in other malaria endemic regions (4; 270; 302; 304). As several studies have now shown, AMA1 based vaccines are effective against *P*. *falciparum* strains that have the same (homologous) or highly similar *pfama1* alleles (68; 85). We identified a total of 88 *pfama1* haplotypes and provide evidence of the extensive *pfama1* diversity present within the DRC that might be considered for the identification of appropriate *P*. *falciparum* haplotypes to include in an AMA1 based vaccine in the DRC and surrounding countries. Understanding the *pfama1* haplotype

110

diversity in a malaria endemic region is critical for developing an AMA1 based vaccine that would successfully protect against disease from circulating endemic *P. falciparum* strains (68; 85; 238; 281; 294; 295).

Second, the 2007 EDS-RDC samples were collected prior to the scale-up of malaria intervention strategies aimed at reducing malaria transmission and disease in the DRC. The DRC became a PMI focus country in 2010 and has since seen an increase in intervention strategies to reduce malaria (12; 13). Five "key intervention areas" are now part of the DRC Malaria Operation Plan (MOP) under the National Malaria Control Program (NMCP) and PMI, including: insecticide treated bed-nets, malaria Rapid Diagnostic Tests (mRDTs), artemisinin-based combination therapy (ACTs), intermittent preventive treatment in pregnancy (IPTp) with sulfadoxine-pyrimethamine (SP), and training heath care workers for malaria treatment and diagnosis (12; 13). A second Demographic and Health Survey was conducted in the DRC (EDS-RDC II) between November 2013 and February 2014. The EDS-RDC II reported 70% of households own at least one ITN, 56% children under five and 60% of pregnant women slept under an ITN the night before (193). Microscopic diagnosis for malaria in children (age 6-59 months) was performed and showed 23% of children positive for malaria parasitemia (193). The EDS-RDC II results highlight the success of malaria control interventions in the DRC since initiation of the PMI program compared to the data reported in the 2007 EDS-RDC. As *P. falciparum* COI has been shown to decline in malaria endemic regions where malaria transmission is reduced, our results can be utilized as a baseline to compare to samples collected in the 2013-2014 EDS-RDC II (72; 212; 263). Future studies are planned to compare *P. falciparum* COI and haplotype population genetics

between the 2007 and 2013-2014 EDS-RDCs in order to understand the changing malaria genetic landscape in the DRC and also monitor for changes in transmission due to enhanced, sustained, or lapsed malaria control campaigns (72).

As we approach the goal of malaria control and elimination, sensitive detection methods, such as amplicon-based deep sequencing technologies have the potential to increase our understanding of the antigen heterogeneity of malaria vaccine candidates, monitor for changes in malaria transmission intensity, and provide additional information about malaria population genomics.

FUTURE DIRECTIONS

We plan to utilize the *pfama1* amplicon-based deep sequencing technique and SeekDeep analysis platform to characterize *P. falciparum* COI and detect minor variants in three future epidemiology studies. First, we are presently collecting cross-sectional samples from a malaria holoendemic region in Western Kenya from individuals with asymptomatic malaria. These cross-sectional samples will also be tested for HIV, which will allow for further investigation into the impact of HIV on the *P. falciparum* COI and from a larger sample number. Additionally, we will also be conducting a longitudinal study that will allow for investigation into the changes in *P. falciparum* COI in HIV positive individuals as they initiate antiretroviral therapy and analyze changes in *P. falciparum* COI based on CD4 T cell counts. Second, we will be receiving both cross-sectional and longitudinal samples from malaria asymptomatic individuals in Nigeria. These individuals will also be tested for HIV and a subset of HIV positive individuals will be followed longitudinally. Finally, as mentioned in the discussion, we are interested in comparing the *P. falciparum* COI and *pfama1* haplotype data between the 2007 EDS-

112

RDC and the 2013-2014 EDS-RDC II samples to see if we can detect changes in *P.*

falciparum COI following the scale-up of malaria intervention strategies in the DRC.

Table 15. *Pfama1* heminested primer sequences and PCR conditions.

PCR Primers	*Primer Sequence (5'→3')	
Ama1OF	GCTGAAGTAGCTGGAACTCAA	
Ama1F	XXXXXXXXXXCCATCAGGGAAATGTCCAGT	
Ama1R	TTTCCTGCATGTCTTGAACA	
PCR Reagents	**Final concentration in PCR**	
Platinum PCR SuperMix High Fidelity (Life Technologies)	0.83X	
Ama1OF/F primer (20μM)	170 nM	
Ama1R primer (20μM)	170 nM	
MgCl₂ (25mM)	420 nM	
Template DNA volume	6.5 μl	
Total volume	30 μl	
PCR Conditions	**Time**	**Temperature**
Step 1: Initial denaturation	2 min	94°C
Step 2: Denaturation	30 sec	94°C
Step 3: Annealing	30 sec	55°C
Step 4: Elongation	1 min	68°C
Step 5: Cycling	Repeat steps 2-5 for 40 cycles total	
Step 6: Final elongation	10 min	68°C
Platform	**BIO-RAD T100 Thermal Cycler**	

*Xs represent the10 nucleotide MID sequence added to the 5' end of primer Ama1F.

Table 16. *Pfama1* Multiplex identifying (MID) sequences.

MID	Sequence (5'→3')	MID	Sequence (5'→3')
MID1	ACGAGTGCGT	MID14	CGAGAGATAC
MID2	ACGCTCGACA	MID15	ATACGACGTA
MID3	AGACGCACTC	MID16	TCACGTACTA
MID4	AGCACTGTAG	MID17	CGTCTAGTAC
MID5	ATCAGACACG	MID18	TCTACGTAGC
MID6	ATATCGCGAG	MID19	TGTACTACTC
MID7	CGTGTCTCTA	MID20	ACGACTACAG
MID8	CTCGCGTGTC	MID21	CGTAGACTAG
MID10	TCTCTATGCG	MID22	TACGAGTATG
MID11	TGATACGTCT	MID23	TACTCTCGTG
MID13	CATAGTAGTG	MID24	TAGAGACGAG

Table 17. Individual sample *pfldh* qPCR, *pfama1* PCR, and *pfama1* sequencing results

Sample	*Pfldh* qPCR	*Pfldh* qPCR ng/ml	*Pfama1* PCR	*Pfama1* sequence
A0D1E	Positive	<0.10	Negative	Negative
A5F6Q	Positive	0.43	Positive	Positive
A6U7I	Positive	<0.10	Positive	Positive
A8F4R	Positive	<0.10	Negative	Negative
A8L0T	Positive	<0.10	Positive	Positive
B0E1G	Positive	0.21	Positive	Positive
B0R5X	Positive	<0.10	Negative	Negative
B3H5Q	Positive	0.93	Positive	Positive
B4Y1F	Positive	<0.10	Positive	Positive
B6H1P	Positive	0.43	Positive	Positive
B8C7S	Positive	0.35	Positive	Positive
B9H9B	Positive	0.40	Positive	Positive
B9N0X	Negative	n/a	Negative	Negative
C4B1I	Positive	0.11	Positive	Positive
C5W9N	Positive	0.24	Positive	Positive
C7U8M	Positive	21.34	Positive	Positive
C7X1I	Positive	>1000	Positive	Negative
D1N7Y	Positive	0.18	Positive	Positive
E4U6J	Positive	<0.10	Negative	Negative
E5K5Y	Positive	<0.10	Positive	Positive
E5O3B	Positive	<0.10	Positive	Positive
E7Q3F	Positive	0.10	Positive	Positive
F1Q8F	Negative	n/a	Negative	Negative
F8M5F	Negative	n/a	Negative	Negative
G0N4X	Positive	<0.10	Positive	Positive
G1K9B	Positive	<0.10	Positive	Positive
G3U03	Positive	0.17	Positive	Positive
G3U5J	Positive	<0.10	Negative	Negative
G9U2M	Positive	<0.10	Negative	Negative
H1O7E	Positive	<0.10	Positive	Positive
H2F2Q	Positive	<0.10	Negative	Negative
H5R5J	Positive	0.17	Positive	Positive
I0Y1I	Positive	<0.10	Positive	Positive
I1Y4M	Positive	3.52	Positive	Positive
I2H8Z	Positive	0.39	Positive	Positive
I4T0G	Positive	<0.10	Negative	Negative
I7J7G	Positive	0.33	Positive	Positive
J1B3O	Positive	0.12	Positive	Positive
J1K6B	Positive	6.57	Positive	Positive

J2D8W	Positive	0.20	Negative	Negative
J3U0H	Positive	0.20	Positive	Positive
J7J4E	Positive	0.34	Positive	Positive
J8O6M	Positive	<0.10	Positive	Positive
K2J7D	Positive	<0.10	Negative	Negative
K4V5P	Positive	0.15	Positive	Positive
K5C3U	Positive	<0.10	Positive	Positive
K5D8B	Positive	<0.10	Positive	Positive
K6U5Q	Positive	<0.10	Positive	Positive
K8B5Y	Positive	<0.10	Negative	Negative
L2L6F	Positive	<0.10	Negative	Negative
L3V5P	Negative	n/a	Negative	Negative
L5V7T	Positive	<0.10	Positive	Positive
L6H9I	Positive	0.51	Positive	Positive
L7F6E	Positive	<0.10	Positive	Positive
L9C8F	Negative	n/a	Positive	Positive
M0Y0L	Positive	0.90	Positive	Positive
M2P4I	Negative	n/a	Positive	Positive
M2R0G	Positive	<0.10	Negative	Negative
M5X6V	Negative	n/a	Negative	Negative
M7E2Z	Positive	<0.10	Positive	Positive
M9R6T	Positive	<0.10	Positive	Positive
N2D2U	Positive	0.82	Positive	Positive
N2Q9P	Positive	<0.10	Negative	Negative
N5U0N	Positive	<0.10	Positive	Positive
O0M4F	Positive	<0.10	Positive	Positive
O1Q9P	Negative	n/a	Negative	Negative
O1V6R	Positive	>1000	Positive	Positive
O3R5O	Positive	0.55	Positive	Positive
O6I9M	Positive	<0.10	Positive	Positive
P1M6J	Positive	0.11	Positive	Positive
P5K0F	Positive	<0.10	Positive	Positive
Q0K1C	Positive	0.30	Negative	Negative
Q1K1D	Positive	<0.10	Positive	Positive
Q3H7I	Positive	<0.10	Positive	Positive
Q6X5Z	Positive	0.31	Positive	Positive
Q7N8T	Positive	<0.10	Positive	Positive
Q8X4B	Positive	<0.10	Positive	Positive
R7I5M	Positive	0.17	Negative	Negative
R7Y9H	Positive	0.66	Positive	Positive
S1H1C	Positive	<0.10	Positive	Positive
S3P2N	Positive	<0.10	Positive	Positive
S5N7S	Negative	n/a	Negative	Negative

S9C5J	Positive	0.63	Negative	Negative
S9F7O	Positive	0.35	Positive	Positive
T1Y2V	Positive	<0.10	Negative	Negative
T5F8M	Positive	15.06	Positive	Positive
T5S6X	Positive	<0.10	Positive	Positive
T7U2X	Positive	0.30	Negative	Negative
T9D7N	Positive	<0.10	Positive	Positive
T9L3R	Positive	1.73	Positive	Positive
T9L5T	Positive	0.44	Positive	Positive
T9X4F	Positive	1.04	Positive	Positive
U3D5G	Positive	<0.10	Negative	Negative
U4X6D	Positive	<0.10	Positive	Positive
V3D4G	Positive	<0.10	Positive	Positive
V3T6Y	Positive	0.55	Positive	Positive
V6H9S	Positive	0.14	Positive	Positive
V9I1O	Positive	<0.10	Positive	Positive
W3O0O	Negative	n/a	Negative	Negative
W7E9R	Negative	n/a	Negative	Negative
W8D8Q	Positive	1.01	Positive	Positive
W9R6E	Negative	n/a	Positive	Negative
X2L4P	Negative	n/a	Negative	Negative
X2P1Q	Positive	<0.10	Positive	Positive
X3P4U	Negative	n/a	Positive	Positive
X5X71	Negative	n/a	Negative	Negative
X5Z8L	Positive	<0.10	Negative	Negative
X7A7M	Positive	<0.10	Positive	Positive
X9E6R	Positive	<0.10	Positive	Positive
Y1O9X	Positive	0.17	Positive	Positive
Y7U3E	Negative	n/a	Negative	Negative
Y9L6Z	Positive	<0.10	Positive	Positive
Y9Q2B	Positive	<0.10	Positive	Positive
Z6T2C	Positive	<0.10	Positive	Positive
Z9M6C	Positive	<0.10	Negative	Negative

Table 18. Demographic characteristics of individual samples with *pfama1* sequence data

Characteristic	Individual samples (n=79)
Female (%)	53 (67%)
Mean age (range)	32.7 (16-58)
HIV positive (%)	21 (27%)

Table 19. Population genetics analyses of *pfama1* haplotypes from individual samples

	n	# of haplotypes	Mean COI	Total # of polymorphic sites (S)	Nucleotide diversity (π)	Haplotype diversity (Hd)
All	79	68	2.43	38	0.04192	0.998
HIV positive	21	27	2.29	22	0.04146	1.000
HIV negative	58	63	2.47	38	0.04215	0.998
Bandundu	11	21	2.64	23	0.04086	1.000
Bas-Congo	4	7	2.00	20	0.04197	1.000
Équateur	11	21	2.36	24	0.04369	1.000
Kasaï-Occidental	9	14	1.89	20	0.04367	1.000
Kasaï-Oriental	8	19	3.50	23	0.04407	1.000
Katanga	4	6	1.75	15	0.03897	1.000
Kinshasa	7	13	2.14	21	0.04196	1.000
Maniema	4	6	1.50	17	0.03932	1.000
Nord-Kivu	2	5	3.00	17	0.04462	1.000
Orientale	14	28	2.93	24	0.04242	1.000
Sud-Kivu	5	7	1.60	21	0.04935	1.000

Table 20. Population pairwise comparisons between DRC provinces from individual samples.

	Bandundu	Bas-Congo	Equateur	Kasaï-Occidental	Kasaï-Oriental	Katanga	Kinshasa	Maniema	Nord-Kivu	Orientale
Bas-Congo	0.00250									
Equateur	0.00604	0.01372								
Kasaï-Occidental	0.00000	0.01231	0.0066							
Kasaï-Oriental	0.00266	0.02485	0.00466	0.01171						
Katanga	0.01885	0.02412	0.02228	0.01493	0.04312					
Kinshasa	0.00697	0.00000	0.00000	0.00000	0.00361	0.02278				
Maniema	0.00000	0.00000	0.00000	0.00265	0.00000	0.00063	0.01056			
Nord-Kivu	0.02426	0.00919	0.00000	0.01954	0.01003	0.03387	0.01838	0.00000		
Orientale	0.00000	0.00000	0.00000	0.00575	0.00799	0.01447	0.00000	0.00000	0.00000	
Sud-Kivu	0.00000	0.00000	0.00404	0.03287	0.02924	0.00605	0.02681	0.00000	0.00000	0.00000

MRKLYCVLLLSAFEFTYMINFGRGQNYWEHPYQNSDVYRPINEHREHPKEYEYPLHQEHTYQQEDSGEDENTLQHAYPIDHEGAEPAPQEQNLFSSIEIV 100

ERSNYMGNPWTEYMAKYDIEEVHGSGIRVDLGEDAEVAGTQYRLPSGKCPVFGKGIIIENSNTTFLTPVATGNQYLKDGGFAFPPTEPLMSPMTLDEMRH 200

FYKDNKYVKNLDELTLCSRHAGNMIPDNDKNSNYKYPAVYDDKDKKCHILYIAAQENNGPRYCNKDESKRNSMFCFRPAKDISFQNYTYLSKNVVDNWEK 300

VCPRKNLQNAKFGLWVDGNCEDIPHVNEFPAIDLFECNKLVFELSASDQPKQYEQHLTDYEKIKEGFKNKNASMIKSAFLPTGAFKADRYKSHGKGYNWG 400

NYNTETQKCEIFNVKPTCLINNSSYIATTALSHPIEVENNFPCSLYKDEIMKEIERESKRIKLNDNDDEGNKKIIAPRIFISDDKDSLKCPCDPEMVSNS 500

TCRFFVCKCVERRAEVTSNNEVVVKEEYKDEYADIPEHKPTYDKMKIIIASSAAVAVLATILMVYLYKRKGNAEKYDKMDEPQDYGKSNSRNDEMLDPEA 600

SFWGEEKRASHTTPVLMEKPYY 622

Figure 12. *P. falciparum* apical membrane antigen 1 (*pfama1*) amino acid (aa) sequence based on the 3d7 reference strain (XM_001347979.1).

