DETECTION OF SPECIES-SPECIFIC *PLASMODIUM* INFECTION USING UNMAPPED READS FROM HUMAN WHOLE GENOME SEQUENCES

by

JASMINE MARIE OLVANY

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Genetics and Genome Sciences

CASE WESTERN RESERVE UNIVERSITY

May 2023

CASE WESTERN RESERVE UNIVERSITY

SCHOOL OF GRADUATE STUDIES

We hereby approve the dissertation of

Jasmine Marie Olvany

Candidate for the degree of Doctor of Philosophy*

Committee Chair

Thomas LaFramboise, Ph.D.

Thesis Advisor

Scott M. Williams, Ph.D.

Thesis Advisor

Peter A. Zimmerman, Ph.D.

Committee Member

Dana C. Crawford, Ph.D.

Committee Member

Arlene E. Dent, MD, Ph.D.

Date of Defense

February 13th, 2023

*We also certify that written approval has been obtained for any proprietary material contained therein

DEDICATION

I dedicate this dissertation to my two late grandmothers, Sharon Marie Kosmer and Carol Sue Olvany, who passed during my doctoral training. They spent their lives showering me with love and support, and with each day that passes I miss them more.

Table of Contents

List of Tables

CHAPTER II

CHAPTER III

List of Figures

CHAPTER I

CHAPTER II

CHAPTER III

List of Abbreviations

ACTs: artemisinin based combination therapies

AFRO: Africa Regional Office (branch of the WHO)

API: apicoplast

BDC: BioData Catalyst

BMI: body mass index

COVID-19: coronavirus disease of 2019

CQ: chloroquine

***CRT*:** chloroquine resistance transport

CWL: common workflow language

DDT: dichloro-diphenyl-trichloroethane

***DHFR*:** dihydrofolate reductase

***DHPS*:** dihydropteroate synthase

DNA: deoxyribose nucleic acid

FN: false negatives

FP: false positives

GMEP: Global Malaria Elimination Program

GTS: Global Technical Strategy

HbC: hemoglobin C

HbD: hemoglobin D

HbE: hemoglobin E

HbO-Arab: hemoglobin O

HbS: hemoglobin S

HDL-C: high density lipoprotein cholesterol

HRP2: histidine-rich protein 2

INTs: insecticide treated nets

IPTp: intermittent chemoprevention for pregnant mothers

IRS: indoor residual spraying

***K13*:** Kelch-13

kb: kilo-bases

LCBs: local collinear blocks

LOD: limit of detection

MalariaGEN: The Malaria Genomic Epidemiology Network

MAP: Malaria Atlas Project

***MDR1*:** multidrug resistant 1

MOI: multiplicity of infection

MSP1: merozoite surface protein 1

MSP2: merozoite surface protein 2

mtDNA: mitochondrial DNA

NAATs: nucleic acid amplification technologies

NGOs: non-governmental organizations

NPV: negative predictive value

PCR: polymerase chain reaction

PF: *Plasmodium falciparum*

PK: *Plasmodium knowlesi*

PM: *Plasmodium malariae*

PO: *Plasmodim ovale*

PPV: positive predictive value

PV: *Plasmodium vivax*

RBCs: red blood cells

RDTs: rapid diagnostic test

rRNA: ribosomal ribonucleic acid

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

SNPs: single nucleotide polymorphisms

SP: sulfadoxine-pyrimethamine

TN: true negatives

TOPMed: Trans-Omics for Precision Medicine

TP: true positives

UMRs: unmapped reads

WES: whole exome sequencing

WGS: whole genome sequencing

WHO: World Health Organization

Acknowledgements

There are many people who, without them, this dissertation would not have been possible. First and foremost, I would like to thank my support system: To my parents Chris and Kevin, I will never be able to thank you enough for, your love, everything you've done and provided for me, and instilling in me the strength to weather all storms. To my brother Kevin, for bringing levity and perspective to all aspects of my life. To my significant other Blake Martin, for being my constant companion in this journey and for your willingness to assist me in all struggles. To other friends and family, especially Pap, Peach, Terry, Jenna, Kelly, Danielle, Dan, Brianne, and Tanya, for all the ways you enrich and provide foundation to my life. And finally, to my bunnies, Nyuszi, Nellie, and Nibbles, for providing innumerous moments of joy in every single day.

As for my academic progress, I want to thank my PhD advisors, Scott Williams and Peter Zimmerman, for their extensive knowledge and input that helped developed this project to its fullest potential. Thanks to my committee members, Tom, Arlene, and Dana, for their support, encouragement, and unique perspectives on this research. And to all my lab members and people of the CWRU community who provided inspiration, motivation, and friendship throughout this entire process.

Detection of Species-Specific *Plasmodium* Infection Using Unmapped Reads From Human Whole Genome Sequences

Abstract

by

JASMINE MARIE OLVANY

Whole genome sequencing (WGS) is an ever expanding tool in the field of genetics, and is widely used to characterize human genetic variation. There are multiple large-scale sequencing studies being conducted today worldwide, like All of Us, Three Million African Genomes, and GenomeAsia 100k. The addition of these diverse datasets alone can be transformative to our understanding of genetics, but the increase in the diversity of populations sampled also has the potential to reveal additional and novel information relevant to health and disease. Specifically, whole genome sequence (WGS) analyses of DNA from human whole blood may be able to capture genetic variation in other species that can affect both individual and public health.

The research detailed in this dissertation aims to illustrate the utility of human WGS data for infectious disease, by showing that the malaria causing parasite *Plasmodium* can be sensitively detected from unmapped reads (UMRs) from WGS data. Malaria has a significant global health burden, and elimination of the

disease has been a goal since the 1950s. Recently, there have been roadblocks in the progress of malaria elimination that can only be resolved through additional research efforts. Development of this detection methodology could be the tool required to better define the parasite population, identify problematic populations, and solve the roadblocks limiting elimination success.

CHAPTER I: Malaria epidemiology and elimination in Africa: the past, present, and potential futures

1.1 The parasite, host, and disease

Malaria, literally translated as “bad” or “evil air,” is a disease that predates modern humans and is thought to have evolved in Ethiopia[1; 2]. It has traveled and colonized the world alongside its human host[1]. Prior to the discovery of the disease agent, many things were attributed to being the cause of these periodic fevers and enlarged spleens. Naturally, before science provided an evidence-based explanation, many attributed the disease to the supernatural. A common myth was to assign a demon to each of the symptoms now associated with malaria[1]. For example, in China, the headache, shudders, and fever were described as three demons with weapons of a hammer, a pail of water, and a stove that afflicted the sick. Another interesting myth assigned to this disease was that in the highly affected city of Rome, people got the sickness from pestilent air exhaled from a dragon living in a cave under the city[1]. The idea that bad air conveyed the disease was found in observational hypotheses as well, as it was thought that breathing in or consuming marshy air caused a complex chemical or small invisible animal to enter the body and cause disease[1; 3]. While misguided, both of these ideas are rooted in reality, as the marsh likely played a role in the transmission of the disease as a breeding ground for mosquitos. The earliest theory that came close to the truth was identified in China around 500 BC, which connected malaria with the bite of an insect[1; 2].

In reality, the disease-causing parasite is a single cellular protozoan of the *Plasmodium* genus, which infects a whole range of different vertebrae hosts[4]. The parasite's transmission depends on the passage through a viable mosquito

vector from the genus *Anopheles*[5]. There are five species currently known to infect humans, *Plasmodium falciparum* (PF), *Plasmodium vivax* (PV), *Plasmodium malariae* (PM), *Plasmodium ovale* (PO), and *Plasmodium knowlesi* (PK)[5]. The parasite lifecycles vary slightly between species but generally is as follows: an infected mosquito takes a blood meal out of a host and deposits an infective sporozoite. This sporozoite then migrates to the liver and differentiates into a merozoite that can invade human red blood cells (RBCs). The merozoite can reproduce asexually through immature trophozoite to mature trophozoite to schizonts then back to merozoite to infect more cells, or it can differentiate into male/female gametocytes that a mosquito will take up to develop into the infectious form again[4]. The only deviation from this cycle occurs in PV and PO, which can form dormant hypnozoites in the liver cells to reactivate at a later date[6]. Generally understanding the life cycle is vital because detecting the parasite in a human infection occurs mainly during the asexual blood stage when the parasite circulates in peripheral blood and can be detected through conventional methods[6]. Currently, three different types of diagnostic methods are regularly employed either in the clinic or research field: light microscopy, rapid diagnostic tests (RDTs), or nucleic acid amplification technologies (NAATs)[7]. Each of these methods has different levels of sensitivity and specificity in detecting the parasite or differentiating the species of infection, which will be detailed later in section 1.5[8].

Determining the species of infection is extremely important when dealing with malaria because not only does each species have different clinical implications,

but the treatment protocol differs based on if the parasite can or cannot form hypnozoites[6]. There are many ways to classify malaria clinically. For example, the disease can be described by the presence or lack of symptoms (symptomatic or asymptomatic), the density of the parasite in the blood (normal or submicroscopic), or the severity of the disease (uncomplicated or severe), each of which has different implications about the disease[8-10]. Clinically, the most important descriptor is the severity because malaria can cause any of the following complications: cerebral malaria, severe malarial anemia, renal failure, or respiratory distress[10; 11]. PF is the parasite most frequently associated with severe malaria outcomes, causing 95% of all malaria deaths, and children under five are the most affected demographic for malaria mortalities worldwide[9; 12]. Multiple publications theorize that PF parasites cause the most severe disease and mortality because of a species' unique ability to cause RBCs to stick to capillary walls, called sequestration[9; 12; 13]. There are known cases of severe malaria from other species, increasingly PV, but these are less common[12; 14].

Additionally, severe malaria although considered dangerous, only leads to death in 0.2% of all clinical malaria cases[10; 12]. The treatment recommendation for all levels of malaria is intervention as quickly as possible with artemisinin-based combination therapies (ACTs) for parasites with no dormant stage and with the addition of primaquine with dormant stage parasites[6; 11]. Treatment recommendations have changed throughout the years in response to parasites gaining resistance to interventions, which will be detailed in later sections[15-17].

Understanding that the parasite has evolved in response to selective pressure via elimination interventions is extremely important. However, it is equally important to acknowledge that the persistence of this disease throughout human history has stimulated evolutionary changes in humans as well. The human genome has evolved several genes to combat this parasite's efficacy and survive through childhood, such as *HBA, HBB*, *G6PD*, and Duffy antigen receptor for chemokines (DARC) variants[18; 19]. Many genes that protect against malaria affect the RBC, which was noted above as a major part of the parasitic life cycle.

HBB is one of the most well-known loci connected to protection against malaria, and single nucleotide polymorphisms (SNPs) in this gene cause structural modification of β-globin. There are currently five known variants of *HBB* that have been associated with protection against malaria, hemoglobin S (HbS | rs334), hemoglobin E (HbE | rs33950507), hemoglobin C (HbC | rs33930165), hemoglobin D (HbD | c.364G>C), and hemoglobin O (HbO-Arab | c.364G>A)[20; 21]. The Malaria Genomic Epidemiology Network (MalariaGEN) found that the HbS variant provided greater than 80% protection against severe malaria for heterozygotes in Africa. In contrast, HbC reduces the risk of severe malaria by 29% per gene copy[19]. The three other hemoglobin variants have no reported prevalence in sub-Saharan Africa or are not connected to resistance to malaria[19; 21-23].

Another hemoglobin-relevant gene, *HBA,* has multiple variants that cause abnormal or absent alpha-globin chains resulting in the disease α-thalassemia, which is highly prevalent in malaria endemic regions[24]. For example, in one study

in Africa, *HBA* had a specific 3.7kb deletion variant that causes the disease $\alpha^{-3.7}$-thalassemia, which was found to reduce risk by 40% in homozygotes[19; 21; 25]. While these hemoglobin variants evolved for protection against malaria, inheriting more than one copy of these alleles can lead to hemoglobin disorders, sickle cell trait (HbS), and a type of α-thalassemia[18; 20; 25].

An additional locus of importance is the Duffy blood group antigens (codominant alleles FYA or FYB) encoded by *DARC* (or *ACKR1*). This gene produces two protein isoforms to create a glycoprotein on RBCs[26]. There are four major phenotypes of this locus: Fy(a+b+), Fy(a+b-), Fy(a-b+), and Fy(a-b-)[26]. This locus is important to malaria because the receptor is directly involved in invasion of the cell by PV[26]. An individual with the phenotype (Fya-b-) is considered Duffy–negative because it causes the absence of the glycoprotein on the membrane wall, making individuals who carry this phenotypr highly resistant to PV infection[26; 27]. Many of these variants mentioned above were originally annotated in African populations and have affected the distribution and burden of each parasite species worldwide.

According to an estimate published by the World Health Organization (WHO), in 2021, nearly half of the world is still at risk for malaria[28]. The most recent estimates of the burden of malaria is 247 million cases and 619 thousand deaths in 2021, with 95% of the cases coming from the 45 countries of the WHO Africa region[29]. These numbers highlight that this disease is still a monumental problem for the global health community, and strides need to be taken to reduce the burden, especially in Africa.

1.2 Malaria Epidemiology across sub-Saharan Africa

Each *Plasmodium* species has a different geographical distribution based on a combination of the *Anopheles* mosquito species available and the above-mentioned human genetic variation. As this thesis is focused on sub-Saharan Africa, there are only four species of consequence to this research (PF, PV, PM, and PO). PK is excluded because it is currently limited to Southeast Asian countries, based on its need for a macaque host to coexist with the human population. Thus, it will not be mentioned in any following sections[30]. Other distribution patterns will affect the frequency of each species in our data, such as the Fya-b- phenotype, which dramatically reduces the prevalence of PV infections in sub-Saharan Africa[27; 31; 32]. Only recently has the field accepted that Duffy-negative individuals can be infected with PV parasites; thus, it is crucial to understand the general landscape of infection across the years in sub-Saharan Africa to contextualize the following research[33; 34].

PF is the dominant species of infection (99.7%) and mortality in sub-Saharan Africa, and it is the species that all of the estimates of infection prevalence are based upon[35]. Non-PF infections have significantly fewer reports and coverage , but there are many indications that PM, PO, and PV exist across sub-Saharan Africa[36]. There have even been indications of surges of non-falciparum species becoming increasingly more prevalent as elimination efforts become more successful in treating PF[35-39]. The lack of focus on the other species is not necessarily surprising because many country-wide reports on infection are diagnosed by RDTs designed to only detect PF[36]. Research conducted country

by country is the only way to unveil the epidemiology of the other species because a molecular-based probe is the most reliable way to identify non-*falciparum* species, or a highly trained microscopist that can differentiate them[40]. Regardless of these challenges, and while historically the epidemiology surrounding malaria in sub-Saharan Africa has focused on PF, it is essential to evaluate all species moving forward to understand the total burden of malaria[38; 40].

The general coverage of epidemiology in Africa pre-2000s is often country-specific and sparse, but decent evidence that with the discovery and use of chloroquine as a treatment, there was a decline in the parasite, followed by an aggressive resurgence of strains with resistance[15; 41; 42]. A publication by Thomas C. Nchinda cites that the parasite mortality was 1.5 to 2.7 million per year in September of 1998, with 90% of that burden being on Africa[41]. In the 1990s, as discussed below, elimination efforts were finally being focused on sub-Saharan Africa, and thus more research and better coverage of *Plasmodium* epidemiology started with these efforts. Since this refocus of elimination efforts to Africa, there have been measurable reductions in the burden of malaria. For example, between 2000 and 2015 there was a 18% decrease in cases globally, a 37% decrease in malaria incidence, and a 48% decrease in mortality[43]. While these numbers technically cover malaria globally, because Africa bore 88% and 90% of cases and mortality, respectively, some reductions would have to be achieved in Africa to be reflected in these mentioned global trends[44].

The era of malaria elimination the field is currently in technically started in 1997, with the known elimination techniques being partnered with research to help provide evidence-based feedback to efforts[45; 46]. However, the most recent benchmark year is 2015, which was the year that the WHO and many financers recommitted to elimination with the Global Technical Strategy[47]. Despite the renewed efforts, many countries in sub-Saharan Africa have plateaued in the progress toward elimination, with only a reported 2% decrease in incidence since 2015[48]. Similarly, there has been only a 2.9% decrease in deaths in populations at risk since 2015, compared to the original 2000 baseline[28]. Collectively, these data suggest there has been a slowdown in the alteration of *Plasmodium* prevalence across Africa, and it appears as if there are roadblocks, detailed later, causing a marked reduction in elimination success worldwide.

1.3 Elimination efforts through the years

Malaria elimination has been a goal science has actively been working towards since the 1950s, with one of the major pushes being the WHO Global Malaria Eradication Program (GMEP est. 1955)[49]. This initiative managed to facilitate the elimination of malaria in multiple countries (n=15) through mass spraying of dichloro-diphenyl-trichloroethane (DDT) and administering of chloroquine (CQ)[49; 50]. However, GMEP was abandoned by 1969 due to rising resistance to DDT and CQ and logistical challenges. One challenge of note recognized that this type of short-term program could not flourish in Africa, the highest-burden region, because it lacked the infrastructure for surveillance[50].

It is important to note that the original goal for elimination, highlighted in the name GMEP, was eradication of the parasite worldwide. Eradication would mean no human-infecting *Plasmodium* exist in the natural world for transimission. There has been one historic example of eradication actually occurring, with the smallpox causing virus, variola. Smallpox was officially declared eradicated in 1980, shortly after the last natural case in Somalia in 1977[51]. The majority of the progress of smallpox eradication was due to aggressive immunization, infection containment, and surrvelliance programs against variola, lead by the WHO from 1966-1980[51]. Achieving malaria eradication is a much greater undertaking for several reasons, like multiple infecting species, potential zoonotic infections, more complex vaccine evasion mechanisms, and limited treatment options. There has been plenty of research into developing a malaria vaccine, with some marked successes, culminating in the first WHO recommended vaccine (RTS,S/AS01) for children in moderate to high transmission regions in 2021[52; 53]. This is implementation is the first major step towards full malaria vaccination, but the current vaccination does not alter the infectivity of gametocytes so endemicity is likely to continue undisrupted. Thus, true eradication has not been considered a viable goal for the field since the 1950s, and currently the goal is elimination of *Plasmodium.* In contrast to the strict definitions of eradication, elimination has more limited definition of no local transmission occurring in a defined geographic region. The slight difference between the two definitions allows for a country-by-country tackling of the disease, which allows for smaller, more attainable goals to be set towards the progress of ridding the world of *Plasmodium.*

Along with revealing that eradication was not viable, the GMEP highlighted that any additional strategies needed to be adaptable and realistic for local capabilities. This perceived “failure” of a program and economic struggles led to no other large-scale attempts for elimination and scaling back of control efforts in the 1970s and 1980s[49]. This lull in support, paired with rising CQ resistance, was highly detrimental to the progress of elimination, leading to an extreme resurgence of the parasite that peaked in 2000[15; 54].

The 1990s was when governmental bodies and non-governmental organizations (NGOs) rekindled the efforts for global-scale elimination (rather than eradication) with initiatives like Global Malaria Control Strategy in 1993 (WHO), Roll Back Malaria in 1998, the Presidents Malaria Initiative, and more[55]. One of the major benefits of the renewed effort is the recognition that the region with the highest burden should be the region leading efforts. This was solidified when the WHO Africa Regional Office (AFRO) was tasked with heading the new Roll Back Malaria campaign to direct the initiative in ways that would especially benefit Africa[55]. The focus of this initiative for underserved populations and developing countries was to heighten existing healthcare systems to provide more access to services regardless of income status, better education to the general public, and more trained staff[55]. One particularly positive outcome of AFRO leading the initiative was they understood the populations at highest risk and refocused the efforts on them. For example, a strategic framework AFRO published in 2004 outlined all of the challenges and presented solutions for malaria in pregnant women[56]. Malaria during pregnancy can be dangerous for

both the mother and child because, in highly endemic regions, mothers can be asymptomatically infected, which can cause severe maternal anemia for her and low birth weight for the baby[9]. Both conditions are known to contribute to higher mortality in these populations. This document then suggested interventions such as insecticide-treated nets (INTs) and intermittent chemoprevention for pregnant mothers (IPTp), which are techniques still used in elimination today[56].

Several of these renewed elimination efforts made remarkable progress in the field, guided mainly by the documents mentioned above, which caused malaria incidence to decrease by 37% between 2000 and 2015[43]. One of the most successful initiatives was INT distribution in sub-Saharan Africa, which expanded access to bed nets from 7% in 2005 to 67% in 2015[57]. This increase indicates substantial progress in efforts like the distribution of INTs in the regions historically ignored by elimination programs. Still, no countries in this region managed to eliminate the parasite entirely. Seventeen countries lost endemic status outside of this region from 2000 to 2015, and four were considered malaria-free by the WHO[57]. Despite this highlighted progress, the original goal for the Roll Back Malaria initiative was to half the malaria burden by 2010; this goal was not reached, especially in the Africa region[58]. However, this initiative was much more successful than the previous attempts and established the framework for the current iteration of elimination that started in 2015.

1.4 Current state of elimination and roadblocks

Elimination is currently guided on a global scale by the WHO Global Technical Strategy 2016-2030 (ratified in 2015). This section will discuss the currently

active document updated in 2021[47]. This document outlines the steps needed to reduce local transmission of the disease globally by 90% by 2030. The process of elimination is broken down into 5-year goals, where mortality and case rates decrease by a specified amount, a certain number of countries eliminate the parasite entirely, and all countries prevent reestablishment. Three pillars and two supporting elements guide all efforts of the strategy and focus on the following: the pillars include access to malaria care via universal healthcare, intensify efforts for elimination, highlight malaria surveillance as an intervention method, supported by the ideas to expand and utilize research, and expand efforts in a sustainable and equitable manner[47]. The document seems to call for two things: to support and push interventions we know work like IPTp and INTs and scale up research to understand malaria better on a global scale. Another note pervasive throughout the document is acknowledging that the world did not meet the first milestone in 2020, and was not projected to meet it even before the coronavirus disease of 2019 (COVID-19) pandemic[47]. This point is highly evident in the most recent WHO World Malaria report as well, and several theories have been put forth as to why interventions are not functioning as well as expected[29].

Antimalarial resistance is a major potential roadblock to elimination because, unless handled with care, resistance to ACTs could cause a significant resurgence of the parasite, undoing years of intervention work. The field has an example of the damage wide-spread resistance can do because CQ resistance was the primary cause of malaria deaths spiking to over 1 million in the 2000s[59]. The mechanism of ACTs is that the artemisinin derivative acts as a quick

parasite clearance drug with a short half-life in the blood, paired with a longer half-life drug to kill any remaining parasites[6; 60]. This type of combination therapy was adapted to reduce the parasite's ability to evolve resistance after the 2000s crisis of CQ. Regardless, there has been a rise in artemisinin resistance, which started at the Thai-Cambodian borders in the 2000s[60]. When resistance is detected or attempted to be contained, there are a few alternative options, like Atovaquone-proguanil or oral quinine. However, both options are non-scalable, have tolerance issues, and cannot be adopted widely[6; 61].

Resistance to ACT treatments in the parasite generally evolves in genes implicated in the drug mechanism of action. Specific examples of genes that have mutations which confer resistance to at least one drug used in an ACT treatment are as follows: PF Kelch-13 gene (*pf-k13)* to artemisinin, PF dihydrofolate reductase gene (*pf-dhfr)* to pyrimethamine, PF dihydropteroate synthase gene (*pf-dhps)* to sulfadoxine*,* and PF multidrug resistant 1 gene (*pf-mdr1)* to mefloquine[62-65]. As mentioned in later sections, genes that confer resistance to current treatment (ACTs) started to evolve resistance from older monotherapies either because one of the two drugs in the combination is an older monotherapy or the genes involved in new therapies have similar underlying mechanisms for resistance. For example, *pf-mdr1* was initially implicated in resistance to CQ monotherapy, which was used until 2004 as a first-line treatment, and now also confers resistance to mefloquine[16; 17]. The cross resistance in this case is likely due to the gene encoding a digestive vacole which has been shown to transport multiple antimalarial therapies and reduces efficacy

with certain mutations[66]. Another explanation is that this overlap occurs because of the small number of effective drug classes against the parasite, thus many combination therapies reuse older antimalarial drugs in combination with artemisinin. In this case, the parasite already has resistance to one of the therapies, which can circumvent the second stage in ACT action mechanism leaving parasites alive in the bloodstream[61]. Due to the pressure of no alternative drugs for treating malaria ready for wide-scale production and distribution, the WHO has called for more research on antimalarial resistance and surveillance techniques to attempt to contain resistance until new therapies are available.

Paired with the constant evolution of the *Plasmodium* parasite, the rise of evolving *Anopheles* vectors poses a risk to elimination efforts. Indoor residual spraying (IRS) and ITNs are the two main mechanisms for vector control[67]. According to the most recent WHO Malaria Report, all four major classes of insecticides are showing evidence of resistance in multiple countries, pyrethroids, carbamates, organophosphates, and organochlorines[29]. The pyrethroids are a particular problem because this class is the only approved option for ITNs treatment, and ITNs have been the most widely used vector control[29]. Resistance to this insecticide, detected in 87% of all countries at least once, arises in a voltage-gate sodium channel gene. This locus has also been shown to confer resistance to the spraying compound DDT, making it an extremely problematic variant[68]. Appearance of resistance to the other three classes of insecticides has occurred in 60%-82% of countries, with the lowest amount of resistance being to organophosphates[29]. Overall, rising resistance must continue to be researched

and documented because it undermines the efficacy of one of the major methods of intervention for elimination, ITNs.

Another highly prevalent issue that impacts all stages of elimination is the existence of a parasite reservoir in human populations, primarily as asymptomatic carriers. Many of the canonical interventions put forth by the WHO do not adequately address this reservoir population in endemic regions because surveillance and treatment policies are generally focused on quickly identifying, confirming, and treating symptomatic infections[47]. While some of the more passive interventions like ITNs and IRS may reduce the burden of asymptomatic individuals by preventing infections in general, there are many reports that asymptomatic infection is a significant problem for much of sub-Saharan Africa[69-73]. Such individuals go completely undetected in many situations and act as a constant resource for infecting more mosquitos, potentially increasing the number of symptomatic cases, thus showing the slowing or halting of the elimination process. We have highlighted how this entire cycle would look in Figure 1.1.

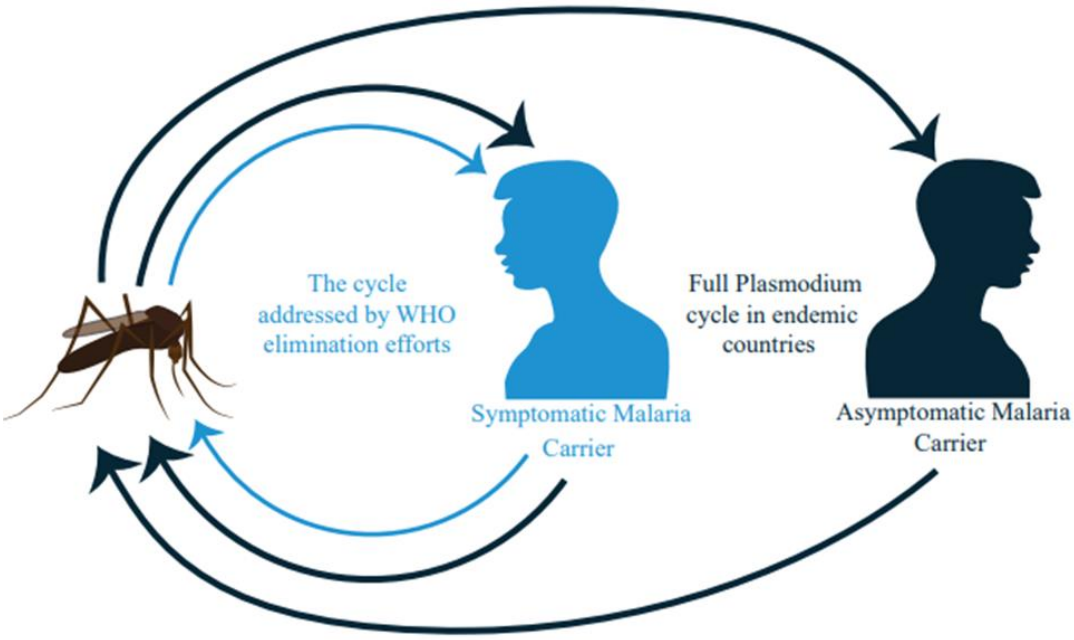


Figure 1.1 Illustration of the malaria transmission cycle with asymptomatic carriers

Visual representation of the two cycles of malaria transmission, light blue: illustrates the general symptomatic to mosquito cycle that has been of most concern to the WHO, as intervention and most surveillance is only concerned with this population, dark blue: the full cycle in most endemic regions where the ignored asymptomatic population is contributing to infecting mosquitos which in turn create more symptomatic cases.

There has been some discourse on whether asymptomatic individuals contribute to ongoing transmission, or if their gametocyte density is too low. In support of the former hypothesis, several studies have proven through feeding assays that they contribute to transmission[72; 74]. Asymptomatic infections have also been shown to be able to transition to active symptomatic infection or cause

other detrimental health outcomes like placental infection or anemia in populations at risk (children, pregnant women, or the elderly)[69; 75-77]. This roadblock was the only one not discussed in detail in the most recent WHO report; the burden of this problem seems to be overlooked Progress towards elimination will never reach its full potential until this population is thoroughly evaluated and intervention methods are developed to address it.

It is impossible to discuss concerns for elimination without acknowledging the COVID-19 pandemic and how it may have affected elimination measures for years to come[78]. Major setbacks in elimination efforts have occurred throughout the COVID-19 pandemic. For example, the pandemic-caused disruption of distribution of insecticide-treated bed nets and RDTs could lead to more malaria cases through more infections and persistence of the parasite in endemic populations from lack of detection[79]. While a portion of this disruption was due to reduced human contact and travel, the COVID-19 crisis also diverted a significant amount of malaria funding to dealing with this emergent health challenge and severely reduced access to needed materials[47].

Another unexpected detriment of the COVID-19 pandemic was the implication of antimalarials in the potential treatment and prevention of the disease. Some studies have shown the effectiveness of certain artemisinin-derived compounds against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) *in vitro,* but have not been confirmed by clinical trial[80; 81]. Regardless, simply claiming potential protection against SARS-CoV-2 could have spurred the usage of the drug for this purpose, similar to the adoption of ivermectin in the US from

unsubstantiated claims, which can severely increase antimalarial resistance to current therapies[82]. If the off-label use of artemisinin therapies could spur a rise in resistance to the drug in endemic countries, the results could be disastrous and worsen the state of elimination for years. Currently, the state of treatment options is already limited, and new drugs are not in the immediate future, thus artemisinin must remain effective[61; 83]. While no novel roadblocks have occurred through the pandemic, countries working toward elimination may have the amount they are off track from the WHO GTS goals exacerbated by COVID-19. This exacerbation caused by COVID-19 has already been noted in the most recent World Malaria Report, which indicated 13.4 million cases were from different facets of pandemic disruptions[29].

As long as the effects of COVID-19 on malaria continue to be uncovered, and the elimination goals are off track, the field will need to rise to the WHO call for more research. Scientists need to understand all aspects of the disease and parasite on the local, country, and global levels to truly define the problem and implement the right interventions to solve them. As it stands right now, providing that information to the WHO is a cumbersome process that requires extensive financial support and manpower to overcome because recruiting and conducting large-scale malaria projects is resource intensive. One way to circumvent the cost of conducting the aforementioned research, and perhaps better define the problems contributing to these roadblocks to elimination would be to develop a method to retrospectively analyze populations for malaria.

1.5 Diagnostic methods and how they shifted research

Each new era of malaria understanding and research was ushered in by a new method of detecting the parasite in humans. Three main methods are used for detecting *Plasmodium* in both the clinical and research setting; blood smear microscopy and RDTs are used in the clinic, and RDTs or NAATs are used in research.

The development of the blood smear method of diagnosing malaria cannot be attributed to any one person. The obvious first step in the development was the introduction of the microscope to the scientific community by Robert Hooke and Antonj van Leeuwenhoek and the improvements introduced by Ernst Abbe and August Köhler, which made light microscopy of biological samples possible[84]. This invention paved the way for the careful study of differently shaped organisms with “pigment” in the blood sample of 200 military patients by Alphonse Laveran in 1880[3]. With this observation, Laveran uncovered that the parasite resided within blood cells and determined that the parasite was responsible for the disease[3]. The final step in curating the entire diagnostic process was a more efficient stain that better differentiated the parasite from the RBC cellular component, which Gustav Giemsa developed in 1902[85]. These three components are the basis of the thin blood smear methodology, where peripheral blood is smeared on a slide, the number of parasitized RBCs is counted, and the parasite species is determined[6; 86]. Obviously, before the introduction of modern technologies, this methodology was the only way to confirm an illness was due to *Plasmodium* infection. Until the 1990s, this method

was the only way we researched, diagnosed, and tracked malaria worldwide. One of the functions that microscopy still serves today is identifying neutrophils with ingested pigment, which has prognostic value, along with identifying the stage and species of parasite better than RDTs in the case of non-*falciparum* infections[9]. As mentioned before, each parasite species has its own clinical implications; thus, knowing which species is infecting an individual can help determine which presentation of severe malaria is possible[9]. Complete species identification paired with standard tests, like quantifying the blood count and determining hemoglobin concentration, better determines if malaria is the root of the illness or if the patient needs to be tested for alternative causes[9]. This differential diagnosis is essential because, in the highly endemic regions of sub-Saharan Africa, many people can be sick and positive with malaria, but the underlying cause of their illness (and potential death) isn't from severe malaria but co-infected bacteria. Despite what this method has provided for the field, it has several downfalls, like the need for trained staff, its unreliability for quantifying PF infections because of sequestering, and relatively high limit of detection (LOD)[8; 86]. The need for trained staff has been highlighted as particularly problematic for sub-Saharan Africa, as many rural areas do not have the capital to maintain someone with these skills[86].

The invention of RDTs in the late 1990s ushered in the next wave of understanding and progress for malaria because it eliminated the need for highly trained microscopists to diagnose the parasite[87]. Most RDTs function by determining the presence of a PF specific protein called histidine-rich protein 2

(PfHRP2) or a pan-malaria antigen like lactate dehydrogenase or adolase[87; 88]. RDTs are the crux of understanding malaria epidemiology today because they are the most readily used tool in diagnosing the disease across sub-Saharan Africa[29]. According to the WHO, 3.5 billion RDTs were sold between 2010 and 2021, and 2.1 billion of those RDTs were distributed in sub-Saharan Africa by national malaria programs, thus showing the substantial amount of utilization of this tool[29]. The advent of this device is one of the major interventions that made tackling malaria in sub-Saharan Africa possible and continues to contribute to the reduction of mortality of this disease today. The limitations to RDTs come from the effectiveness of the test, its constrained ability to determine species of infection, and its higher LOD[8; 86].

Up to this point in the technology, most cases detected by RDTs and blood smears needed a high infection density because both methods have been shown to miss submicroscopic and asymptomatic infections in endemic regions[89]. It was not until the introduction of NAATs could the field push the limits of detection into fully characterizing the burden of malaria in sub-Saharan Africa. This fact is becoming increasingly evident as countries continue showing a substantial asymptomatic population or low-density infections missed by normal surveillance procedures[90]. The research sector mainly utilizes it because the polymerase chain reaction (PCR) assay does not need to happen at the collection site. Deoxyribose nucleic acid (DNA) can be isolated in the field and frozen or dried whole blood spots can be transported back to laboratories for processing later, eliminating the need to transport laboratory equipment to remote locations[91; 92].

NAATs are also significantly more specific with respect to differentiating species of infection and even calling mixed infections with high accuracy[8; 93]. While NAATs are not used in point-of-care situations, it is one of the only reliable methods to detect parasite other than PF or PV, because current RDTs for other species have less than 50% sensitivity to PM and PO[94; 95]. Thus, a major finding of the assay was that the lesser-discussed parasites PO and PM are prevalent across sub-Saharan Africa and need to be better detected and addressed by elimination efforts alongside PF[96].

Additionally, because the assay targets the DNA instead of actively expressed genes, this technique is also less sensitive to evolutionary changes by the parasite in an attempt to avoid detection, unlike RDTs[97]. Overall, NAATs are the optimal tool for defining the composition and scope of *Plasmodium* species worldwide. Unfortunately, it cannot be widely applied to the clinical sector because it is considered prohibitively expensive compared to RDTs and requires a laboratory setting and trained technicians.

1.6 Potential role of genomics in malaria

As highlighted above, the introduction of DNA-based methodologies was revolutionary in the malaria research field. Outside of sensitive parasite detection, genomics has played a much greater role in our understanding of *Plasmodium*. For example, genomics has played a part in identifying where this parasite may have evolved from, uncovering the two different species of *Plasmodium ovale* (*curtisi* and *wallikeri*), determining which mutations confer antimalarial resistance, and understanding multiplicity of infection and its impact

on disease[98-103]. And as sequencing studies become more prevalent and the technology continues to become more affordable, there is no doubt that genomic methodologies will continue to contribute to our understanding of malaria.

Recently, there has been an increase in publications focused on developing new methods of detecting the parasite. There have been papers published this year on electrochemical sensors, deep and machine learning processing of microscopy slides, and metagenomics approaches queried for their ability to detect the parasite without using traditional methods[104-107]. Genomics could play an integral role in the next step of malaria detection, specifically from human whole genome sequencing data that is processed via a PCR-free methodology. Modern shotgun sequencing technology does not select for species, meaning that any sequence present I isolated DNA has an equal chance to be captured by the sequencer. There are two basic whole genome sequencing (WGS) types, one with a final PCR amplification step before sequencing and one without[108; 109]. Choosing the type of sequencing is of particular consequence to malaria for the following reason: PCR-based sequencing has been shown to be poor at capturing sequences with C+G-bias. This initial difficulty in capturing extreme genomes is generally worsened by the final PCR amplification step of a general flow, even though the universal primers are designed to this and capture all DNA in a sample indiscriminately. This inability to capture extreme genomes has historically made capturing the malaria parasite sequence difficult. However, the original paper proves this issue has been solved in PCR-free sequencing by testing G+C poor PF sequence as part of the method's proof of concept, and

highlighting how much better PCR-free performs in extreme conditions[109]. With the advent of PCR-free technology, the relative coverage of the PF (3D7) genome went from missing 4.8% to 19.9% of its sequencing in coverage to less than 4% having under 10x coverage[109]. Many sequencing studies are now utilizing this PCR-free method of sequencing for humans. For example, all studies originating with the Trans-Omics for Precision Medicine (TOPMed) project use WGS technology for analyzing their human populations[110]. Therefore, all populations that are sequenced in this manner are now available to mine for *Plasmodium* DNA, and are far more likely to capture the parasite when present in endemic populations, making WGS studies a perfect candidate for developing a computational detection pipeline.

1.7 Dissertation aims

In summary, malaria is a significant infectious disease that disproportionately affects underserved populations in Africa, and the progress toward elimination has slowied significantly in recent years. Most roadblocks highlighted in this chapter have called for more surveillance and research to fully untangle the intricacies of malaria on a global scale. Here I set out to develop a tool that will facilitate the necessary research.

As studies of human genetic variation expand beyond a Eurocentric focus to include a wider range of populations such as the Three Million African Genomes, TOPMed initiatives, and the African Genome Variation Project, we will begin to fill in gaps in our knowledge of patterns of human genetic diversity and how extant variation affects disease risks and distributions as well as our evolutionary

histories[110-114]. The addition of these diverse datasets alone can be transformative understanding human diversity, evolution, and discovery of genetic variation underlying human health disorders, but the increase in the diversity of populations sampled also has the potential to reveal additional and novel information relevant to health and disease. Specifically, whole genome sequence (WGS) analyses of DNA from human whole blood may be able to capture genetic variation in other species that can affect both individual and public health[115-117]. However, most analyses from human next-generation sequencing ignore this latter possibility.

The typical pipeline for processing human whole genome sequence data is to align reads to the human reference genome(s) and discard all other or unmapped reads (UMR)[118]. However, it is possible that sequence from many potentially infectious agents can be evaluated using the UMRs. Here we hypothesized that in malaria-endemic regions, UMRs would align with the genomes of one of the parasites that can cause malaria in humans (*Plasmodium falciparum*, *P. malariae*, *P. ovale,* and *P. vivax*)[28]. As malaria is highly prevalent in Africa and this region carries the majority of the malaria disease burden, it is an ideal locale to see if human WGS can be used to detect *Plasmodium* infections[28; 93] I hypothesize that a combination of efficient genomic analyses tools can adequately harness the information in these UMRs to detect the parasite epidemiology from general human WGS studies, and even provide information on relevant parasite antimalarial resistance genes. Determining if *Plasmodium* sequence is present in the UMRs may provide insights on how to

learn more about host-parasite interactions. Also, with further exploitation, investigation of UMRs could provide insight for development of additional molecular epidemiological approaches to study malaria epidemiology.

Therefore, my dissertation aims to identify the appropriate parasite target and genomic tools to detect *Plasmodium* DNA in human WGS UMRs and demonstrate that this novel detection method adequately captures the malaria epidemiology across several populations in sub-Saharan Africa. To achieve this goal, I will address the following specific aims:

Aim 1 - Develop a species-specific *Plasmodium* detection pipeline that utilizes human WGS data and mines the UMRs.

1.) Determine which parasite genome (mitochondrial, apicoplast, or total) is the best target for mining the UMRs.
2.) Compare the infection status indicated by this novel approach in Ghana (n=1904) to a previously accepted NAATs diagnostic method to determine assay specificity, sensitivity, and overall agreement.
3.) Check the following assumptions to ensure no bias exists in the novel pipeline: a. sequencing across the *Plasmodium* genome is occurring randomly with no hot spots, b. the number of unmapped reads that enter the pipeline does not determine the infection call outcome, and c. each species found is likely present in the sample.

Aim 2 – Investigate how accurately our novel detection method characterizes malaria epidemiology in different endemicity levels across seven sub-Saharan African populations.

1.) Extend the pipeline to all other TOPMed Africa6k Populations to determine the level of infection and species composition in the following populations: Botswana, Cameroon, Ethiopia, Guinea-Bissau, Kenya, Tanzania
2.) Compare found values and estimates to well-known malaria epidemiology databases, WHO and the Malaria Atlas Project (MAP), to determine if the novel assay is performing reliably
3.) Investigate the utility of the mined *Plasmodium* genomic data to query for known antimalarial resistance loci in the following genes: *pf-mdr1,* PF chloroquine resistance transport gene (*pf-crt), pf-dhps, pf-dhfr,* and *pf-k13.*

CHAPTER II: Development of a species-specific *Plasmodium* assay using human whole genome sequencing reads from Ghana

Portions of this Chapter are under revision in the following manuscript:

Olvany JM**,** Chan ER, Martin WB, Hansen MEB, Harris DN, Tishkoff SA, Williams SM, Zimmerman PA. Detection of species-specific Plasmodium infection using unmapped reads from human whole genome sequences.

2.1 Introduction

PCR-free sequencing methodologies offer a unique data mining opportunity, as they can capture all DNA in a sample, regardless of the species of origin. The typical pipeline for processing human whole genome sequence data is to align reads to the human reference and discard all other or unmapped reads (UMR). However, it is possible that sequences from many potentially infectious agents can be evaluated using the UMRs. The potential for this approach was recently demonstrated, as *Plasmodium* could be detected at the genus, and sometimes species level, using metagenomics shotgun sequencing of DNA from blood if host DNA was removed[119]. A few other attempts have taken a metagenomics approach to detect *Plasmodium*, but those publications did not use general WGS data meant to sequence the human genome without selecting for the parasite genome prior to sequencing[120]. Here we show that using only human whole genome sequencing data, without selecting for or amplifying the parasite genome, can inform infection status even in presumably low parasitemia and asymptomatic infection. We do this by demonstrating the presence of *Plasmodium* in Ghanaian UMRs from the greater TOPMed human whole genome sequence project.

2.2 Methods for assay development

2.2.1 Rationale

It is essential to determine which species of *Plasmodium* is in an infection to both the epidemiology of malaria and to design clinical treatment. There are three different genomes in each *Plasmodium* parasite: the nuclear, the

apicoplast, and the mitochondrial, all of which offer a potential target for capturing sequencing reads[121]. We tested for the presence of these three genomes independently to determine which was most sensitive for the resolution of the species composition: mitochondria genome only, apicoplast genome only, and the combination genome (all three combined). Once the best target was chosen, we filtered each of the four *Plasmodium* species references through an artificial reads methodology, detailed below, which eliminated any region which could identify sequence from more than one species. We assessed the performance of the resulting assay through comparison with an accepted methodology that uses PCR amplification of the 18s ribosomal ribonucleic acid (rRNA) parasitic gene to diagnose individual samples.

2.2.2 Unmapped reads isolation

All individuals (n=6,457) included in this study constitute the Trans-Omics for Precision Medicine (TOPMed) sequencing project, Africa6k[110]. Individuals were selected for sequencing from seven countries, and several pre-existing studies collected by Dr. Scott Williams (Ghana) and Dr. Sarah Tishkoff (Botswana, Cameroon, Ethiopia, Guinea-Bissau, Kenya, Tanzania). The selected samples were whole genome sequenced (WGS) to a median depth of 30x using DNA isolated from blood. For all samples, DNA isolation was performed on 5 mL of whole blood drawn into EDTA-coated tubes (BD Biosciences) using the Gentra Puregene Blood kit (QIAGEN) in accordance with the manufacturer's protocol. This protocol captures all DNA present in the sample. Sequencing was completed using PCR-free library construction and Illumina HiSeq X technology,

described fully in previous publications[110; 122-125]. All cram files were provided by The Broad Institute and connected to the cloud computation ecosystem, NHLBI BioData Catalyst (BDC) *Powered by* Seven Bridges, for computation[126].

A portion (n=1904, Ghana, 29.5%) of the aforementioned Africa6k population was used for the data discussed in this chapter based on the availability of frozen DNA samples. These samples could be evaluated for *Plasmodium* infection using accepted methodologies, detailed below, serving as a point of comparison for assay performance. All adult participants were recruited as part of a larger (n= 3,782) cardiovascular study in and around Sunyani, the capital of the Brong-Ahafo region of Ghana[127]. Individuals were collected between 2002 and 2006; each participant was measured for cardiovascular variables, medical history, current medications, and demographic and socio-economic data. Exclusion criteria included acute illness in the previous two weeks or a first or second-degree relative already enrolled in the study. A full description of recruitment and the resulting study can be found in previous publications (refs). Participants were asked to read and sign broad informed consent forms prior to sample collection, which detailed the possibility of further research being done on their collected samples. Institutional review boards at Vanderbilt University, Dartmouth College, Regional Hospital, Sunyani, and Case Western Reserve University approved all protocols

Prior to the cram files being joint called for human variation, all unaligned UMRs were isolated using SAMtools v 1.6 on *BDC-Seven Bridges* using common workflow language (CWL) as the base workflow language[128].

2.2.3 PCR-based detection of Plasmodium infection

A preliminary Plasmodium infection screen on a portion of these samples (n=1,090) was performed using the previously published PCR-based method that amplifies species-specific 18S rRNA parasite gene from human whole blood isolates, then uses a ligase detection reaction-fluorescent microsphere assay to determine infection composition. This confirmed the presence of the parasite prior to developing a computational detection protocol. The prior method can detect all four species that cause disease, PF, PV, PO, or PM infections. Full descriptions of the PCR primers, fluorescent probes, and other details can be found McNamara et al[93].

2.2.4 Mitochondrial reference genome

We chose the mitochondrial genome as the first target for the following reasons: the mitochondrial DNA (mtDNA) is at a higher copy number than the autosomal DNA, and it would have the lowest computational burden as it is the smallest genome. We used the following mitochondrial references from GenBank: PF (M76611), PM (LT594637), PO *curtisi* (HQ712052), PO *wallikeri* (HQ712053), PV (NC_007243). We combined all of the mitochondrial reference sequences of the four species into one NCBI MagicBLAST database[129]. A small pilot study (n=600) was conducted by BLASTing UMRs against this new database. To determine the viability of this target, we compared the small sample to the PCR-based diagnostic methodology described above. To investigate the cross-reactivity of the sequences, we evaluated the similarity of the mitochondrial sequences through comparative genetics. Pairwise comparisons between the

species were done using 200 bp sliding windows with 20 bp steps and calculated using Hammings distance in Simplot[130]. We estimated the total difference between the genomes through Clustal Omega, using percent identity[131]. Evolutionarily conserved regions, called local collinear blocks (LCBs), were identified through Mauve multiple genome alignment[132]. Two parameters can be changed in the Mauve progressive alignment process, the LCB weight and seed weight, which can alter how conservatively the algorithm calls sequence blocks as a conserved region. We used the following parameters: seed weight at 11 and minimum local collinear blocks LBCs weight at 477. The results of these analyses were plotted using Circos[133].

2.2.5 Apicoplast reference genome

Because NCBI MagicBLAST aligner does not allow for customization of the alignment parameters, from this point forward, we used Bowtie2 aligner[134]. We queried for the apicoplast genome of all four *Plasmodium* species to assess the feasibility of detecting *Plasmodium* in UMRs without prior amplification using the new alignment program. The apicoplast is an apicomplexan organelle homologous to plant chloroplasts, with no human counterpart[121]. Therefore, the detection of this genome is absolutely of non-human origin and demonstrates the presence of *Plasmodium*. We combined all four species into one reference file and used Bowtie2 to align with the thresholds mentioned below. Any apicoplast (API) read, regardless of species, was considered likely positive for Plasmodium infection. To reduce the likelihood of poor-quality reads aligning to the reference

indices, we set the minimum length for alignment to 75 bp and used the “very-sensitive” alignment scoring parameters described in the Bowtie2 manual.

2.2.6 Total reference genome

Our final goal was to develop a species-specific computational detection methodology was the total *Plasmodium* genome (nuclear, apicoplast, and mitochondrial). We used the following Plasmodium reference files (fasta) for each of the four known species that exist in sub-Saharan Africa: PlasmoDB-51_Pfalciparum3D7_Genome (GCA_000002765.3), PlasmoDB-51_PmalariaeUG01_Genome(GCA_900090045.1), PlasmoDB-51_PovalecurtisiGH01_Genome (GCA_900090035.2), and PlasmoDB-55_PvivaxP01_Genome(GCA_900093555.2), for PF, PM, PO, and PV, respectively[135]. The human-infecting species of *Plasmodium* diverged relatively recently and are therefore closely related enough to have high sequence similarity in some regions based on 150 bp reads. Because the species of infection is relevant to clinical outcomes and research, any detection method must be able to resolve infection composition. For this purpose, we developed a method to eliminate regions of high similarity across species (Figure 2.1). Using the example of *Plasmodium falciparum* (PF), the methodology worked as follows: all contigs in the PF reference file are broken down into 150 bp “reads” to mimic the data coming from the sequencer. These 150 fragments are then sequentially aligned to every other species of *Plasmodium* (Pv, Pm, Po) and human. We aligned the fabricated reads to each alternative reference file using Bowtie2 aligner. Any reads that mapped to any other reference file were discarded at

each step, leaving only portions of the original PF reference genome that are unique to PF. The unique “reads” are then remapped to the original genome to reassemble a PF reference file. This resulting filtered reference was used to query the unmapped reads isolated from the Ghanaian participants.

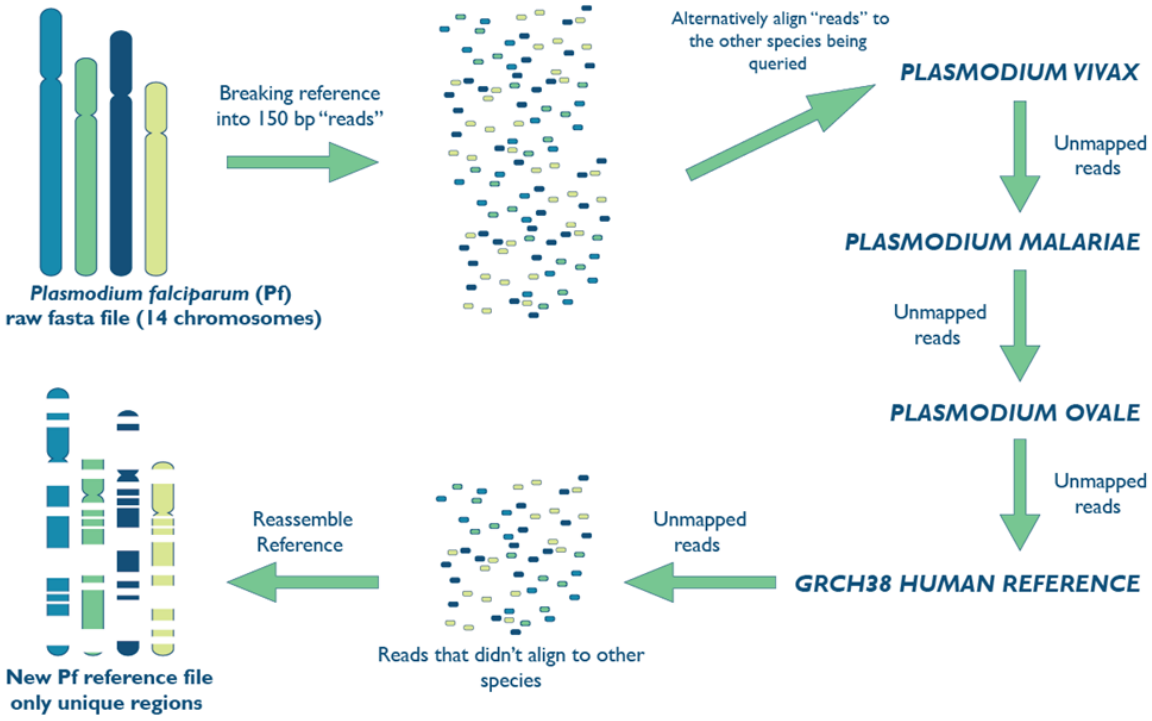


Figure 2.1 Illustration of Plasmodium reference genome filtering.

Visual representation of segmenting the reference files into 150 bp fabricated reads, aligning them sequentially to all other species, and removing any reads that align to another species. After filtering, the pieces are reassembled to create a species-specific reference file.

For any species that still showed non-specific binding, meaning an unexpectedly high level of positivity, we used Mauve multiple genome aligner to identify evolutionarily conserved LCBs and eliminated those regions from the

reference file. To ensure elimination of all regions that could potentially produce a false positive, we evaluated several different values for the LCB weight (10kb, 5kb, 2.5kb, and 1kb) parameter of Mauve to assess which one eliminated the most regions. We found that an LCB weight of 1kb culled the most regions when used with a seed weight of 15. After the LCBs were removed from the original, unaltered reference files, the remaining genomic regions were subjected to the fabricated reads methodology as described above. The complete list of retained regions for each species of Plasmodium is in Supplemental Table 1. Each of the four species were built into a DNA index using Bowtie2, for use as reference files for mapping in future alignments. We used a threshold of 50 reads aligned anywhere in the genome to consider a species positive.

As a negative control, the alignment process was repeated on the 1000 Genomes WGS samples on the BDC-Seven Bridges. This dataset covers several endemic malaria populations that should be positive; however, the DNA was isolated from cells that had been cultured prior to sequencing. Therefore, even endemic regions should remain negative for *Plasmodium* DNA if the assay performed reliably and in a parasite-specific manner.

2.2.7 Comparison to accepted methodologies

To assess the accuracy of our WGS analytical pipeline, results from the sequencing data were compared to the PCR-based assay, the gold standard for infection sensitivity. Using binary infected/uninfected coding, we calculated percent agreement, sensitivity, specificity, positive predictive value, and negative predictive value. After binary comparisons, each *Plasmodium* species was

considered a single outcome, and we calculated species concordance using a kappa-weighted percent agreement calculation, where a weight of 1.00 indicated tests that were 100% concordant. Weights of 0.75, 0.5, and 0.25 were applied when the tests agreed ¾, ½, and ¼ of the time.

There were a handful of infection calls that seemed non-specific (i.e., all species were considered positive), and we reevaluated the calls by calling the species with the highest number of reads positive, and any other species with at least 10% of the highest reads also positive. For example, if a Mixed (all) infection had the following composition: PF 10,000 reads aligned, PM 200 reads aligned, PO 1,000 reads aligned, and PV 100 reads aligned, then that infection would be alternatively called as Mixed(Pf and Po). To investigate if the presence of these infections was due to a difference in reads mapped to a single species, we compared the mean amount of total reads mapped for the Mixed(all) infections to all other infection types using a Welch's t-test.

2.2.8 Assay Validation

Because there was a range in the number of mapped reads per individual, it is possible that an individual with fewer unmapped reads would be more likely to be assessed as negative. To ensure that this bias was not present we compared the number of UMRs in the original unaligned files to the two-by-two table designation (true positive, true negative, false positive, false negative) and used an ANOVA to check for a significant difference in mean UMRs.

After assessing the methodology using the above methods, we determined the patterns of aligned reads in each *Plasmodium* genome. We first determined which species had the highest coverage to evaluate the global binding landscape. In our data, PF was the only parasite infection that occurred frequently enough and had enough read coverage to evaluate the entire genome. If reads were random, each region should have the same number of reads based on the region length, thus indicating non-specific and random sequencing. To understand what the alignment looked like across the genome, we used Circos to plot the number of reads mapped to each of the filtered regions of *Plasmodium falciparum*. We plotted the number of reads aligned versus the size of the region to which the reads aligned and calculated the correlation of these two variables using Pearson's coefficient for PF. Finally, we performed one more Bowtie alignment using an unfiltered Plasmodium falciparum reference genome (PlasmoDB-51_Pfalciparum3D7_Genome). On the individual level, no single infection captured enough reads to pass 1x coverage of the genome based on coverage calculations. Therefore, we combined all aligned reads across the entire population sample into a single global alignment file to visualize the alignment landscape across the entire parasite genome. We used WGSCoveragePlotter to generate a total coverage plot.

2.2.9 Data availability

All human genomic data utilized in this study can be accessed through the TOPMed data exchange site under the Africa6k project if authorized. The UMRs from this particular study are not a part of this publicly available data as they

were retrieved from pre-processed files and are owned by Dr. Scott Williams and Dr. Sarah Tishkoff. All index files needed to replicate this detection methodology on alternative data are stored on Github, the pipeline published on GitHub, and available for public use on the BioData Catalyst® (BDC) *Powered by* Seven Bridges. The GitHub can be found at: https://github.com/jolvany/plasmodium_diagnostic.

2.3 Results

2.3.1 Mitochondrial genome results

In our sample from Ghana (n=1904), between 0.14% and 7.95% of the total reads were unmapped. These represent from 1.01 million to 65.3 million reads that could be queried for *Plasmodium*. Of the 600 (31.5%) that were BLASTed against the mitochondrial references, 280 had PCR data. We found that the mitochondrial genome target compared to the sensitivity gold standard as such: 128 were positive with both assays and deemed true positives (TP), 92 were true negatives (TN), 21 were false positives (FP), and 39 were false negatives (FN) for the overall binary infected/uninfected call. In comparison to the gold standard, using UMR BLASTed to the mitochondrial genome was found to have the following performance: 78.6% agreement for binary calls, 76.6% sensitivity, 81.4% specificity, 85.9% positive predictive value (PPV), and 70.2% negative predictive value (NPV). When we transitioned to looking at the infection composition individual by individual, we began to doubt the efficacy of this method. We found that in using the mtDNA, there were both more uninfected

individuals (131mtDNA vs 113 PCR) than the PCR and more Mixed infections (76 mtDNA vs 42 PCR) (Table 2.1).

New Method	Gold Standard PCR+	PCR-	**Total**
DNA+	128	21	149
DNA-	39	92	131
Total	167	113	*n=280*

Infection status	**mtDNA (+)**	**PCR**
Uninfected	131	113
PF	63	116
PV	0	0
PM	6	3
PO	4	6
Mixed	76	42

Table 2.1 Performance of mtDNA target vs PCR-based assay

TOP: two-by-two comparison of the mitochondrial DNA (mtDNA) target and the PCR assay. Bottom: Shorthand for the species is as follows: PF *Plasmodium falciparum,* PV *P. vivax,* PM *P. malariae*, PO *P. ovale*, and mixed means a combination of two or more of said species.

The increase in Mixed infections could indicate non-specific binding between the different species mtDNA; thus, we compared sequence similarity to investigate this possibility.

In 200 bp increments and pairwise comparisons, we found that the similarity across the mitochondrial genomes ranged from 72% to 100%. The PF

mtDNA genome was the most divergent of the species with a percent identity with other species from 87.5% with PV to 88.6% with PM. The most similar sequence was found between the two subspecies of PO, which were so similar that at no point in their mitochondrial genome were they less than 97% similar. From this observation forward, we treated all species of PO the same and simply reported PO positivity. Full details on the percent identity matrix can be found in Chapter 2 Appendix: Supplemental Figure 1. We pushed this analysis further to try and determine if any regions in the mtDNA could potentially be informative for the species by identifying which blocks of the mtDNA were considered evolutionarily conserved through Mauve. If there was a potentially informative region, we would ideally see it somewhere where the LCBs do not touch, as it would indicate a region that is not conserved. Once the Mauve data was plotted, it became evident that the mtDNA is not a viable target for differentiating the *Plasmodium* sequence, as the entire swath of the sequence is considered mathematically conserved (Figure 2.2).