> The first 24 aa (first grey box) contain the signal sequence (ss) followed by Domain I (blue shaded, aa 25-320), Domain II (red shaded, aa 321-442), Domain III (green shaded, aa 443-546), and the transmembrane domain (grey boxed, aa 547-622). The heminsted primers (Ama1OF, Ama1F, Ama1R) are shown in black boxes.

119
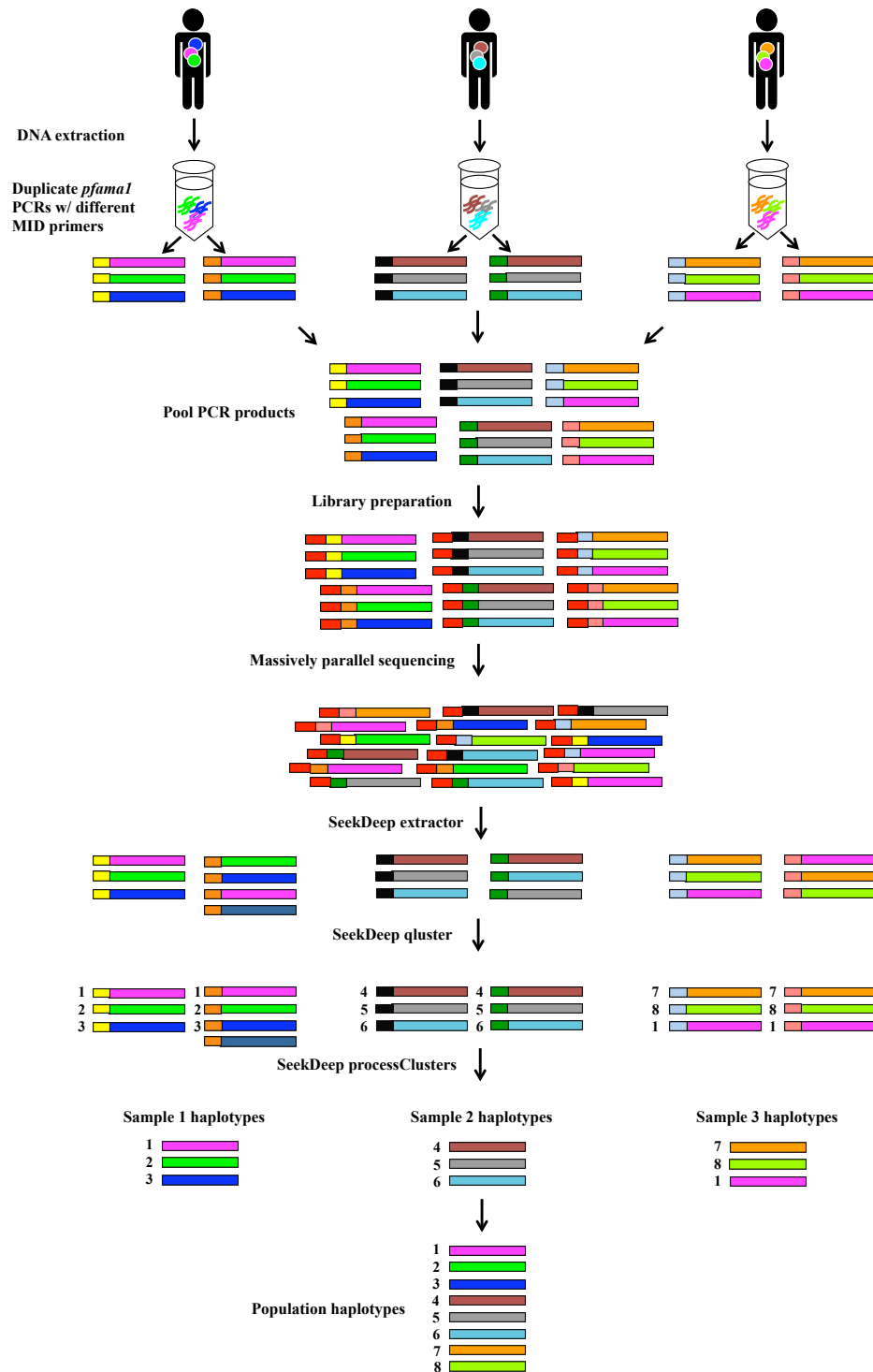
Figure 13. Schematic representation of the sample processing, library preparation, and SeekDeep pipeline for sequence analysis.

Circles represent distinct *P. falciparum* strains within an individual. Different color large rectangles indicate distinct *pfama1* haplotypes. Small rectangles represent MID sequences. Red rectangles represent the barcoded index sequence.
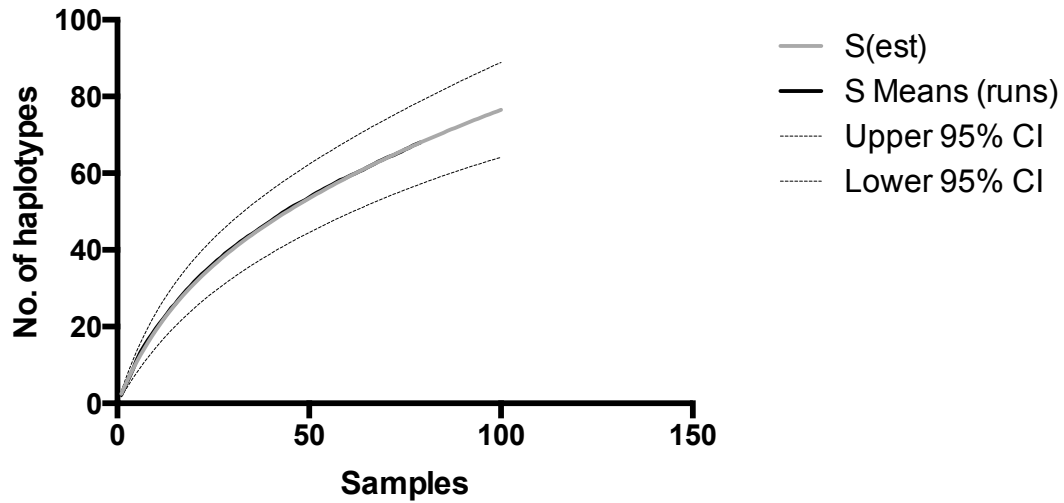
Figure 14. Rarefaction curve of *pfama1* haplotypes from individual samples.
Rarefaction curve of the actual number of samples compared to the number of haplotypes (S Means (runs)) compared to the estimated number of haplotypes (S(est)) with continued sampling.
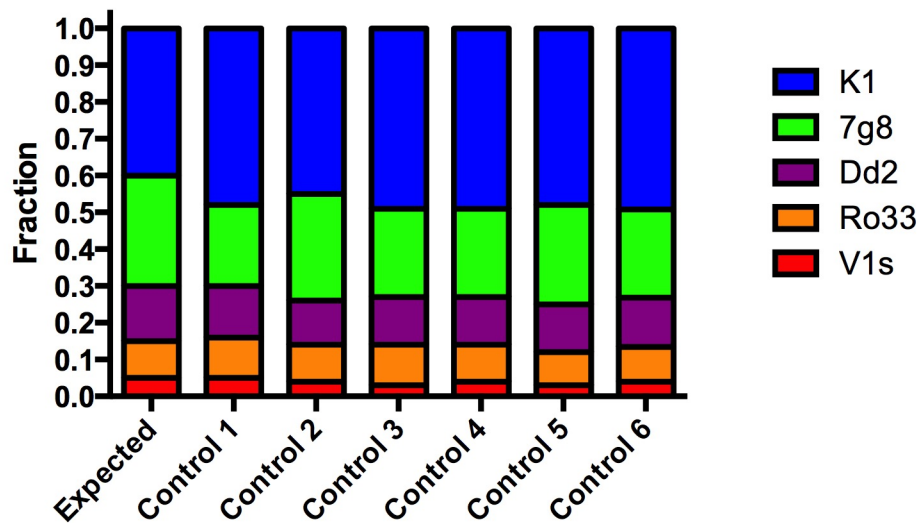
Figure 15. Deep sequencing control samples.
A sequencing control with known concentrations of *P. falciparum* strains (K1, 7g8, Dd2, Ro33, V1s) was utilized in duplicate for three separate PCRs and downstream sequencing reactions.
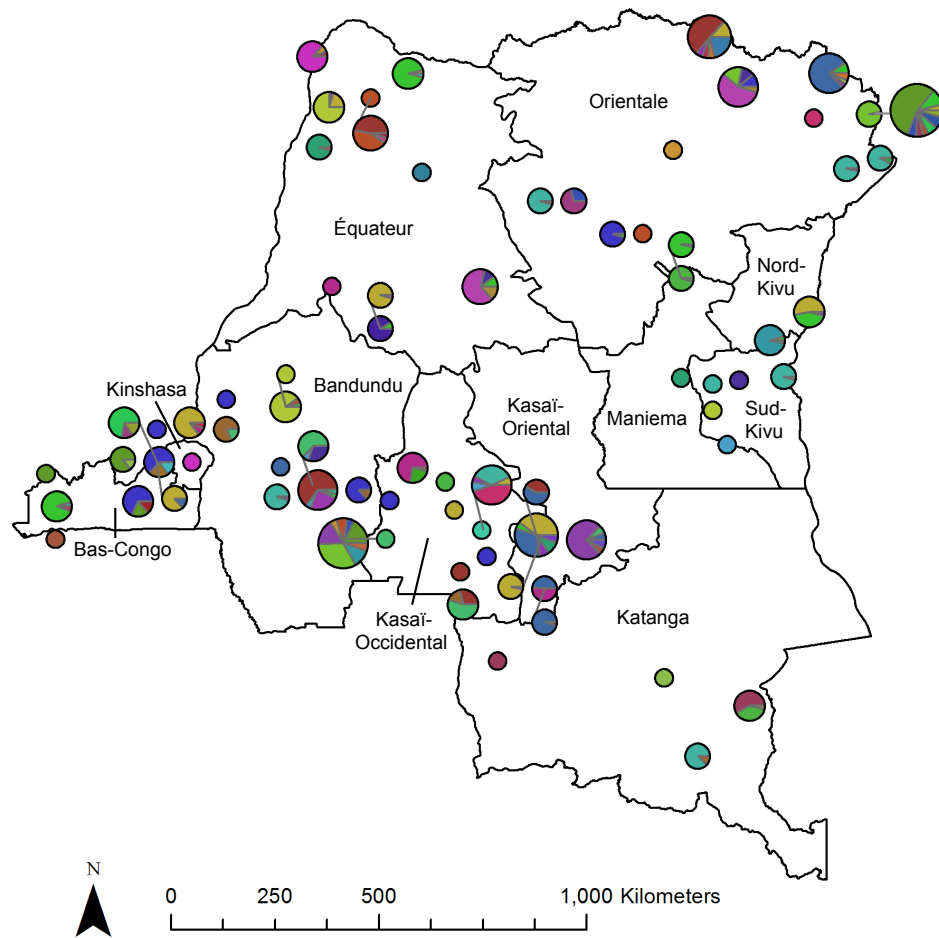
Figure 16. Individual sample COIs and haplotype frequencies based on geographical location in the DRC.

Size of pie chart represents the number of different haplotypes within a single individual. Colors indicate different haplotypes.

Figure 17. *P. falciparum* COI stratified by age in individual samples.
No differences were observed in *P. falciparum* COI based on age (ANOVA
P=0.9668).

Figure 18. *P. falciparum* COI based on sex in individual samples.
No differences were observed in *P. falciparum* COI between males and females
(two-tailed, unpaired t-test P=0.9017).
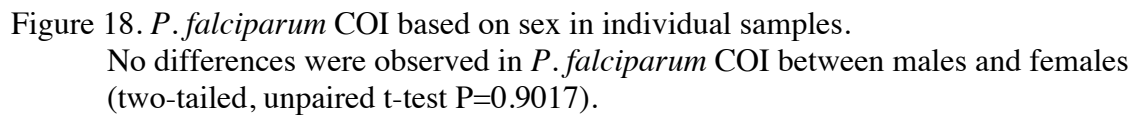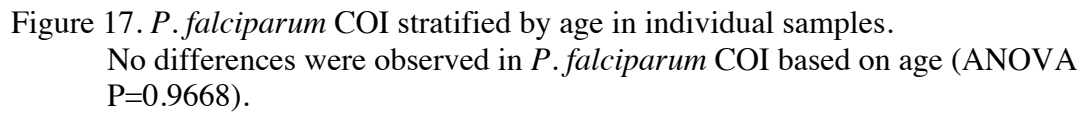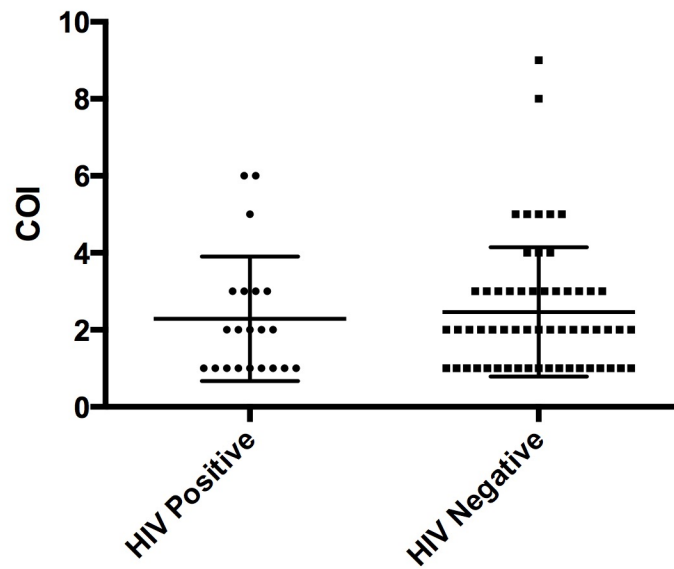
Figure 19. *P. falciparum* COI based on HIV status.
    No differences were observed in *P. falciparum* COI between HIV positive and HIV negative individuals (two-tailed, Fisher's exact test, P=0.6722).
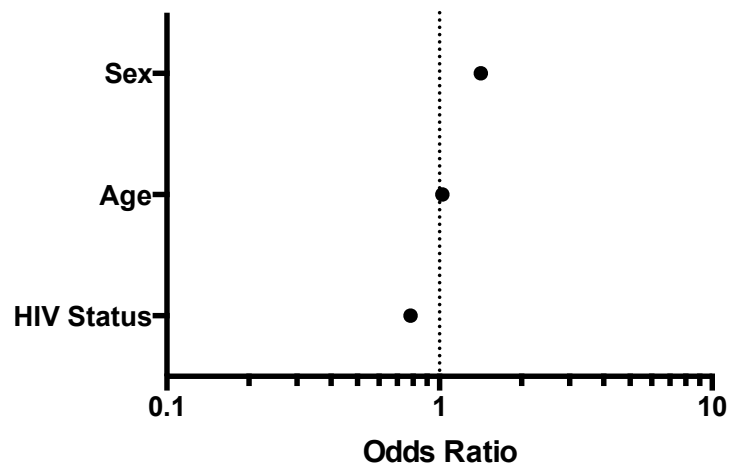
Figure 20. Estimated odds ratios using logistic regression analysis.
    Logistic regression to estimate the odds ratio of a high COI (COI>2) while controlling for age (OR=1.025, CI: 0.977-1.075, p=0.312), sex (OR=1.419, CI: 0.536-3.754, p=0.481), and HIV status (OR=0.783, CI: 0.270-2.272, p=0.652) was not significant.
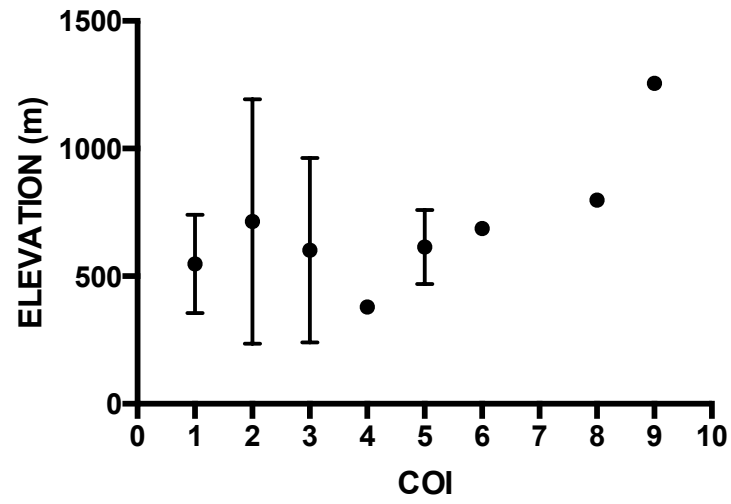
Figure 21. *Pfama1* COI based on elevation in meters.
Bars represent standard deviation (SD).