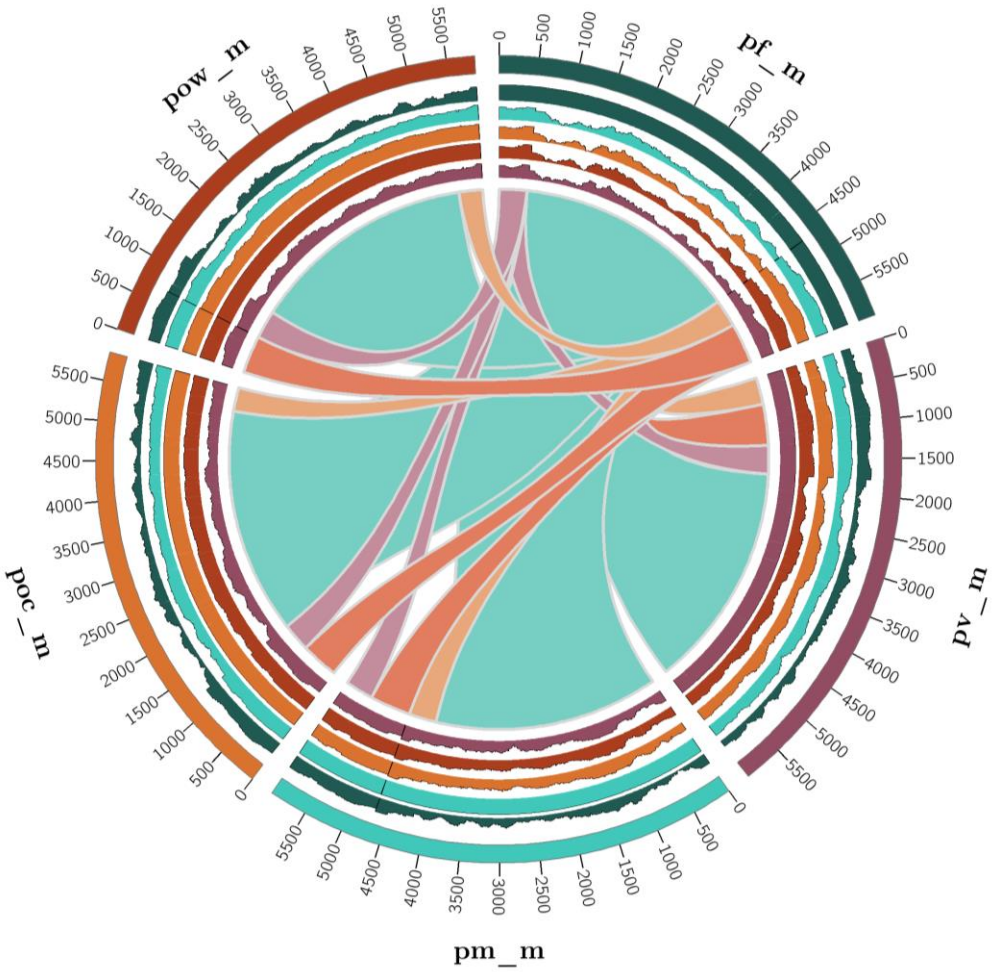


Figure 2.2 Circos plot of mtDNA sequence comparison

Circle legend is reported from the outside in: first circle is species labels and are as follows: pf_m *Plasmodium falciparum* mitochondria, pv_m *Plasmodium vivax* mitochondria, pm_m *Plasmodium malariae* mitochondria, poc_m *Plasmodium ovale curtisi* mitochondria, and pow_m *Plasmodium ovale wallikeri* mitochondria. The second is a generic representation of each of the species mitochondrial genome, with the size labels in bp, each species was assigned a color, and they are as follows: dark blue (PF), pink (PV), teal (PM), orange (PO *curtisi*), red (PO *wallikeri*). The third through sixth circles are a histogram plot of the pairwise

percent similarity calculated by Simplot, the color represents the secondary species that that column is compared against. The eighth and internal circle is a ribbon plot showing where each of the evolutionarily conserved LCBs can be found in each species mtDNA.

2.3.2 Apicoplast genome results

Through the above exploration of the mtDNA, we determined we needed an alignment mechanism with more customization capabilities. We anticipated that through the stringent parameters discussed above, bowtie2 would provide more resolution power between the species than a general aligner like MagicBLAST. All four apicoplast genomes were used for our initial scan in the entire Ghanaian population, based on the knowledge that this organelle is unique to the Apicomplexan phyla and has relatively low similarity among the seven human infecting parasite types (three major genera being Plasmodium, Toxoplasmosis, and Cryptosporidium). We found that of the 1904 individuals in the Ghanaian population, 706 individuals had evidence of the API genome in the UMRs. This query provided proof that Bowtie2 could adequately detect parasite genomes in the sample using UMRs from whole blood DNA isolates as several individuals had at least one read align to the API genomes.

2.3.3 Full Plasmodium reference creation

While we originally pursued the other two genomes to reduce computational burden based on reference size, we ultimately concluded that using the entire

genome was the most accurate only way to make a species-specific detection methodology.

We came to this conclusion because, compared to the API DNA, each nuclear genome is nearly 1000x larger than its API genome (~35 Kb) and have approximately 277X the number of SNPs (11.6 kb vs 3.2 Mb), thus defining many more variable regions that can be used for species designation. Additionally, Bowtie2 could denote a large portion of each of the genomes as unique. Specifically, 99.8% of the PF genome, 99.3% of PV genome on the first pass, 99.9% of PM genome, and 99.6% of the PO genome were designated as unique and remained for subsequent analyses (Figure 2.3). For our purposes, the reference genome refers to the unique sequence unless otherwise stated.

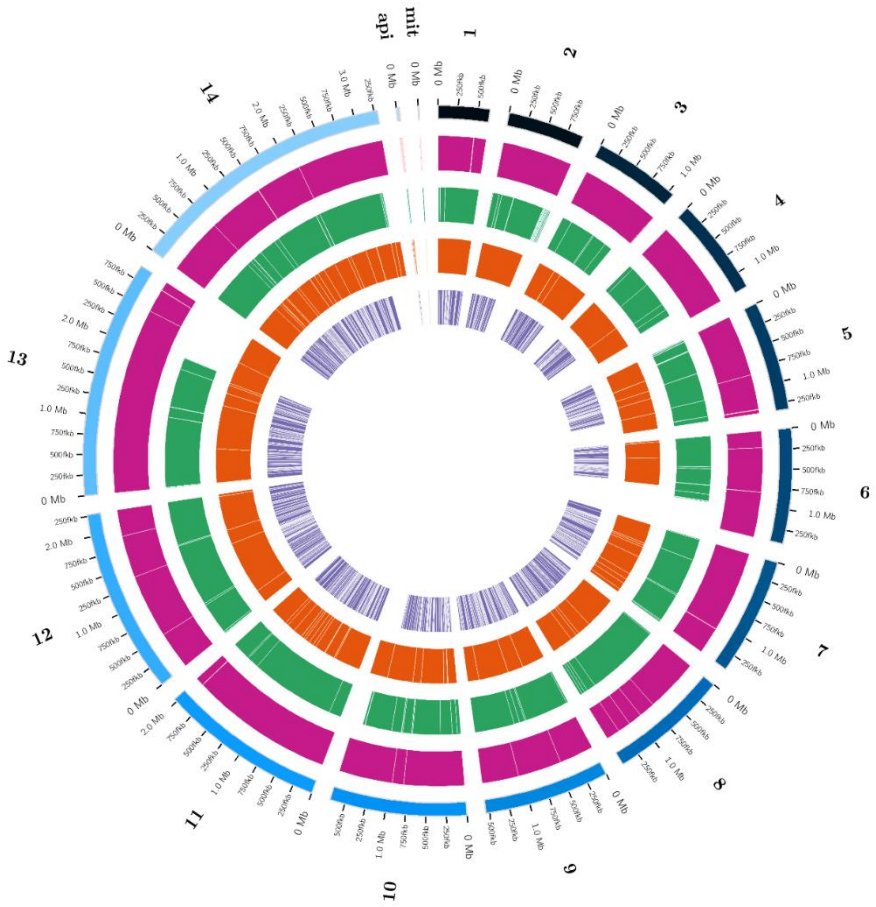


Figure 2.3 Representation of all retained regions in the assembled chromosomes of the four *Plasmodium* genomes

The concentric circles represent the following going from the outer circle: name of organelle or chromosome, general representation of pan-*Plasmodium* assembled chromosomes with size markers in 250kb increments (blue), *Plasmodium falciparum* (pink), *Plasmodium vivax* (green), *Plasmodium malariae* (orange), and *Plasmodium ovale* (purple). For each genome, the regions removed are represented in white.

Coordinates of regions for each species that were used to detect species-specific infection can be found in Chapter 2 Appendix: Supplemental Table 1. The initial estimates of infection status were within expected ranges in three of the four species, but PV had unexpectedly high levels of positivity, nearing 100%. Such a high level of PV sequence and presumed infection could be due to regions in PV with high similarity to other Plasmodium species. To reduce false positive assignment to PV, we removed all unassembled contigs and performed Mauve multiple genome alignment. After applying these procedures to create a conservative reference genome for PV, the majority of the putatively unique PV sequence was removed (only 42.5% of the genome remained and was called *Pv after refinement*, Figure 2.4)

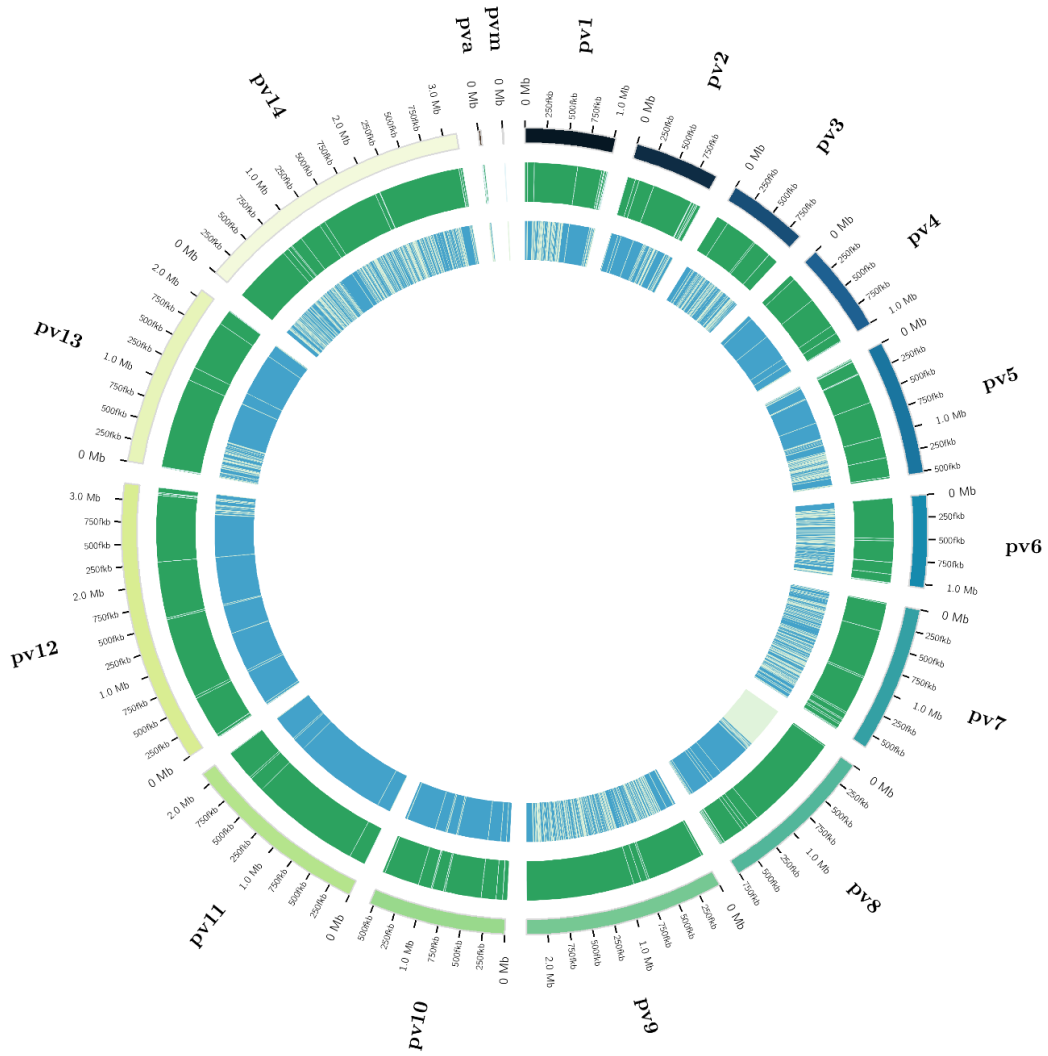


Figure 2.4 Representation of the change in retained regions before removing Mauve identified LCBs and after refinement

The concentric circles represent the following going from the outer circle: name of organelle or chromosome attached with general species (PV), general representation of pan-*Plasmodium vivax* assembled chromosomes with size markers in 250kb increments (blue to white gradient), *Plasmodium vivax* before

refinement (green), *Plasmodium vivax* after refinement (light blue). For each genome, the regions removed are represented in white.

2.3.4 Comparison to PCR diagnoses

Using the reference genomes established above, infection prevalence estimates were 40.5% (n=772) across all years. Of the individuals who were positive based on the total reference genomes, 672 (87.0%) were in agreement with the initial apicoplast status; 100 new individuals were determined to be infection positive, and using a minimum 50 *Plasmodium* read threshold (versus the original 1 read aligned to apicoplast) 34 individuals were considered uninfected.

Among individuals who had both PCR and NGS results, with PCR being considered the “sensitivity gold standard”, (n=1,090) 531 were positive with both assays and deemed TP, 409 were TN) 66 were FP, and 84 were FN for any *Plasmodium* species (Global estimates). In comparison to the gold standard, using UMR from WGS performed reasonably well, with 86.2% agreement for binary calls, 86.3% sensitivity, 86.1% specificity, 88.9% PPV, and 83.0% NPV. The sequence-based assay was marginally better at detecting true negatives versus true positives. We expected the rate of false positives and false negatives to be approximately equal but found that there was a 30% higher rate of false negatives than false positives (Table 2.2). Species agreement was evaluated using a weighted kappa percent agreement. Concordance occurred 88.7% of the time.

Global two-by-two using 50 read threshold

PCR calls (Gold Standard)

DNA Read Calls	Positive	Negative	Total:
Positive	531	66	597
Negative	84	409	493
Total:	615	475	1,090

Assay performance by species

Measures	Global	PF	PM	PO	PV
Percent agreement	86.2	87.3	81.5	83.9	Could not be calculated based on the presence of 0 cells
Sensitivity	86.3	84.6	93.9	75.4	
Specificity	86.1	90.7	80.0	84.3	
PPV	88.9	91.6	35.4	21.0	
NPV	83.0	83.0	99.1	98.4	

Table 2.2 Performance measures of the new methodology on the global and individual species scale

Analyses for PV were not possible because of the extremely low incidence of this species in Ghana. There are no prior publications that report PV in Ghana. As there has not been a validated case of PV infection in Ghana, and we detected it as part of mixed infections, we evaluated the Mixed(all) (n=44) infection read composition. We observed that it was possible that these results were from cross-hybridization between the species that had not been completely corrected out by our methodology. We therefore developed an alternative calling method to deal with these non-specific infections detailed in the methods, and the new calls being found in Chapter 2 Appendix: Supplemental Table 2. We found that the threshold indicated above replicated the original PCR diagnoses in the Ghanaian population, with a weighted kappa statistic for species concordance of 97.1% between the new alternative calls and PCR. We hypothesized that these infection types occurred more frequently in the infections with an abnormally large number of reads binding to the reference, based on observation of the infections. We found that the distribution of the number of reads mapped in the Mixed(all) infections vs all other types of infection was significantly different using a welch's t-test ($p=0.03$), supporting our hypothesis. This shows that both the global calling, and our alternative culling mechanism for problematic infections, reliably replicate a previously accepted methodology.

Finally, we repeated the detection process on the 1000 Genomes Project samples but did not find a single *Plasmodium* read, regardless of the endemicity from which the sample was recruited. As all sequencing samples for the 1000 Genomes Project are cultured prior to sequencing, which would eliminate the

parasite, these results support the conclusion that the sequence-based assay is detecting *Plasmodium* only when truly present in the sequenced sample.

2.3.5 Assay validation

To assess the possibility of systematic bias in our assessment of Plasmodium infection, we addressed if the number of UMRs isolated from a WGS file influenced the results with respect to being false negative, false positive, true negative, or true positive. Specifically, the expectation is that there would be a significant difference in the average number of reads from a sample (positive samples being higher than negative) across categories if there was a bias. If this bias was present in our sample, the two positive categories (FP and TP) would be shifted up on the y-axis indicating higher in average read distribution than the negative categories (FN, and TN), especially in the case of false negatives.To test if this bias is present, an ANOVA analysis was performed to test if the number of UMRs in the initial file was different between the four two-by-two designations. There was no statistically significant difference in the mean number of original UMRs among the categories (Figure 2.5).

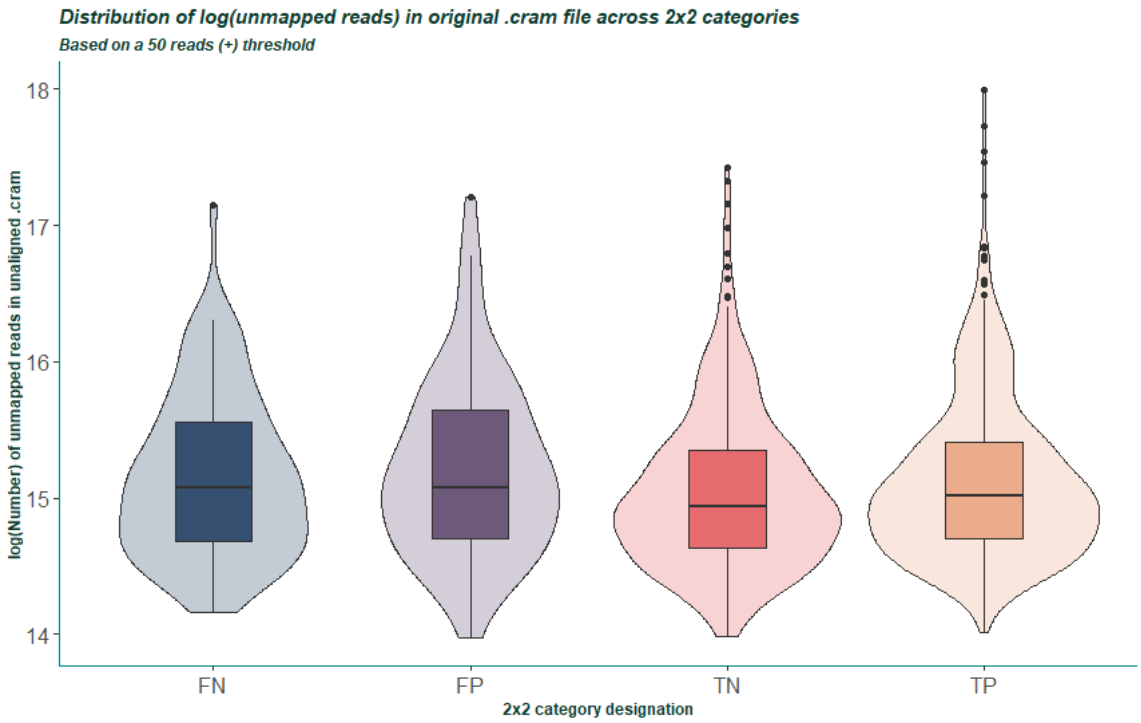


Figure 2.5 Violin plot of two-by-two designation versus log(number original UMRs)

Plot created using ggplot in RStudio v3.6.21. Shorthand of the two-by-two categories are as follows in order of the plot: true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). The data was transformed using log10 to illustrate the shape of the violins but not in the ANOVA analysis.

We found that the distribution of the initial UMRs was not differentially distributed between the categories, meaning that the UMR number is not influencing the eventual category designation of a sample. After validating that the negative individuals did not associate with initial UMR numbers, we validated that the reference genomes were aligning reads in a random manner. Under the null hypothesis, the reference should be aligning reads to Pla*smodium reference*

genomes randomly (i.e., without having significant hot spots for read depth); if not, it may indicate a sink that is picking up non-specific or poor-quality reads. To evaluate the global binding, we chose the species with the highest amount of coverage in Ghana when all aligned reads were combined into one file, which was PF at 103.2x coverage (PM=7.87x, PO=2.68x, and PV=0.074x). This new file allows us to evaluate where the binding occurs to ensure no biases. Because we are working with variable sizes of genomic regions to bind to after reference genome creation, a direct number of reads to number of reads comparison will inadequately represent the alignment across the genome. For example, a region that is only 1000 base pairs after filtering has a much lower chance of aligning a read than a region with 2.4 million base pairs. Therefore, if the binding is random, we expect that regions should align reads proportionately to their size. This can be seen in Figure 2.6, as the longer pieces seemingly have much larger histograms. For example, the largest piece of retained reference is on chromosome 13 (2.5 million bp) captured 1.75 million reads in the whole population, whereas one of the smallest on chromosome 5 (1133 bp) captured only 692.

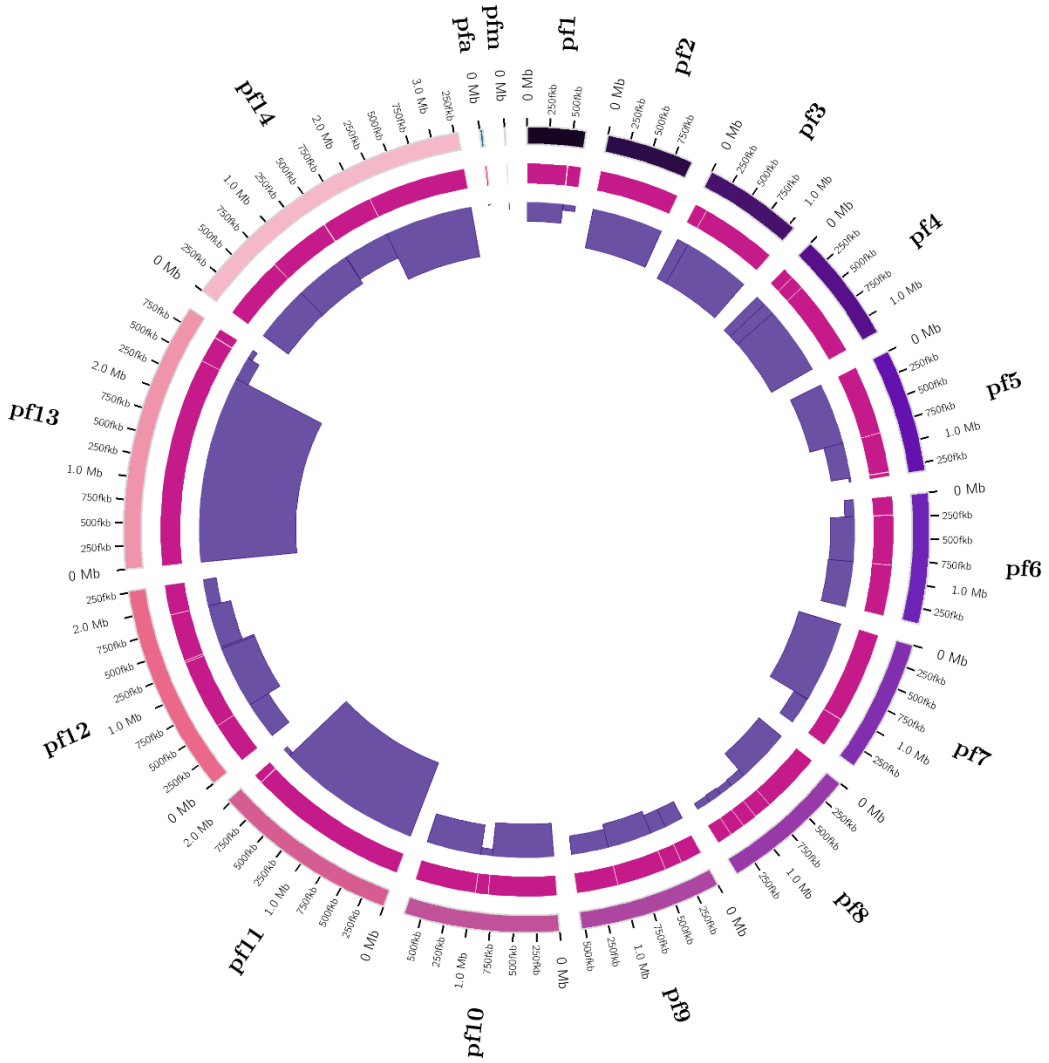


Figure 2.6 Representation of the number of reads captured by each piece of the *Plasmodium falciparum* reference genome

The concentric circles represent the following going from the outer circle: name of organelle or chromosome attached with general species (PF), general representation of pan-*Plasmodium falciparum* assembled chromosomes with size markers in 250kb increments (purple to pink gradient), *Plasmodium falciparum* unique regions genome (dark pink), histogram of the number of reads captured

by each genomic “piece.” White striped in the third circle represent removed regions.

We tested this assumption computationally by correlating the two variables. A plot of the number of reads a region aligned to versus the size in the PF genome found the relationship to be highly linear with a Pearson’s correlation coefficient of r=0.99. This indicates that these variables are highly related to each other, and the assay is likely performing without obvious bias (Figure 2.7)

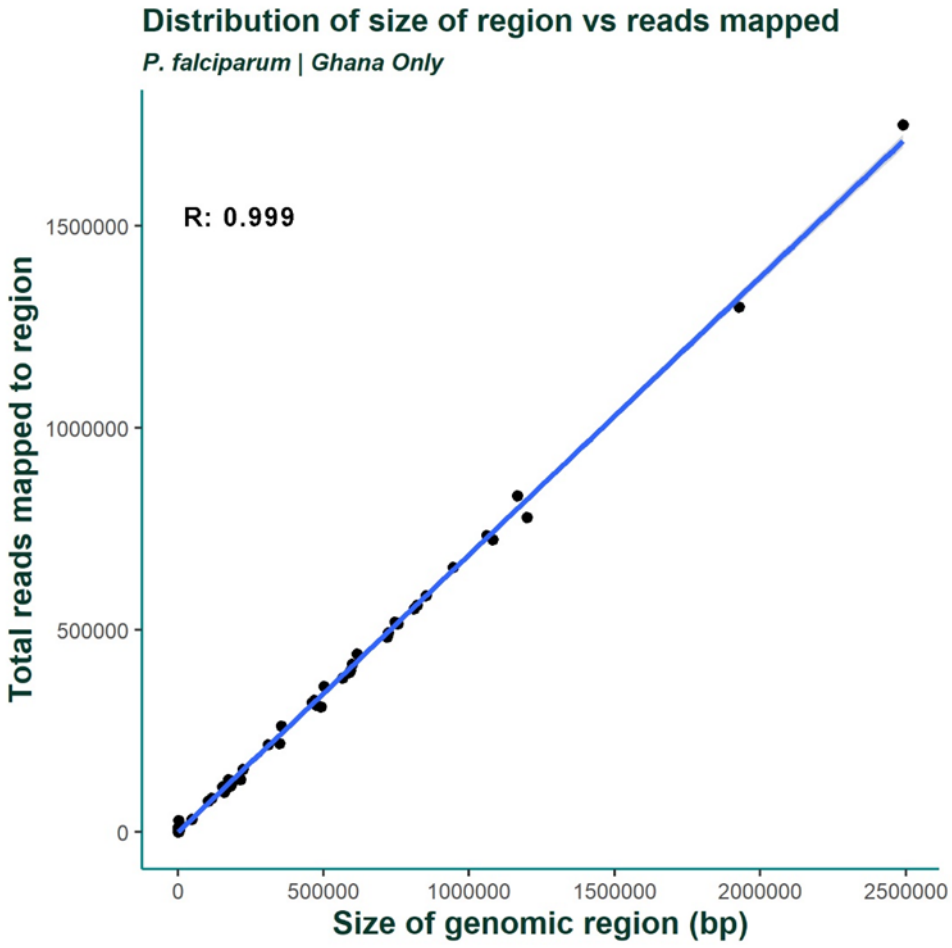


Figure 2.7 Correlation plot between number of reads aligned vs length of region in the *Plasmodium falciparum* genome

Plot was generated in RStudio using ggplot2, using only the *Plasmodium falciparum* aligned reads from Ghana, all correlation coefficients were generated using Pearson's[136].

Finally, to elucidate the true alignment landscape, we performed the aforementioned Bowtie2 alignment with an unfiltered PF reference genome to make plotting the entire genome possible. Using WGSCoveragePlotter, we found that the median numbers of reads aligned to each chromosome were approximately equal, and the only real dips in coverage happened near the telomeric and centromeric regions which are known to be hard to sequence, (Figure 2.8).

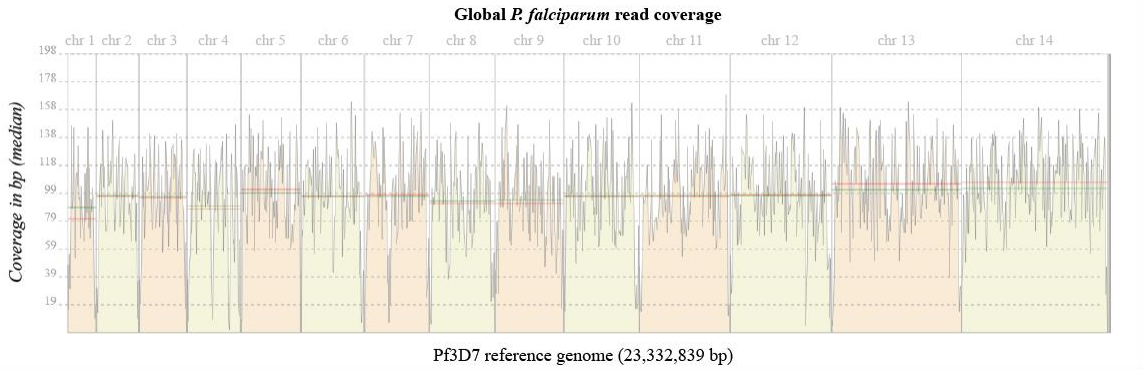


Figure 2.8 Global read coverage across the Plasmodium falciparum genome

Plot was generated using WGSCoveragePlotter which is a function of the Jvarkit Java tool. Each box across the x-axis represents one of the 14 nuclear chromosomes, y-axis is number of aligned reads. Red line represents mean read coverage, green represent median coverage.

2.4 Discussion

Our results highlight the power of whole genome sequencing from whole human blood to detect blood-borne *Plasmodium* infection. These results indicate we can use previously ignored WGS data to improve knowledge of malaria epidemiology even in the absence of malaria symptoms. There have been other known attempts at detecting *Plasmodium* from originally human samples, like Manske et al in 2013[137]. However, to the best of our knowledge, all other attempts have either taken steps to eliminate human sequence prior to processing or testing to ensure the *Plasmodium* read depths hit a certain level. Our pipeline is the first of its kind to prove that infection information can be gleaned from normal 30x human whole genome sequencing with a performance comparable to previously accepted detection methodologies. Using untargeted sequencing to detect the parasite can improve surveillance across Africa on the species level. The one exception for several countries in our population was *Plasmodium vivax* (PV), which showed potential false positivity in infections with extremely high amounts of reads aligning to another species. This species was problematic to resolve because there were not enough PCR-confirmed infections in our pilot country, Ghana, to compare the two methods because PV is extremely rare in this region of Africa[138].

Finding a genomic target that allows for sensitive species-specific detection of *Plasmodium* was one of the most important goals of this research. Each species has its own clinical presentation and research implication, and while assays of this type will likely never be used at point of care, it is still

important for surveillance efforts[139]. One notable example of why species differentiation is important has been mentioned directly above, the presence of PV infection in many African populations[140]. It was a long-standing belief in the malaria field that Duffy-negative individuals cannot be infected with PV and thus should be essentially absent from Africa, where the Duffy-negative (*FYES*) allele originated and is highly prevalent[33; 141]. This belief was challenged with the rise of PCR-based diagnostic methods, which were shown to be more sensitive to species identification and uncovered the parasite's ability to infect these "immune" individuals[141]. Since the initial discovery, this topic is one of great concern to the field as more and more evidence has contributed to overturning this assumption. Our pipeline must then be at least as good at resolving between species as the 18s rRNA methodology to be able to serve the field in this manner. This is one of the reasons we chose to compare our results to the PCR "gold standard".

A limitation of this study is that because our population lacked an actual PV infection when using only the Ghanaians, we cannot confirm that this pipeline can accurately detrect this species[138]. However, the high concordance that the pipeline shows with the PCR-based assay (81.5%-87.3% agreement), gives us confidence that it will accurately detect the parasite when present. Additionally, because this mechanism works with the discarded UMRs of human sequencing, every individual tested will automatically have data on their Duffy status. We expect a high level of Duffy-negative alleles in Ghana, as previous reports have estimated the prevalence of the allele to be near fixation in the country[138]. Taken

together, the lack of detection of PV infection in this population is highly expected.

While the density of information we found around PV infections through our assay is poor due to the population's geographic location, having over 103x coverage of the PF genome sequenced globally represents a potential source of other knowledge we could glean from this assay. Understanding that this research is meant to be a tool to further elimination goals leads us to the possibility of querying thse data for antimalarial resistance genes. An obvious roadblock in eliminating the parasite is its ability to circumvent traditional treatment options and persist in humans. The first occurrence of this would be the mutations in the *pf-crt* gene, which conferred resistance to chloroquine and was extremely prevalent in Africa. One study found that the resistant allele persists in up to 95% of cases queried in Ghana still in 2017, twelve years after the removal of chloroquine monotherapy[16; 142]. Thus, this gene would be an excellent candidate to explore in this population in future studies, as it is likely to be highly present in our collection years. Based on treatment recommendations during recruitment from 2002-2006, other genes to potentially query would be known antimalarial resistance loci for sulfadoxine-pyrimethamine combination therapy like *pf-dhps* and *pf-dhfr*[143; 144.] This therapy was used as a second-line treatment until 2004 and is still used as chemoprevention in pregnant women today[143]. Thus, we would expect to see resistance to this therapy in our parasite population if the data could be utilized in this manner. Outside of important genes for antimalarial resistance, it is also theoretically possible that we could query

genes that are used for evaluating multiplicity of infection (MOI), like merozoite surface protein 1 (*MSP)* and merozoite surface protein 2 (*MSP2)*[145]*.* Looking at the diversity of infection in our population is potentially interesting because some previous publications have tied the MOI to clinical outcomes in humans[100; 146; 147].

Implicating the human data in correlation to this newfound infection status will likely open up many other avenues of research outside of simply contributing to surveillance and epidemiology projects. One particular point of interest in the utilized Ghanaian population, which is outside the scope of this research, is that every individual collected was not thought to be acutely ill[127]. As part of enrollment into the original study, individuals had to verbally confirm they were not currently and had not in the past two weeks been infected with anything. Yet, the assay found 40.5% (n=772) of the population was infected with the parasite. One possible explanation for this is that some of these infections may be individuals may have been recently inoculated with the parasite and have not reached symptomatic parasite density. Or, some individuals were perhaps missed by this exclusion criteria, both of which would contribute to the presence of positivity in our population. We posit a third and more likely explanation: most of these infected individuals were asymptomatic carriers of *Plasmodium*.

Asymptomatic carriers represent a real threat to the process of malaria elimination because they act as a reservoir for the parasite that cannot easily be targeted through normal interventions[148-150]. While Ghana has made substantial progress towards elimination by meeting the 2020 GTS milestone of 40% incidence reduction in 2021, the country has seen a plateau in reduction since

2018. There have been many explanations as to why countries are struggling to continue towards 95% incidence reduction, and the presence of this aforementioned asymptomatic reservoir is one of them. Specifically, publications have shown that asymptomatic infections can replenish *Anopheles* mosquitos and that traditional diagnostic techniques (RDT and microscopy) are not always sensitive enough to pick these individuals up as positive[150]. These techniques that miss these individuals show that continuing the research using DNA-based detection, like PCR and our assay, is imperative to uncovering the full infection scope of the population. A sizable asymptomatic adult population (73% of tested individuals) has already been noted as problematic in a nearby region, Ashanti, in Ghana previously[71]. Another study by Okyere et al., shows that asymptomatic infections exist within the younger population, albeit at lower levels (36.8%)[151]. Because this study shows that children under five can maintain asymptomatic status, the relationship between the ability to suppress symptoms and human variation is likely more complex than just age-related acquired immunity, which has been suggested in the past[152]. Therefore, creating a mechanism to intimately tie the parasite with the entire human genome variation, as our assay does, to investigate this relationship is genuinely beneficial to understanding this population and potentially eliminating it.

In this discussion, we've highlighted the potential avenues of further research that could be spurred from evaluating the UMRs from human WGS data for *Plasmodium* sequence. This pipeline's successful construction highlights how much useful information is discarded in best-practice WGS analyses. As more

sequencing studies become readily available, like the Three Million African Genomes and All of US, we may be able to understand the consequence of this research[113; 153]. In the 2016 GTS, the WHO highlighted that finding ways to utilize previously collected data is one avenue to contributing to our understanding of malaria because we can investigate on a much larger scale precisely what is happening with the parasite and why elimination roadblocks are occurring[47]. Through this pipeline, we have provided the field with an elegant way to accomplish this and open up all present and future human WGS projects to contribute to the end goal of global malaria elimination.

CHAPTER III: Exploring the epidemiology of TOPMed Africa6k

Portions of this Chapter are under revision in the following manuscript:

Olvany JM**,** Chan ER, Martin WB, Hansen MEB, Harris DN, Tishkoff SA, Williams SM, Zimmerman PA. Detection of species-specific Plasmodium infection using unmapped reads from human whole genome sequences.

3.1 Introduction

Having a functional pipeline to detect all Plasmodium infecitons opens the possibility of exploring the *Plasmodium* epidemiology across the entire Africa6K population. Africa continues to carry the majority of the malaria burden, making it the prime place to study the parasite to further elimination goals. Each country in Africa is differently affected by *Plasmodium* as there are several types of endemicity and stages of elimination that result in highly variable epidemiology of the parasite across the continent. As seven countries are represented in the Africa6k population, all with unique malaria fingerprints, the appearance of the parasite should differ among the populations we tested, if our new detection methodology is functioning correctly. Here we show that our novel pipeline can accurately evaluate *Plasmodium* in many different contexts, exposing the parasite's epidemiology across sub-Saharan Africa. As a further benefit to the public health sector, this methodology can assess the antimalarial drug resistance in the detected parasite population, which is of enormous consequence to malaria elimination goals. We evaluated five PF genes of relevance in Africa from 2000-2015 (*dhps*, *dhfr*, *mdr1*, *crt,* and *k13*) that confer resistance to sulfadoxine-pyrimethamine (SP), chloroquine, and artemisinin, respectively.

3.2 Methods for evaluating the epidemiology

3.2.1 Rationale

One of the core principles of many of the WHO malaria documents is developing new tools to help further research. This was noted in the 2016 Global

Technical Strategy, apropos to malaria epidemiology in general and in the most recent World Malaria Report for antimalarial resistance genes[29; 47]. Our pipeline offers utility to both of these calls. However, until methodologies of this nature, as our pipeline and those prroposed in this recent Nature Microbiology perspective, are popularized, the results will likely receive intense scrutiny[117]. Therefore, we explored the epidemiology of our population by contextualizing our estimates with available databases for malaria prevalence, like the WHO or the Malaria Atlas Project (MAP)[154]. Through these comparisons, we can argue that our results reasonably represent the parasite state in our collected populations, thus making the tool much more helpful in the field beyond our specific study. We extend this estimate and apply the model to our antimalarial query by making country-by-country comparisons to previously reported data from the years around our collection time points. Our findings highlight the utility of evaluating parasite/microbe populations through human WGS and could be of incredible utility for public health researchers moving forward.

3.2.2 Checking assay assumptions

While the random binding appeared reliable in the Ghanaian population where we explored the PF genome, we extended the analysis using all four species and all populations. To test that all four genomes have random sequencing patterns and no non-specific binding, using ggplot2 we plotted the number of reads aligned to a region vs. the size of the region. We tested the correlation between the variables using a Pearson's coefficient[155].

To ensure any differences in infection prevalence was due to actual differences in endemicity and not the density of UMRs available for query, we isolated the total UMRs in the original .cram file for each individual and compared the mean UMRs by country using ANOVA to check for significant differences. To examine the distribution of UMRs across our population, we graphed log(UMRs), stratified by country of origin, using a violin plot in ggplot2[155]. A summary of each population, basic demographics, UMRs distribution, and infection prevalence can be found in Table 3.1.

3.2.3 Epidemiology of Africa6k

Once the detection assay was constructed and validated using the Ghanaian population, we ran the remaining (n=4,553) TOPMed Africa6K populations through the pipeline detailed above[110]. We called infections using the same 50-read threshold for each *Plasmodium* species. Infection prevalence was determined for each country separately, and collection sites were located at the regional level plotted against infection data from the Malaria Atlas Project, using the R package *malariaAtlas* v.1.0.1[136]. Regions of collection were highlighted using longitude and latitude coordinates isolated from GADM in Rstudio and plotted using ggplot2 geom_polygon[156]. Generated values were also compared to the historical prevalence reported by country in World Malaria Reports from 2012 and 2016, based on which report covered the entire window of collection[157; 158]. We evaluated relative parasite species proportion through normalization of percent prevalence on the population level and graphed each summary using ggplot2 in Rstudio. The greater Africa map was also created through data built

into ggplot2, and the infection prevalence was sourced through the previously mentioned WHO World Malaria Reports. An average malaria prevalence was calculated for any country's collection period that spanned over more than one calendar year.

Any mixed infection which called all four species queried was evaluated more stringently to account for potentially missed cross-hybridization between the species and were alternatively called using the following parameter: the species with the highest amount of reads aligned was considered positive, and any other species that had at least 10% aligned reads of the highest species was considered positive

3.2.4 Evaluating antimalarial resistance genes

The coverage of the PF genome allowed for the examination of several genes of clinical interest, specifically those that confer resistance to antimalarials in use at the time of sample collection across sub-Saharan Africa (2000-2015). We looked for variation in two chloroquine resistance genes, *pf-mdr1* and *pf-crt*, and two for sulfadoxine-pyrimethamine (SP) resistance, *pf-dhfr*, and *pf-dhps* using SAMtools mpileup [63; 159-161]. We queried +/-500 bp around regions of interest in each gene. We scanned each infected individual for coverage at each position of interest. "Coverage" was defined as at least one aligned read covering codons of interest based on variants in the literature. For each sample with coverage, we counted each person with coverage only once towards the total, regardless of read depth for each subject. If any individual had evidence of a mutation, we counted them as a mutant individual at that position, regardless if

some of that individual's reads were the wild-type amino acid. For example, an infected person who appeared to have a heterogeneous infection (wt/mut) at a relevant site, one individual was added to the "Individuals with coverage" tally, and one "Individuals with mutations" was added to the total. All regions of variation were evaluated separately, other than amino acid positions 51 and 59 in *pf-dhfr,* amino acid positions 436 and 437 in *pf-dhps,* and amino acid positions 72-76 in *pf-crt,* which were close enough to evaluate collectively in assigning haplotypes [63; 160]. In these cases, we estimated the prevalence of the mutant alleles at both the individual and haplotype levels. Allele frequency estimates included only mutations at the positions reported in the literature, and all evidence of novel unreported mutations is noted separately. Variants in these genes are well characterized, and details of the expected wild type and missense mutations amino acids are presented in Chapter 3 Appendix: Supplemental Table 1. In the case of *pf-dhfr* (51 and 59) and *pf-dhps* (436 and 437), we investigated the collective inheritance of the alleles by calculating linkage disequilibrium, D', between the two codons in populations with coverage over 30 reads aligned[162; 163]. To determine if the haplotypes we observed significantly differed from the expected proportions, we performed a chi-squared test of independence (df=3).

Genes were chosen based on the predominant clinical care suggested for Africa for malaria during the early 2000s; chloroquine was the first-line treatment, and SP combination was the second-line treatment for uncomplicated malaria up until 2005[16; 164]. We also queried the *pf-k13* gene for artemisinin resistance, as

the drug was being transitioned into usage in the middle of the total collection period and was recommended as a first-line treatment by the WHO starting in 2006[17].

3.3 Results

3.3.1 Checking assay assumptions

Our original Ghanaian plot of the number of reads a region aligned versus its size in the *Plasmodium falciparum* genome found the relationship to be highly linear with a Pearson's correlation coefficient of r=0.990. To see that these validation mechanisms apply to all of our data, we extended this correlation analysis to all four species across the entire TOPMed Africa6K population. We found that the correlation between these two variables increases once expanded, with PF now having a Pearson's coefficient of 0.999. Both PM and PO also had a highly linear relationship between genomic fragment size and number of reads captures, at r=0.990 and r=0.831, respectively. Only PV had a low correlation coefficient. We hypothesize that the observed low correlation might be due to the extremely low density of infection information that makes the relationship noisy (Figure 3.1).

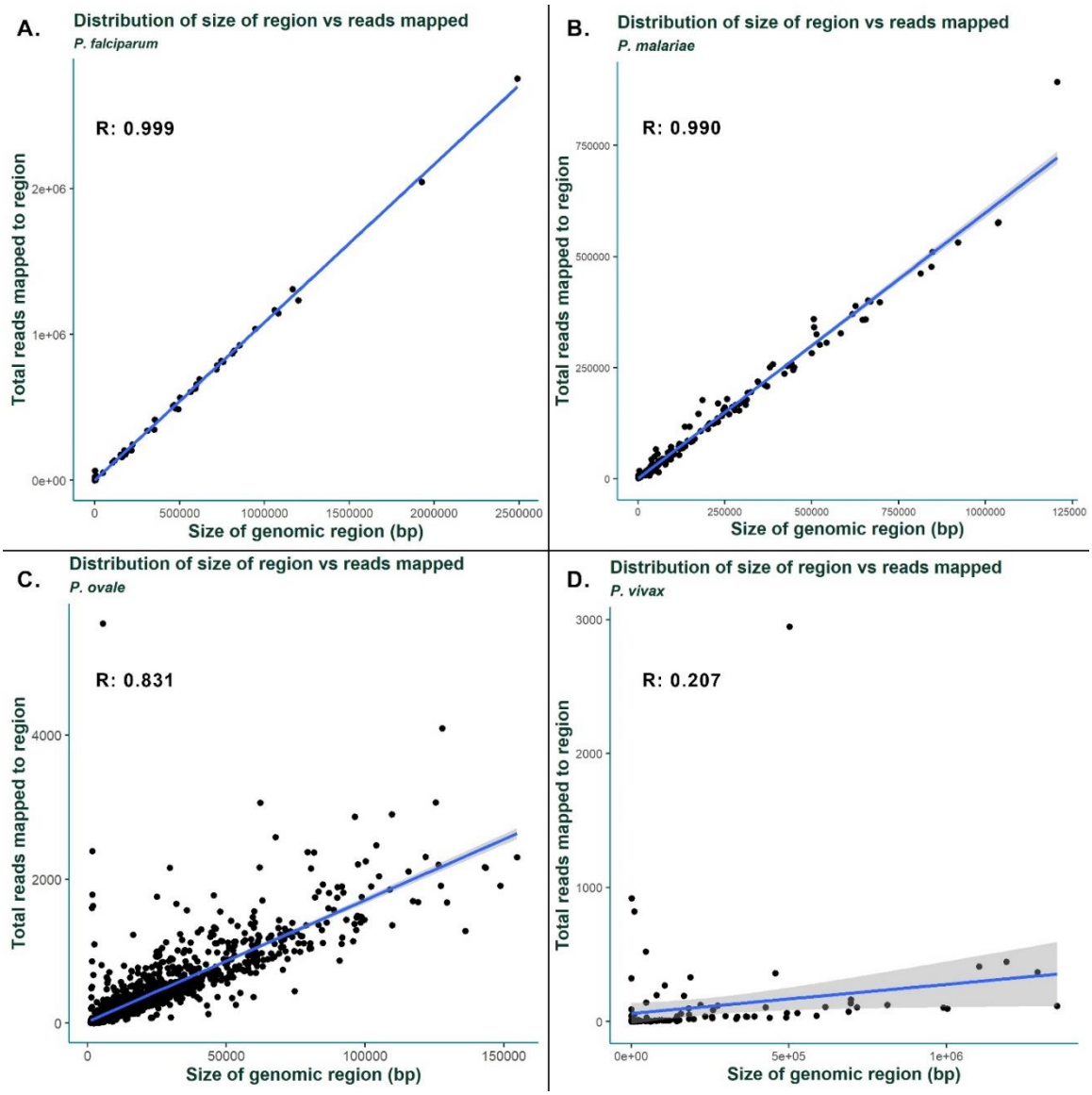


Figure 3.1 Correlation between size of genomic region used as a reference and number of reads aligned, stratified by *Plasmodium* species

Plots were generated in RStudio using ggplot2, all correlation coefficients were generated using Pearson's[136]. Species are indicated below the title and are presented in the following order starting from the top left: *Plasmodium falciparum* (A), *Plasmodium malariae* (B), *Plasmodium ovale* (C), *Plasmodium vivax* (D).

We found a variable range of infection prevalence between our populations (1.4%-40.5%). We noted in the rationale that this is not unexpected; nevertheless, we sought to computationally confirm that this variation is not due to a difference in the number of UMRs influencing the estimated infection rate. That is, it could be that populations with a higher number of UMRs going into the pipeline have higher levels of infection just because there were more sequencing reads to query for *Plasmodium.* To ensure that the difference in population measures was due to actual differences in inter-country variation in infection and not introduced by the difference in initial amount of UMRs per country, we determined the total UMRs available for query and plotted their distributions against Infection positivity (Figure 3.2).

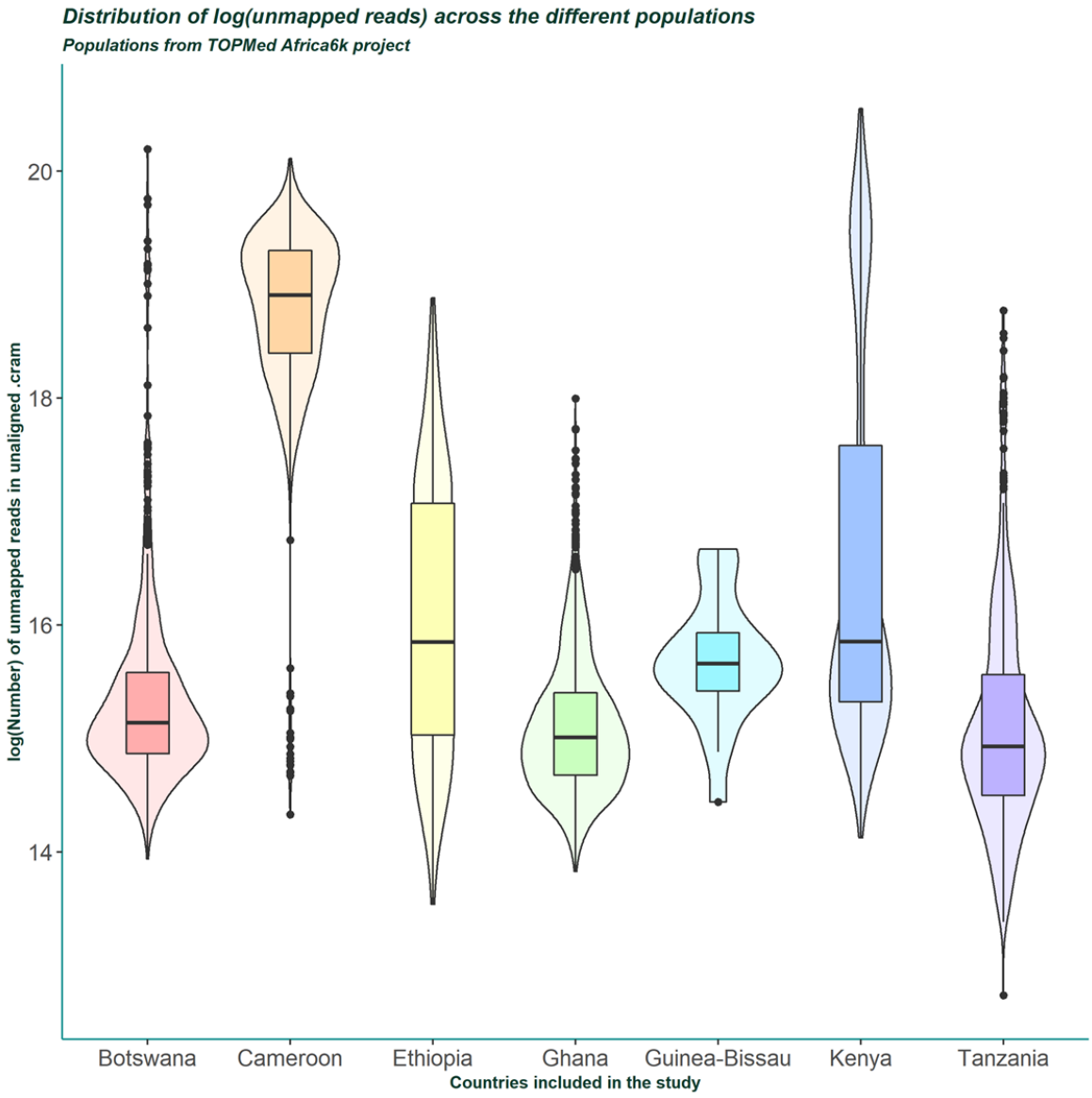


Figure 3.2 Violin plot of log(total UMRs) available for query vs country of collection.

Plots were generated in RStudio using ggplot2.Boxplots inside the violins have a middle line centered on each population's average UMRs in the original .cram file.

As seen in the figure the number of reads, and distribution, vary by population. There was a significant difference in the number of UMRs between the populations ($p<2x10-16$). However, this is likely due to variations in the DNA isolation process rather than just a larger number of reads (Cameroon and Kenya). This assumption is supported by the fact that the lowest average number of total UMRs (avg= 4,547,026) was the population with the highest Plasmodium infection rate (Ghana)(Table 3.1) (Chapter 3 Appendix: Supplemental Table 2).

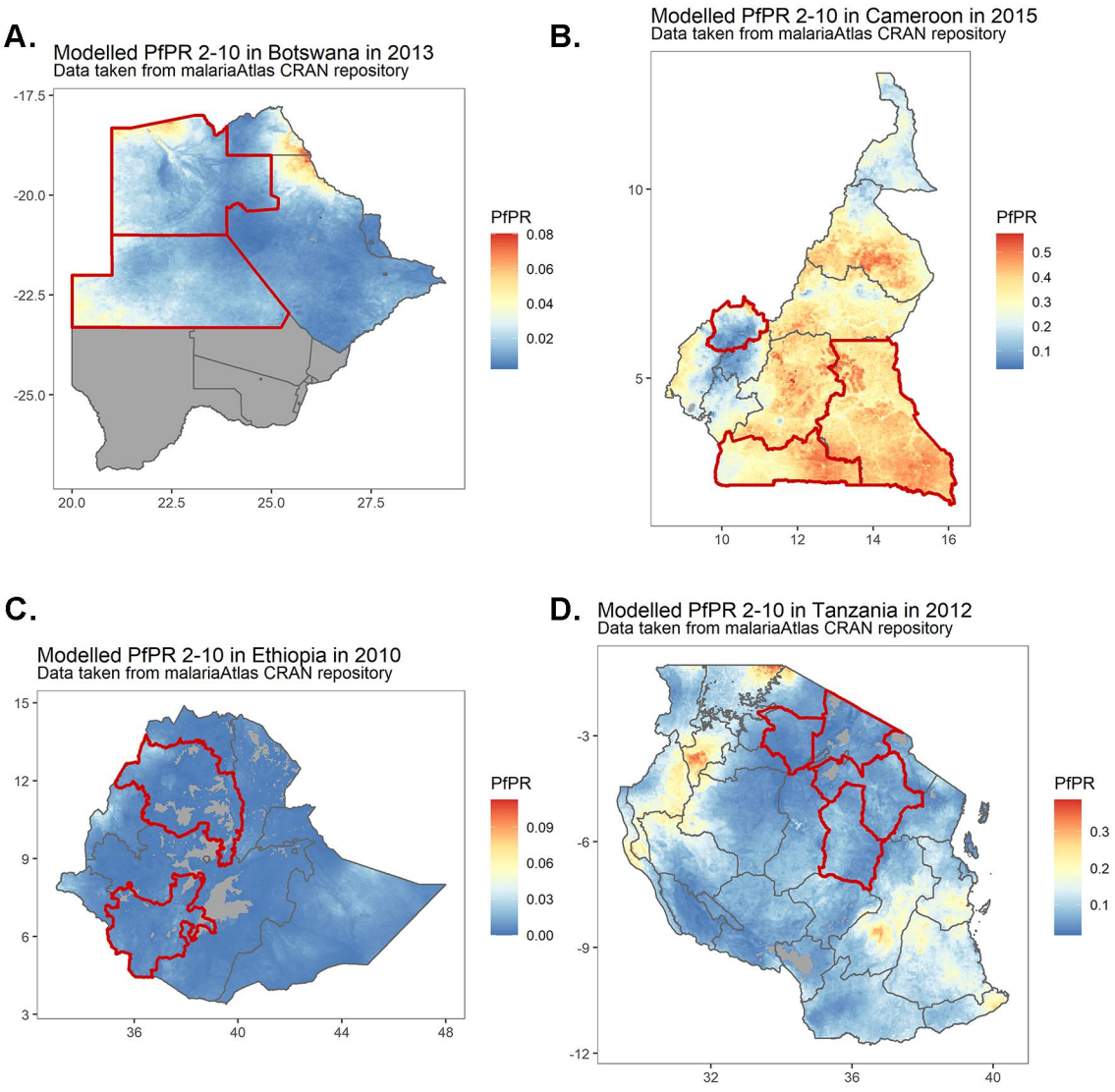
TOPMed Country Designation	**Number of Individuals *(n(% infected))***	**Number of unmapped reads *(mean(range))***	**Years of sample collection**	**Years of age *(mean±SD)***	**Female Sex (n(%))**
Botswana	653 (1.4%)	10,745,922 (1,131,964-588,201,938)	2013	38.3±15.4	401 (61.4%)
Cameroon	1,110 (25.7%)	172,816,097 (1,672,016-541,892,254)	2015	42.3±18.8	703 (63.3%)
Ethiopia	620 (7.3%)	18,634,438 (758,702-158,646,922)	2010	36.0±14.3	269 (43.4%)
Ghana	1904 (40.5%)	4,547,026 (1,012,360-65,293,224)	2002-2006	43.8±14.5	1108 (58.2%)
Guinea-Bissau	20 (20%)	7,408,687 (1,871,744-17,346,606)	*NA*	30.9±17.9	7 (35%)
Kenya	1,485 (9.0%)	74,644,620 (1,360,994-842,171,966)	2000-2007	42.2±17.5	754 (50.8%)
Tanzania	665 (8.9%)	6,899,562 (340,734-141,996,086)	2011-2012	43.4±18.1	335 (50.4%)

Table 3.1 Summary of Africa6K population demographics

NA Indicates that information was missing on all individuals in the population.

3.3.2 Epidemiology of Africa6k

The Africa6k population was collected originally as multiple separate studies, which were focused on ancestral population genetic surveys or cardiovascular studies, and not for the purpose of studying malaria. As mentioned before, endemicity varies greatly country-by-country, and even within countries depending on elevation, standing water, and seasonality. Since all collection was done without regard to malaria, it is possible that individuals collected may not live in a region where the parasite is present. Therefore, to confirm that all populations we collected hail from endemic regions we plotted the region of collection against the MAP PrPf raster maps provided by the *malariaAtlas* package in Rstudio. Through this we found that all individuals in our population are eligible to be utilized in this study (Figure 3.3).