A)

COI
☐	1.5 - 2.0
▨	>2.0 - 2.5
▨	>2.5 - 3.0
■	>3.0 - 3.5

0   250   500   1,000 Kilometers

N
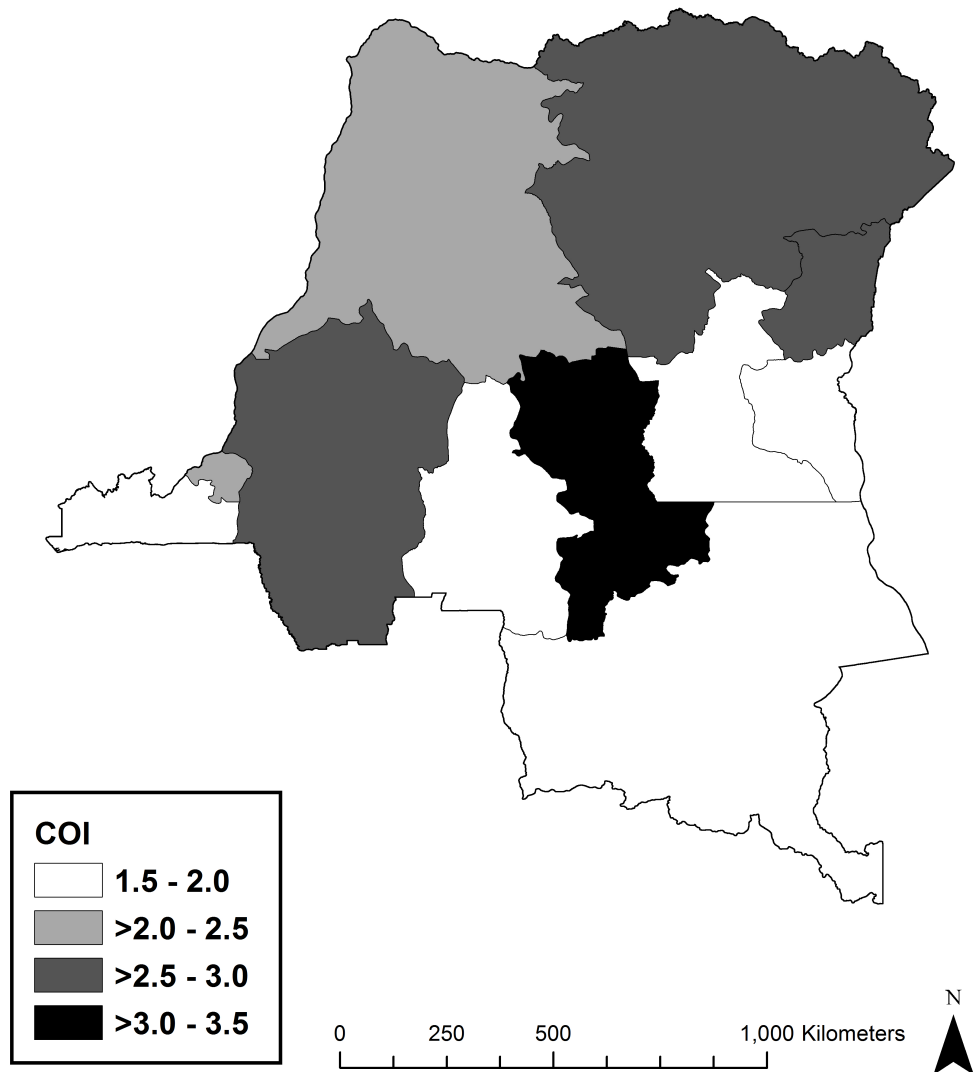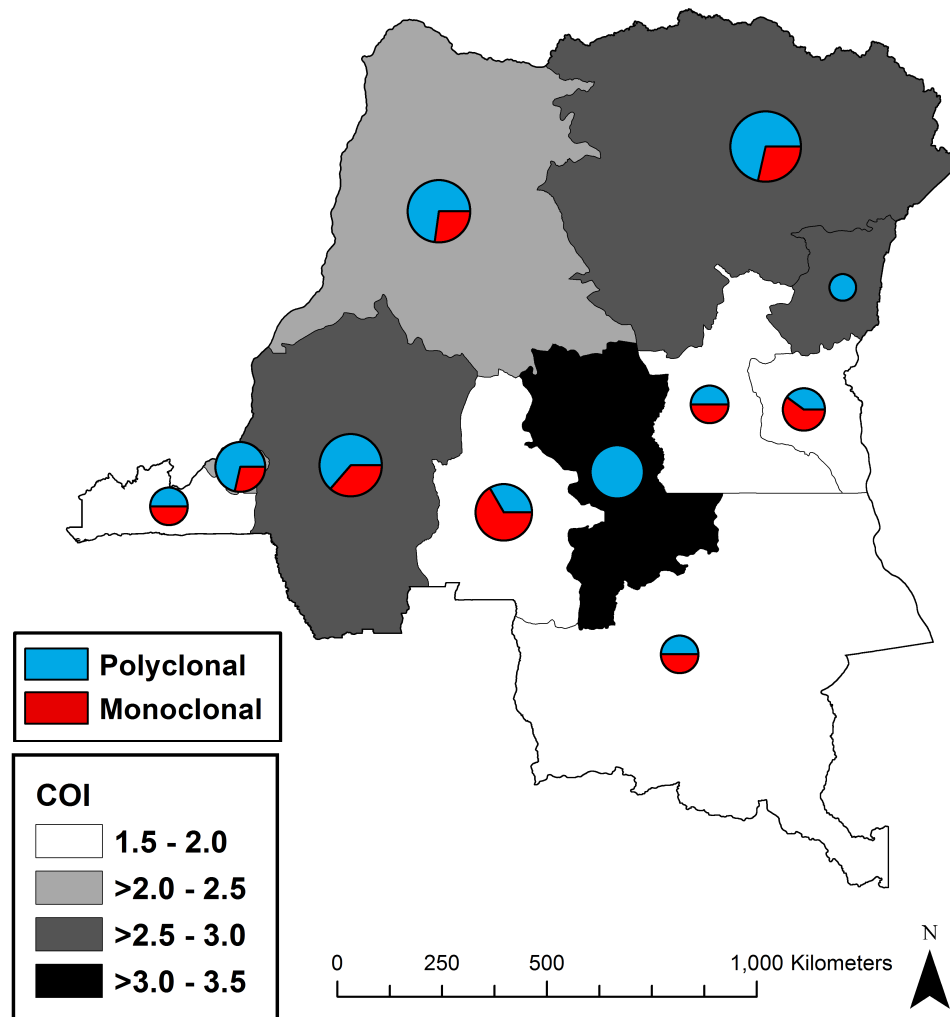
Figure 22. *P. falciparum* COI by DRC Province.
A) Average *P. falciparum* COI from individuals in each of the 11 DRC provinces.
B) Pie charts represent the fraction of individuals with monoclonal (COI=1) or
polyclonal (COI>1) from each province. The size of the pie chart reflects the
number of individual samples form each province.
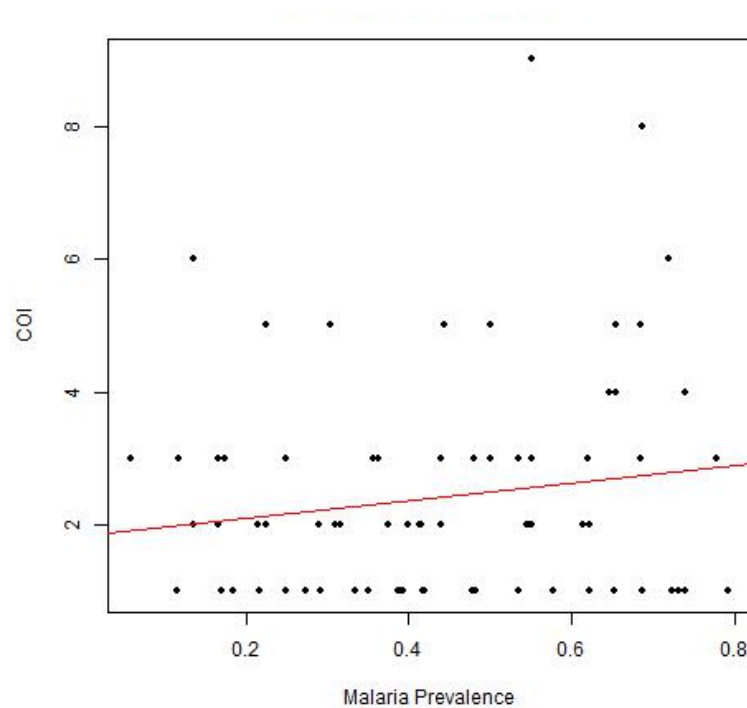
Figure 23. *P. falciparum* COI versus prevalence in the DRC.
　　　No significant relationship was found between *P. falciparum* COI and malaria
　　　prevalence in individual samples (Coefficient of determination, $R^2 = 0.01$,
　　　p=0.18). Figure prepared by Mark Janko, Departments of Biostatistics &
　　　Geography, University of North Carolina Chapel Hill.
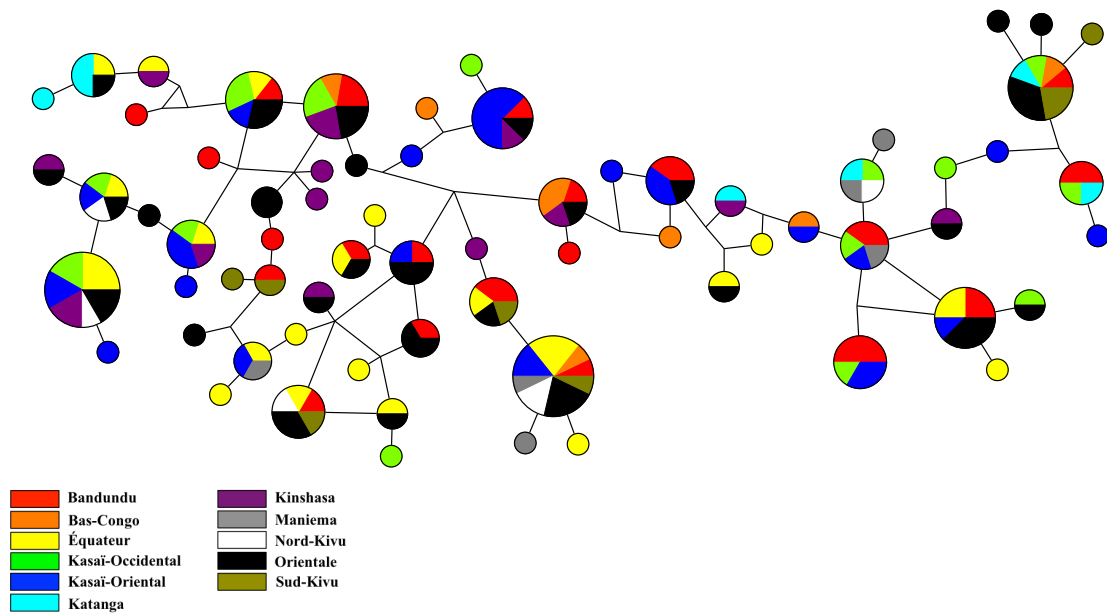
Figure 24. Median-joining Network Diagram representing the 68 *pfama1* haplotypes sequenced from individual samples (n=79).

> Each circle represents a different haplotype, the size of the circle reflects the number of samples containing that particular haplotype, and the colors indicate different DRC provinces. Each line between the circles indicates a nucleotide difference between related haplotypes.
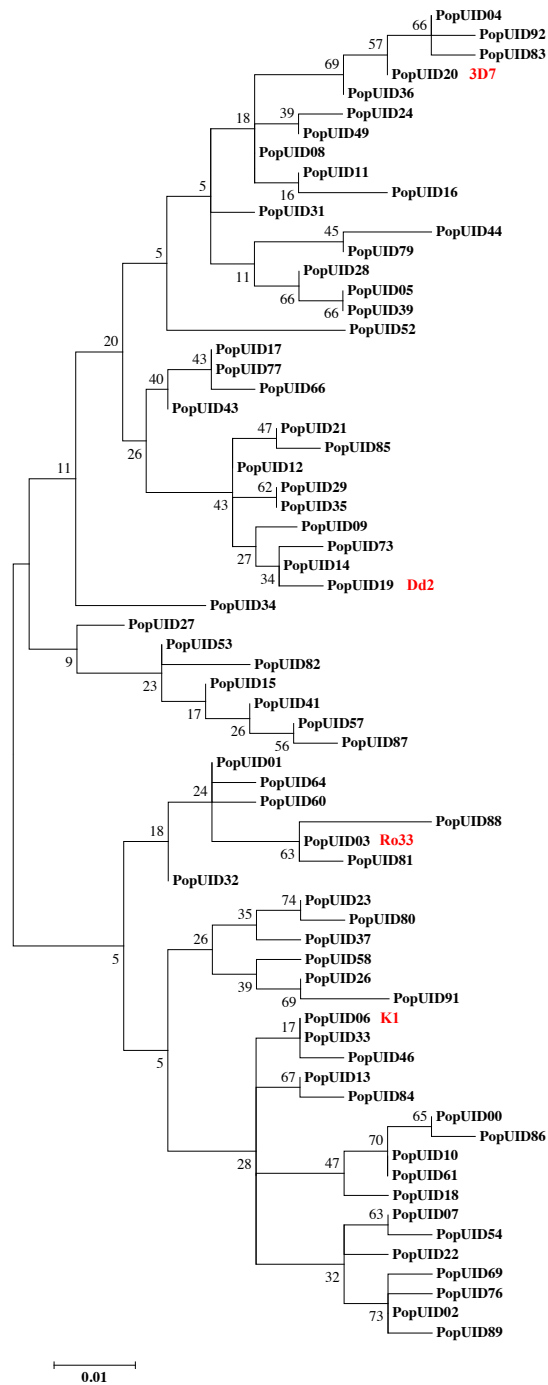
Figure 25. Phylogenetic tree of 68 *pfama1* haplotypes from individual samples. We used MEGA6 to generate a maximum-likelihood tree (1,000 bootstrap replicates) based on the Tamura-Nei model (284).