A.
Modelled PfPR 2-10 in Botswana in 2013
Data taken from malariaAtlas CRAN repository
-17.5
-20.0
-22.5
-25.0
20.0
22.5
25.0
27.5
PfPR
0.08
0.06
0.04
0.02
B.
Modelled PfPR 2-10 in Cameroon in 2015
Data taken from malariaAtlas CRAN repository
10
5
10
12
14
16
PfPR
0.5
0.4
0.3
0.2
0.1
C.
Modelled PfPR 2-10 in Ethiopia in 2010
Data taken from malariaAtlas CRAN repository
15
12
9
6
3
36
40
44
48
PfPR
0.09
0.06
0.03
0.00
D.
Modelled PfPR 2-10 in Tanzania in 2012
Data taken from malariaAtlas CRAN repository
-3
-6
-9
-12
32
36
40
PfPR
0.3
0.2
0.1

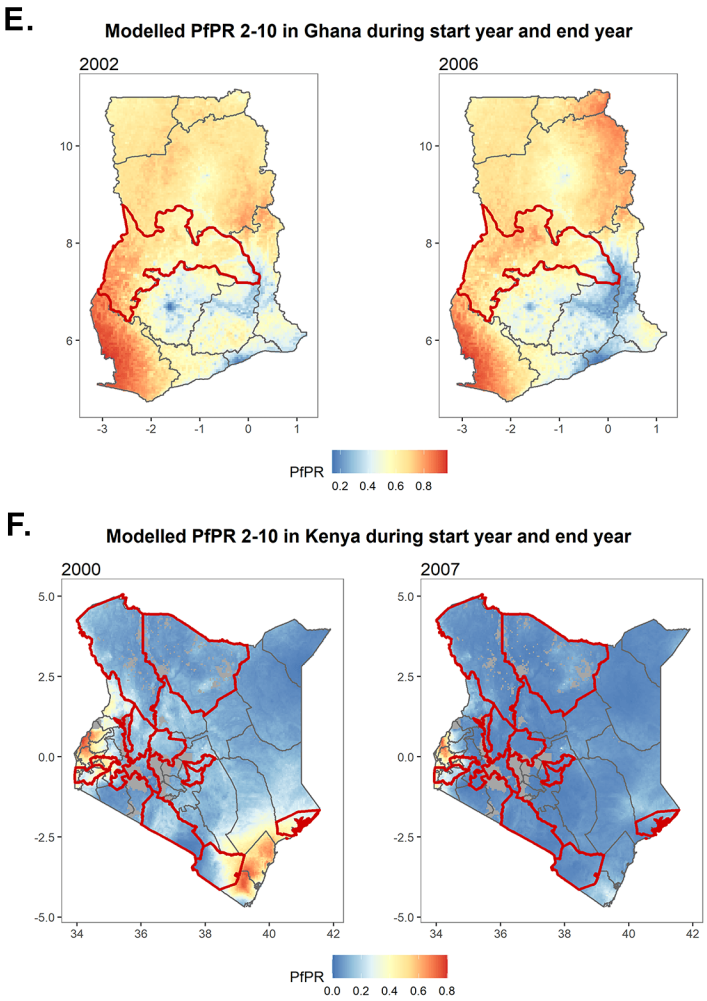


Figure 3.3 Visualization of collection sites for the TOPMed Africa6k population with *Plasmodium falciparum* prevalence data

The red-outlined regions are where collections took place. Malaria Atlas data was sourced from RStudio, the heat map, and resulting PfPR (*Plasmodium falciparum* parasite rate) scaled automatically with each map generation. For any collection that spanned 2+ years, we plotted the first and final years of the collection. The counties highlighted are as follows: A. Botswana (North-West and Ghanzi), B. Cameroon (Nord-Ouest, Est, and Sud), C. Ethiopia (Amhara and Southern Nations, Nationalities and Peoples), D. Tanzania (Arusha, Dodoma, Manyara,

and Simiyu), E. Ghana (Brong-Ahafo), and F.Kenya (Baringo, Bomet, Elgeyo-Marakwet, Embu, Homa Bay, Kajiado, Kisumu, Laikipia, Lamu, Marsabit, Nakuru, Taita Taveta, and Turkana)

Of the 6,457 individuals across all sites, 1,307 were infected (20.2%). There was a wide range of prevalence across countries, ranging from 1.4% (Botswana) to 40.5% (Ghana). In five of the seven countries, the PF was the dominant species (range= 62.2%-92.4%), followed by PM (range=42.2%-54.2%), then either PO (range=8.9%-31.3%) or PV (range=3.39%-14.3%). All countries, except Guinea-Bissau, had a mixture of both single-species infections and mixed (double, triple, quadruple) species infections. The complete distribution of *Plasmodium* positivity by country is found in Figure 3.4. Each species was called individually for infection positivity (>50 reads threshold) and then concatenated together to call final composition. Prevalence for each species in the tables are calculated as follows (# of that species infections/total infections)*100. As they are calculated species by species, the percentages will not add up to 100%. Species representation in the greater Africa map are relative proportions of each species so that these values will total 100%.

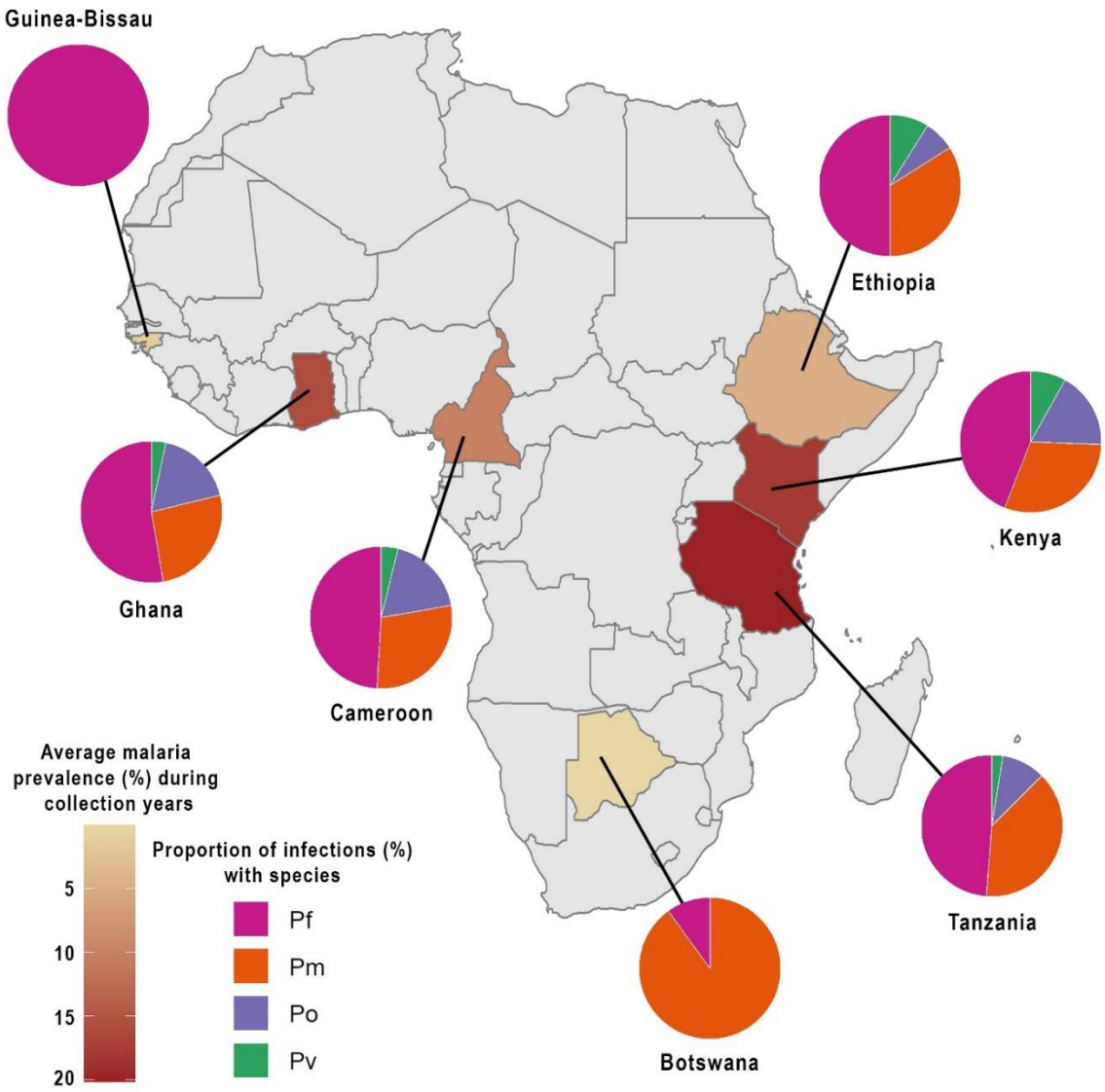


Figure 3.4 Visualization of *Plasmodium* species prevalence found in the TOPMed Africa6k UMRs

Each of the generated pie charts were created from the proportion of species infections we found in our data, normalized to 100, by counting each species positivity as one infection. For example, a single species Pf+ is one infection contributing to total infections, but a mixed infection (Pf+ and Pm+) counts as two towards the total, then the percentage was calculated as # of species

occurrences/total *100. Each country is colored according to the average reported infection prevalence across the entire collection window, found in either the 2012 WHO World malaria report (Ghana and Kenya) or the 2016 WHO World malaria report (Botswana, Cameroon, Ethiopia, Guinea-Bissau, and Tanzania). Prevalence was calculated from *"Reported malaria cases by method of confirmation"* and population totals from worldbank.org[157; 158].

Our lowest prevalence found was in Botswana, where total infection prevalence was 1.4%. In Botswana, there was an unusually high proportion of PM infections, as it was found in all infections, but with only 9 positive individuals in total, it could be by chance alone that this species was estimated to be common (Table 3.2). Not all regions in Botswana are malaria endemic, but all individuals in this population were collected from two regions that are naturally malarious (Figure 3.3A). There is no occurrence of PV infection found in Botswana. We found that our prevalence estimation for Botswana did fall within the range between the MAP and WHO (1.63 and 0.024, respectively) (Table 3.3). Unless otherwise noted, all following prevalence estimates for other countries fell within expected ranges.

Botswana (infected: 1.4%)
Jan 2013-Apr 2013

Distribution of Infections (DNA-based assay)	
Type of Infection	**Number of individuals**
Pf only	0
Po only	0
Pm only	8
Pv only	0
Mixed (Pf and Po)	0
Mixed (Pf and Pm)	1
Mixed (Pm and Po)	0
Mixed (Pf, Pm, Po)	0
Total Infected	9
Uninfected	644
Total	653
Parasite Species Prevalence in Infections	
Plasmodium species	Prevalence
P. falciparum	11.1%
P. vivax	0.0%
P. malariae	100.0%
P. ovale	0.0%

Table 3.2 Infection Composition Data for Botswana

Malaria Prevalence Comparison Table				
Country	Years	Our Assay Prevalence	MaP Prevalence (avg)[°]	WHO Report Prevalence (avg)[±]
Botswana	2013	1.4	1.63	0.024
Cameroon	2015	25.7	32.9	9.97
Ethiopia	2010	7.3	0.5955453	4.64
Ghana	2002-2006	40.5	58.1	16.04
Guinea-Bissau	UNK	20	UNK	UNK
Kenya	2000-2007	9.0	10.3	17.85
Tanzania	2011-2012	8.9	9.29	20.14

Table 3.3 Malaria Prevalence Comparison between our assay and two known infection repositories

±Calculated from "Reported malaria cases by method of confirmation" from WHO World Malaria Report 2012 (Ghana and Kenya) and World Malaria Report 2016 and population totals from worldbank.org : °Taken from *malariaAtlas "Plasmodium falciparum PR2-10"* average of the rasters used in the Figure 3.3. All prevalences are percentages.

Based on the collection sites, we expected Cameroon to be among the most infected populations in our population (Figure 3.3B). Our data agreed with that assumption because Cameroon was the second-highest prevalence at 25.7%. Cameroon had a mix of single and multiple species infections at 56.5% and 43.5%, respectively. Overall, the most common infection was single species PF infections (n=109) (Table 3.4). As there were no single species PV infections, and there were 17 of the Mixed (all) infections, we assumed they might be the

result of non-specific binding and alternatively called these infections using the method detailed above (Chapter 3 appendix: Supplemental table 3).

Cameroon (infected: 25.7%)
Feb 2015-Aug 2015

Distribution of Infections (DNA-based assay)	
Type of Infection	**Number of individuals**
Pf only	109
Po only	16
Pm only	36
Pv only	0
Mixed (Pf and Po)	22
Mixed (Pf and Pm)	55
Mixed (Pm and Po)	5
Mixed (Pf, Pm, Po)	25
Mixed (All)	17
Total Infected	285
Uninfected	825
Total	1,110
Parasite Species Prevalence in Infections	
Plasmodium species	Prevalence
P. falciparum	80.1%
P. vivax	6.3%
P. malariae	48.6%
P. ovale	30.1%

Table 3.4 Infection Composition Data for Cameroon

Ethiopia was an important country for the implications of our assay, as it was the population that provided validity of PV detection. There were 3

occurrences of single species PV infections, two mixed infections, and no occurrences of the problematic Mixed(all) infections. Thees data pointed to the ability of the pipeline to detect real PV infections when the parasite is truly there and not just non-specific binding noise. Overall, there was a higher prevalence of single species infections in Ethiopia versus the mixed species (80% vs. 20%), and as most other populations in our data, single-species PF infections were the most common. The proportions of each species deviated from the trend in this country, as PV infections were more common than PO infections (Table 3.5). Finally, we found a higher infection prevalence level than the WHO and MAP estimates (Table 3.3).

Ethiopia (infected: 7.3%)
Feb 2010-June 2010

Distribution of Infections (DNA-based assay)	
Type of Infection	**Number of individuals**
Pf only	21
Po only	1
Pm only	11
Pv only	3
Mixed (Pf and Po)	0
Mixed (Pf and Pm)	4
Mixed (Pf and Pv)	1
Mixed (Pm and Po)	2
Mixed (Pf, Pm, Pv)	1
Mixed (Pf, Pm, Po)	1
Mixed (All)	0
Total Infected	45
Uninfected	575
Total	620
Parasite Species Prevalence in Infections	
Plasmodium species	Prevalence
P. falciparum	62.2%
P. vivax	11.1%
P. malariae	42.2%
P. ovale	8.9%

Table 3.5 Infection Composition Data for Ethiopia

The entire assay was built on this population; however, the epidemiology of it generally was not discussed in the last chapter. Ghana was our most infected population, which aligns with what is seen in the collection graph

because that graph had the highest upper limit for the heat map and dense infection coloration (Figure 3.3E). The data found in Ghana mimicked much of the other information we found, where the split between mixed infections and single-species was similar (46.5% and 53.5%, respectively), and single-species PF infections were the most common. Ghana also had the highest prevalence of Mixed(all) infections but was the only country where we could validate the alternative calls with PCR. As stated, the weighted kappa statistic for species concordance was 97.1% between the alternative WGS calls and PCR. Thus, the high correlation provides evidence that these problematic infections can be better understood through alternative calling. Overall, the ability to discern what these infections contain is of little consequence because all occurrences of Mixed(all) infections represent only 1.2% of all individuals in the entire Africa6k population (total n=77)(Table 3.6).

Ghana (infected 40.5%)
2002-2006

Distribution of Infections (DNA-based assay)	
Type of Infection	**Number of individuals**
Pf only	364
Po only	17
Pm only	32
Pv only	0
Mixed (Pf and Po)	38
Mixed (Pf and Pm)	134
Mixed (Pm and Po)	10
Mixed (Pf, Pm, Po)	133
Mixed (all)	44
Total Infected	772
Uninfected	1132
Total	1904
Parasite Species Prevalence in Infections	
Plasmodium species	Prevalence
P. falciparum	92.40%
P. vivax	5.70%
P. malariae	45.70%
P. ovale	31.30%

Table 3.6 Infection Composition Data for Ghana

Guinea-Bissau had 4 individuals that were *Plasmodium* positive, and all were PF infections (Table 3.7). We could not compare this population to other known databases as collection information was missing year data. As there were only 20 individuals from Guinea-Bissau, and all were from the Fulani population,

this would likely not reflect the actual epidemiology of the greater country regardless. Previous publications have shown that the Fulani have some natural resistance to infection, making them a poor representation of the general public, and infection prevalence estimated from a small group is likely to be more volatile[165].

Guinea-Bissau (infected: 20%)

Distribution of Infections (DNA-based assay)	
Type of Infection	**Number of individuals**
Pf only	4
Total Infected	4
Uninfected	16
Total	20
Parasite Species Prevalence in Infections	
Plasmodium species	Prevalence
P. falciparum	100.0%
P. vivax	0.0%
P. malariae	0.0%
P. ovale	0.0%

Table 3.7 Infection Composition Data for Guinea-Bissau

Kenya was our second largest population, and as it was collected in the early to mid-2000s, we initially expected it to be our second most infected population. However, once we plotted the collection sites, we noticed that most of our individuals were recruited from areas of lower malaria endemicity (Figure

3.3F). Thus, it was not unexpected that our prevalence estimation was under both the MAP and WHO estimates for the same years. Kenya also had slightly more single-species infections than the other two larger populations (Ghana and Cameroon) at 59.4%. However, as in other countries, PF single-species infections were the most prevalent. PV infections in Kenya were the most unique out of all populations because, other than Ethiopia, it is the only population where the parasite was present in infections other than Mixed(all). There were no occurrences of single-species PV infections, so these infections do not provide as much proof the assay is adequately picking up PV as those in Ethiopia.

Kenya (infected: 9.0%)
Mar 2000-July 2007

Distribution of Infections (DNA-based assay)	
Type of Infection	**Number of individuals**
Pf only	58
Po only	3
Pm only	18
Pv only	0
Mixed (Pf and Po)	2
Mixed (Pf and Pm)	15
Mixed (Pm and Po)	5
Mixed(Pm,Po, and Pv)	4
Mixed(Pf, Pm, and Pv)	1
Mixed (Pf, Pm, Po)	13
Mixed (All)	14
Total Infected	133
Uninfected	1352
Total	1485
Parasite Species Prevalence in Infections	
Plasmodium species	Prevalence
P. falciparum	77.4%
P. vivax	14.3%
P. malariae	52.6%
P. ovale	30.8%

Table 3.8 Infection Composition Data for Kenya

Our final population, Tanzania, also had an infection prevalence lower than the MAP and WHO. Single-species infections in this population were more prevalent than mixed-species infections at 76.3% and 23.7%, respectively. Of the

14 mixed infections, 2 were Mixed(all), and there were no other occurrences of PV infections in this population. All mixed infections contained PF; the most common infection was single species PF infections. A low prevalence of infection aligns with expectations based on the collection map generated in Figure 3.3D, as all individuals were collected in the middle of the country where malaria is less endemic.

Tanzania (infected: 8.9%)
Nov 2011-Mar 2012

Distribution of Infections (DNA-based assay)	
Type of Infection	**Number of individuals**
Pf only	25
Po only	1
Pm only	19
Pv only	0
Mixed (Pf and Po)	2
Mixed (Pf and Pm)	7
Mixed (Pm and Po)	0
Mixed (Pf, Pm, Po)	3
Mixed (All)	2
Total Infected	59
Uninfected	606
Total	665
Parasite Species Prevalence in Infections	
Plasmodium species	Prevalence
P. falciparum	66.1%
P. vivax	3.39%
P. malariae	54.2%
P. ovale	13.6%

Table 3.9 Infection Composition Data for Tanzania

PV is of particular interest in sub-Saharan Africa because the Duffy negative allele that confers resistance to infection with this species is very common[138]. As stated above, Ethiopia was the only country where we detected pure PV infection (n=3). We observed that it was possible that these results were from cross-hybridization between the species that had not been completely corrected out by our methodology[138]. Data from a region where PV infections are common and extremely endemic would be the easiest way to identify the regions of our reference genome that are non-specific. If the 1000 Genomes Project populations were sequenced directly from the sample, this would have been an excellent addition to these analyses. Alternative calls for the Africa6K population can be found in Chapter 3 Appendix: Supplemental Table 3.

3.3.3 Evaluating antimalarial resistance genes

We found that we can evaluate antimalarial resistance allele frequencies among our PF-positive individuals, especially in countries with higher infection burdens. Alleles that confer resistance to the combination sulfadoxine-pyrimethamine (SP) and chloroquine were expected among our samples as these were the standard of care either during (Ghana and Kenya) or directly before sample collection (Botswana, Cameroon, Ethiopia, Guinea-Bissau, and Tanzania)[166]. Because the WHO recommended the use of artemisinin-based therapies starting in 2006, it was possible that some of the later collected samples carry alleles known to confer resistance to these drugs. The following genes were queried in the UMRs: *p-dhps* (sulfadoxine), *pfdhfr* (pyrimethamine), *pfcrt* and *pfmdr1*

(chloroquine), and *pfk13* (artemisinin)[63; 160; 161]. Only three of our populations had >30 reads across the queried genes and therefore are the three populations with the most reliable measures of resistance (Cameroon, Kenya, and Ghana). The Guinea-Bissau population had only one read of partial coverage in the *p-mdr1* gene and thus was excluded from further consideration. A detailed report of coverage across the genes can be found in Chapter 3 Appendix: Supplemental Tables 4-9.

The SP treatment genes showed the highest frequencies of resistance alleles, with only one position out of nine known sites, *pfdhfr* I164L, having no evidence of mutant alleles across the populations (Table 3.10). Mutant alleles were found in all positions across the *pfdhps* gene in at least one population. The highest frequency mutant allele proportion was A437G, ranging from 0.50-1, and the three populations with the most coverage had mutant frequencies above 0.80. The haplotype frequency in the three populations of high coverage was also assessed for codons close enough to be on the same read, i.e., *pfdhps* 436 and 437 and *pfdhfr* 51 and 59. For Cameroon and Ghana, the double mutant haplotype of *pfdhps* 436/437 was the most prevalent at 0.538 and 0.665, respectively. In contrast, Kenya had a low frequency of the mutant allele at position 436 despite high frequencies at 437, making the mut/mut haplotype rare at 0.083. Conversely, for the haplotype at *pfdhfr* 51/59 the mut/mut haplotype was the most frequent for Cameroon and Kenya (0.90 and 0.80), while it was only the second most prevalent haplotype in Ghana at 0.275. For any position that had a mixture of both wild-type and mutant alleles, we investigated further by

determining the linkage disequilibrium across those two positions and tested for deviation from the expected haplotype distribution[162; 163]. We found extremely high levels of co-inheritance for the alleles at *pfdhps* for Ghana and Kenya (D'=0.91 and 1.0, respectively), but neither passed the threshold for significance using the chi-squared test. In contrast, haplotypes in Ghana at 51/59 *pfdhfr,* revealed evidence of linkage disequilibrium (D'=0.98, $p < 0.005$).

Gene	Codon	Wild-type amino acid	Mutant amino acid	Resistant allele frequency (%) by country					
				Cameroon	*Ethiopia*	*Ghana*	*Guinea-Bissau*	*Kenya*	*Tanzania*
sulfadoxine-pyrimethamine									
pfdhps	436	S	A	53.8	0	80.4	0	9.1	0
	437	A	G	100	50	82.7	0	90.9	57.1
	540	K	E	0	0	0	0	100	100
	581	A	G	18.2	0	0	0	0	25
	613	A	S	15.4	0	17.1	0	0	0
pfdhfr	51	N	I	100	100	43.8	0	100	100
	59	C	R	90	0	62.7	0	80	66.7
	108	S	N	100	100	74	0	100	100
	164	I	L	0	0	0	0	0	0
chloroquine									
pfmdr1	86	N	Y	50	50	25.3	0	61.5	25
	184	Y	F	55	100	61.6	0	57.1	33.3
	1034	S	C	0	0	0	0	0	0
	1042	N	D	0	0	0	0	0	0
	1246	D	Y	0	33.3	4.7	100	47.1	0
pfcrt	72	C	S	0	0	0	0	0	0
	73	V	^	0	0	0	0	0	0
	74	M	I	12.5	0	2.9	0	33.3	0
	75	N	E	0	0	0	0	0	0
	76	K	T	0	0	0	0	33.3	0
artemisinin									
pfk13	446	F	I	0	0	0	0	0	0
	476	M	I	0	0	0	0	0	0
	493	Y	H	0	0	0	0	0	0
	539	R	T	0	0	0	0	4.76	0
	543	I	T	0	0	0	0	0	0
	553	P	L	0	0	1.35	0	0	0

Table 3.10 Summary of antimalarial resistance allele prevalence across all relevant genes in the Africa6k populations

Our expectation of high resistance frequency in chloroquine-relevant genes, based on it being the first-line treatment until 2004, was not necessarily reflected in the data. There was some evidence of chloroquine resistance found in *pfmdr1* at positions N86Y (range= 0.25-0.615) and Y184F (range= 0.333-1). We found no evidence of resistance at positions S1034C and N1042D across all Africa6k populations. Surprisingly, the wild-type haplotype at *pfcrt,* CVMNK, was the most frequent haplotype across all populations at 0.957, with the most evidence of resistance happening at position M74I in three samples. The overall estimation of the mutant haplotypes at these positions may be underestimated because the gene overall had low coverage at positions 74-76, which we have detailed in Chapter 3 Appendix: Supplemental Table 10.

As expected, we also saw minimal variation in the artemisinin gene *pfk13,* regardless of population collection years. Only two of the variants of known consequence R539T and P553L were detected, and both of these mutations occurred in only one individual in the entire dataset.

Several of our genes in multiple populations had novel variations in the positions we queried. If novel variation was seen in an individual with no other evidence of previously noted alleles, meaning that all aligned reads had only this previously unreported, novel variation, that individual was not included in the summary table above, and they were noted in Chapter 3 Appendix: Supplemental Table 11. If an individual had a mixture of known allele variants and novel variants, the known

variation was included in the table above, and the novel variation was noted in Chapter 3 Appendix: Supplemental Table 12. The lack of density of coverage made stringent filters for calling variants impossible to apply, therefore, novel variation could not be validated for reporting as a true result.

3.4 Discussion

Extending our pipeline to the greater Africa6K population provided much-needed validity to our assay and revealed the utility of this analysis. One of the main benefits this data brought is the confirmation that the pipeline can detect true PV infection because in Ethiopia where the species was detected in an infection other than Mixed(all). The presence of PV in this population is known, thus showing that our assay can detect true infections, if present. We also showed that we could detect the allele frequencies of anti-malarial resistance alleles among infecting *Plasmodium falciparum* (PF). This observation is important for policy with respect to drug recommendations at the population level.

The majority of our prevalence estimates fall within the range between two major sources of malaria epidemiology, the WHO and the MAP. It is extremely encouraging that the numbers are within realistic bounds[154; 157; 158]. We would not expect either end of this range of prevalence (MAP generally high | WHO generally low) to match exactly with our estimates, because malaria infection rates are highly influenced by several factors that cannot be completely corrected for in this study. Our measures were unlikely to reach the level of infection rate reported by the Malaria Atlas Project, because their R package models their raster maps off of infection surveys in children 2-10[154]. Children are the

demographic in endemic regions who carry the most morbidity and mortality of malaria, and as people age they gain acquired immunity to the parasite, which generally reduces the amount of infection in the population, specifically symptomatic. Since our population focuses on adults, we would expect to find less infection than the MAP, which is how most of our data trends. We had one population that was lower than the WHO estimate for the same year, Tanzania, which was shown in our data to have a prevalence of infection at 8.9% (reported ranges: 9.29%-20.1%). A potential culprit of this lower estimate is the location and time of collection (November-March) because the individuals from this country were collected mainly from the north/eastern part of this country which has a low transmission season during most of the collection window starting in December and ending in March[167]. Conversely, the only higher prevalence outlier is Ethiopia, which had higher measures of infection than both the WHO and MAP estimate for the same year. There are several reasons why this might be true: Ethiopia has seasonal transmission of malaria, with the spring peak being from March to May, only a portion of the country is considered to be malarious, and the possibility of collecting a large portion of asymptomatic individuals who are often missed in traditional surveys[168]. The Ethiopia collection was done not only through the entire seasonal window but also partially in the region of the highest prevalence of at least one species (SNNPR), according to a recent meta-analysis[169]. All of these factors could contribute to why our prevalence was higher than both other sources. Our three largest populations, which fell within the range between the MAP and WHO, were collected over the longest spans of

time, from 6 months to several years. A more consistent collection process is likely to better capture what a population looks like throughout an entire year, which could better estimate important epidemiological variables like level of asymptomatic infection, differences in seasonal transmission, a more detailed look into species composition, and unusual outbreaks of infection.

Interestingly, our data seems to highlight that the lesser discussed species, such as *Plasmodium malariae* (PM) and *Plasmodium ovale* (PO)*,* exist in quite high levels in the adult population. Botswana, in particular in our data, had a high proportion of PM infection at 9/9 having the species in the composition. A deviation from expected ratios of *Plasmodium* species is documented previously in Botswana, as the prevalence of all species varies substantially between publications. For example, in two studies in children, one study showed a PF prevalence of 52.4% and a PM prevalence of 6.78%, whereas a similarly conducted study found the prevalences to be 2.23% and 3.35%, respectfully[170; 171]. Both papers also noted a much higher rate of PV infection in their study, which was absent from this country in our study entirely. With such low infection rates in our population (prevalence= 1.4%), however, the high prevalence of one species versus another could be by chance. Additionally, one of the two regions our samples were collected from, Ghanzi, which is considered epidemic rather than endemic. Thus, we could have captured a local outbreak which may have further inflated the prevalence of PM in our population[171]. Botswana was not the only country with unexpected levels of the lesser studied PM and PO*;* Ghana also showed significant deviation from the

levels noted by other malaria resources. In our data, the relative frequency of PM and PO was 45.7% and 31.3%, respectively, which is much higher than the National Malaria Programme numbers (2.7% and 0.7%, respectively)[172]. This could be due to molecular and nucleic methods of detection providing better resolution to the species composition than the commonly used RDTs or blood smears in the clinic[8]. This potential underestimate based on detection methodology is highlighted when comparing our data to an epidemiology study done in the same Ahafo region of Ghana in 2006, which found the adult prevalence to be only 22.8% by smear, whereas we had a prevalence of 52.1% in that same year[173]. Not all results differed with previously reported literature; however, one highly encouraging trend in our data that agrees with the general consensus of malaria epidemiology in Africa, is that in all countries except Botswana, PF is the dominant infecting species (range=62.2%-100%)[174]. Our results, and cited papers, seem to point to the fact that our understanding of parasite prevalence is highly variable and needs additional research in all affected populations. Expanding these types of casual estimates to larger populations through mechanisms such as our assay may help in clarifying the ratio of each species present in populations, which is incredibly important to moving forward in the process of malaria elimination.

Another equally important outcome of this research is the ability to evaluate possible anti-malarial resistance in parasites on a global scale. It is important to note that this type of analysis seems to be only possible in populations with a high enough density of infection (n~100) based on our data. It

seems that this threshold is easily overcome as long as at least 1,000 individuals are collected in the country of interest, which we successfully passed in Cameroon, Kenya, and Ghana. Our expectation for what mutations will be present in each of these countries changes based on years of collection and the corresponding therapy recommendations. There was a considerable shift in treatment recommendation that occurred in 2004 (enacted 2006), which moved chloroquine (CQ) out of the first-line treatment to artemisinin-based combination therapy (ACTs)[175]. Two of our populations were collected mostly before this switch, and possibly entirely before the recommendation was put into practice, Ghana and Kenya. In these populations, we expect a high level of resistance accumulating in both the CQ and SP drugs as they were the primary drugs utilized in these years. These mutations may also be present in the later collected Cameroon population as well, especially in SP-related genes, which is still used as a preventative for pregnancies. Still, we expect genes like *pf-crt* to revert to wild-type as it has been reported to do so after pressure is removed[176].

As expected, we saw a significant amount of resistance in SP-relevant genes, *pfdhps* and *pfdhfr*, across our populations. And generally, we found that our results are in high concordance with previous studies that used more traditional methods of querying antimalarial resistance, like PCR amplification. Our results for Cameroon closely mimic another study done three years after our collection, which has a major collection site, Mfou, in the Equatorial Facies where all our samples were recruited[177]. This study genotyped all four SNPs of interest that we evaluated in *pfdhfr* (51, 59, 108, and 164) and found that the first three

had mutation rates at or approaching 1.0, and residue 164 had only wild-type alleles. A result almost completely replicated in our own data; the only slight variation is in position 59, we found only 0.90 mutant allele prevalence versus their reported 0.993. As for the other SP-related gene *pfdhps* this article only evaluated three of the five alleles of interest we queried (437, 540, and 581). Their report generally agrees with our findings, with the highest amount of mutation accumulated on 437, then 581, then 540 at 0.948, 0.0185, and 0.0012 respectively, whereas we found 1.0, 0.182, and 0.0 mutant alleles at that same positions[177]. As with Cameroon, both Ghana and Kenya mutant allele frequencies fall in line with previously reported literature in these two genes. A publication by Duah et al. estimated the prevalence of mutant alleles in the same capital city, Sunyani, during a very similar timeframe, from 2003-2008[144]. The only results we had that differed from the estimates in this publication were we found a slightly higher level of mutation for both *pfdhfr* 108 and *pfdhps* 437, but both estimates were less than 0.05 different[144]. For Kenya, a study was done in two cities, Kombewa and Kakamega, over many years, focused on several of the same resistance alleles we have in this study (*pfdhfr* 51,59, and 108 | and *pfdhps* 437 and 540). All mutant allele prevalences we found fell within their reported range from 2003-2008, except one *pfdhps* 540, where we found a 1.0 prevalence of the mutant allele, and they found only up to 0.971 in Kakamega in 2005[178]. Several reports show the absence of the *pfdhfr* 164 mutation in Western Africa, which is replicated across all three of our populations. There were some variants that we chose to evaluate that were consistently unreported in *pfdhps* (436 and

613), but this could be because they arose as alleles of interest more recently in the field, as evidenced by their presence in a publication from Ghana in 2020[142].

Unlike the SP-relevant genes, our results for CQ resistance differ from the historically reported prevalence. In the most dramatic case, our near total absence of *pfcrt* mutant haplotypes, specifically K76T, was shown to be at nearly 1.0 prevalence in the late 1990s[15]. This allele is often used as a marker for treatment failure because it has been shown to be a highly sensitive marker of resistance. Thus, many earlier publications use this allele exclusively rather than looking at the entire haplotype over the 72-76 codon window. In our data, we found the presence of this K76T mutation only in the Kenyan population at 0.333 prevalence. Our estimation was substantially lower than both the 0.50-0.912 prevalence found in Hemming-Schroeder et al. and the >0.80 estimates by Frosch et al. in the same time frame[178; 179]. Outside of Kenya, this marker is not found in any of our other populations, despite reports showing the allele existed at greater than 0.60 frequency in the early 2000s Ghana and 0.115 prevalence in 2018 Cameroon[177; 180]. We assume that, based on our reported systematic sequencing errors across the haplotype, our estimates are deflated, but the lack of mutation seen in what data we do have is unexpected. Issues genotyping this locus are known to the field, as it can be highly polymorphic, so while our lack of coverage in this region is discouraging, it is not entirely unexpected[161]. Conversely, *pfmdr1* showed sustained resistance across our populations, specifically at positions 86,184, and 1246. In our population, position 184 has the most evidence of mutant allele prevalence. At this position, our prevalence in

Cameroon was smaller than previous reports, 0.55 versus 0.72, but the later study was collected several years earlier[181]. Similarly, while position 184 had the highest frequency of the mutation in Ghana in our study, our estimated 0.616 was lower than a similar study done in 2007 in Ghana, which reported 0.879[182]. Our Kenya estimates did fall within the range reported in other articles[178]. Despite our estimates not aligning perfectly with other reports, continued CQ resistance was not expected to be present in similar levels in populations collected later in the 2000s because of the removal of the drug as a treatment. However, according to several of the publications discussed above, the persistence of mutation in the *pfmdr1* is not unexpected as certain combinations of mutant genotypes (N86**Y**-**Y**184F-D1246**Y** triple) are being connected to resistance to modern therapies like lumefantrine[176; 178; 182]. Thus, even if the initial pressure of chloroquine was removed long before our collection of this population, these loci are still likely to be selected for and be present.

Here we have shown that you can accurately estimate the prevalence of all four species of *Plasmodium* and uncover potentially relevant parasite gene mutations through human whole genome sequencing data. Our results fit well within expected values for multiple levels of endemicity, transmission, and infection compositions of malaria across sub-Saharan Africa. We believe that it is incredibly important to revisit data from endemic malaria regions- because the analyses can fill in local data gaps on malaria prevalence and antimalarial resistance gene transmission that may have been missed on the country-level reports or may not have been well documented at all. In reaching for a goal of a

world free of malaria, the field and the governmental bodies taking on the task, must know exactly what they are up against, and providing more information can only help fill in the picture further. While our study focused solely on creating a mechanism for specifically detecting *Plasmodium* parasites and their anti-malaria genes, this type of data mining could be applied to the entirety of the microbiome of any sequenced tissue. Other diseases of consequence to Africa which could potentially be assessed in this way are tuberculosis, sleeping sickness, and filariasis[183-185]. Taken together, our results highlight how successful utilization of previous data can be an extraordinary tool for malaria, and public health, moving forward.

CHAPTER IV: Discussion, Conclusions, and Future Directions

4.1 Discussion

4.1.1 Overview

As human genome sequencing technology continues to develop and decrease in cost, the number of individuals sequenced will continue to grow. With each individual sequenced, researchers will uncover more about human genetic variation and global health. One particular global health issue that could benefit from this type of data that was highlighted in this thesis is malaria; a disease that half of the world's population is still at risk for despite many continuing efforts to reduce malaria burden. Consequentially, since 2015 the elimination of malaria has seemingly reached a standstill in progress. To better understand why these roadblocks are occurring, the WHO determined that more research needs to be conducted on the problem. In this dissertation, we aimed to utilize the expanding field of genetics to develop a tool capable of fostering more malaria research. We proposed that the discarded UMRs that result from normal WGS harnessed the ability to detect malaria in endemic populations in a species-specific manner retrospectively. My results demonstrate that PCR-free WGS data can be mined adequately to determine species of *Plasmodium* infection compared to already accepted methodologies and can evaluate the parasite for relevant antimalarial genes. Thus opening a whole new type of data in previously collected endemic populations to elucidate the true burden of malaria and contribute to elimination efforts.

4.1.2 Summary

In Chapter 2, we identified which genomic target from *Plasmodium* (mitochondrial, API, or total) performed the best compared to the 18s rRNA NAAT assay. Using a portion of the TOPMed Africa6k population (Ghana - n=1904) and a threshold of 50 reads aligned to call positivity, we found that the total genome distinguished between the four species the best and agreed with the accepted methodology 86.2% of the time. The novel diagnostic assay had no obvious biases, such as infection outcome being determined by the raw number of UMRs, and reliably detected both single species and mixed infections in our population. One exception to the performance was no validated PV infections, which are known to be absent in Ghana, and therefore were expected to be missing. Additionally, a small portion of the population (n=44 - 2.3%) had evidence of remnant non-specific binding, as they were called Mixed(all) infections, but we found these could be corrected by using an alternative calling mechanism.

This study highlights the underappreciated UMRs in sequencing studies and their utility in furthering malaria research. Comparing our novel assay to the sensitive NAAT showed that we have successfully developed a retrospective method in determining *Plasmodium* infection from WGS data. Despite these advances, the utilization of this method is highly dependent on the expansion of sequencing studies worldwide, which could be a challenge in underserved populations. The pipeline has to be expanded to other populations to better characterize its performance, specifically in the case of PV infections. Regardless

of these limitations, advances in the tools to assess the true malaria burden is the only way to understand the roadblocks to elimination better and adjust efforts.

In Chapter 3, we extended the novel detection assay to the entire TOPMed Africa6k population and investigated its ability to capture malaria epidemiology in seven countries. We found that our estimated parasite prevalence generally fell within previously reported ranges defined by the WHO and MAP. As expected, the novel assay also detected PF as the dominant species of infection across Africa, which aligned with expectations. We also uncovered a significant portion of the underrepresented PM and PO infections, therefore adding to the burden of proof these parasites need to be added to general surveillance mechanisms. The last conclusion that could be made about our pipeline in the extended population is that it can adequately detect PV infections, which could not be validated in our first chapter because Ethiopia had three single species infections of PV. Overall, our pipeline seems to produce results that fit within the historical context of malaria and offer a novel tool to extend to other sequencing studies in endemic regions.

Along with the parasite prevalence, we proved the assay has the additional utility of evaluating antimalarial resistance genes in our three largest populations (Cameroon, Ghana, Kenya). We found significant evidence of resistance to SP-therapies in related genes, which aligned with our expectations because the drug was used through the entire collection timeframe thus keeping these regions under evolutionary pressure. Conversely, our findings for the CQ-related genes were unexpected, as we found less resistance than expected in our populations.

However, especially in the case of *pfcrt,* this locus has been historically difficult to characterize. This known difficulty may have contributed to the low coverage of the gene via sequencing.

Further studies are necessary to fully understand the burden of malaria worldwide to develop the best methods of elimination and potentially reaching malaria eradication.

4.2 Future Directions

4.2.1 Extend the novel detection method to other malaria-endemic regions

This dissertation showed that our novel computational method could adequately detect malaria in several different levels of endemicity; however, all data used in this study were from sub-Saharan Africa. Certain limitations are placed on this research based on the geographic distribution of several *Plasmodium* species. In particular, there are still concerns about how efficiently the methodology can detect PV infections, as we only found 3 single-species infections in Ethiopia in our data. The original purpose of including the 1000 Genomes Project in this research was to evaluate populations that may have a different distribution of the parasite species, but the passage of the cells used for sequencing through culture eliminated the possibility of assessing *Plasmodium.* Including studies from South America or Asia, where PV is more prevalent, will better elucidate the assay's ability to detect this species and identify regions that are still binding non-specifically. Additionally, it would be possible to include the detection of zoonotic species, like PK, in the methodology to inform malaria burden in populations where the species is relevant.

Expanding to other African sequencing studies will also have utility within the continent. For example, where this research had inconsistencies with WHO and MAP reported prevalence values, it mostly came from the collection being from a low-burden region or during the off-season in seasonally endemic countries. Including large-scale studies such as the Three Million African genomes will better elucidate the performance of the assay and allow for more in-depth comparisons of the data. One question that could be investigated in more extensive studies is the difference in infection composition between areas of high endemicity versus low endemicity because there have been some recent studies suggesting that PM and PO infections become problematic when PF is controlled[186-188]. A higher density of information could also allow for more granular comparisons in the antimalarial resistance genes further to understand the regional transmission and spread of these genes[189]. Extending the scope of the assay by including diverse populations from different endemic regions, both inside and outside of sub-Saharan Africa, will only improve its utility as a tool in research for the malaria field.

4.2.2 Further assess Plasmodium genes relevant to malaria epidemiology

Throughout this dissertation, I've discussed other genes that may be relevant to understanding *Plasmodium* epidemiology, like the *msp1* and *msp2* genes for MOI. While we focused on antimalarial-resistant genes for a proof of concept, with the type of data captured in the pipeline, there is no limit to the number of genes that can be queried using the same methodology highlighted in Chapter 3. One gene family of interest that could be evaluated are the *var* genes, such as

erythrocyte membrane protein 1 (*PfEMP1)*, which are often characterized for their high level of diversity in single infections and populations[190]. These genes are important in understanding the pathogenesis of PF infection because some variants are associated with different severities of infection[190]. Another extension of this research could occur if the pipeline was paired with malaria disease outcome because a recent study by Band et al. discovered that SNPs in the parasite genome correlate with the parasite overcoming HbS human resistance and causing disease (termed PF sickle-associated alleles), thus indicating more ways the parasite is co-evolving in response to the human host[191]. Adding to the burden of proof of how widespread these problematic variants are could identify more populations at risk for malaria than previously thought.

As more populations are evaluated with this mechanism, especially in regions with different densities of parasites, the number of genes consequential to the greater malaria field which can be evaluated will grow to include other species. In this dissertation, it was only prudent to investigate genes important to PF infections because that was the only species prevalent enough to have reasonable genome coverage. However, where PV infections are more prevalent, like in Asia and South America, that genome could be better covered and open for investigation. Several candidate genes of interest in antimalarial resistance have known orthologs to the PF loci in alternative species. For example, in PV the ones that would be important to investigate are PV *kelch-12 (pvk12)*, *pvcrt, pvmdr1*, *pvdhps,* and *pvdhfr*[192; 193]. Paired with better malaria data collected with the original population, the parasite genetics could theoretically be

investigated for novel resistance alleles through an association study using a resistant/susceptible outcome. This is important as it has been noted that less is known about resistance in PV infections because PF infections have always received more focus in research[192]. Although it is important to note that variant discovery analyses could be at the data limits, because there would have to be a high coverage of the PV genome for associations to be reasonably detected.

4.2.3 Evaluate the utility of this assay to capture the total blood microbiome

The natural progression of this research would first query the UMRs for known co-infections of malaria to determine which portion of the population was actively carrying other parasites. The greatest concern is co-infections that can complicate diagnoses or facilitate severe disease outcomes, like *Streptococcus pneumoniae, Leptospira,* and schistosomiasis[194-196]. Taking a closer look at the prevalence of these co-infections is important clinically because many of these diseases have overlapping symptoms, but divergent treatments, and treating only one infection can lead to unexpected mortality[9]. In addition to understanding the frequency of these co-infections, our data may help expose which co-infections are less harmful because most of our participants were collected under the assumption of no acute infection. Despite this thesis focusing only on the benefits UMRs can provide to malaria research and elimination, the pipeline can be adapted to detect any microbe. As mentioned previously in the discussion of Chapter 3, there are plenty of other parasitic diseases that circulate within the blood system that could be investigated in this manner.

The connection between a WGS pipeline and infectious disease is established in this research; however, there could be other interesting implications about the blood microbiome uncovered through UMRs. While the blood has historically been considered a “sterile” environment, there have been some indications that non-parasitic microbes may also circulate within the system and have consequences for non-communicable diseases[197]. In multiple recent publications, bacteria like *Porphyromonas gingivalis, Firmicutes, Proteobacteria,* were discovered in tissue traditionally considered to lack a microbiome and were connected with diseases like atherosclerosis, obesity, and cardiovascular disease[198-200]. Because at least one of the Africa6k original populations (Ghana) was collected as a cardiovascular study, these data would be ideal for exploring the potential for the novel detection method and adding to the burden of proof for the mentioned associations. This revolution in thinking surrounding the blood microbiome was led by molecular-based methods, making this type of analysis perfect to further our new understanding of the total human microbiota[198].

4.2.4 Utilize this new infection status to elucidate the correlation between human genetic variation and asymptomatic infection status

Prior to the completed sequencing of the Africa6K population, I conducted a small pilot study using the PCR-generated *Plasmodium* data as an infection outcome for logistic models in Ghana. As mentioned previously, the Ghanaian population was collected under the presumption that subjects had no acute illness at the time of enrollment. Therefore, a large portion of the n=696 PCR+ individuals in that population are likely asymptomatic parasite carriers. I

investigated the association between PCR infection status and several demographic and cardiovascular variables.

A first look into the data compared the distribution of the variables between the infected and uninfected populations and found (using chi-squared tests for categorical variables and ANOVA for continuous) that the following variables were significantly different ($p<0.01$) between the two populations: age, body mass index (BMI) category, blood pressure, fasting glucose, cholesterol, and high density lipoprotein cholesterol (HDL-C). Additionally, I found that many of the logistic regressions tested, being found categorically “healthy” in the cardiovascular variables mentioned made an individual more likely to be asymptomatically infected. Our theory of why this could be true is that these individuals who have no other cardiovascular health challenges, like diabetes or high blood pressure, have immune systems that can suppress the canonical symptoms of malaria, and therefore are more likely to be carrying the parasite aymptomatically. This could indicate that people who are not “healthy” are more likely to become symptomatic when challenged with the parasite, thus leading to rapid treatment. If that speculation is true then unhealthy individuals would statistically appear protected against asymptomatic infections, but this type of assumption cannot be adequately assessed in this data. The breakdown of the associations is as follows: protective against being PCR+ in our population ($p<0.05$): BMI category underweight (OR=0.623), BMI category overweight (OR=0.523), BMI category obese (OR=0.499), cholesterol over 200 mg/dL (OR= 0.545), glucose over 100 mg/dL (OR= 0.675), and triglycerides over 110 mg/dL

(OR= 0.703). While these findings are preliminary and need to be validated using the novel method across the populations, it highlights the utility of uncovering what allows individuals to remain asymptomatic.

There are two ways these associations can be explored in this population, the first being a complete look into the association with demographic variables, and more consequential to the type of data used here is the investigation into the human genetic variation. Throughout this thesis, several known loci were described which contribute to the susceptibility of disease, like HBB, G6PD, and Duffy, but our understanding of what contributes to asymptomatic disease is still evolving[201; 202]. A recent study between an asymptomatic and uninfected population has associated SNPs in *CD36* and *IL10* with infection status[203]. The sample size for this study was small (n=300); therefore, only candidate genes were assessed. Our population, Africa6k, has 20 times more individuals and is likely more powered to investigate the association landscape on the genome-wide scale. Studies of this type could be very informative in defining the asymptomatic population, and the information could then inform how best to target and treat them.

4.3 Conclusions

All research in this dissertation provides evidence of the utility of discarded UMRs from human WGS data. It shows that UMRs can be harnessed to assess malaria epidemiology across multiple populations and levels of endemicity. Additionally, it shows that the captured parasitic data can be used to investigate important antimalarial resistance genes. Development of this assay was the first

step in truly impacting the malaria field, but the assay needs to be applied to more extensive and geographically diverse populations to understand its full impact. Overall, this assay could be the next tool that spurs the research needed to solve roadblocks in malaria elimination, as called for by the WHO and provide the field with a path forward in finally targeting the parasite and significantly reducing the burden it has on sub-Saharan Africa.

Appendix

Chapter 2 Appendix

Supplemental Figure 1: Percent identity matrix generated by Clustal Omega

```
Percent Identity  Matrix - created by Clustal2.1


  1: PVmitochondria    100.00   93.78   93.69   87.46   92.13
  2: POCmitochondria    93.78  100.00   98.44   87.90   92.86
  3: POWmitochondria    93.69   98.44  100.00   87.65   92.47
  4: PFmitochondra      87.46   87.90   87.65  100.00   88.56
  5: PMmitochondria     92.13   92.86   92.47   88.56  100.00
```

Pairwise comparison of percent identity of each mtDNA sequence compared to all other species. Row number corresponds with column number. Shorthand representations are the same as the body, but POC is *Plasmodium ovale curtisi,* and POW is *Plasmodium ovale wallikeri.*

Supplemental Table 1: Retained Regions of the *Plasmodium* sequence by species

P. falciparum		
Pf3D7_01_v3	199	475065
Pf3D7_01_v3	475153	478298
Pf3D7_01_v3	478370	478654
Pf3D7_01_v3	478669	479902
Pf3D7_01_v3	480109	480493
Pf3D7_01_v3	480873	640538
Pf3D7_02_v3	78	334
Pf3D7_02_v3	583	767
Pf3D7_02_v3	1010	946686
Pf3D7_03_v3	0	305
Pf3D7_03_v3	609	764
Pf3D7_03_v3	1658	1888
Pf3D7_03_v3	3343	1063367
Pf3D7_03_v3	1066536	1066686
Pf3D7_04_v3	249	1199939
Pf3D7_04_v3	1200067	1200426
Pf3D7_05_v3	200	820934
Pf3D7_05_v3	820988	1289934
Pf3D7_05_v3	1290143	1290519
Pf3D7_05_v3	1290612	1290861

Pf3D7_05_v3	1290961	1291512
Pf3D7_05_v3	1291572	1293397
Pf3D7_05_v3	1293491	1293750
Pf3D7_05_v3	1293773	1294906
Pf3D7_05_v3	1295135	1295445
Pf3D7_05_v3	1295893	1343190
Pf3D7_05_v3	1343209	1343471
Pf3D7_06_v3	524	223752
Pf3D7_06_v3	223962	816302
Pf3D7_06_v3	816407	816614
Pf3D7_06_v3	816881	817106
Pf3D7_06_v3	817220	1416628
Pf3D7_06_v3	1417690	1417840
Pf3D7_06_v3	1417975	1418242
Pf3D7_07_v3	985	1218
Pf3D7_07_v3	1516	1084092
Pf3D7_07_v3	1084297	1084677
Pf3D7_07_v3	1084768	1085017
Pf3D7_07_v3	1085115	1085668
Pf3D7_07_v3	1085723	1087554
Pf3D7_07_v3	1087649	1087816
Pf3D7_07_v3	1087926	1089069
Pf3D7_07_v3	1089294	1089602
Pf3D7_07_v3	1090052	1444842
Pf3D7_08_v3	29	718122
Pf3D7_08_v3	718197	933687
Pf3D7_08_v3	933709	1113820
Pf3D7_08_v3	1113866	1288636
Pf3D7_08_v3	1288733	1472379
Pf3D7_08_v3	1472632	1472782
Pf3D7_09_v3	146	308
Pf3D7_09_v3	333	460401
Pf3D7_09_v3	460452	1048737
Pf3D7_09_v3	1048781	1541089
Pf3D7_10_v3	192	423
Pf3D7_10_v3	709	812732
Pf3D7_10_v3	812771	965014
Pf3D7_10_v3	965026	1686963
Pf3D7_11_v3	678	1927321
Pf3D7_11_v3	1927408	1931763
Pf3D7_11_v3	1931900	1932437
Pf3D7_11_v3	1932701	2037267
Pf3D7_11_v3	2037910	2038340
Pf3D7_12_v3	196	346
Pf3D7_12_v3	505	502403
Pf3D7_12_v3	502691	1356012
Pf3D7_12_v3	1356043	1921476
Pf3D7_12_v3	1921514	2271494
Pf3D7_13_v3	633	2490907
Pf3D7_13_v3	2490925	2801332

Pf3D7_13_v3	2801418	2805775
Pf3D7_13_v3	2805913	2806456
Pf3D7_13_v3	2806713	2922604
Pf3D7_13_v3	2922805	2922965
Pf3D7_13_v3	2923585	2924299
Pf3D7_14_v3	885	1036
Pf3D7_14_v3	1254	757526
Pf3D7_14_v3	757760	1503157
Pf3D7_14_v3	1503204	1508916
Pf3D7_14_v3	1508944	2125199
Pf3D7_14_v3	2125217	3291769
Pf3D7_API_v3	61	1234
Pf3D7_API_v3	1520	8534
Pf3D7_API_v3	8824	9219
Pf3D7_API_v3	9383	9888
Pf3D7_API_v3	10077	12615
Pf3D7_API_v3	12845	14501
Pf3D7_API_v3	14701	16725
Pf3D7_API_v3	16791	17009
Pf3D7_API_v3	17269	17609
Pf3D7_API_v3	17843	18013
Pf3D7_API_v3	18170	18707
Pf3D7_API_v3	18755	19219
Pf3D7_API_v3	19445	20438
Pf3D7_API_v3	20536	23442
Pf3D7_API_v3	23478	23695
Pf3D7_API_v3	23804	24068
Pf3D7_API_v3	24670	24868
Pf3D7_API_v3	25213	25500
Pf3D7_API_v3	26042	26297
Pf3D7_API_v3	26320	26832
Pf3D7_API_v3	27078	27428
Pf3D7_API_v3	27457	27718
Pf3D7_API_v3	28436	28620
Pf3D7_API_v3	28764	29427
Pf3D7_API_v3	29570	29822
Pf3D7_API_v3	30472	30735
Pf3D7_API_v3	30762	31112
Pf3D7_API_v3	31359	31869
Pf3D7_API_v3	31893	32149
Pf3D7_API_v3	32689	32977
Pf3D7_API_v3	33321	33521
Pf3D7_MIT_v3	103	302
Pf3D7_MIT_v3	575	1840
Pf3D7_MIT_v3	1909	4851
P.vivax before further refinement		
PvP01_01_v2	0	3206
PvP01_01_v2	3207	6976
PvP01_01_v2	6977	31364
PvP01_01_v2	31664	34472

PvP01_01_v2	34774	115127
PvP01_01_v2	115427	642734
PvP01_01_v2	643034	942527
PvP01_01_v2	942854	964260
PvP01_01_v2	964560	974163
PvP01_01_v2	974288	976017
PvP01_01_v2	976020	995782
PvP01_01_v2	995783	997578
PvP01_01_v2	999997	1011915
PvP01_01_v2	1012215	1017531
PvP01_01_v2	1017831	1019642
PvP01_01_v2	1020316	1021662
PvP01_02_v2	0	2342
PvP01_02_v2	2379	99670
PvP01_02_v2	99970	146372
PvP01_02_v2	148009	335626
PvP01_02_v2	335645	337079
PvP01_02_v2	337110	829342
PvP01_02_v2	829735	832181
PvP01_02_v2	832483	835724
PvP01_02_v2	836024	845659
PvP01_02_v2	845662	851180
PvP01_02_v2	851184	861062
PvP01_02_v2	861064	868604
PvP01_02_v2	868907	881783
PvP01_02_v2	881784	892300
PvP01_02_v2	892603	894689
PvP01_02_v2	894714	906901
PvP01_02_v2	906986	908337
PvP01_02_v2	908338	911330
PvP01_02_v2	911394	927319
PvP01_02_v2	927320	928814
PvP01_02_v2	928815	929867
PvP01_02_v2	930850	940342
PvP01_02_v2	940343	954995
PvP01_02_v2	955295	956327
PvP01_03_v2	0	144623
PvP01_03_v2	144645	535558
PvP01_03_v2	535858	554727
PvP01_03_v2	554827	774697
PvP01_03_v2	774997	896704
PvP01_04_v2	0	61229
PvP01_04_v2	61531	394779
PvP01_04_v2	395082	786757
PvP01_04_v2	787057	890633
PvP01_04_v2	890933	1002219
PvP01_04_v2	1002519	1010083
PvP01_04_v2	1010383	1012024
PvP01_05_v2	0	2259
PvP01_05_v2	2559	13976

PvP01_05_v2	15718	35237
PvP01_05_v2	35717	37955
PvP01_05_v2	39263	41585
PvP01_05_v2	42278	194891
PvP01_05_v2	196195	197720
PvP01_05_v2	198198	207882
PvP01_05_v2	208890	576707
PvP01_05_v2	577007	1016299
PvP01_05_v2	1016604	1272901
PvP01_05_v2	1273201	1508416
PvP01_05_v2	1508716	1524814
PvP01_06_v2	2	494205
PvP01_06_v2	494505	524678
PvP01_06_v2	524978	784156
PvP01_06_v2	784456	928771
PvP01_06_v2	928783	1034761
PvP01_06_v2	1035120	1038056
PvP01_06_v2	1038356	1040072
PvP01_06_v2	1040073	1042791
PvP01_07_v2	0	3841
PvP01_07_v2	3842	7960
PvP01_07_v2	8260	344789
PvP01_07_v2	345089	933225
PvP01_07_v2	933525	955895
PvP01_07_v2	956195	1413401
PvP01_07_v2	1413701	1433164
PvP01_07_v2	1433464	1481187
PvP01_07_v2	1481188	1502509
PvP01_07_v2	1502809	1517429
PvP01_07_v2	1518270	1550285
PvP01_07_v2	1550287	1562284
PvP01_07_v2	1562584	1583193
PvP01_07_v2	1583493	1637334
PvP01_07_v2	1637635	1647486
PvP01_07_v2	1647819	1652210
PvP01_08_v2	0	12805
PvP01_08_v2	12832	1204120
PvP01_08_v2	1204420	1346297
PvP01_08_v2	1346597	1395034
PvP01_08_v2	1395334	1440643
PvP01_08_v2	1440943	1698198
PvP01_08_v2	1698498	1702052
PvP01_08_v2	1702352	1729093
PvP01_08_v2	1729496	1745673
PvP01_08_v2	1745973	1756760
PvP01_08_v2	1757097	1761288
PvP01_09_v2	0	20756
PvP01_09_v2	20775	717172
PvP01_09_v2	717472	740398
PvP01_09_v2	740678	822600

PvP01_09_v2	822900	912677
PvP01_09_v2	912977	2201900
PvP01_09_v2	2201901	2232517
PvP01_09_v2	2232518	2235644
PvP01_09_v2	2235771	2237066
PvP01_10_v2	0	49000
PvP01_10_v2	49300	130266
PvP01_10_v2	130658	312501
PvP01_10_v2	312537	760454
PvP01_10_v2	760759	775516
PvP01_10_v2	775814	914665
PvP01_10_v2	914822	932461
PvP01_10_v2	932761	1090051
PvP01_10_v2	1090151	1516346
PvP01_10_v2	1517240	1522151
PvP01_10_v2	1522451	1524306
PvP01_10_v2	1524372	1525498
PvP01_10_v2	1525499	1548844
PvP01_11_v2	20	198598
PvP01_11_v2	198932	1550516
PvP01_11_v2	1550820	1733918
PvP01_11_v2	1734273	1757183
PvP01_11_v2	1757483	2121405
PvP01_11_v2	2121406	2124095
PvP01_11_v2	2124096	2126418
PvP01_11_v2	2126419	2128606
PvP01_11_v2	2129668	2131221
PvP01_12_v2	0	1957
PvP01_12_v2	2257	4302
PvP01_12_v2	7665	10380
PvP01_12_v2	10382	35836
PvP01_12_v2	35837	38078
PvP01_12_v2	38378	39804
PvP01_12_v2	40104	542310
PvP01_12_v2	542610	573880
PvP01_12_v2	574180	1565091
PvP01_12_v2	1565391	1583103
PvP01_12_v2	1583403	2272899
PvP01_12_v2	2273199	3085860
PvP01_12_v2	3086160	3087397
PvP01_12_v2	3087697	3105785
PvP01_12_v2	3106086	3120412
PvP01_12_v2	3120715	3182761
PvP01_13_v2	0	1290
PvP01_13_v2	1721	11323
PvP01_13_v2	11623	1115322
PvP01_13_v2	1115722	1282576
PvP01_13_v2	1282876	1898691
PvP01_13_v2	1898991	2082519
PvP01_13_v2	2082820	2090613

PvP01_13_v2	2090614	2092018
PvP01_13_v2	2092318	2093554
PvP01_14_v2	0	716327
PvP01_14_v2	716627	790517
PvP01_14_v2	790817	841487
PvP01_14_v2	841865	851713
PvP01_14_v2	852013	994293
PvP01_14_v2	994593	1269383
PvP01_14_v2	1269728	1345193
PvP01_14_v2	1345493	2042700
PvP01_14_v2	2043000	2102789
PvP01_14_v2	2103089	2104681
PvP01_14_v2	2104981	3108753
PvP01_14_v2	3109055	3131164
PvP01_14_v2	3131464	3133798
PvP01_14_v2	3133800	3153342
PvP01_API_v2	1621	3146
PvP01_API_v2	3499	8204
PvP01_API_v2	10945	12367
PvP01_API_v2	13025	14919
PvP01_API_v2	15194	17222
PvP01_API_v2	21232	22887
Transfer.PvP01_00_1.final	37	1118
Transfer.PvP01_00_1.final	1437	3513
Transfer.PvP01_00_1.final	3514	11253
Transfer.PvP01_00_1.final	11255	190162
Transfer.PvP01_00_1.final	190464	219766
Transfer.PvP01_00_1.final	220067	273193
Transfer.PvP01_00_1.final	273493	275231
Transfer.PvP01_00_1.final	275531	311703
Transfer.PvP01_00_1.final	312003	345722
Transfer.PvP01_00_1.final	345723	347400
Transfer.PvP01_00_1.final	347401	349295
Transfer.PvP01_00_1.final	349296	351937
Transfer.PvP01_00_1.final	351938	359765
Transfer.PvP01_00_1.final	359877	377976
Transfer.PvP01_00_1.final	378826	380073
Transfer.PvP01_00_1.final	380074	381091
Transfer.PvP01_00_1.final	381094	402819
Transfer.PvP01_00_10.final	4	19361
Transfer.PvP01_00_10.final	19662	43496
Transfer.PvP01_00_10.final	43497	52299
Transfer.PvP01_00_10.final	52600	55171
Transfer.PvP01_00_10.final	55172	56838
Transfer.PvP01_00_10.final	56843	63316
Transfer.PvP01_00_10.final	63618	65462
Transfer.PvP01_00_10.final	65463	76913
Transfer.PvP01_00_10.final	76915	80323
Transfer.PvP01_00_10.final	80681	90932
Transfer.PvP01_00_10.final	91232	102884