CHAPTER 5: Summary and General Conclusions

DISSERTATION SUMMARY

The research objectives for this dissertation were to 1) utilize geospatial tools to estimate the JEV vector *Cx. tritaeniorhynchus* prevalence in Asia and 2) develop highly sensitive molecular tools to detect malaria parasite species and strains in complex malaria infections. To address our research objectives, we utilized a geospatial modeling program to estimate *Cx. tritaeniorhynchus* prevalence throughout the entire JE endemic region, developed a novel real-time PCR (qPCR) assay to detect the neglected malaria parasite species *P. ovale,* and utilized next-generation deep sequencing to detect multiclonal *P. falciparum* infections and investigate the *P. falciparum* Complexity of Infection (COI) in the Democratic Republic of Congo (DRC). We utilized GIS to map the geographical location of *Cx. tritaeniorhynchus* and *P. falciparum* in endemic regions for JE and malaria, respectively, thus highlighting the utility of GIS to map the distribution of important vectors and pathogens that cause vector-borne diseases in humans. The approaches we used to achieve the research objectives in this dissertation are distinct from each other, and yet each was developed and conducted to address the overarching goal of improving the detection and surveillance of vector-borne diseases. Detection of both the pathogen and vector are critical for understanding vector-borne pathogen transmission dynamics and for controlling vector-borne diseases.

CHAPTER SUMMARIES

Chapter 2 Summary

The aim of chapter 2 (Aim 1) was to develop an ecological niche model to estimate the JEV vector *Cx. tritaeniorhynchus* prevalence in Japanese encephalitis endemic regions. To address this aim, we utilized the Maxent program and GIS software to develop an ecological niche model based on known occurrences of *Cx. tritaeniorhynchus* and preferred environmental conditions of the vector. Our ecological niche model showed that the majority of JE clinical cases were located within regions predicted to have high probability of *Cx. tritaeniorhynchus* prevalence. Since JEV remains a significant cause of viral encephalitis in Asia and the endemic region for JE has expanded dramatically into new geographical locations, our ecological niche model could potentially be utilized to allocate resources, such as vector control measures or vaccination campaigns, to areas with high estimated vector prevalence.

Limitations

Our ecological niche model to estimate *Cx. tritaeniorhynchus* prevalence is limited by several factors. First, sampling bias likely influenced our model as we only included a snapshot of *Cx. tritaeniorhynchus* occurrence data that was reported in the literature and did not take into account environmental changes that could increase or decrease vector prevalence over time. Also, extensive *Cx. tritaeniorhynchus* sampling data is not available for all regions in Southeast Asia and the Western Pacific and therefore our model could be improved with additional occurrence points from more geographical locations. Our study is further limited by including only JE case data available through published literature and ProMED data and does not reflect the actual

132

number of JEV infections, since the majority of infections are asymptomatic. Finally, ecological niche models only provide an initial framework for identifying regions that could be targeted by vector control measures, as additional entomological field surveys and community engagement are critical for successful control of vector populations.

Future Directions

Several future studies could be conducted based on our initial ecological niche model for *Cx. tritaeniorhynchus* distribution. For example, incorporation of recent *Cx. tritaeniorhynchus* occurrence data into the model and comparison to recent JE cases and outbreaks would provide an updated assessment of vector distribution. In fact, we recently provided the data used in our model to researchers at the University of Oxford who are developing an updated JE risk map based on several covariates, including vector prevalence. Another potential future study is to utilize environmental and climatic variables based on the climate change predictions to investigate whether *Cx. tritaeniorhynchus* estimated prevalence and geographical distribution is expected to change due to global warming.

Chapter 3 Summary

The aim of chapter 3 (Aim 2) was to improve detection of *P. ovale* by targeting a conserved genetic region between *P. ovale curtisi* and *P. ovale wallikeri* subspecies using a real-time PCR (qPCR) approach. Previously developed *P. ovale* PCR assays unintentionally detected only one of the two *P. ovale* subspecies due to genetic polymorphisms in the primer binding regions that limited amplification and detection by qPCR. Our novel *P. ovale*-specific qPCR assay is based on a conserved region between the two subspecies within the reticulocyte binding protein 2 (*rbp2)* gene and detects low-

levels of *P. ovale* parasites with high specificity even in the context of mixed malaria species infections. We validated our *P. ovale*-specific assay based several parameters such as defining the limit of detection and quantification, specificity, repeatability, and reproducibility. Our *P. ovale*-specific assay successfully detected both *P. ovale curtisi* and *P. ovale wallikeri* with a range of parasitemias from malaria asymptomatic samples in Kenya. Finally, we utilized multilocus genotyping to demonstrate for the first time that both *P. ovale curtisi* and *P. ovale wallikeri* circulate in a malaria holoendemic region in Western Kenya. In conclusion, our novel *P. ovale*-specific assay is a useful tool for the simultaneous detection of *P. ovale* subspecies that can be utilized to improve detection of this neglected malaria parasite species in malaria endemic regions.

Limitations

This research study is limited by several factors. First, the paucity of genetic information published for *P. ovale* limits our ability to ensure the complete lack of polymorphisms in the primer and probe regions from all *P. ovale* subspecies and strains worldwide. Therefore, we could still potentially miss *P. ovale* infections using our qPCR assay due to genetic polymorphisms within the primer and probe binding regions that would subsequently inhibit PCR amplification. Second, we only had access to a small sample set of *P. ovale* samples from one malaria endemic region, limiting our ability to fully validate our *P. ovale*-specific assay from larger and more globally representative sample set that also includes sub-microscopic *P. ovale* infections. Finally, we were unable to obtain pure *P. ovale* positive samples for generation of a standard curve, and thus were limited to validating our assay using a plasmid containing the *P. ovale rbp2* target.

Future Directions

Our *P. ovale*-specific qPCR assay is currently being used in a cross-sectional study in Western Kenya to determine *P. ovale* prevalence from several hundred samples collected from malaria asymptomatic adults. We also plan to utilize our *P. ovale*-specific assay for the detection of *P. ovale* infections in Nigeria in collaboration with the Walter Reed Program-Nigeria. Another possibility for future research is multilocus genotyping to determine *P. ovale* subspecies prevalence in Kenya and Nigeria, as the *P. ovale* subspecies prevalence in these countries has not yet been characterized. Additionally, we have plans to obtain CAP/CLIA certification for our *P. ovale*-specific qPCR assay to allow malaria parasite species detection in returned soldiers with malaria at the Walter Reed National Military Medical Center hospital in collaboration with the Walter Reed Army Institute of Research (WRAIR). In addition to detecting both *P. ovale* subspecies at the same time with a single qPCR assay, we are also interested in developing qPCR assays that can differentiate *P. ovale curtisi* and *P. ovale wallikeri* as we attempt to understand the potential clinical and epidemiological differences between these two subspecies.

Chapter 4 Summary

The aim of chapter 4 (Aim 3) was to utilize a deep sequencing approach to detect multiclonal *P. falciparum* infections and also explore the impact of several demographic factors such as age, sex, HIV status, and geographic location on the *P. falciparum* Complexity of Infection (COI). We utilized an amplicon-based deep sequencing approach targeting the polymorphic *P. falciparum* apical membrane antigen *(pfama1)* gene and analyzed the deep sequence data using the SeekDeep targeted amplicon analysis

pipeline. Overall, we successfully sequenced a region of the *pfama1* and found 88 unique *pfama1* haplotypes in the Democratic Republic of Congo (DRC). Also, the majority of individuals included in our study had multiclonal (COI>1) *P. falciparum* infections. We found no association between age, sex, HIV status, or geographical location and *P. falciparum* COI. Using several population genetics tools, we explored the *pfama1* diversity and report high *pfama1* genetic diversity from *P. falciparum* malaria parasites in the DRC. In conclusion, the research included in this chapter showed the utility of amplicon-based deep sequencing based on the *pfama1* target for the detection of multiclonal *P. falciparum* infections.

Limitations

The research included in this chapter is limited by several factors. First, we only included asymptomatic malaria infections in adults collected from a subset of samples from the 2007 EDS-RDC and therefore did not include malaria symptomatic infections or malaria infections in children. Second, we may have inadvertently failed to detect some *P. falciparum* strains due to polymorphisms in the *pfama1* primer binding regions that would prevent PCR amplification or *P. falciparum* strains that occur below the limit of *pfama1* PCR detection. Additionally, as we used a 2.5% cut off for inclusion of *pfama1* haplotypes for sequence analysis, we may have missed haplotypes that occurred below the cut off frequency level. Finally, we were unable find any difference in the *P. falciparum* COI based HIV status, age, sex, malaria prevalence, or geographic location, although this may have been the result of a relatively small sample size.

Future Directions

We are currently preparing the research included in this chapter of the dissertation for publication. We are also planning to perform additional analyses of *pfama1* sequences from the DRC, Republic of Congo, and Malawi in collaboration with Dr. Steven Meshnick and Dr. Jon Juliano at the University of North Carolina Chapel Hill that we hope will also be included in the publication. Additionally, we plan to investigate *P. falciparum* COI from samples collected as part of the 2013-2014 EDS-RDC II. This will allow for comparison of *P. falciparum* COI and prevalence after PMI's implementation of malaria intervention strategies in 2010. Finally, we also plan to use the *pfama1* amplicon-based deep sequencing approach to investigate multiclonal *P. falciparum* infections and the impact of HIV on *P. falciparum* COI in two malaria/HIV co-infections studies currently being conducted in Kenya and Nigeria using both cross-sectional and longitudinal samples.

OVERALL CONCLUSIONS

The overall goal of the research presented in this dissertation is to improve the detection of pathogens and vectors in order to advance the control and potential elimination of vector-borne diseases. We describe three distinct approaches for improving vector-borne disease detection, utilizing computational, geospatial, and molecular tools. Although we focused on the JEV vector and malaria parasites, the approaches described in this dissertation could potentially be applied to improve the detection other vector-borne and non-vector-borne infectious diseases. Control of vector-borne diseases requires a multifaceted approach that incorporates knowledge of the pathogen, host, vector, and environment while also maintaining sustainable invention

137

strategies that prevent disease reintroduction and work towards disease elimination. In conclusion, we end with the inspiring words of Dr. Margaret Chan, Director-General of the World Health Organization: " No one in the 21st century should die from the bite of a mosquito, a sandfly, a blackfly, or a tick."

REFERENCES

1. *Culex tritaeniorhynchus*. http://www.wrbu.org/SpeciesPages_non-ANO/non-ANO_A-hab/CXtri_hab.html

2. Democratic Republic of the Congo Demographic and Health Survey 2007 Key Findings, The Demographic and Health Surveys Program, USAID, Macro International Inc, Calverton, Maryland, USA

3. HIV Prevalence and HIV Testing. Rockville, Maryland, USA: The DHS Program Demographic and Health Surveys, ICF International

4. Malaria Vaccine Rainbow Tables. World Health Organization

5. Spatial Data Repository. The DHS Program

6. 1988. Malaria diagnosis: memorandum from a WHO meeting. *Bull World Health Organ* 66:575-94

7. 2008. Vector-Borne Diseases: Understanding the Environmental, Human Health, and Ecological Connections, Workshop Summary, National Academies Press, Washington, DC

8. 2009. Methods for Surveillance of Antimalarial Drug Efficacy, World Health Organization Geneva, Switzerland

9. 2009. *Using Multiplex Identifier (MID) Adaptors for the GS FLX Titanium Chemistry - Basic MID Set* http://454.com/downloads/my454/documentation/technical-bulletins/TCB-09004_UsingMultiplexIdentifierAdaptorsForTheGSFLXTitaniumSeriesChemistry-BasicMIDSet.pdf

10. 2012. Disease Surveillance for Malaria Elimination: An Operational Manual, Global Malaria Programme, World Health Organization, Geneva, Switzerland

11. 2012. Management of Severe Malaria: A Practical Handbook, World Health Organization, Geneva, Switzerland

12. 2013. Democratic Republic of the Congo, President's Malaria Initiative, Centers for Disease Control and Prevention, USAID

13. 2014. Democratic Republic of Congo Malaria Operational Plan FY 2014, President's Malaria Initiative, USAID, CDC, DRC NMCP

14. 2014. A Global Brief on Vector-Borne Diseases. Geneva, Switzerland: World Health Organization

15.   2014. *Japanese encephalitis*.
       http://www.who.int/mediacentre/factsheets/fs386/en/

16.   2014. Malaria Rapid Diagnostic Test Performance: Summary Results of WHO
       Product Testing of Malaria RDTs: Round 1-5 (2008-2013), World Health
       Organization, Geneva, Switzerland

17.   2014. Vector-borne diseases. Geneva, Switzerland: World Health Organization

18.   2014. World Malaria Report 2014, World Health Organization Geneva,
       Switzerland

19.   2015. International Data Base. United Status Census Bureau

20.   2015. *Japanese Encephalitis*.
       http://www.cdc.gov/japaneseencephalitis/symptoms/index.html

21.   2015. Japanese Encephalitis Vaccines: WHO position paper, World Health
       Organization, Geneva, Switzerland

22.   2015. Surveillance and Control of Aedes aegypti and Aedes albopictus in the
       United States. Atlanta, GA: Center for Disease Control and Prevention, National
       Center for Emerging and Zoonotic Infectious Diseases (NCEZID)

23.   Abu-Raddad LJ, Patnaik P, Kublin JG. 2006. Dual infection with HIV and
       malaria fuels the spread of both diseases in sub-Saharan Africa. *Science*
       314:1603-6

24.   Achee NL, Gould F, Perkins TA, Reiner RC, Jr., Morrison AC, et al. 2015. A
       critical assessment of vector control for dengue prevention. *PLoS Negl Trop Dis*
       9:e0003655

25.   Ahmed Ismail H, Ribacke U, Reiling L, Normark J, Egwang T, et al. 2013.
       Acquired antibodies to merozoite antigens in children from Uganda with
       uncomplicated or severe *Plasmodium falciparum* malaria. *Clin Vaccine Immunol*
       20:1170-80

26.   al-Yaman F, Genton B, Reeder JC, Anders RF, Smith T, Alpers MP. 1998.
       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* 91:602-5

27.   Alemu A, Fuehrer HP, Getnet G, Tessema B, Noedl H. 2013. *Plasmodium ovale
       curtisi* and *Plasmodium ovale wallikeri* in North-West Ethiopia. *Malar J* 12:346

28.   Alexander DL, Mital J, Ward GE, Bradley P, Boothroyd JC. 2005. Identification of the moving junction complex of *Toxoplasma gondii*: a collaboration between distinct secretory organelles. *PLoS pathogens* 1:e17

29.   Alves FP, Gil LH, Marrelli MT, Ribolla PE, Camargo EP, Da Silva LH. 2005. Asymptomatic carriers of *Plasmodium spp.* as infection source for malaria vector mosquitoes in the Brazilian Amazon. *J Med Entomol* 42:777-9

30.   Arnott A, Wapling J, Mueller I, Ramsland PA, Siba PM, et al. 2014. Distinct patterns of diversity, population structure and evolution in the AMA1 genes of sympatric *Plasmodium falciparum* and *Plasmodium vivax* populations of Papua New Guinea from an area of similarly high transmission. *Malar J* 13:233

31.   Auburn S, Campino S, Miotto O, Djimde AA, Zongo I, et al. 2012. Characterization of Within-Host *Plasmodium falciparum* Diversity Using Next-Generation Sequence Data. *PLoS One* 7:e32891

32.   Ayala D, Costantini C, Ose K, Kamdem GC, Antonio-Nkondjio C, et al. 2009. Habitat suitability and ecological niche profile of major malaria vectors in Cameroon. *Malar J* 8:307