Transfer.PvP01_00_10.final	102886	104069
Transfer.PvP01_00_10.final	105186	116979
Transfer.PvP01_00_10.final	117279	124415
Transfer.PvP01_00_10.final	124417	126652
Transfer.PvP01_00_10.final	126654	132426
Transfer.PvP01_00_100.final	0	4184
Transfer.PvP01_00_101.final	2598	4144
Transfer.PvP01_00_102.final	59	1475
Transfer.PvP01_00_102.final	2986	4062
Transfer.PvP01_00_103.final	0	1918
Transfer.PvP01_00_104.final	0	1549
Transfer.PvP01_00_104.final	1870	3914
Transfer.PvP01_00_105.final	0	1071
Transfer.PvP01_00_105.final	2454	3825
Transfer.PvP01_00_106.final	0	1468
Transfer.PvP01_00_106.final	1836	3796
Transfer.PvP01_00_107.final	0	3768
Transfer.PvP01_00_108.final	0	2233
Transfer.PvP01_00_108.final	2533	3767
Transfer.PvP01_00_109.final	1540	3181
Transfer.PvP01_00_11.final	546	118980
Transfer.PvP01_00_11.final	118984	123358
Transfer.PvP01_00_110.final	0	3685
Transfer.PvP01_00_111.final	0	3501
Transfer.PvP01_00_113.final	0	1881
Transfer.PvP01_00_113.final	1955	3319
Transfer.PvP01_00_114.final	0	3094
Transfer.PvP01_00_115.final	0	1372
Transfer.PvP01_00_115.final	1782	2818
Transfer.PvP01_00_117.final	1107	2649
Transfer.PvP01_00_118.final	0	2482
Transfer.PvP01_00_119.final	0	2537
Transfer.PvP01_00_12.final	0	27562
Transfer.PvP01_00_12.final	27563	31387
Transfer.PvP01_00_12.final	31390	49392
Transfer.PvP01_00_12.final	49750	57544
Transfer.PvP01_00_12.final	57844	109966
Transfer.PvP01_00_12.final	110267	117582
Transfer.PvP01_00_12.final	117865	118908
Transfer.PvP01_00_120.final	0	2521
Transfer.PvP01_00_121.final	0	2382
Transfer.PvP01_00_122.final	0	2370
Transfer.PvP01_00_123.final	0	2164
Transfer.PvP01_00_124.final	0	2157
Transfer.PvP01_00_125.final	0	2093
Transfer.PvP01_00_127.final	0	2028
Transfer.PvP01_00_129.final	0	1859
Transfer.PvP01_00_13.final	0	80335
Transfer.PvP01_00_13.final	80635	103576
Transfer.PvP01_00_13.final	103578	106628

Transfer.PvP01_00_13.final	106629	108648
Transfer.PvP01_00_13.final	109645	111388
Transfer.PvP01_00_130.final	0	1902
Transfer.PvP01_00_131.final	0	1881
Transfer.PvP01_00_132.final	0	1873
Transfer.PvP01_00_133.final	1	1856
Transfer.PvP01_00_134.final	0	1846
Transfer.PvP01_00_135.final	0	1827
Transfer.PvP01_00_136.final	0	1820
Transfer.PvP01_00_137.final	1	1818
Transfer.PvP01_00_138.final	0	1793
Transfer.PvP01_00_139.final	0	1780
Transfer.PvP01_00_14.final	1104	2881
Transfer.PvP01_00_14.final	3904	6536
Transfer.PvP01_00_14.final	6836	12377
Transfer.PvP01_00_14.final	13702	16025
Transfer.PvP01_00_14.final	16250	17671
Transfer.PvP01_00_14.final	17934	25951
Transfer.PvP01_00_14.final	25954	27996
Transfer.PvP01_00_14.final	28369	29386
Transfer.PvP01_00_14.final	29388	78103
Transfer.PvP01_00_14.final	81042	101449
Transfer.PvP01_00_141.final	0	1736
Transfer.PvP01_00_143.final	0	1668
Transfer.PvP01_00_145.final	0	1648
Transfer.PvP01_00_147.final	0	1626
Transfer.PvP01_00_148.final	0	1615
Transfer.PvP01_00_15.final	1	1320
Transfer.PvP01_00_15.final	1620	7839
Transfer.PvP01_00_15.final	7840	11351
Transfer.PvP01_00_15.final	11352	28389
Transfer.PvP01_00_15.final	28689	33434
Transfer.PvP01_00_15.final	33435	34614
Transfer.PvP01_00_15.final	34615	38141
Transfer.PvP01_00_15.final	38142	41833
Transfer.PvP01_00_15.final	41834	54836
Transfer.PvP01_00_15.final	54839	56423
Transfer.PvP01_00_15.final	57661	77500
Transfer.PvP01_00_15.final	77502	86204
Transfer.PvP01_00_15.final	86208	88185
Transfer.PvP01_00_15.final	89710	90796
Transfer.PvP01_00_15.final	92546	94220
Transfer.PvP01_00_15.final	94520	96105
Transfer.PvP01_00_15.final	96473	98687
Transfer.PvP01_00_150.final	0	1613
Transfer.PvP01_00_151.final	0	1586
Transfer.PvP01_00_152.final	0	1538
Transfer.PvP01_00_153.final	0	1520
Transfer.PvP01_00_154.final	0	1522
Transfer.PvP01_00_155.final	0	1518

Transfer.PvP01_00_156.final	0	1516
Transfer.PvP01_00_157.final	72	1514
Transfer.PvP01_00_158.final	0	1437
Transfer.PvP01_00_159.final	0	1499
Transfer.PvP01_00_16.final	2	22111
Transfer.PvP01_00_16.final	22474	26334
Transfer.PvP01_00_16.final	26335	30950
Transfer.PvP01_00_16.final	31012	33361
Transfer.PvP01_00_16.final	33571	35567
Transfer.PvP01_00_16.final	35568	71373
Transfer.PvP01_00_16.final	71722	80673
Transfer.PvP01_00_16.final	80975	87719
Transfer.PvP01_00_16.final	87720	88816
Transfer.PvP01_00_16.final	89116	96962
Transfer.PvP01_00_16.final	96966	98334
Transfer.PvP01_00_160.final	1	1479
Transfer.PvP01_00_161.final	0	1499
Transfer.PvP01_00_162.final	0	1469
Transfer.PvP01_00_163.final	1	1457
Transfer.PvP01_00_164.final	0	1442
Transfer.PvP01_00_166.final	0	1425
Transfer.PvP01_00_167.final	0	1342
Transfer.PvP01_00_169.final	1	1315
Transfer.PvP01_00_17.final	1	49081
Transfer.PvP01_00_17.final	49381	52173
Transfer.PvP01_00_17.final	52536	54698
Transfer.PvP01_00_17.final	55471	61460
Transfer.PvP01_00_17.final	61465	92190
Transfer.PvP01_00_17.final	93478	94527
Transfer.PvP01_00_170.final	0	1405
Transfer.PvP01_00_171.final	0	1390
Transfer.PvP01_00_172.final	0	1382
Transfer.PvP01_00_173.final	0	1381
Transfer.PvP01_00_174.final	0	1353
Transfer.PvP01_00_175.final	0	1359
Transfer.PvP01_00_176.final	1	1357
Transfer.PvP01_00_177.final	0	1373
Transfer.PvP01_00_178.final	1	1353
Transfer.PvP01_00_179.final	1	1351
Transfer.PvP01_00_18.final	1	1620
Transfer.PvP01_00_18.final	3007	5038
Transfer.PvP01_00_18.final	5338	9174
Transfer.PvP01_00_18.final	9474	11918
Transfer.PvP01_00_18.final	11921	43002
Transfer.PvP01_00_18.final	43003	44559
Transfer.PvP01_00_18.final	44859	56593
Transfer.PvP01_00_18.final	57818	88940
Transfer.PvP01_00_18.final	88942	91632
Transfer.PvP01_00_180.final	1	1346
Transfer.PvP01_00_181.final	0	1335

Transfer.PvP01_00_182.final	0	1267
Transfer.PvP01_00_183.final	1	1308
Transfer.PvP01_00_184.final	2	1306
Transfer.PvP01_00_185.final	0	1295
Transfer.PvP01_00_186.final	60	1290
Transfer.PvP01_00_187.final	0	1289
Transfer.PvP01_00_188.final	0	1287
Transfer.PvP01_00_189.final	0	1274
Transfer.PvP01_00_19.final	0	24617
Transfer.PvP01_00_19.final	24917	55053
Transfer.PvP01_00_19.final	55353	65296
Transfer.PvP01_00_19.final	65596	70829
Transfer.PvP01_00_19.final	70830	73884
Transfer.PvP01_00_19.final	73885	86105
Transfer.PvP01_00_190.final	0	1271
Transfer.PvP01_00_191.final	0	1263
Transfer.PvP01_00_192.final	0	1257
Transfer.PvP01_00_193.final	0	1243
Transfer.PvP01_00_194.final	0	1241
Transfer.PvP01_00_195.final	1	1248
Transfer.PvP01_00_196.final	60	1212
Transfer.PvP01_00_197.final	0	1209
Transfer.PvP01_00_198.final	0	1203
Transfer.PvP01_00_199.final	0	1200
Transfer.PvP01_00_2.final	0	196673
Transfer.PvP01_00_2.final	196676	203316
Transfer.PvP01_00_2.final	203617	329406
Transfer.PvP01_00_20.final	0	41518
Transfer.PvP01_00_20.final	41822	47543
Transfer.PvP01_00_20.final	47843	68798
Transfer.PvP01_00_20.final	69098	72855
Transfer.PvP01_00_200.final	0	1200
Transfer.PvP01_00_201.final	0	1191
Transfer.PvP01_00_202.final	0	1191
Transfer.PvP01_00_203.final	0	1191
Transfer.PvP01_00_204.final	0	1173
Transfer.PvP01_00_206.final	0	1142
Transfer.PvP01_00_207.final	0	1139
Transfer.PvP01_00_208.final	1	1104
Transfer.PvP01_00_209.final	0	1062
Transfer.PvP01_00_21.final	0	13453
Transfer.PvP01_00_21.final	13454	17434
Transfer.PvP01_00_21.final	17436	38896
Transfer.PvP01_00_21.final	39196	43704
Transfer.PvP01_00_21.final	43705	61642
Transfer.PvP01_00_21.final	61942	71880
Transfer.PvP01_00_210.final	3	1136
Transfer.PvP01_00_211.final	0	1131
Transfer.PvP01_00_212.final	0	1130
Transfer.PvP01_00_213.final	2	1054

Transfer.PvP01_00_214.final	0	1124
Transfer.PvP01_00_215.final	2	1113
Transfer.PvP01_00_216.final	0	1107
Transfer.PvP01_00_217.final	0	1100
Transfer.PvP01_00_218.final	0	1092
Transfer.PvP01_00_219.final	0	1088
Transfer.PvP01_00_22.final	2393	3927
Transfer.PvP01_00_22.final	4981	9101
Transfer.PvP01_00_22.final	9404	12838
Transfer.PvP01_00_22.final	13138	22781
Transfer.PvP01_00_22.final	22800	39773
Transfer.PvP01_00_22.final	40551	44421
Transfer.PvP01_00_22.final	44722	46400
Transfer.PvP01_00_22.final	46700	61457
Transfer.PvP01_00_220.final	0	1084
Transfer.PvP01_00_221.final	0	1073
Transfer.PvP01_00_222.final	0	1072
Transfer.PvP01_00_223.final	0	1070
Transfer.PvP01_00_224.final	1	1070
Transfer.PvP01_00_225.final	0	1063
Transfer.PvP01_00_226.final	2	1061
Transfer.PvP01_00_228.final	1	1056
Transfer.PvP01_00_229.final	1	1055
Transfer.PvP01_00_23.final	0	7463
Transfer.PvP01_00_23.final	7465	8913
Transfer.PvP01_00_23.final	9213	11063
Transfer.PvP01_00_23.final	11064	16317
Transfer.PvP01_00_23.final	16320	19233
Transfer.PvP01_00_23.final	19533	24947
Transfer.PvP01_00_23.final	25247	31074
Transfer.PvP01_00_23.final	31078	57863
Transfer.PvP01_00_23.final	58820	62799
Transfer.PvP01_00_23.final	63448	64741
Transfer.PvP01_00_230.final	0	1053
Transfer.PvP01_00_231.final	0	1053
Transfer.PvP01_00_232.final	0	1041
Transfer.PvP01_00_233.final	2	1039
Transfer.PvP01_00_235.final	0	1020
Transfer.PvP01_00_236.final	1	1017
Transfer.PvP01_00_237.final	0	1015
Transfer.PvP01_00_238.final	0	1011
Transfer.PvP01_00_24.final	252385	
Transfer.PvP01_00_24.final	2592	7514
Transfer.PvP01_00_24.final	7816	53855
Transfer.PvP01_00_24.final	54430	60102
Transfer.PvP01_00_24.final	60103	61480
Transfer.PvP01_00_24.final	62208	64543
Transfer.PvP01_00_25.final	0	7680
Transfer.PvP01_00_25.final	7980	39993
Transfer.PvP01_00_25.final	39994	50290

Transfer.PvP01_00_25.final	50291	53740
Transfer.PvP01_00_25.final	54040	56991
Transfer.PvP01_00_25.final	57503	59933
Transfer.PvP01_00_26.final	0	1387
Transfer.PvP01_00_26.final	1388	31581
Transfer.PvP01_00_26.final	31883	33975
Transfer.PvP01_00_26.final	33976	36171
Transfer.PvP01_00_26.final	36471	41030
Transfer.PvP01_00_26.final	41331	44452
Transfer.PvP01_00_26.final	45382	51895
Transfer.PvP01_00_26.final	53477	54748
Transfer.PvP01_00_26.final	54750	58046
Transfer.PvP01_00_27.final	71	45609
Transfer.PvP01_00_27.final	45610	47156
Transfer.PvP01_00_27.final	47456	49643
Transfer.PvP01_00_27.final	49646	55477
Transfer.PvP01_00_28.final	1	1857
Transfer.PvP01_00_28.final	1858	4327
Transfer.PvP01_00_28.final	4677	16538
Transfer.PvP01_00_28.final	16838	19225
Transfer.PvP01_00_28.final	19226	29513
Transfer.PvP01_00_28.final	29815	39072
Transfer.PvP01_00_28.final	39372	47196
Transfer.PvP01_00_29.final	68	2567
Transfer.PvP01_00_29.final	2568	41113
Transfer.PvP01_00_29.final	41413	42965
Transfer.PvP01_00_29.final	43265	45364
Transfer.PvP01_00_3.final	0	6969
Transfer.PvP01_00_3.final	7286	38421
Transfer.PvP01_00_3.final	38422	51276
Transfer.PvP01_00_3.final	51577	67972
Transfer.PvP01_00_3.final	68332	72986
Transfer.PvP01_00_3.final	72987	110375
Transfer.PvP01_00_3.final	110675	148024
Transfer.PvP01_00_3.final	148324	175959
Transfer.PvP01_00_3.final	176274	201925
Transfer.PvP01_00_3.final	201927	204615
Transfer.PvP01_00_3.final	204916	206383
Transfer.PvP01_00_3.final	206696	213262
Transfer.PvP01_00_30.final	0	36707
Transfer.PvP01_00_30.final	36708	38171
Transfer.PvP01_00_30.final	38398	40765
Transfer.PvP01_00_30.final	41065	43143
Transfer.PvP01_00_300.final	3	1519
Transfer.PvP01_00_300.final	1819	5951
Transfer.PvP01_00_300.final	5952	7924
Transfer.PvP01_00_300.final	8276	29214
Transfer.PvP01_00_300.final	29591	30778
Transfer.PvP01_00_300.final	31078	35187
Transfer.PvP01_00_300.final	35188	43799

Transfer.PvP01_00_301.final	8	1608
Transfer.PvP01_00_31.final	0	10755
Transfer.PvP01_00_31.final	10756	18679
Transfer.PvP01_00_31.final	18680	22353
Transfer.PvP01_00_31.final	22354	25395
Transfer.PvP01_00_31.final	25699	28800
Transfer.PvP01_00_31.final	29100	33305
Transfer.PvP01_00_31.final	34243	39828
Transfer.PvP01_00_31.final	39831	41255
Transfer.PvP01_00_32.final	0	14903
Transfer.PvP01_00_32.final	14904	16880
Transfer.PvP01_00_32.final	16881	18292
Transfer.PvP01_00_32.final	20488	24488
Transfer.PvP01_00_32.final	24490	32124
Transfer.PvP01_00_32.final	32127	35428
Transfer.PvP01_00_32.final	35787	40940
Transfer.PvP01_00_33.final	0	6422
Transfer.PvP01_00_33.final	6783	23270
Transfer.PvP01_00_33.final	23350	31015
Transfer.PvP01_00_33.final	31017	36946
Transfer.PvP01_00_33.final	37419	39245
Transfer.PvP01_00_34.final	0	2388
Transfer.PvP01_00_34.final	2428	18856
Transfer.PvP01_00_34.final	19156	39018
Transfer.PvP01_00_35.final	0	10838
Transfer.PvP01_00_35.final	11138	19071
Transfer.PvP01_00_35.final	19074	22473
Transfer.PvP01_00_35.final	22474	24539
Transfer.PvP01_00_35.final	24540	26958
Transfer.PvP01_00_35.final	27561	32699
Transfer.PvP01_00_35.final	32700	35135
Transfer.PvP01_00_35.final	35506	36874
Transfer.PvP01_00_36.final	0	1302
Transfer.PvP01_00_36.final	1303	11171
Transfer.PvP01_00_36.final	11471	15893
Transfer.PvP01_00_36.final	15894	18022
Transfer.PvP01_00_36.final	18023	25778
Transfer.PvP01_00_36.final	26078	34801
Transfer.PvP01_00_37.final	0	34237
Transfer.PvP01_00_38.final	1	1542
Transfer.PvP01_00_38.final	1546	7294
Transfer.PvP01_00_38.final	7295	10551
Transfer.PvP01_00_38.final	10573	19913
Transfer.PvP01_00_38.final	20213	26487
Transfer.PvP01_00_38.final	26789	32429
Transfer.PvP01_00_38.final	32735	33746
Transfer.PvP01_00_39.final	0	1726
Transfer.PvP01_00_39.final	1727	30627
Transfer.PvP01_00_4.final	0	2645
Transfer.PvP01_00_4.final	2647	3793

Transfer.PvP01_00_4.final	3794	12468
Transfer.PvP01_00_4.final	12471	50985
Transfer.PvP01_00_4.final	51286	54000
Transfer.PvP01_00_4.final	54300	75982
Transfer.PvP01_00_4.final	76282	77955
Transfer.PvP01_00_4.final	77959	81756
Transfer.PvP01_00_4.final	81757	83141
Transfer.PvP01_00_4.final	83143	116230
Transfer.PvP01_00_4.final	116530	153529
Transfer.PvP01_00_4.final	153530	162249
Transfer.PvP01_00_4.final	162550	167607
Transfer.PvP01_00_40.final	1	14208
Transfer.PvP01_00_40.final	14508	19705
Transfer.PvP01_00_40.final	20006	27412
Transfer.PvP01_00_41.final	0	8713
Transfer.PvP01_00_41.final	9013	12449
Transfer.PvP01_00_41.final	12749	16589
Transfer.PvP01_00_41.final	16592	25912
Transfer.PvP01_00_41.final	25913	27035
Transfer.PvP01_00_42.final	0	18410
Transfer.PvP01_00_42.final	18770	22086
Transfer.PvP01_00_42.final	22088	24174
Transfer.PvP01_00_42.final	24474	26582
Transfer.PvP01_00_43.final	0	1209
Transfer.PvP01_00_43.final	1510	8804
Transfer.PvP01_00_43.final	9104	12477
Transfer.PvP01_00_43.final	12490	26325
Transfer.PvP01_00_44.final	59	2258
Transfer.PvP01_00_44.final	2525	21120
Transfer.PvP01_00_44.final	21121	22962
Transfer.PvP01_00_44.final	22963	25707
Transfer.PvP01_00_45.final	1	1807
Transfer.PvP01_00_45.final	2108	4879
Transfer.PvP01_00_45.final	6474	9655
Transfer.PvP01_00_45.final	9955	24787
Transfer.PvP01_00_46.final	0	1751
Transfer.PvP01_00_46.final	1753	6887
Transfer.PvP01_00_46.final	6976	11515
Transfer.PvP01_00_46.final	11816	15557
Transfer.PvP01_00_46.final	15558	19147
Transfer.PvP01_00_46.final	19148	22475
Transfer.PvP01_00_46.final	22775	24221
Transfer.PvP01_00_47.final	0	2248
Transfer.PvP01_00_47.final	2548	4189
Transfer.PvP01_00_47.final	4247	5675
Transfer.PvP01_00_47.final	5975	19938
Transfer.PvP01_00_47.final	19957	21500
Transfer.PvP01_00_48.final	0	5809
Transfer.PvP01_00_48.final	6109	8759
Transfer.PvP01_00_48.final	9059	12520

Transfer.PvP01_00_48.final	13700	17032
Transfer.PvP01_00_48.final	17333	21425
Transfer.PvP01_00_49.final	0	8683
Transfer.PvP01_00_49.final	8986	10228
Transfer.PvP01_00_49.final	10528	19783
Transfer.PvP01_00_49.final	20083	21386
Transfer.PvP01_00_5.final	45	13980
Transfer.PvP01_00_5.final	14280	54160
Transfer.PvP01_00_5.final	54460	57156
Transfer.PvP01_00_5.final	58263	66995
Transfer.PvP01_00_5.final	66996	70416
Transfer.PvP01_00_5.final	70417	143287
Transfer.PvP01_00_5.final	143290	146543
Transfer.PvP01_00_5.final	146545	148947
Transfer.PvP01_00_5.final	149247	150508
Transfer.PvP01_00_5.final	150509	152244
Transfer.PvP01_00_5.final	152245	153550
Transfer.PvP01_00_50.final	0	1709
Transfer.PvP01_00_50.final	1710	20857
Transfer.PvP01_00_51.final	77	1334
Transfer.PvP01_00_51.final	1984	3910
Transfer.PvP01_00_51.final	5341	9199
Transfer.PvP01_00_51.final	9499	10571
Transfer.PvP01_00_51.final	10871	20797
Transfer.PvP01_00_52.final	0	1957
Transfer.PvP01_00_52.final	3181	5243
Transfer.PvP01_00_52.final	5784	6818
Transfer.PvP01_00_52.final	6819	8139
Transfer.PvP01_00_52.final	8504	13837
Transfer.PvP01_00_52.final	14137	15852
Transfer.PvP01_00_52.final	16153	20050
Transfer.PvP01_00_53.final	60	1595
Transfer.PvP01_00_53.final	1596	5889
Transfer.PvP01_00_53.final	5893	15594
Transfer.PvP01_00_53.final	15595	19057
Transfer.PvP01_00_54.final	0	17847
Transfer.PvP01_00_55.final	0	2518
Transfer.PvP01_00_55.final	4422	5522
Transfer.PvP01_00_55.final	5821	8560
Transfer.PvP01_00_55.final	8562	10811
Transfer.PvP01_00_56.final	112	16781
Transfer.PvP01_00_57.final	0	2532
Transfer.PvP01_00_57.final	2533	5340
Transfer.PvP01_00_57.final	5640	9856
Transfer.PvP01_00_57.final	9862	16268
Transfer.PvP01_00_59.final	0	14706
Transfer.PvP01_00_6.final	0	15171
Transfer.PvP01_00_6.final	15172	43678
Transfer.PvP01_00_6.final	43728	49610
Transfer.PvP01_00_6.final	49912	50954

Transfer.PvP01_00_6.final	50959	97014
Transfer.PvP01_00_6.final	97015	128822
Transfer.PvP01_00_6.final	128825	131930
Transfer.PvP01_00_6.final	132233	148257
Transfer.PvP01_00_60.final	0	14647
Transfer.PvP01_00_61.final	0	6655
Transfer.PvP01_00_61.final	6656	9005
Transfer.PvP01_00_61.final	9006	13203
Transfer.PvP01_00_63.final	0	5760
Transfer.PvP01_00_64.final	45	1389
Transfer.PvP01_00_64.final	1689	3049
Transfer.PvP01_00_64.final	3428	5153
Transfer.PvP01_00_64.final	5454	6526
Transfer.PvP01_00_64.final	6859	8105
Transfer.PvP01_00_64.final	8106	10566
Transfer.PvP01_00_64.final	10568	12076
Transfer.PvP01_00_65.final	97	10230
Transfer.PvP01_00_65.final	10303	11608
Transfer.PvP01_00_66.final	0	6261
Transfer.PvP01_00_66.final	6561	9704
Transfer.PvP01_00_67.final	68	1107
Transfer.PvP01_00_67.final	1963	3823
Transfer.PvP01_00_67.final	3824	9635
Transfer.PvP01_00_67.final	9636	10706
Transfer.PvP01_00_69.final	0	1743
Transfer.PvP01_00_69.final	1745	4096
Transfer.PvP01_00_69.final	4396	7203
Transfer.PvP01_00_69.final	8566	10408
Transfer.PvP01_00_7.final	0	3376
Transfer.PvP01_00_7.final	3450	42260
Transfer.PvP01_00_7.final	42560	45539
Transfer.PvP01_00_7.final	45839	128805
Transfer.PvP01_00_7.final	129601	131297
Transfer.PvP01_00_7.final	131298	135793
Transfer.PvP01_00_7.final	136970	138489
Transfer.PvP01_00_70.final	1317	3897
Transfer.PvP01_00_70.final	4200	10338
Transfer.PvP01_00_71.final	2027	7215
Transfer.PvP01_00_71.final	7220	8751
Transfer.PvP01_00_71.final	8752	10042
Transfer.PvP01_00_73.final	0	2105
Transfer.PvP01_00_73.final	2405	7485
Transfer.PvP01_00_73.final	7786	9681
Transfer.PvP01_00_74.final	0	2111
Transfer.PvP01_00_74.final	2412	6044
Transfer.PvP01_00_74.final	6045	8944
Transfer.PvP01_00_75.final	0	7203
Transfer.PvP01_00_75.final	7828	8855
Transfer.PvP01_00_77.final	0	1018
Transfer.PvP01_00_77.final	1019	4425

Transfer.PvP01_00_77.final	4727	8140
Transfer.PvP01_00_78.final	0	1626
Transfer.PvP01_00_78.final	1926	3852
Transfer.PvP01_00_78.final	4152	5468
Transfer.PvP01_00_78.final	6856	8043
Transfer.PvP01_00_79.final	69	2258
Transfer.PvP01_00_79.final	2331	4055
Transfer.PvP01_00_79.final	6344	7952
Transfer.PvP01_00_8.final	0	8791
Transfer.PvP01_00_8.final	9091	10761
Transfer.PvP01_00_8.final	10765	19500
Transfer.PvP01_00_8.final	19501	46570
Transfer.PvP01_00_8.final	47790	48835
Transfer.PvP01_00_8.final	49135	116665
Transfer.PvP01_00_8.final	116707	129916
Transfer.PvP01_00_8.final	130216	131973
Transfer.PvP01_00_8.final	131974	136931
Transfer.PvP01_00_80.final	0	1889
Transfer.PvP01_00_80.final	1890	3675
Transfer.PvP01_00_80.final	3684	4985
Transfer.PvP01_00_80.final	4986	7518
Transfer.PvP01_00_81.final	0	2351
Transfer.PvP01_00_81.final	2352	7437
Transfer.PvP01_00_82.final	0	3229
Transfer.PvP01_00_82.final	4474	5562
Transfer.PvP01_00_82.final	5862	7384
Transfer.PvP01_00_83.final	0	3642
Transfer.PvP01_00_83.final	5008	6875
Transfer.PvP01_00_84.final	57	6603
Transfer.PvP01_00_85.final	1	1342
Transfer.PvP01_00_85.final	2340	3959
Transfer.PvP01_00_85.final	4320	6500
Transfer.PvP01_00_86.final	0	5134
Transfer.PvP01_00_87.final	0	1731
Transfer.PvP01_00_88.final	1	1752
Transfer.PvP01_00_88.final	2604	4317
Transfer.PvP01_00_88.final	4618	5908
Transfer.PvP01_00_89.final	0	2901
Transfer.PvP01_00_89.final	3202	4577
Transfer.PvP01_00_9.final	0	1428
Transfer.PvP01_00_9.final	2577	16506
Transfer.PvP01_00_9.final	16808	69338
Transfer.PvP01_00_9.final	69638	98737
Transfer.PvP01_00_9.final	99037	127172
Transfer.PvP01_00_9.final	127472	129210
Transfer.PvP01_00_9.final	129511	132465
Transfer.PvP01_00_90.final	0	1853
Transfer.PvP01_00_90.final	2782	3870
Transfer.PvP01_00_90.final	3871	5572
Transfer.PvP01_00_92.final	2716	5003

Transfer.PvP01_00_93.final	78	2012
Transfer.PvP01_00_93.final	2013	4744
Transfer.PvP01_00_94.final	1	1642
Transfer.PvP01_00_94.final	2915	4551
Transfer.PvP01_00_95.final	0	2865
Transfer.PvP01_00_95.final	3012	4453
Transfer.PvP01_00_96.final	0	1803
Transfer.PvP01_00_96.final	2986	4291
Transfer.PvP01_00_98.final	3	1181
Transfer.PvP01_00_98.final	3157	4312
Transfer.PvP01_00_99.final	0	1240
Transfer.PvP01_00_99.final	1540	4222
P. malariae		
PmUG01_01_v1	0	1035821
PmUG01_01_v1	1035843	1381516
PmUG01_02_v1	0	813519
PmUG01_03_v1	0	315512
PmUG01_03_v1	315536	320552
PmUG01_03_v1	320561	456935
PmUG01_03_v1	456947	1103899
PmUG01_03_v1	1105585	1109031
PmUG01_03_v1	1111668	1341934
PmUG01_04_v1	0	655294
PmUG01_04_v1	655350	1076789
PmUG01_05_v1	0	388125
PmUG01_05_v1	388310	390062
PmUG01_05_v1	390747	595498
PmUG01_05_v1	595681	748463
PmUG01_05_v1	748472	1039880
PmUG01_05_v1	1039926	1887012
PmUG01_06_v1	0	25910
PmUG01_06_v1	25978	338986
PmUG01_06_v1	339297	862290
PmUG01_07_v1	0	379469
PmUG01_07_v1	379595	417547
PmUG01_07_v1	417550	464502
PmUG01_07_v1	464532	665384
PmUG01_07_v1	665421	667248
PmUG01_07_v1	667355	1039372
PmUG01_07_v1	1039381	1159499
PmUG01_07_v1	1159534	1255677
PmUG01_07_v1	1256011	1384395
PmUG01_07_v1	1384439	2080591
PmUG01_08_v1	0	617251