33.   Bailey JA, Pablo J, Niangaly A, Travassos MA, Ouattara A, et al. 2015. Seroreactivity to a large panel of field-derived *Plasmodium falciparum* apical membrane antigen 1 and merozoite surface protein 1 variants reflects seasonal and lifetime acquired responses to malaria. *Am J Trop Med Hyg* 92:9-12

34.   Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37-48

35.   Banerjee K, Deshmukh PK, Ilkal MA, Dhanda V. 1978. Transmission of Japanese encephalitis virus by *Culex bitaeniorhynchus* Giles. *Indian J Med Res* 67:889-93

36.   Bargieri DY, Andenmatten N, Lagal V, Thiberge S, Whitelaw JA, et al. 2013. Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion. *Nat Commun* 4:2552

37.   Bartoloni A, Zammarchi L. 2012. Clinical aspects of uncomplicated and severe malaria. *Mediterr J Hematol Infect Dis* 4:e2012026

38.   Bauffe F, Desplans J, Fraisier C, Parzy D. 2012. Real-time PCR assay for discrimination of *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* in the Ivory Coast and in the Comoros Islands. *Malar J* 11:307

39.   Bean AG, Baker ML, Stewart CR, Cowled C, Deffrasnes C, et al. 2013. Studying immunity to zoonotic diseases in the natural host - keeping it real. *Nat Rev Immunol* 13:851-61

40.   Beck HP, Felger I, Huber W, Steiger S, Smith T, et al. 1997. Analysis of multiple *Plasmodium falciparum* infections in Tanzanian children during the phase III trial of the malaria vaccine SPf66. *J Infect Dis* 175:921-6

41.   Beck S, Mockenhaupt FP, Bienzle U, Eggelte TA, Thompson WN, Stark K. 2001. Multiplicity of *Plasmodium falciparum* infection in pregnancy. *Am J Trop Med Hyg* 65:631-6

42.   Beier MS, Schwartz IK, Beier JC, Perkins PV, Onyango F, et al. 1988. Identification of malaria species by ELISA in sporozoite and oocyst infected *Anopheles* from western Kenya. *Am J Trop Med Hyg* 39:323-7

43.   Bell D, Wongsrichanalai C, Barnwell JW. 2006. Ensuring quality and access for malaria diagnosis: how can it be achieved? *Nature reviews*. *Microbiology* 4:S7-20

44.   Bendixen M, Msangeni HA, Pedersen BV, Shayo D, Bodker R. 2001. Diversity of *Plasmodium falciparum* populations and complexity of infections in relation to transmission intensity and host age: a study from the Usambara Mountains, Tanzania. *Trans R Soc Trop Med Hyg* 95:143-8

45.   Bereczky S, Liljander A, Rooth I, Faraja L, Granath F, et al. 2007. Multiclonal asymptomatic *Plasmodium falciparum* infections predict a reduced risk of malaria disease in a Tanzanian population. *Microbes Infect* 9:103-10

46.   Betson M, Sousa-Figueiredo JC, Atuhaire A, Arinaitwe M, Adriko M, et al. 2014. Detection of persistent Plasmodium spp. infections in Ugandan children after artemether-lumefantrine treatment. *Parasitology* 16:1-11

47.   Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, et al. 2015. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature* 526:207-11

48.   Bhattacharya S, Chakraborty SK, Chakraborty S, Ghosh KK, Palit A, et al. 1986. Density of *Culex vishnui* and appearance of JE antibody in sentinel chicks and wild birds in relation to Japanese encephalitis cases. *Trop Geogr Med* 38:46-50

49.   Bigaillon C, Fontan E, Cavallo JD, Hernandez E, Spiegel A. 2005. Ineffectiveness of the Binax NOW malaria test for diagnosis of *Plasmodium ovale* malaria. *J Clin Microbiol* 43:1011

50.   Bobanga T, Ayieko W, Zanga M, Umesumbu S, Landela A, et al. 2013. Field efficacy and acceptability of PermaNet(R) 3.0 and OlysetNet(R) in Kinshasa, Democratic Republic of the Congo. *J Vector Borne Dis* 50:206-14

51.   Bousema T, Okell L, Felger I, Drakeley C. 2014. Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nature reviews*. *Microbiology* 12:833-40

52.   Branch OH, Takala S, Kariuki S, Nahlen BL, Kolczak M, et al. 2001. *Plasmodium falciparum* genotypes, low complexity of infection, and resistance to subsequent malaria in participants in the Asembo Bay Cohort Project. *Infect Immun* 69:7783-92

53.   Bruce MC, Macheso A, Kelly-Hope LA, Nkhoma S, McConnachie A, Molyneux ME. 2008. Effect of transmission setting and mixed species infections on clinical measures of malaria in Malawi. *PLoS One* 3:e2775

54.   Bruce MC, Macheso A, McConnachie A, Molyneux ME. 2011. Comparative population structure of *Plasmodium malariae* and *Plasmodium falciparum* under different transmission settings in Malawi. *Malar J* 10:38

55.   Buermans HP, den Dunnen JT. 2014. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta* 1842:1932-41

56.   Buescher EL, Scherer WF, Rosenberg MZ, Gresser I, Hardy JL, Bullock HR. 1959. Ecologic studies of Japanese encephalitis virus in Japan. II. Mosquito infection. *Am J Trop Med Hyg* 8:651-64

57.   Calderaro A, Piccolo G, Gorrini C, Montecchini S, Rossi S, et al. 2012. A New Real-Time PCR for the Detection of *Plasmodium ovale wallikeri*. *PLoS One* 7:e48033

58.   Calderaro A, Piccolo G, Perandin F, Gorrini C, Peruzzi S, et al. 2007. Genetic Polymorphisms Influence *Plasmodium ovale* PCR Detection Accuracy. *J Clin Microbiol* 45:1624-7

59.   Carlton J, Sullivan S, Le Roch G. 2013. *Plasmodium* Genomics and the Art of Sequencing Malaria Parasite Genomes. In *Malaria Parasites: Comparative Genomics, Evolution and Molecular Biology*, ed. J Carlton, S Perkins, K Deitsch. Norfolk, VA: Caister Academic Press

60.   Carrel M, Patel J, Taylor SM, Janko M, Mwandagalirwa MK, et al. 2015. The geography of malaria genetics in the Democratic Republic of Congo: A complex and fragmented landscape. *Soc Sci Med* 133:233-41

61.   Chelimo K, Ofulla AV, Narum DL, Kazura JW, Lanar DE, John CC. 2005. Antibodies to *Plasmodium falciparum* antigens vary by age and antigen in children in a malaria-holoendemic area of Kenya. *Pediatr Infect Dis J* 24:680-4

62.   Coene J. 1993. Malaria in urban and rural Kinshasa: the entomological input. *Med Vet Entomol* 7:127-37

63.   Coene J, Ngimbi NP, Mulumba MP, Wery M. 1989. Ineffectiveness of mosquito coils in Kinshasa, Zaire. *Trans R Soc Trop Med Hyg* 83:568-9

64.   Cohen R, Feghali K, Alemayehu S, Komisar J, Hang J, et al. 2013. Use of qPCR and Genomic Sequencing to Diagnose *Plasmodium ovale wallikeri* Malaria in a Returned Soldier in the Setting of a Negative Rapid Diagnostic Assay. *Am J Trop Med Hyg* 89:501-6

65.   Collins WE, Jeffery GM. 2005. *Plasmodium ovale*: parasite and disease. *Clin Microbiol Rev* 18:570-81

66.   Colwell RK. 2013. EstimateS: Statistical estimation of species richness and shared species from samples.

67.   Conway DJ, Roper C, Oduola AM, Arnot DE, Kremsner PG, et al. 1999. High recombination rate in natural populations of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 96:4506-11

68.   Cortes A, Mellombo M, Masciantonio R, Murphy VJ, Reeder JC, Anders RF. 2005. Allele specificity of naturally acquired antibody responses against *Plasmodium falciparum* apical membrane antigen 1. *Infect Immun* 73:422-30

69.   Cortes A, Mellombo M, Mueller I, Benet A, Reeder JC, Anders RF. 2003. Geographical structure of diversity and differences between symptomatic and asymptomatic infections for *Plasmodium falciparum* vaccine candidate AMA1. *Infect Immun* 71:1416-26

70.   Costa J, Peterson AT. 2012. Ecological niche modeling as a tool for understanding distributions and interactions of vectors, hosts, and etiologic agents of Chagas disease. *Adv Exp Med Biol* 710:59-70

71.   Daniels R, Chang HH, Sene PD, Park DC, Neafsey DE, et al. 2013. Genetic surveillance detects both clonal and epidemic transmission of malaria following enhanced intervention in Senegal. *PLoS One* 8:e60780

72.   Daniels RF, Schaffner SF, Wenger EA, Proctor JL, Chang HH, et al. 2015. Modeling malaria genomics reveals transmission decline and rebound in Senegal. *Proc Natl Acad Sci U S A* 112:7067-72

73.   de Laval F, Oliver M, Rapp C, Pommier de Santi V, Mendibil A, et al. 2010. The challenge of diagnosing *Plasmodium ovale* malaria in travellers: report of six clustered cases in French soldiers returning from West Africa. *Malar J* 9:358

74.   de Laval F, Simon F, Bogreau H, Rapp C, Wurtz N, et al. 2014. Emergence of *Plasmodium ovale* malaria among the French Armed Forces in the Republic of Ivory Coast: 20 years of clinical and biological experience. *Clin Infect Dis* 58:e122-8

75. de Souza AM, de Araujo FC, Fontes CJ, Carvalho LH, de Brito CF, de Sousa TN. 2015. Multiple-clone infections of *Plasmodium vivax*: definition of a panel of markers for molecular epidemiology. *Malar J* 14:330

76. Demas A, Oberstaller J, DeBarry J, Lucchi NW, Srinivasamoorthy G, et al. 2011. Applied genomics: data mining reveals species-specific malaria diagnostic targets more sensitive than 18S rRNA. *J Clin Microbiol* 49:2411-8

77. Diakite NR, Guindo-Coulibaly N, Adja AM, Ouattara M, Coulibaly JT, et al. 2015. Spatial and temporal variation of malaria entomological parameters at the onset of a hydro-agricultural development in central Cote d'Ivoire. *Malar J* 14:340

78. Dicker R. 2006. *Principles of Epidemiology in Public Health Practice*. Atlanta, GA Centers for Disease Control and Prevention

79. Dinko B, Oguike MC, Larbi JA, Bousema T, Sutherland CJ. 2013. Persistent detection of *Plasmodium falciparum*, *P. malariae*, *P. ovale curtisi* and *P. ovale wallikeri* after ACT treatment of asymptomatic Ghanaian school-children. *Int J Parasitol Drugs Drug Resist*. 3:45-50

80. Doderer-Lang C, Atchade PS, Meckert L, Haar E, Perrotey S, et al. 2014. The ears of the African elephant: unexpected high seroprevalence of *Plasmodium ovale* and *Plasmodium malariae* in healthy populations in Western Africa. *Malar J* 13:240

81. Dolmazon V, Matsika-Claquin MD, Manirakiza A, Yapou F, Nambot M, Menard D. 2008. Genetic diversity and genotype multiplicity of *Plasmodium falciparum* infections in symptomatic individuals living in Bangui (CAR). *Acta Trop* 107:37-42

82. Doolan DL, Dobano C, Baird JK. 2009. Acquired immunity to malaria. *Clin Microbiol Rev* 22:13-36

83. Dowdle WR. 1998. The principles of disease elimination and eradication. *Bull World Health Organ* 76 Suppl 2:22-5

84. Dowling MA, Shute GT. 1966. A comparative study of thick and thin blood films in the diagnosis of scanty malaria parasitaemia. *Bulletin of the World Health Organization* 34:249-67

85. Drew DR, Hodder AN, Wilson DW, Foley M, Mueller I, et al. 2012. Defining the antigenic diversity of *Plasmodium falciparum* apical membrane antigen 1 and the requirements for a multi-allele vaccine against malaria. *PLoS One* 7:e51023

86. Druilhe P, Daubersies P, Patarapotikul J, Gentil C, Chene L, et al. 1998. A primary malarial infection is composed of a very wide range of genetically diverse but related parasites. *J Clin Invest* 101:2008-16

87. Duan J, Mu J, Thera MA, Joy D, Kosakovsky Pond SL, et al. 2008. Population structure of the genes encoding the polymorphic *Plasmodium falciparum* apical membrane antigen 1: implications for vaccine design. *Proc Natl Acad Sci U S A* 105:7857-62

88. Dutta S, Dlugosz LS, Drew DR, Ge X, Ababacar D, et al. 2013. Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical membrane antigen-1. *PLoS pathogens* 9:e1003840

89. Dutta S, Lee SY, Batchelor AH, Lanar DE. 2007. Structural basis of antigenic escape of a malaria vaccine candidate. *Proc Natl Acad Sci U S A* 104:12488-93

90. Eisen DP, Marshall VM, Billman-Jacobe H, Coppel RL. 1999. A *Plasmodium falciparum* apical membrane antigen-1 (AMA-1) gene apparently generated by intragenic recombination. *Mol Biochem Parasitol* 100:243-6

91. Enosse S, Dobano C, Quelhas D, Aponte JJ, Lievens M, et al. 2006. RTS,S/AS02A malaria vaccine does not induce parasite CSP T cell epitope selection and reduces multiplicity of infection. *PLoS Clin Trials* 1:e5

92. Erlanger TE, Weiss S, Keiser J, Utzinger J, Wiedenmayer K. 2009. Past, present, and future of Japanese encephalitis. *Emerg Infect Dis* 15:1-7

93. Escalante AA, Grebert HM, Chaiyaroj SC, Magris M, Biswas S, et al. 2001. Polymorphism in the gene encoding the apical membrane antigen-1 (AMA-1) of *Plasmodium falciparum*. X. Asembo Bay Cohort Project. *Mol Biochem Parasitol* 113:279-87

94. Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564-7

95. Farnert A, Arez AP, Babiker HA, Beck HP, Benito A, et al. 2001. Genotyping of *Plasmodium falciparum* infections by PCR: a comparative multicentre study. *Trans R Soc Trop Med Hyg* 95:225-32

96. Farnert A, Snounou G, Rooth I, Bjorkman A. 1997. Daily dynamics of *Plasmodium falciparum* subpopulations in asymptomatic children in a holoendemic area. *Am J Trop Med Hyg* 56:538-47

97. Farnert A, Williams TN, Mwangi TW, Ehlin A, Fegan G, et al. 2009. Transmission-dependent tolerance to multiclonal *Plasmodium falciparum* infection. *J Infect Dis* 200:1166-75

98. Ferreira MU, Liu Q, Kaneko O, Kimura M, Tanabe K, et al. 1998. Allelic diversity at the merozoite surface protein-1 locus of *Plasmodium falciparum* in clinical isolates from the southwestern Brazilian Amazon. *Am J Trop Med Hyg* 59:474-80

99. Flateau C, Le Loup G, Pialoux G. 2011. Consequences of HIV infection on malaria and therapeutic implications: a systematic review. *Lancet Infect Dis* 11:541-56

100. Francesconi P, Fabiani M, Dente MG, Lukwiya M, Okwey R, et al. 2001. HIV, malaria parasites, and acute febrile episodes in Ugandan adults: a case-control study. *AIDS* 15:2445-50

101. Fuehrer HP, Habler VE, Fally MA, Harl J, Starzengruber P, et al. 2012. *Plasmodium ovale* in Bangladesh: genetic diversity and the first known evidence of the sympatric distribution of *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* in southern Asia. *Int J Parasitol* 42:693-9

102. Fuehrer HP, Noedl H. 2014. Recent advances in detection of *Plasmodium ovale*: implications of separation into the two species *Plasmodium ovale wallikeri* and *Plasmodium ovale curtisi*. *J Clin Microbiol* 52:387-91

103. Fuehrer HP, Stadler MT, Buczolich K, Bloeschl I, Noedl H. 2012. Two techniques for simultaneous identification of *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* by use of the small-subunit rRNA gene. *J Clin Microbiol* 50:4100-2

104. Galinsky K, Valim C, Salmier A, de Thoisy B, Musset L, et al. 2015. COIL: a methodology for evaluating malarial complexity of infection using likelihood from single nucleotide polymorphism data. *Malar J* 14:4

105. Gandhi K, Thera MA, Coulibaly D, Traore K, Guindo AB, et al. 2012. Next Generation Sequencing to Detect Variation in the *Plasmodium falciparum* Circumsporozoite Protein. *Am J Trop Med Hyg* 86:775-81

106. Garg S, Alam MT, Das MK, Dev V, Kumar A, et al. 2007. Sequence diversity and natural selection at domain I of the apical membrane antigen 1 among Indian *Plasmodium falciparum* populations. *Malar J* 6:154

107. Getachew S, To S, Trimarsanto H, Thriemer K, Clark TG, et al. 2015. Variation in Complexity of Infection and Transmission Stability between Neighbouring Populations of *Plasmodium vivax* in Southern Ethiopia. *PLoS One* 10:e0140780

108.   Go YY, Balasuriya UB, Lee CK. 2014. Zoonotic encephalitides caused by arboviruses: transmission and epidemiology of alphaviruses and flaviviruses. *Clin Exp Vaccine Res* 3:58-77

109.   Gonzalez C, Wang O, Strutz SE, Gonzalez-Salazar C, Sanchez-Cordero V, Sarkar S. 2010. Climate change and risk of leishmaniasis in north america: predictions from ecological niche models of vector and reservoir species. *PLoS Negl Trop Dis* 4:e585

110.   Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, et al. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38:W695-9

111.   Grobusch MP, Hanscheid T, Zoller T, Jelinek T, Burchard GD. 2002. Rapid immunochromatographic malarial antigen detection unreliable for detecting *Plasmodium malariae* and *Plasmodium ovale*. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 21:818-20

112.   Gubler DJ. 1998. Resurgent vector-borne diseases as a global health problem. *Emerg Infect Dis* 4:442-50

113.   Gubler DJ, Reiter P, Ebi KL, Yap W, Nasci R, Patz JA. 2001. Climate variability and change in the United States: potential impacts on vector- and rodent-borne diseases. *Environ Health Perspect* 109 Suppl 2:223-33

114.   Guitard J, Andersen P, Ermont C, Gnidehou S, Fievet N, et al. 2010. *Plasmodium falciparum* population dynamics in a cohort of pregnant women in Senegal. *Malar J* 9:165

115.   Gurgel-Goncalves R, Galvao C, Costa J, Peterson AT. 2012. Geographic distribution of chagas disease vectors in Brazil based on ecological niche modeling. *J Trop Med* 2012:705326

116.   Hadler JL, Patel D, Nasci RS, Petersen LR, Hughes JM, et al. 2015. Assessment of Arbovirus Surveillance 13 Years after Introduction of West Nile Virus, United States. *Emerg Infect Dis* 21:1159-66

117.   Halldórsson B, Bafna V, Edwards N, Lippert R, Yooseph S, Istrail S. 2004. A Survey of Computational Methods for Determining Haplotypes. In *Computational Methods for SNPs and Haplotype Inference*, 2983:26-47. Heidelberg, Germany: Springer Berlin Heidelberg

118.   Hanna JN, Ritchie SA, Phillips DA, Lee JM, Hills SL, et al. 1999. Japanese encephalitis in north Queensland, Australia, 1998. *Med J Aust* 170:533-6

119. Harris I, Sharrock WW, Bain LM, Gray KA, Bobogare A, et al. 2010. A large proportion of asymptomatic *Plasmodium* infections with low and sub-microscopic parasite densities in the low transmission setting of Temotu Province, Solomon Islands: challenges for malaria diagnostics in an elimination setting. *Malar J* 9:254

120. Hastings IM, D'Alessandro U. 2000. Modelling a predictable disaster: the rise and spread of drug-resistantmalaria. *Parasitol Today* 16:340-7

121. Hay SI, Okiro EA, Gething PW, Patil AP, Tatem AJ, et al. 2010. Estimating the global clinical burden of *Plasmodium falciparum* malaria in 2007. *PLoS medicine* 7:e1000290

122. Hehl AB, Lekutis C, Grigg ME, Bradley PJ, Dubremetz JF, et al. 2000. *Toxoplasma gondii* homologue of plasmodium apical membrane antigen 1 is involved in invasion of host cells. *Infect Immun* 68:7078-86

123. Hills S, Dabbagh A, Jacobson J, Marfin A, Featherstone D, et al. 2009. Evidence and rationale for the World Health Organization recommended standards for Japanese encephalitis surveillance. *BMC infectious diseases* 9:214

124. Hills S, Nett R, Fischer M. 2012. Infectious Diseases Related to Travel: Japanese encephalitis. In *CDC Health Information for International Travel* ed. G Burnette: Oxford University Press

125. Hodder AN, Crewther PE, Matthew ML, Reid GE, Moritz RL, et al. 1996. The disulfide bond structure of *Plasmodium* apical membrane antigen-1. *The Journal of biological chemistry* 271:29446-52

126. Homan T, Maire N, Hiscox A, Di Pasquale A, Kiche I, et al. 2016. Spatially variable risk factors for malaria in a geographically heterogeneous landscape, western Kenya: an explorative study. *Malar J* 15:1

127. Hou Y, Zhang H, Miranda L, Lin S. 2010. Serious overestimation in quantitative PCR by circular (supercoiled) plasmid standard: microalgal pcna as the model gene. *PLoS One* 5:e9545

128. Hsiang MS, Hwang J, Kunene S, Drakeley C, Kandula D, et al. 2012. Surveillance for malaria elimination in Swaziland: a national cross-sectional study using pooled PCR and serology. *PLoS One* 7:e29550

129. Hupalo DN, Bradic M, Carlton JM. 2015. The impact of genomics on population genetics of parasitic diseases. *Curr Opin Microbiol* 23:49-54

130. Ichimori K, King JD, Engels D, Yajima A, Mikhailov A, et al. 2014. Global programme to eliminate lymphatic filariasis: the processes underlying programme success. *PLoS Negl Trop Dis* 8:e3328

131. Iroezindu MO, Agaba EI, Okeke EN, Daniyam CA, Obaseki DO, et al. 2012. Prevalence of malaria parasitaemia in adult HIV-infected patients in Jos, North-central Nigeria. *Nigerian journal of medicine : journal of the National Association of Resident Doctors of Nigeria* 21:209-13

132. Jiang H, Li N, Gopalan V, Zilversmit MM, Varma S, et al. 2011. High recombination rates and hotspots in a *Plasmodium falciparum* genetic cross. *Genome biology* 12:R33

133. Joshi K, Witcombe J. 2003. The impact of participatory plant breeding (PPB) on landrace diversity: A case study for high-altitude rice in Nepal. *Euphytica* 134:117-25

134. Juliano JJ, Porter K, Mwapasa V, Sem R, Rogers WO, et al. 2010. Exposing malaria in-host diversity and estimating population diversity by capture-recapture using massively parallel pyrosequencing. *Proc Natl Acad Sci U S A* 107:20138-43

135. Kalluri S, Gilruth P, Rogers D, Szczur M. 2007. Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review. *PLoS pathogens* 3:1361-71

136. Kamau E, Alemayehu S, Feghali KC, Saunders D, Ockenhouse CF. 2013. Multiplex qPCR for detection and absolute quantification of malaria. *PLoS One* 8:e71539

137. Kamya MR, Gasasira AF, Yeka A, Bakyaita N, Nsobya SL, et al. 2006. Effect of HIV-1 infection on antimalarial treatment outcomes in Uganda: a population-based study. *J Infect Dis* 193:9-15

138. Kang JM, Moon SU, Kim JY, Cho SH, Lin K, et al. 2010. Genetic polymorphism of merozoite surface protein-1 and merozoite surface protein-2 in *Plasmodium falciparum* field isolates from Myanmar. *Malar J* 9:131

139. Kawamoto F, Liu Q, Ferreira MU, Tantular IS. 1999. How prevalent are *Plasmodium ovale* and *P. malariae* in East Asia? *Parasitol Today* 15:422-6

140. Kawamoto F, Miyake H, Kaneko O, Kimura M, Nguyen TD, et al. 1996. Sequence variation in the 18S rRNA gene, a target for PCR-based malaria diagnosis, in *Plasmodium ovale* from southern Vietnam. *J Clin Microbiol* 34:2287-9

141. Keiser J, Maltese MF, Erlanger TE, Bos R, Tanner M, et al. 2005. Effect of irrigated rice agriculture on Japanese encephalitis, including challenges and opportunities for integrated vector management. *Acta Trop* 95:40-57

142. Kelly-Hope L, Ranson H, Hemingway J. 2008. Lessons from the past: managing insecticide resistance in malaria control and eradication programmes. *Lancet Infect Dis* 8:387-9

143. Kim HC, Klein TA, Takhampunya R, Evans BP, Mingmongkolchai S, et al. 2011. Japanese encephalitis virus in culicine mosquitoes (Diptera: Culicidae) collected at Daeseongdong, a village in the demilitarized zone of the Republic of Korea. *J Med Entomol* 48:1250-6

144. Kiszewski A, Mellinger A, Spielman A, Malaney P, Sachs SE, Sachs J. 2004. A global index representing the stability of malaria transmission. *Am J Trop Med Hyg* 70:486-98

145. Kiwanuka GN. 2009. Genetic diversity in *Plasmodium falciparum* merozoite surface protein 1 and 2 coding genes and its implications in malaria epidemiology: a review of published studies from 1997-2007. *J Vector Borne Dis* 46:1-12

146. Kiwuwa MS, Ribacke U, Moll K, Byarugaba J, Lundblom K, et al. 2013. Genetic diversity of *Plasmodium falciparum* infections in mild and severe malaria of children from Kampala, Uganda. *Parasitol Res* 112:1691-700

147. Kocken CH, Narum DL, Massougbodji A, Ayivi B, Dubbeld MA, et al. 2000. Molecular characterisation of *Plasmodium reichenowi* apical membrane antigen-1 (AMA-1), comparison with *P. falciparum* AMA-1, and antibody-mediated inhibition of red cell invasion. *Mol Biochem Parasitol* 109:147-56

148. Kocken CH, Withers-Martinez C, Dubbeld MA, van der Wel A, Hackett F, et al. 2002. High-level expression of the malaria blood-stage vaccine candidate *Plasmodium falciparum* apical membrane antigen 1 and induction of antibodies that inhibit erythrocyte invasion. *Infect Immun* 70:4471-6

149. Konate L, Zwetyenga J, Rogier C, Bischoff E, Fontenille D, et al. 1999. Variation of *Plasmodium falciparum* msp1 block 2 and msp2 allele prevalence and of infection complexity in two neighbouring Senegalese villages with different transmission conditions. *Trans R Soc Trop Med Hyg* 93 Suppl 1:21-8

150. Konishi E, Shoda M, Yamamoto S, Arai S, Tanaka-Taya K, Okabe N. 2006. Natural infection with Japanese encephalitis virus among inhabitants of Japan: a nationwide survey of antibodies against nonstructural 1 protein. *Vaccine* 24:3054-6

151.   Korenromp EL, Williams BG, de Vlas SJ, Gouws E, Gilks CF, et al. 2005. Malaria attributable to the HIV-1 epidemic, sub-Saharan Africa. *Emerg Infect Dis* 11:1410-9

152.   Kwiek JJ, Alker AP, Wenink EC, Chaponda M, Kalilani LV, Meshnick SR. 2007. Estimating true antimalarial efficacy by heteroduplex tracking assay in patients with complex *Plasmodium falciparum* infections. *Antimicrob Agents Chemother* 51:521-7

153.   Laishram DD, Sutton PL, Nanda N, Sharma VL, Sobti RC, et al. 2012. The complexities of malaria disease manifestations with a focus on asymptomatic malaria. *Malar J* 11:29

154.   Langhorne J, Ndungu FM, Sponaas AM, Marsh K. 2008. Immunity to malaria: more questions than answers. *Nat Immunol* 9:725-32

155.   Lara-Silva Fde O, Michalsky EM, Fortes-Dias CL, Fiuza Vde O, Pessanha JE, et al. 2015. Epidemiological aspects of vector, parasite, and domestic reservoir in areas of recent transmission and no reported human cases of visceral leishmaniasis in Brazil. *Acta Trop* 148:128-36

156.   Larson SR, DeGroote JP, Bartholomay LC, Sugumaran R. 2010. Ecological niche modeling of potential West Nile virus vector mosquito species in Iowa. *J Insect Sci* 10:110

157.   Lau YL, Lee WC, Tan LH, Kamarulzaman A, Syed Omar SF, et al. 2013. Acute respiratory distress syndrome and acute renal failure from *Plasmodium ovale* infection with fatal outcome. *Malar J* 12:389

158.   Le Flohic G, Porphyre V, Barbazan P, Gonzalez JP. 2013. Review of climate, landscape, and viral genetics as drivers of the Japanese encephalitis virus ecology. *PLoS Negl Trop Dis* 7:e2208

159.   Leke RF, Bioga JD, Zhou J, Fouda GG, Leke RJ, et al. 2010. Longitudinal studies of *Plasmodium falciparum* malaria in pregnant women living in a rural Cameroonian village with high perennial transmission. *Am J Trop Med Hyg* 83:996-1004

160.   Lemon S, Sparling P, Hamburg M, Relman D, Choffnes E, Mack A. 2008. Vector-Borne Diseases: Understanding the Environmental, Human Health, and Ecological Connections, Board on Global Health, Institute of Medicine Washington, DC

161.   Li M, Xia Z, Yan H. 2014. New type of SSUrDNA sequence was detected from both *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* samples. *Malar J* 13:216

162.   Li MH, Fu SH, Chen WX, Wang HY, Guo YH, et al. 2011. Genotype v Japanese encephalitis virus is emerging. *PLoS Negl Trop Dis* 5:e1231

163.   Li W, Guo Q, Elkan C. 2011. Can we model the probability of presence of species without absence data? *Ecography* 34:1096-105

164.   Li X, Ma SJ, Liu X, Jiang LN, Zhou JH, et al. 2014. Immunogenicity and safety of currently available Japanese encephalitis vaccines: a systematic review. *Hum Vaccin Immunother* 10:3579-93

165.   Li Y, Wang G, Sun D, Meng F, Lin S, et al. 2013. A case of *Plasmodium ovale wallikeri* infection in a Chinese worker returning from West Africa. *Korean J Parasitol* 51:557-62

166.   Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-2

167.   Lim SS, Yang W, Krishnarjuna B, Kannan Sivaraman K, Chandrashekaran IR, et al. 2014. Structure and dynamics of apical membrane antigen 1 from *Plasmodium falciparum* FVO. *Biochemistry* 53:7310-20

168.   Lin CH, Chen YC, Pan TM. 2011. Quantification bias caused by plasmid DNA conformation in quantitative real-time PCR assay. *PLoS One* 6:e29101

169.   Lin JT, Hathaway NJ, Saunders DL, Lon C, Balasubramanian S, et al. 2015. Using amplicon deep sequencing to detect genetic signatures of *Plasmodium vivax* relapse. *J Infect Dis*

170.   Lin JT, Patel JC, Kharabora O, Sattabongkot J, Muth S, et al. 2013. *Plasmodium vivax* isolates from Cambodia and Thailand show high genetic complexity and distinct patterns of *P. vivax* multidrug resistance gene 1 (pvmdr1) polymorphisms. *Am J Trop Med Hyg* 88:1116-23

171.   Lozano-Fuentes S, Elizondo-Quiroga D, Farfan-Ale JA, Lorono-Pino MA, Garcia-Rejon J, et al. 2008. Use of Google Earth to strengthen public health capacity and facilitate management of vector-borne diseases in resource-poor environments. *Bull World Health Organ* 86:718-25

172.   Lysenko AJ, Beljaev AE. 1969. An analysis of the geographical distribution of *Plasmodium ovale*. *Bull World Health Organ* 40:383-94

173.   Mackenzie JS, Johansen CA, Ritchie SA, van den Hurk AF, Hall RA. 2002. Japanese encephalitis as an emerging virus: the emergence and spread of Japanese encephalitis virus in Australasia. *Curr Top Microbiol Immunol* 267:49-73

174.   Mackinnon MJ, Marsh K. 2010. The selection landscape of malaria parasites. *Science* 328:866-71

175.   Mak S, Morshed M, Henry B. 2010. Ecological niche modeling of lyme disease in British Columbia, Canada. *J Med Entomol* 47:99-105

176.   Maltha J, Gillet P, Jacobs J. 2013. Malaria rapid diagnostic tests in endemic settings. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 19:399-407

177.   Manning L, Laman M, Law I, Bona C, Aipit S, et al. 2012. Features and prognosis of severe malaria caused by *Plasmodium falciparum*, *Plasmodium vivax* and mixed *Plasmodium* species in Papua New Guinean children. *PLoS One* 6:e29203

178.   Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, et al. 2012. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* 487:375-9

179.   Marshall VM, Zhang L, Anders RF, Coppel RL. 1996. Diversity of the vaccine candidate AMA-1 of *Plasmodium falciparum*. *Mol Biochem Parasitol* 77:109-13

180.   Martin C, Curtis B, Fraser C, Sharp B. 2002. The use of a GIS-based malaria information system for malaria research and control in South Africa. *Health Place* 8:227-36

181.   Masuoka P, Klein TA, Kim HC, Claborn DM, Achee N, et al. 2010. Modeling the distribution of *Culex tritaeniorhynchus* to predict Japanese encephalitis distribution in the Republic of Korea. *Geospat Health* 5:45-57

182.   Masuoka PM, Burke R, Colaccico M, Razuri H, Hill D, Murrell KD. 2009. Predicted geographic ranges for North American sylvatic *Trichinella* species. *J Parasitol* 95:829-37

183.   Maxwell CA, Chambo W, Mwaimu M, Magogo F, Carneiro IA, Curtis CF. 2003. Variation of malaria transmission and morbidity with altitude in Tanzania and with introduction of alphacypermethrin treated nets. *Malar J* 2:28

184.   Mayor A, Saute F, Aponte JJ, Almeda J, Gomez-Olive FX, et al. 2003. *Plasmodium falciparum* multiple infections in Mozambique, its relation to other malariological indices and to prospective risk of malaria morbidity. *Tropical medicine & international health : TM & IH* 8:3-11

185.   Mellon G, Ficko C, Thellier M, Kendjo E, Aoun O, et al. 2014. Two cases of late *Plasmodium ovale* presentation in military personnel. *J Travel Med* 21:52-4

186. Menner N, Borchert M, Dieckmann S, Ignatius R, Mockenhaupt FP. 2012. Uncommon manifestation of a mixed-species malaria infection: cryptic *falciparum* malaria in a traveler with successfully treated tertian malaria. *J Travel Med* 19:133-5

187. Messina JP. 2011. *Disease Ecology in the Democratic Republic of the Congo: Integration of Spatial Analysis with Population Surveillance*. University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

188. Messina JP, Emch M, Muwonga J, Mwandagalirwa K, Edidi SB, et al. 2010. Spatial and socio-behavioral patterns of HIV prevalence in the Democratic Republic of Congo. *Soc Sci Med* 71:1428-35

189. Messina JP, Mwandagalirwa K, Taylor SM, Emch M, Meshnick SR. 2013. Spatial and social factors drive anemia in Congolese women. *Health Place* 24:54-64

190. Messina JP, Pigott DM, Golding N, Duda KA, Brownstein JS, et al. 2015. The global distribution of Crimean-Congo hemorrhagic fever. *Trans R Soc Trop Med Hyg* 109:503-13

191. Messina JP, Taylor SM, Meshnick SR, Linke AM, Tshefu AK, et al. 2011. Population, behavioural and environmental drivers of malaria prevalence in the Democratic Republic of Congo. *Malar J* 10:161

192. Mideo N, Kennedy DA, Carlton JM, Bailey JA, Juliano JJ, Read AF. 2013. Ahead of the curve: next generation estimators of drug resistance in malaria infections. *Trends Parasitol* 29:321-8

193. Ministère du Plan et Suivi de la Mise en œuvre de la Révolution de la Modernité (MPSMRM) MrdlSPMaI. 2014. Democratic Republic of the Congo Demographic and Health Survey 2013-14 Key Findings, Rockville, Maryland, USA

194. Mital J, Meissner M, Soldati D, Ward GE. 2005. Conditional expression of *Toxoplasma gondii* apical membrane antigen-1 (TgAMA1) demonstrates that TgAMA1 plays a critical role in host cell invasion. *Mol Biol Cell* 16:4341-9

195. Mitchell GH, Thomas AW, Margos G, Dluzewski AR, Bannister LH. 2004. Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells. *Infect Immun* 72:154-8

196. Moore C, McLean R, Mitchell C, Nasci R, Tsai T, et al. 1993. Guidelines for Arbovirus Surveillance Programs in the United States Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, CO

197. Mu J, Myers RA, Jiang H, Liu S, Ricklefs S, et al. 2010. *Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. *Nat Genet* 42:268-71

198. Mueller I, Zimmerman PA, Reeder JC. 2007. *Plasmodium malariae* and *Plasmodium ovale*--the "bashful" malaria parasites. *Trends Parasitol* 23:278-83

199. Muller DA, Charlwood JD, Felger I, Ferreira C, do Rosario V, Smith T. 2001. Prospective risk of morbidity in relation to multiplicity of infection with *Plasmodium falciparum* in Sao Tome. *Acta Trop* 78:155-62

200. Murty US, Rao MS, Arunachalam N. 2010. The effects of climatic factors on the distribution and abundance of Japanese encephalitis vectors in Kurnool district of Andhra Pradesh, India. *J Vector Borne Dis* 47:26-32

201. Nair M, Hinds MG, Coley AM, Hodder AN, Foley M, et al. 2002. Structure of domain III of the blood-stage malaria vaccine candidate, *Plasmodium falciparum* apical membrane antigen 1 (AMA1). *Journal of molecular biology* 322:741-53

202. Narum DL, Thomas AW. 1994. Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of *Plasmodium falciparum* merozoites. *Mol Biochem Parasitol* 67:59-68

203. Nei M. 1987. *Molecular Evolutionary Genetics*. New York, NY USA: Columbia University Press

204. Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U S A* 76:5269-73

205. Nieto P, Malone JB, Bavia ME. 2006. Ecological niche modeling for visceral leishmaniasis in the state of Bahia, Brazil, using genetic algorithm for rule-set prediction and growing degree day-water budget analysis. *Geospat Health* 1:115-26

206. Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, et al. 2013. Population genetic correlates of declining transmission in a human pathogen. *Mol Ecol* 22:273-85

207. Nolder D, Oguike MC, Maxwell-Scott H, Niyazi HA, Smith V, et al. 2013. An observational study of malaria in British travellers: *Plasmodium ovale wallikeri* and *Plasmodium ovale curtisi* differ significantly in the duration of latency. *BMJ open* 3:e002711

208. Noor AM, Kinyoki DK, Mundia CW, Kabaria CW, Mutua JW, et al. 2014. The changing risk of *Plasmodium falciparum* malaria infection in Africa: 2000-10: a spatial and temporal analysis of transmission intensity. *Lancet* 383:1739-47

209. O'Meara WP, Barcus M, Wongsrichanalai C, Muth S, Maguire JD, et al. 2006. Reader technique as a source of variability in determining malaria parasite density by microscopy. *Malar J* 5:118

210. O'Meara WP, Hall BF, McKenzie FE. 2007. Malaria vaccine efficacy: the difficulty of detecting and diagnosing malaria. *Malar J* 6:36

211. Obare P, Ogutu B, Adams M, Odera JS, Lilley K, et al. 2013. Misclassification of *Plasmodium* infections by conventional microscopy and the impact of remedial training on the proficiency of laboratory technicians in species identification. *Malar J* 12:113

212. Ofosu-Okyere A, Mackinnon MJ, Sowa MP, Koram KA, Nkrumah F, et al. 2001. Novel *Plasmodium falciparum* clones and rising clone multiplicities are associated with the increase in malaria morbidity in Ghanaian children during the transition into the high transmission season. *Parasitology* 123:113-23

213. Oguike MC, Betson M, Burke M, Nolder D, Stothard JR, et al. 2011. *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* circulate simultaneously in African communities. *Int J Parasitol* 41:677-83

214. Ohrt C, Obare P, Nanakorn A, Adhiambo C, Awuondo K, et al. 2007. Establishing a malaria diagnostics centre of excellence in Kisumu, Kenya. *Malar J* 6:79

215. Oldham AL, Duncan KE. 2012. Similar gene estimates from circular and linear standards in quantitative PCR analyses using the prokaryotic 16S rRNA gene as a model. *PLoS One* 7:e51931

216. Organization WH. 2011. Countries Using JE Vaccine in National Immunization Schedule, 2010. ed. WMS WHO/IVB database

217. Osier FH, Fegan G, Polley SD, Murungi L, Verra F, et al. 2008. Breadth and magnitude of antibody responses to multiple *Plasmodium falciparum* merozoite antigens are associated with protection from clinical malaria. *Infect Immun* 76:2240-8

218. Osier FH, Weedall GD, Verra F, Murungi L, Tetteh KK, et al. 2010. Allelic diversity and naturally acquired allele-specific antibody responses to *Plasmodium falciparum* apical membrane antigen 1 in Kenya. *Infect Immun* 78:4625-33

219.   Palaniyandi M. 2012. The role of remote sensing and GIS for spatial prediction of vector-borne diseases transmission: a systematic review. *J Vector Borne Dis* 49:197-204

220.   Pandey B, Yamamoto A, Morita K, Kurosawa Y, Rai S, et al. 2003. Serodiagnosis of Japanese encephalitis among Nepalese patients by the particle agglutination assay. *Epidemiol Infect* 131:881-5

221.   Parker DM, Carrara VI, Pukrittayakamee S, McGready R, Nosten FH. 2015. Malaria ecology along the Thailand-Myanmar border. *Malar J* 14:388

222.   PATH W. 2009. Japanese Encephalitis Morbidity, Mortality, and Disability: Reduction and Control by 2015

223.   Patnaik P, Jere CS, Miller WC, Hoffman IF, Wirima J, et al. 2005. Effects of HIV-1 serostatus, HIV-1 RNA concentration, and CD4 cell count on the incidence of malaria infection in a cohort of adults in rural Malawi. *J Infect Dis* 192:984-91

224.   Patterson JE, Bia FJ, Miller K, McPhedran P. 1987. Relapsing malaria infection acquired in Kenya. *Yale J Biol Med* 60:245-53

225.   Peiris JS, Amerasinghe FP, Arunagiri CK, Perera LP, Karunaratne SH, et al. 1993. Japanese encephalitis in Sri Lanka: comparison of vector and virus ecology in different agro-climatic areas. *Trans R Soc Trop Med Hyg* 87:541-8

226.   Peterson AT, Martinez-Campos C, Nakazawa Y, Martinez-Meyer E. 2005. Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Trans R Soc Trop Med Hyg* 99:647-55

227.   Peterson MG, Marshall VM, Smythe JA, Crewther PE, Lew A, et al. 1989. Integral membrane protein located in the apical complex of *Plasmodium falciparum*. *Mol Cell Biol* 9:3151-4

228.   Pfeffer M, Dobler G. 2010. Emergence of zoonotic arboviruses by animal trade and migration. *Parasites & vectors* 3:35

229.   Philip Samuel P, Arunachalam N, Hiriyan J, Tyagi BK. 2008. Host feeding pattern of Japanese encephalitis virus vector mosquitoes (Diptera: Culicidae) from Kuttanadu, Kerala, India. *J Med Entomol* 45:927-32

230.   Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231-59

231.   Phillips SJ, Dudik M, Shapire RE. 2004. A maximum entropy approach to species distribution modeling. *Proceedings of the Twenty-First International Conference on Machine Learning, AMC Press*:655-62

232.   Phuong M, Lau R, Ralevski F, Boggild AK. 2014. Sequence-based optimization of a quantitative real-time PCR assay for detection of *Plasmodium ovale* and *Plasmodium malariae*. *J Clin Microbiol* 52:1068-73

233.   Pickard AL, Wongsrichanalai C, Purfield A, Kamwendo D, Emery K, et al. 2003. Resistance to antimalarials in Southeast Asia and genetic polymorphisms in pfmdr1. *Antimicrob Agents Chemother* 47:2418-23

234.   Plesner AM, Arlien-Soborg P, Herning M. 1998. Neurological complications to vaccination against Japanese encephalitis. *European journal of neurology : the official journal of the European Federation of Neurological Societies* 5:479-85

235.   Polley SD, Chokejindachai W, Conway DJ. 2003. Allele frequency-based analyses robustly map sequence sites under balancing selection in a malaria vaccine candidate antigen. *Genetics* 165:555-61

236.   Polley SD, Mwangi T, Kocken CH, Thomas AW, Dutta S, et al. 2004. Human antibodies to recombinant protein constructs of *Plasmodium falciparum* Apical Membrane Antigen 1 (AMA1) and their associations with protection from malaria. *Vaccine* 23:718-28

237.   Quinde-Calderon L, Rios-Quituizaca P, Solorzano L, Dumonteil E. 2015. Ten years (2004-2014) of Chagas disease surveillance and vector control in Ecuador: successes and challenges. *Tropical medicine & international health : TM & IH*

238.   Remarque EJ, Faber BW, Kocken CH, Thomas AW. 2008. Apical membrane antigen 1: a malaria vaccine candidate in review. *Trends Parasitol* 24:74-84

239.   Reuben R. 1971. Studies on the mosquitoes of North Arcot District, Madras State, India. 5. Breeding places of the *Culex vishnui* group of species. *J Med Entomol* 8:363-6

240.   Richards EE, Masuoka P, Brett-Major D, Smith M, Klein TA, et al. 2010. The relationship between mosquito abundance and rice field density in the Republic of Korea. *Int J Health Geogr* 9:32

241.   Ritchie SA, Phillips D, Broom A, Mackenzie J, Poidinger M, van den Hurk A. 1997. Isolation of Japanese encephalitis virus from *Culex annulirostris* in Australia. *Am J Trop Med Hyg* 56:80-4

242.   Robinson T, Campino SG, Auburn S, Assefa SA, Polley SD, et al. 2011. Drug-resistant genotypes and multi-clonality in *Plasmodium falciparum* analysed by

direct genome sequencing from peripheral blood of malaria patients. *PLoS One* 6:e23204

243.   Rochlin I, Ninivaggi DV, Hutchinson ML, Farajollahi A. 2013. Climate change and range expansion of the Asian tiger mosquito (*Aedes albopictus*) in Northeastern USA: implications for public health practitioners. *PLoS One* 8:e60874

244.   Rodriguez-Perez MA, Fernandez-Santos NA, Orozco-Algarra ME, Rodriguez-Atanacio JA, Dominguez-Vazquez A, et al. 2015. Elimination of Onchocerciasis from Mexico. *PLoS Negl Trop Dis* 9:e0003922

245.   Rojo-Marcos G, Rubio-Munoz JM, Ramirez-Olivencia G, Garcia-Bujalance S, Elcuaz-Romano R, et al. 2014. Comparison of imported *Plasmodium ovale curtisi* and *P. ovale wallikeri* infections among patients in Spain, 2005-2011. *Emerg Infect Dis* 20:409-16

246.   Roper C, Richardson W, Elhassan IM, Giha H, Hviid L, et al. 1998. Seasonal changes in the *Plasmodium falciparum* population in individuals and their relationship to clinical malaria: a longitudinal study in a Sudanese village. *Parasitology* 116 (Pt 6):501-10

247.   Rosenthal J. 2009. Climate change and the geographic distribution of infectious diseases. *Ecohealth* 6:489-95

248.   Roth JM, Korevaar DA, Leeflang MM, Mens PF. 2015. Molecular malaria diagnostics: A systematic review and meta-analysis. *Crit Rev Clin Lab Sci*:1-19

249.   Roucher C, Rogier C, Sokhna C, Tall A, Trape JF. 2014. A 20-year longitudinal study of *Plasmodium ovale* and *Plasmodium malariae* prevalence and morbidity in a West African population. *PLoS One* 9:e87169

250.   Rougemont M, Van Saanen M, Sahli R, Hinrikson HP, Bille J, Jaton K. 2004. Detection of four *Plasmodium* species in blood from humans by 18S rRNA gene subunit-based and species-specific real-time PCR assays. *J Clin Microbiol* 42:5636-43

251.   Rozas J. 2009. DNA sequence polymorphism analysis using DnaSP. *Methods Mol Biol* 537:337-50

252.   Sakamoto H, Thiberge S, Akerman S, Janse CJ, Carvalho TG, Menard R. 2005. Towards systematic identification of *Plasmodium* essential genes by transposon shuttle mutagenesis. *Nucleic Acids Res* 33:e174

253. Scherer WF, Moyer JT, Izumi T, Gresser I, Mc CJ. 1959. Ecologic studies of Japanese encephalitis virus in Japan. VI. Swine infection. *Am J Trop Med Hyg* 8:698-706

254. Schleiermacher D, Rogier C, Spiegel A, Tall A, Trape JF, Mercereau-Puijalon O. 2001. Increased multiplicity of *Plasmodium falciparum* infections and skewed distribution of individual msp1 and msp2 alleles during pregnancy in Ndiop, a Senegalese village with seasonal, mesoendemic malaria. *Am J Trop Med Hyg* 64:303-9

255. Schuller E, Klingler A, Dubischar-Kastner K, Dewasthaly S, Muller Z. 2011. Safety profile of the Vero cell-derived Japanese encephalitis virus (JEV) vaccine IXIARO((R)). *Vaccine* 29:8669-76

256. Senn H, Alattas N, Boggild AK, Morris SK. 2014. Mixed-species *Plasmodium falciparum* and *Plasmodium ovale* malaria in a paediatric returned traveller. *Malar J* 13:78

257. Service M, Townson H. 2002. The *Anopheles* Vector In *Essential Malariology*. London, United Kingdom: Hodder Arnold

258. Shirayama Y, Phompida S, Shibuya K. 2009. Geographic information system (GIS) maps and malaria control monitoring: intervention coverage and health outcome in distal villages of Khammouane province, Laos. *Malar J* 8:217

259. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* 7:539

260. Silvie O, Franetich JF, Charrin S, Mueller MS, Siau A, et al. 2004. A role for apical membrane antigen 1 during invasion of hepatocytes by *Plasmodium falciparum* sporozoites. *The Journal of biological chemistry* 279:9490-6

261. Sinden R, Gilles H. 2002. The Malaria Parasites. In *Essential Malariology* London, United Kingdom: Hodder Arnold

262. Singh R, Jain V, Singh PP, Bharti PK, Thomas T, et al. 2013. First report of detection and molecular confirmation of *Plasmodium ovale* from severe malaria cases in central India. *Tropical medicine & international health : TM & IH* 18:1416-20

263. Sisya TJ, Kamn'gona RM, Vareta JA, Fulakeza JM, Mukaka MF, et al. 2014. Subtle changes in *Plasmodium falciparum* infection complexity following enhanced intervention in Malawi. *Acta Trop* 142c:108-14

264. Smith A, Denholm J, Shortt J, Spelman D. 2011. *Plasmodium* species co-infection as a cause of treatment failure. *Travel Medicine and Infectious Disease* 9:306-9

265. Smith T, Beck HP, Kitua A, Mwankusye S, Felger I, et al. 1999. Age dependence of the multiplicity of *Plasmodium falciparum* infections and of other malariological indices in an area of high endemicity. *Trans R Soc Trop Med Hyg* 93 Suppl 1:15-20

266. Snounou G. 2002. Genotyping of *Plasmodium* spp. Nested PCR. *Methods in molecular medicine* 72:103-16

267. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, et al. 1993. High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. *Mol Biochem Parasitol* 61:315-20

268. Snow R, Gilles H. 2002. The Epidemiology of Malaria. In *Essential Malariology* London, United Kingdom: Hodder Arnold

269. Snow RW, Trape JF, Marsh K. 2001. The past, present and future of childhood malaria mortality in Africa. *Trends Parasitol* 17:593-7

270. Spring MD, Cummings JF, Ockenhouse CF, Dutta S, Reidler R, et al. 2009. Phase 1/2a study of the malaria vaccine candidate apical membrane antigen-1 (AMA-1) administered in adjuvant system AS01B or AS02A. *PLoS One* 4:e5254

271. Srinivasan P, Beatty WL, Diouf A, Herrera R, Ambroggio X, et al. 2011. Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proc Natl Acad Sci U S A* 108:13275-80

272. Steketee RW, Nahlen BL, Parise ME, Menendez C. 2001. The burden of malaria in pregnancy in malaria-endemic areas. *Am J Trop Med Hyg* 64:28-35

273. Stevens JWW. 1922. A New Malaria Parasite of Man. *Ann Trop Med Parasitol* 16:383-8

274. Strydom KA, Ismail F, Frean J. 2014. *Plasmodium ovale*: a case of not-so-benign tertian malaria. *Malar J* 13:85

275. Suaya JA, Shepard DS, Siqueira JB, Martelli CT, Lum LC, et al. 2009. Cost of dengue cases in eight countries in the Americas and Asia: a prospective study. *Am J Trop Med Hyg* 80:846-55

276. Sutherland CJ, Tanomsing N, Nolder D, Oguike M, Jennison C, et al. 2010. Two nonrecombining sympatric forms of the human malaria parasite *Plasmodium ovale* occur globally. *J Infect Dis* 201:1544-50

277. Sutherst RW. 2004. Global change and human vulnerability to vector-borne diseases. *Clin Microbiol Rev* 17:136-73

278. Swets JA. 1988. Measuring the accuracy of diagnostic systems. *Science* 240:1285-93

279. Tachibana M, Tsuboi T, Kaneko O, Khuntirat B, Torii M. 2002. Two types of *Plasmodium ovale* defined by SSU rRNA have distinct sequences for ookinete surface proteins. *Mol Biochem Parasitol* 122:223-6

280. Takahashi M. 1976. The effects of environmental and physiological conditions of *Culex tritaeniorhynchus* on the pattern of transmission of Japanese encephalitis virus. *J Med Entomol* 13:275-84

281. Takala SL, Coulibaly D, Thera MA, Batchelor AH, Cummings MP, et al. 2009. Extreme polymorphism in a vaccine antigen and risk of clinical malaria: implications for vaccine development. *Sci Transl Med* 1:2ra5

282. Takala SL, Plowe CV. 2009. Genetic diversity and malaria vaccine design, testing and efficacy: preventing and overcoming 'vaccine resistant malaria'. *Parasite Immunol* 31:560-73

283. Takhampunya R, Kim HC, Tippayachai B, Kengluecha A, Klein TA, et al. 2011. Emergence of Japanese encephalitis virus genotype V in the Republic of Korea. *Virol J* 8:449

284. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725-9

285. Tangpukdee N, Duangdee C, Wilairatana P, Krudsood S. 2009. Malaria diagnosis: a brief review. *Korean J Parasitol* 47:93-102

286. Tanomsing N, Imwong M, Sutherland CJ, Dolecek C, Hien TT, et al. 2013. Genetic marker suitable for identification and genotyping of *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri*. *J Clin Microbiol* 51:4213-6

287. Tatfeng YM, Ihongbe JC, Okodua M, Oviasogie E, Isibor J, et al. 2007. CD4 count, viral load and parasite density of HIV positive individuals undergoing malaria treatment with dihydroartemisinin in Benin City, Edo state, Nigeria. *J Vector Borne Dis* 44:111-5

288. Taylor SM, Antonia AL, Parobek CM, Juliano JJ, Janko M, et al. 2013. *Plasmodium falciparum* sulfadoxine resistance is geographically and genetically clustered within the DR Congo. *Sci Rep* 3:1165

163

289. Taylor SM, Juliano JJ, Trottman PA, Griffin JB, Landis SH, et al. 2010. High-throughput pooling and real-time PCR-based strategy for malaria detection. *J Clin Microbiol* 48:512-9

290. Taylor SM, Mayor A, Mombo-Ngoma G, Kenguele HM, Ouedraogo S, et al. 2014. A quality control program within a clinical trial Consortium for PCR protocols to detect *Plasmodium* species. *J Clin Microbiol* 52:2144-9

291. Taylor SM, Messina JP, Hand CC, Juliano JJ, Muwonga J, et al. 2011. Molecular malaria epidemiology: mapping and burden estimates for the Democratic Republic of the Congo, 2007. *PLoS One* 6:e16420

292. Taylor SM, van Eijk AM, Hand CC, Mwandagalirwa K, Messina JP, et al. 2011. Quantification of the burden and consequences of pregnancy-associated malaria in the Democratic Republic of the Congo. *J Infect Dis* 204:1762-71

293. Tchuinkam T, Nyih-Kong B, Fopa F, Simard F, Antonio-Nkondjio C, et al. 2015. Distribution of *Plasmodium falciparum* gametocytes and malaria-attributable fraction of fever episodes along an altitudinal transect in Western Cameroon. *Malar J* 14:96

294. Terheggen U, Drew DR, Hodder AN, Cross NJ, Mugyenyi CK, et al. 2014. Limited antigenic diversity of *Plasmodium falciparum* apical membrane antigen 1 supports the development of effective multi-allele vaccines. *BMC Med* 12:183

295. Thera MA, Doumbo OK, Coulibaly D, Laurens MB, Ouattara A, et al. 2011. A field trial to assess a blood-stage malaria vaccine. *N Engl J Med* 365:1004-13

296. Thera MA, Plowe CV. 2012. Vaccines for malaria: how close are we? *Annu Rev Med* 63:345-57

297. Thomas AW, Deans JA, Mitchell GH, Alderson T, Cohen S. 1984. The Fab fragments of monoclonal IgG to a merozoite surface antigen inhibit *Plasmodium knowlesi* invasion of erythrocytes. *Molecular & Biochemical Parasitology* 13:187-99

298. Thomas AW, Narum D, Waters AP, Trape JF, Rogier C, et al. 1994. Aspects of immunity for the AMA-1 family of molecules in humans and non-human primates malarias. *Memorias do Instituto Oswaldo Cruz* 89 Suppl 2:67-70

299. Ting SH, Tan HC, Wong WK, Ng ML, Chan SH, Ooi EE. 2004. Seroepidemiology of neutralizing antibodies to Japanese encephalitis virus in Singapore: continued transmission despite abolishment of pig farming? *Acta Trop* 92:187-91

300. Tobian AA, Mehlotra RK, Malhotra I, Wamachi A, Mungai P, et al. 2000. Frequent umbilical cord-blood and maternal-blood infections with *Plasmodium falciparum*, *P. malariae*, and *P. ovale* in Kenya. *J Infect Dis* 182:558-63

301. Tran TM, Ongoiba A, Coursen J, Crosnier C, Diouf A, et al. 2014. Naturally acquired antibodies specific for *Plasmodium falciparum* reticulocyte-binding protein homologue 5 inhibit parasite growth and predict protection from malaria. *J Infect Dis* 209:789-98

302. Triglia T, Healer J, Caruana SR, Hodder AN, Anders RF, et al. 2000. Apical membrane antigen 1 plays a central role in erythrocyte invasion by *Plasmodium* species. *Mol Microbiol* 38:706-18

303. Tsai TF. 2000. New initiatives for the control of Japanese encephalitis by vaccination: minutes of a WHO/CVI meeting, Bangkok, Thailand, 13-15 October 1998. *Vaccine* 18 Suppl 2:1-25

304. Udhayakumar V, Kariuki S, Kolczack M, Girma M, Roberts JM, et al. 2001. Longitudinal study of natural immune responses to the *Plasmodium falciparum* apical membrane antigen (AMA-1) in a holoendemic region of malaria in western Kenya: Asembo Bay Cohort Project VIII. *Am J Trop Med Hyg* 65:100-7

305. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, et al. 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Res* 40:e115

306. Vafa M, Troye-Blomberg M, Anchang J, Garcia A, Migot-Nabias F. 2008. Multiplicity of *Plasmodium falciparum* infection in asymptomatic children in Senegal: relation to transmission, age and erythrocyte variants. *Malar J* 7:17

307. van den Hurk AF, Smith CS, Field HE, Smith IL, Northill JA, et al. 2009. Transmission of Japanese Encephalitis virus from the black flying fox, *Pteropus alecto*, to *Culex annulirostris* mosquitoes, despite the absence of detectable viremia. *Am J Trop Med Hyg* 81:457-62

308. Van geertruyden JP, D'Alessandro U. 2007. Malaria and HIV: a silent alliance. *Trends Parasitol* 23:465-7

309. Van Geertruyden JP, Mulenga M, Mwananyanda L, Chalwe V, Moerman F, et al. 2006. HIV-1 immune suppression and antimalarial treatment outcome in Zambian adults with uncomplicated malaria. *J Infect Dis* 194:917-25

310. Volkman SK, Neafsey DE, Schaffner SF, Park DJ, Wirth DF. 2012. Harnessing genomics and genome biology to understand malaria biology. *Nat Rev Genet* 13:315-28

311. Wang H, Liang G. 2015. Epidemiology of Japanese encephalitis: past, present, and future prospects. *Ther Clin Risk Manag* 11:435-48

312. Wang JL, Pan XL, Zhang HL, Fu SH, Wang HY, et al. 2009. Japanese encephalitis viruses from bats in Yunnan, China. *Emerg Infect Dis* 15:939-42

313. Warren DL, Seifert SN. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl* 21:335-42

314. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189-91

315. Weaver SC, Reisen WK. 2010. Present and future arboviral threats. *Antiviral Res* 85:328-45

316. Whitworth J, Morgan D, Quigley M, Smith A, Mayanja B, et al. 2000. Effect of HIV-1 and increasing immunosuppression on malaria parasitaemia and clinical episodes in adults in rural Uganda: a cohort study. *Lancet* 356:1051-6

317. Wilson PE, Kazadi W, Kamwendo DD, Mwapasa V, Purfield A, Meshnick SR. 2005. Prevalence of pfcrt mutations in Congolese and Malawian P*lasmodium falciparum* isolates as determined by a new Taqman assay. *Acta Trop* 93:97-106

318. Win TT, Jalloh A, Tantular IS, Tsuboi T, Ferreira MU, et al. 2004. Molecular analysis of *Plasmodium ovale* variants. *Emerg Infect Dis* 10:1235-40

319. Wirth DF, Rogers WO, Barker R, Dourado H, Suesbang L, Albuquerque B. 1989. Leishmaniasis and malaria: DNA probes for diagnosis and epidemiologic analysis. *Ann N Y Acad Sci* 569:183-92

320. Wongsrichanalai C, Barcus MJ, Muth S, Sutamihardja A, Wernsdorfer WH. 2007. A review of malaria diagnostic tools: microscopy and rapid diagnostic test (RDT). *Am J Trop Med Hyg* 77:119-27

321. Yap A, Azevedo MF, Gilson PR, Weiss GE, O'Neill MT, et al. 2014. Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cell Microbiol* 16:642-56

322. Zimmerman PA, Mehlotra RK, Kasehagen LJ, Kazura JW. 2004. Why do we need to know more about mixed *Plasmodium* species infections in humans? *Trends Parasitol* 20:440-7