PmUG01_08_v1	617297	879253
PmUG01_08_v1	879561	939195
PmUG01_08_v1	939210	1068928
PmUG01_08_v1	1068931	1612055
PmUG01_08_v1	1612088	2818518
PmUG01_09_v1	0	441567
PmUG01_09_v1	441572	747653
PmUG01_09_v1	747703	1331298
PmUG01_09_v1	1331346	2252078
PmUG01_09_v1	2252096	2283786
PmUG01_09_v1	2283933	2312277
PmUG01_10_v1	0	159404
PmUG01_10_v1	162042	165547
PmUG01_10_v1	167233	238593
PmUG01_10_v1	238596	685119
PmUG01_10_v1	685125	842485
PmUG01_10_v1	842627	1342291
PmUG01_10_v1	1342349	1788676
PmUG01_10_v1	1788713	2219075
PmUG01_11_v1	0	506323
PmUG01_11_v1	506396	532276
PmUG01_11_v1	532351	732030
PmUG01_11_v1	732057	742432
PmUG01_11_v1	742437	949073
PmUG01_11_v1	949080	1011644
PmUG01_11_v1	1011700	1097559
PmUG01_11_v1	1097595	1186889
PmUG01_11_v1	1186897	1230414
PmUG01_11_v1	1230464	1281303
PmUG01_11_v1	1281348	1287390
PmUG01_11_v1	1287496	1468290
PmUG01_11_v1	1468299	1629830
PmUG01_11_v1	1629837	1694848
PmUG01_11_v1	1695788	2145503
PmUG01_11_v1	2145505	2376120
PmUG01_11_v1	2376330	2721162
PmUG01_12_v1	0	245939
PmUG01_12_v1	246022	1283254
PmUG01_12_v1	1283264	1562601
PmUG01_12_v1	1562613	2231838
PmUG01_12_v1	2231842	2557787
PmUG01_12_v1	2557815	2921088
PmUG01_12_v1	2922083	3436066

PmUG01_13_v1	1	661992
PmUG01_13_v1	662001	1506937
PmUG01_13_v1	1507164	1756911
PmUG01_13_v1	1756945	1920857
PmUG01_13_v1	1920882	1962451
PmUG01_13_v1	1962469	2397222
PmUG01_13_v1	2397295	3023686
PmUG01_14_v1	0	506697
PmUG01_14_v1	506712	625478
PmUG01_14_v1	625485	667429
PmUG01_14_v1	667493	674936
PmUG01_14_v1	674969	792351
PmUG01_14_v1	792397	808349
PmUG01_14_v1	808351	1050843
PmUG01_14_v1	1050969	1108964
PmUG01_14_v1	1109069	1380598
PmUG01_14_v1	1380610	1383834
PmUG01_14_v1	1383849	1487217
PmUG01_14_v1	1487250	1574454
PmUG01_14_v1	1574457	1717092
PmUG01_14_v1	1717124	1946514
PmUG01_14_v1	1946569	2228404
PmUG01_14_v1	2228417	2295865
PmUG01_14_v1	2295954	2604312
PmUG01_14_v1	2604345	2822028
PmUG01_14_v1	2822103	3113433
PmUG01_14_v1	3113754	3194890
PmUG01_14_v1	3194915	3563816
PmUG01_00_v1_archived_contig_1	1	127265
PmUG01_00_v1_archived_contig_10	0	16069
PmUG01_00_v1_archived_contig_11	0	94367
PmUG01_00_v1_archived_contig_12	0	23069
PmUG01_00_v1_archived_contig_13	0	78380
PmUG01_00_v1_archived_contig_14	0	186045
PmUG01_00_v1_archived_contig_15	0	64661
PmUG01_00_v1_archived_contig_16	0	95107
PmUG01_00_v1_archived_contig_17	0	78840
PmUG01_00_v1_archived_contig_18	0	45685
PmUG01_00_v1_archived_contig_19	0	59998
PmUG01_00_v1_archived_contig_2	0	86103
PmUG01_00_v1_archived_contig_20	0	56448
PmUG01_00_v1_archived_contig_21	0	134984
PmUG01_00_v1_archived_contig_22	0	94729

PmUG01_00_v1_archived_contig_23	0	80088
PmUG01_00_v1_archived_contig_24	0	31906
PmUG01_00_v1_archived_contig_25	0	148497
PmUG01_00_v1_archived_contig_26	0	58220
PmUG01_00_v1_archived_contig_27	0	224922
PmUG01_00_v1_archived_contig_28	1	102704
PmUG01_00_v1_archived_contig_29	0	38860
PmUG01_00_v1_archived_contig_3	0	249876
PmUG01_00_v1_archived_contig_30	0	39249
PmUG01_00_v1_archived_contig_31	0	79282
PmUG01_00_v1_archived_contig_32	0	20830
PmUG01_00_v1_archived_contig_33	0	23913
PmUG01_00_v1_archived_contig_34	0	52785
PmUG01_00_v1_archived_contig_35	349	27403
PmUG01_00_v1_archived_contig_36	0	120189
PmUG01_00_v1_archived_contig_37	0	51630
PmUG01_00_v1_archived_contig_38	0	31656
PmUG01_00_v1_archived_contig_39	0	310300
PmUG01_00_v1_archived_contig_4	0	35798
PmUG01_00_v1_archived_contig_40	0	32467
PmUG01_00_v1_archived_contig_41	0	95717
PmUG01_00_v1_archived_contig_42	0	118421
PmUG01_00_v1_archived_contig_43	0	33409
PmUG01_00_v1_archived_contig_44	0	43071
PmUG01_00_v1_archived_contig_45	0	51766
PmUG01_00_v1_archived_contig_46	0	48291
PmUG01_00_v1_archived_contig_47	0	174589
PmUG01_00_v1_archived_contig_5	0	39871
PmUG01_00_v1_archived_contig_6	1019	41200
PmUG01_00_v1_archived_contig_7	314	37286
PmUG01_00_v1_archived_contig_8	0	256666
PmUG01_00_v1_archived_contig_9	0	94428
PmUG01_API_v1	6338	7827
PmUG01_API_v1	8181	12880
PmUG01_API_v1	14827	16845
PmUG01_API_v1	17662	19457
PmUG01_API_v1	19871	21786
PmUG01_API_v1	25442	27453
P. vivax after refinement		
PvP01_01_v2	0	3206
PvP01_01_v2	3207	6976
PvP01_01_v2	6977	31364

PvP01_01_v2	31664	34470
PvP01_01_v2	34774	115127
PvP01_01_v2	115427	642734
PvP01_01_v2	643034	942527
PvP01_01_v2	942854	964260
PvP01_01_v2	964560	974163
PvP01_01_v2	974288	976017
PvP01_01_v2	976021	995782
PvP01_01_v2	995783	997578
PvP01_01_v2	997701	998563
PvP01_01_v2	999997	1011915
PvP01_01_v2	1012218	1017531
PvP01_01_v2	1017831	1019642
PvP01_01_v2	1020318	1021661
PvP01_02_v2	0	2342
PvP01_02_v2	2379	99670
PvP01_02_v2	99970	146372
PvP01_02_v2	146416	147013
PvP01_02_v2	147106	147274
PvP01_02_v2	147440	147820
PvP01_02_v2	148009	335626
PvP01_02_v2	335645	337079
PvP01_02_v2	337110	829342
PvP01_02_v2	829735	832181
PvP01_02_v2	832483	835724
PvP01_02_v2	836025	845659
PvP01_02_v2	845662	851180
PvP01_02_v2	851184	861061
PvP01_02_v2	861066	868604
PvP01_02_v2	868907	881783
PvP01_02_v2	881784	892300
PvP01_02_v2	892603	894689
PvP01_02_v2	894714	906901
PvP01_02_v2	906986	908337
PvP01_02_v2	908338	911330
PvP01_02_v2	911394	927319
PvP01_02_v2	927320	928814
PvP01_02_v2	928817	929867
PvP01_02_v2	929868	930845
PvP01_02_v2	930851	940342
PvP01_02_v2	940343	954995
PvP01_02_v2	955295	956327
PvP01_03_v2	0	144623

PvP01_03_v2	144645	535558
PvP01_03_v2	535858	554727
PvP01_03_v2	554827	774697
PvP01_03_v2	774997	896704
PvP01_04_v2	0	61228
PvP01_04_v2	61531	394779
PvP01_04_v2	395082	786757
PvP01_04_v2	787057	890633
PvP01_04_v2	890933	1002219
PvP01_04_v2	1002519	1010083
PvP01_04_v2	1010383	1012024
PvP01_05_v2	0	2259
PvP01_05_v2	2559	13976
PvP01_05_v2	13977	14728
PvP01_05_v2	14729	15642
PvP01_05_v2	15718	35237
PvP01_05_v2	35717	37955
PvP01_05_v2	39263	41585
PvP01_05_v2	41586	42204
PvP01_05_v2	42278	194891
PvP01_05_v2	195659	195968
PvP01_05_v2	196195	197720
PvP01_05_v2	197744	198178
PvP01_05_v2	198198	207882
PvP01_05_v2	207912	208253
PvP01_05_v2	208260	208740
PvP01_05_v2	208890	576707
PvP01_05_v2	577007	1016299
PvP01_05_v2	1016604	1272901
PvP01_05_v2	1273201	1508416
PvP01_05_v2	1508716	1524814
PvP01_06_v2	2	494205
PvP01_06_v2	494505	524678
PvP01_06_v2	524978	784156
PvP01_06_v2	784456	928771
PvP01_06_v2	928783	1034761
PvP01_06_v2	1035120	1038056
PvP01_06_v2	1038356	1040072
PvP01_06_v2	1040073	1042791
PvP01_07_v2	0	3841
PvP01_07_v2	3842	7960
PvP01_07_v2	8260	344789
PvP01_07_v2	345089	933225

PvP01_07_v2	933525	955895
PvP01_07_v2	956195	1413401
PvP01_07_v2	1413701	1433164
PvP01_07_v2	1433464	1481187
PvP01_07_v2	1481188	1502507
PvP01_07_v2	1502809	1517429
PvP01_07_v2	1517487	1517970
PvP01_07_v2	1518270	1550283
PvP01_07_v2	1550293	1562284
PvP01_07_v2	1562584	1583193
PvP01_07_v2	1583493	1637330
PvP01_07_v2	1637635	1647486
PvP01_07_v2	1647819	1652210
PvP01_08_v2	0	12805
PvP01_08_v2	12832	1204120
PvP01_08_v2	1204420	1346297
PvP01_08_v2	1346597	1395034
PvP01_08_v2	1395334	1440643
PvP01_08_v2	1440943	1698198
PvP01_08_v2	1698498	1702052
PvP01_08_v2	1702352	1729093
PvP01_08_v2	1729496	1745673
PvP01_08_v2	1745973	1756760
PvP01_08_v2	1757097	1761288
PvP01_09_v2	0	20756
PvP01_09_v2	20775	717172
PvP01_09_v2	717472	740397
PvP01_09_v2	740678	822600
PvP01_09_v2	822900	912677
PvP01_09_v2	912977	2201900
PvP01_09_v2	2201901	2232517
PvP01_09_v2	2232518	2235644
PvP01_09_v2	2235771	2237066
PvP01_10_v2	0	49000
PvP01_10_v2	49303	130266
PvP01_10_v2	130658	312501
PvP01_10_v2	312537	760454
PvP01_10_v2	760759	775516
PvP01_10_v2	775814	914665
PvP01_10_v2	914822	932461
PvP01_10_v2	932761	1090051
PvP01_10_v2	1090151	1516346
PvP01_10_v2	1516348	1517239

PvP01_10_v2	1517240	1522150
PvP01_10_v2	1522451	1524306
PvP01_10_v2	1524372	1525498
PvP01_10_v2	1525499	1548844
PvP01_11_v2	20	198598
PvP01_11_v2	198932	1550516
PvP01_11_v2	1550820	1733918
PvP01_11_v2	1734273	1757183
PvP01_11_v2	1757483	2121405
PvP01_11_v2	2121406	2124095
PvP01_11_v2	2124096	2126418
PvP01_11_v2	2126419	2128606
PvP01_11_v2	2129668	2131221
PvP01_12_v2	0	1957
PvP01_12_v2	2257	4302
PvP01_12_v2	5528	6444
PvP01_12_v2	6445	7328
PvP01_12_v2	7665	10377
PvP01_12_v2	10382	35836
PvP01_12_v2	35837	38078
PvP01_12_v2	38380	39804
PvP01_12_v2	40104	542310
PvP01_12_v2	542610	573880
PvP01_12_v2	574180	1565091
PvP01_12_v2	1565391	1583103
PvP01_12_v2	1583403	2272899
PvP01_12_v2	2273199	3085860
PvP01_12_v2	3086160	3087397
PvP01_12_v2	3087697	3105785
PvP01_12_v2	3106086	3120412
PvP01_12_v2	3120715	3182761
PvP01_13_v2	0	1290
PvP01_13_v2	1727	11323
PvP01_13_v2	11623	1115322
PvP01_13_v2	1115722	1282576
PvP01_13_v2	1282876	1898691
PvP01_13_v2	1898991	2082519
PvP01_13_v2	2082820	2090613
PvP01_13_v2	2090614	2092018
PvP01_13_v2	2092318	2093554
PvP01_14_v2	0	716327
PvP01_14_v2	716627	790517
PvP01_14_v2	790817	841487

PvP01_14_v2	841865	851713
PvP01_14_v2	852013	994293
PvP01_14_v2	994593	1269383
PvP01_14_v2	1269728	1345193
PvP01_14_v2	1345493	2042700
PvP01_14_v2	2043000	2102789
PvP01_14_v2	2103089	2104681
PvP01_14_v2	2104981	3108753
PvP01_14_v2	3109055	3131164
PvP01_14_v2	3131464	3133798
PvP01_14_v2	3133800	3153342
PvP01_API_v2	13	231
PvP01_API_v2	239	629
PvP01_API_v2	641	1324
PvP01_API_v2	1621	3146
PvP01_API_v2	3153	3491
PvP01_API_v2	3499	8204
PvP01_API_v2	8207	8689
PvP01_API_v2	8922	9387
PvP01_API_v2	9539	10057
PvP01_API_v2	10298	10937
PvP01_API_v2	10945	12367
PvP01_API_v2	12552	13023
PvP01_API_v2	13025	14919
PvP01_API_v2	15194	17222
PvP01_API_v2	17263	17425
PvP01_API_v2	17494	17724
PvP01_API_v2	17828	18008
PvP01_API_v2	18370	18716
PvP01_API_v2	19013	19359
PvP01_API_v2	19362	19539
PvP01_API_v2	19756	20372
PvP01_API_v2	20460	20635
PvP01_API_v2	20826	21203
PvP01_API_v2	21232	22887
PvP01_API_v2	22902	23143
PvP01_API_v2	23376	23941
PvP01_API_v2	24044	24232
PvP01_API_v2	25178	25333
PvP01_API_v2	25721	25933
PvP01_API_v2	26248	26402
PvP01_API_v2	26463	26729
PvP01_API_v2	26812	27330

PvP01_API_v2	27586	27876
PvP01_API_v2	29233	29582
PvP01_MIT_v2	1848	2219
PvP01_MIT_v2	2231	2759
PvP01_MIT_v2	3234	4004
PvP01_MIT_v2	4150	4333
PvP01_MIT_v2	4366	4650
PvP01_MIT_v2	5416	5602
P.ovale		
PocGH01_01_v2	0	13146
PocGH01_01_v2	13248	46100
PocGH01_01_v2	46211	119342
PocGH01_01_v2	119441	123227
PocGH01_01_v2	123341	130686
PocGH01_01_v2	130743	142800
PocGH01_01_v2	142900	188405
PocGH01_01_v2	188510	199091
PocGH01_01_v2	199198	201585
PocGH01_01_v2	201694	219008
PocGH01_01_v2	219118	226679
PocGH01_01_v2	226789	255086
PocGH01_01_v2	255195	276177
PocGH01_01_v2	276274	289048
PocGH01_01_v2	289156	307255
PocGH01_01_v2	307335	310482
PocGH01_01_v2	310592	362376
PocGH01_01_v2	362480	461292
PocGH01_01_v2	461394	476859
PocGH01_01_v2	476956	482705
PocGH01_01_v2	482800	581150
PocGH01_01_v2	581250	608522
PocGH01_01_v2	608630	639025
PocGH01_01_v2	639142	648383
PocGH01_01_v2	648501	666026
PocGH01_01_v2	666117	671996
PocGH01_01_v2	672099	739435
PocGH01_01_v2	739557	750489
PocGH01_01_v2	750607	755941
PocGH01_01_v2	756050	760950
PocGH01_01_v2	761059	762909
PocGH01_01_v2	762969	784857
PocGH01_02_v2	969	47047

PocGH01_02_v2	47149	95005
PocGH01_02_v2	95091	108276
PocGH01_02_v2	108373	111625
PocGH01_02_v2	111719	120544
PocGH01_02_v2	120652	200689
PocGH01_02_v2	200800	220707
PocGH01_02_v2	220720	221849
PocGH01_02_v2	221958	252067
PocGH01_02_v2	252187	264660
PocGH01_02_v2	264769	309149
PocGH01_02_v2	309227	357633
PocGH01_02_v2	357744	409470
PocGH01_02_v2	409580	417191
PocGH01_02_v2	417299	430264
PocGH01_02_v2	430373	445692
PocGH01_02_v2	445804	448950
PocGH01_02_v2	449062	489298
PocGH01_02_v2	489386	549381
PocGH01_02_v2	549489	575926
PocGH01_02_v2	575986	607860
PocGH01_03_v2	0	51197
PocGH01_03_v2	51309	60755
PocGH01_03_v2	60863	83659
PocGH01_03_v2	83769	96047
PocGH01_03_v2	96158	106132
PocGH01_03_v2	106241	128913
PocGH01_03_v2	130086	132374
PocGH01_03_v2	132493	194731
PocGH01_03_v2	194841	239480
PocGH01_03_v2	239575	241913
PocGH01_03_v2	242017	281068
PocGH01_03_v2	281080	284545
PocGH01_03_v2	284655	373332
PocGH01_03_v2	373435	399296
PocGH01_03_v2	399405	428096
PocGH01_03_v2	428205	436997
PocGH01_03_v2	437103	486085
PocGH01_03_v2	487283	587472
PocGH01_03_v2	587568	591115
PocGH01_03_v2	591223	596099
PocGH01_03_v2	596208	656826
PocGH01_03_v2	656934	672350
PocGH01_03_v2	672458	677021

PocGH01_03_v2	677133	685710
PocGH01_03_v2	685814	702762
PocGH01_03_v2	702868	737804
PocGH01_03_v2	737914	758190
PocGH01_03_v2	758300	759718
PocGH01_03_v2	759826	785200
PocGH01_03_v2	785260	820896
PocGH01_04_v2	0	50612
PocGH01_04_v2	50721	62743
PocGH01_04_v2	62852	76216
PocGH01_04_v2	76317	81698
PocGH01_04_v2	81808	103487
PocGH01_04_v2	103609	128423
PocGH01_04_v2	128532	154442
PocGH01_04_v2	154560	175484
PocGH01_04_v2	175599	191546
PocGH01_04_v2	191663	198478
PocGH01_04_v2	198587	206800
PocGH01_04_v2	206912	220463
PocGH01_04_v2	220581	238589
PocGH01_04_v2	238697	249662
PocGH01_04_v2	249768	269508
PocGH01_04_v2	269622	271590
PocGH01_04_v2	271699	280167
PocGH01_04_v2	280289	333440
PocGH01_04_v2	333549	366357
PocGH01_04_v2	366467	393096
PocGH01_04_v2	393192	396682
PocGH01_04_v2	396805	523261
PocGH01_04_v2	523341	563584
PocGH01_04_v2	563692	580293
PocGH01_04_v2	580402	608839
PocGH01_04_v2	608939	611654
PocGH01_04_v2	611763	643646
PocGH01_04_v2	643757	649429
PocGH01_04_v2	649489	684040
PocGH01_05_v2	0	23710
PocGH01_05_v2	23818	45051
PocGH01_05_v2	45179	46653
PocGH01_05_v2	46703	52321
PocGH01_05_v2	52334	53340
PocGH01_05_v2	54335	68961
PocGH01_05_v2	69077	166329

PocGH01_05_v2	166438	237935
PocGH01_05_v2	237944	271282
PocGH01_05_v2	271390	277151
PocGH01_05_v2	277258	289734
PocGH01_05_v2	289844	302355
PocGH01_05_v2	302470	324768
PocGH01_05_v2	324873	369444
PocGH01_05_v2	369548	376106
PocGH01_05_v2	376216	380260
PocGH01_05_v2	380269	387162
PocGH01_05_v2	387277	391076
PocGH01_05_v2	391196	414507
PocGH01_05_v2	414748	429250
PocGH01_05_v2	429272	481672
PocGH01_05_v2	481790	508782
PocGH01_05_v2	508908	527815
PocGH01_05_v2	527925	536979
PocGH01_05_v2	537087	587424
PocGH01_05_v2	587532	600227
PocGH01_05_v2	600328	612839
PocGH01_05_v2	612948	619965
PocGH01_05_v2	620078	623689
PocGH01_05_v2	623791	632531
PocGH01_05_v2	632649	661288
PocGH01_05_v2	661331	668612
PocGH01_05_v2	668721	684798
PocGH01_05_v2	684906	692361
PocGH01_05_v2	692479	710498
PocGH01_05_v2	710606	722184
PocGH01_05_v2	722295	736575
PocGH01_05_v2	736685	759683
PocGH01_05_v2	759793	795200
PocGH01_05_v2	795304	838716
PocGH01_05_v2	838818	892313
PocGH01_05_v2	892423	921589
PocGH01_05_v2	921699	934482
PocGH01_05_v2	934605	961693
PocGH01_05_v2	961798	966471
PocGH01_05_v2	967298	1028244
PocGH01_05_v2	1028304	1039375
PocGH01_06_v2	1042	19916
PocGH01_06_v2	20026	32397
PocGH01_06_v2	32503	56237

PocGH01_06_v2	56347	122804
PocGH01_06_v2	122917	130898
PocGH01_06_v2	130988	135013
PocGH01_06_v2	135108	164389
PocGH01_06_v2	164494	184034
PocGH01_06_v2	184135	185540
PocGH01_06_v2	185648	192870
PocGH01_06_v2	192941	290314
PocGH01_06_v2	290326	338032
PocGH01_06_v2	338141	352763
PocGH01_06_v2	352871	360276
PocGH01_06_v2	360393	370252
PocGH01_06_v2	370360	378365
PocGH01_06_v2	378475	396456
PocGH01_06_v2	396569	414337
PocGH01_06_v2	414447	430559
PocGH01_06_v2	430669	491923
PocGH01_06_v2	492031	494191
PocGH01_06_v2	494314	501556
PocGH01_06_v2	501666	513496
PocGH01_06_v2	513596	528424
PocGH01_06_v2	528507	544580
PocGH01_06_v2	544698	547874
PocGH01_06_v2	547971	558273
PocGH01_06_v2	558380	567423
PocGH01_06_v2	567533	581727
PocGH01_06_v2	582487	603126
PocGH01_06_v2	603234	627157
PocGH01_06_v2	627271	653623
PocGH01_06_v2	653720	656071
PocGH01_06_v2	657161	681814
PocGH01_06_v2	681924	695609
PocGH01_06_v2	695719	702529
PocGH01_06_v2	702639	784610
PocGH01_06_v2	784718	804103
PocGH01_06_v2	804874	836784
PocGH01_06_v2	836844	840843
PocGH01_07_v2	0	56624
PocGH01_07_v2	56732	184575
PocGH01_07_v2	184686	198611
PocGH01_07_v2	198641	200668
PocGH01_07_v2	200778	218242
PocGH01_07_v2	218351	228834

PocGH01_07_v2	228944	230383
PocGH01_07_v2	230492	237949
PocGH01_07_v2	238066	252916
PocGH01_07_v2	253025	275193
PocGH01_07_v2	275302	298804
PocGH01_07_v2	298915	307460
PocGH01_07_v2	307570	342093
PocGH01_07_v2	342194	350995
PocGH01_07_v2	351103	360874
PocGH01_07_v2	360984	372633
PocGH01_07_v2	372649	377660
PocGH01_07_v2	377768	412228
PocGH01_07_v2	412336	421674
PocGH01_07_v2	421771	444349
PocGH01_07_v2	444442	456243
PocGH01_07_v2	456357	475123
PocGH01_07_v2	475209	496693
PocGH01_07_v2	496803	509803
PocGH01_07_v2	509913	535671
PocGH01_07_v2	535779	651405
PocGH01_07_v2	651513	673482
PocGH01_07_v2	673596	694078
PocGH01_07_v2	694186	703685
PocGH01_07_v2	703694	740295
PocGH01_07_v2	740403	773786
PocGH01_07_v2	773894	833462
PocGH01_07_v2	833585	846900
PocGH01_07_v2	846996	854266
PocGH01_07_v2	854505	857988
PocGH01_07_v2	858092	874034
PocGH01_07_v2	874141	906176
PocGH01_07_v2	906285	927263
PocGH01_07_v2	927367	963755
PocGH01_07_v2	963861	976971
PocGH01_07_v2	977082	982569
PocGH01_07_v2	982677	987977
PocGH01_07_v2	988085	1005448
PocGH01_07_v2	1005543	1021314
PocGH01_07_v2	1021431	1051853
PocGH01_07_v2	1051963	1061519
PocGH01_07_v2	1061629	1083588
PocGH01_07_v2	1083696	1104970
PocGH01_07_v2	1105057	1122486

PocGH01_07_v2	1122589	1139122
PocGH01_07_v2	1139230	1162660
PocGH01_07_v2	1162768	1173965
PocGH01_07_v2	1174072	1175661
PocGH01_07_v2	1175770	1201688
PocGH01_07_v2	1201793	1218034
PocGH01_07_v2	1218153	1223420
PocGH01_07_v2	1223529	1228398
PocGH01_07_v2	1228506	1302973
PocGH01_07_v2	1303085	1317475
PocGH01_07_v2	1317587	1325107
PocGH01_07_v2	1325216	1335824
PocGH01_07_v2	1335929	1351212
PocGH01_07_v2	1351321	1375886
PocGH01_07_v2	1375946	1397181
PocGH01_08_v2	0	41423
PocGH01_08_v2	41537	60181
PocGH01_08_v2	60291	68245
PocGH01_08_v2	68355	100939
PocGH01_08_v2	101044	103965
PocGH01_08_v2	104088	120265
PocGH01_08_v2	120375	127264
PocGH01_08_v2	127372	128920
PocGH01_08_v2	129016	158472
PocGH01_08_v2	158589	168078
PocGH01_08_v2	168186	177113
PocGH01_08_v2	177221	233099
PocGH01_08_v2	233209	254535
PocGH01_08_v2	254646	264828
PocGH01_08_v2	264939	325516
PocGH01_08_v2	325626	382032
PocGH01_08_v2	382141	428795
PocGH01_08_v2	428904	447072
PocGH01_08_v2	447186	450943
PocGH01_08_v2	451053	529774
PocGH01_08_v2	529868	566519
PocGH01_08_v2	566628	580441
PocGH01_08_v2	580554	588352
PocGH01_08_v2	588476	611376
PocGH01_08_v2	611520	617371
PocGH01_08_v2	617479	637326
PocGH01_08_v2	637438	673601
PocGH01_08_v2	673711	696824

PocGH01_08_v2	696932	734056
PocGH01_08_v2	734166	740714
PocGH01_08_v2	740822	746131
PocGH01_08_v2	746248	754810
PocGH01_08_v2	754910	779521
PocGH01_08_v2	779636	871818
PocGH01_08_v2	871930	910613
PocGH01_08_v2	910693	935305
PocGH01_08_v2	935397	974868
PocGH01_08_v2	974971	978795
PocGH01_08_v2	978904	1031703
PocGH01_08_v2	1031811	1049818
PocGH01_08_v2	1050025	1066322
PocGH01_08_v2	1066442	1088335
PocGH01_08_v2	1088449	1099744
PocGH01_08_v2	1099852	1114049
PocGH01_08_v2	1114159	1119732
PocGH01_08_v2	1119842	1126103
PocGH01_08_v2	1126212	1145401
PocGH01_08_v2	1145509	1158431
PocGH01_08_v2	1158539	1204811
PocGH01_08_v2	1204927	1230822
PocGH01_08_v2	1230929	1239159
PocGH01_08_v2	1239268	1242078
PocGH01_08_v2	1242187	1302442
PocGH01_08_v2	1302556	1308124
PocGH01_08_v2	1308231	1341692
PocGH01_08_v2	1341801	1369428
PocGH01_08_v2	1369538	1379318
PocGH01_08_v2	1379432	1402936
PocGH01_08_v2	1403050	1406507
PocGH01_08_v2	1406624	1474425
PocGH01_08_v2	1474485	1535399
PocGH01_09_v2	2	47970
PocGH01_09_v2	48089	81666
PocGH01_09_v2	81774	120101
PocGH01_09_v2	120212	159329
PocGH01_09_v2	159437	175403
PocGH01_09_v2	175512	186918
PocGH01_09_v2	187028	200844
PocGH01_09_v2	200954	225025
PocGH01_09_v2	225027	231454
PocGH01_09_v2	231564	247242

PocGH01_09_v2	247349	250175
PocGH01_09_v2	250299	272764
PocGH01_09_v2	272854	290778
PocGH01_09_v2	290886	293187
PocGH01_09_v2	293295	295734
PocGH01_09_v2	295842	302537
PocGH01_09_v2	302646	321762
PocGH01_09_v2	321872	367740
PocGH01_09_v2	367848	383110
PocGH01_09_v2	383218	409499
PocGH01_09_v2	409600	444403
PocGH01_09_v2	444511	447410
PocGH01_09_v2	447499	524768
PocGH01_09_v2	524878	546377
PocGH01_09_v2	546485	561645
PocGH01_09_v2	561758	576411
PocGH01_09_v2	576528	583325
PocGH01_09_v2	583435	618709
PocGH01_09_v2	618817	623930
PocGH01_09_v2	624025	660997
PocGH01_09_v2	661058	687975
PocGH01_09_v2	688085	698987
PocGH01_09_v2	699100	712277
PocGH01_09_v2	712409	727628
PocGH01_09_v2	727746	824109
PocGH01_09_v2	824225	836851
PocGH01_09_v2	836952	843979
PocGH01_09_v2	844081	860864
PocGH01_09_v2	860973	879216
PocGH01_09_v2	879323	886788
PocGH01_09_v2	886903	901221
PocGH01_09_v2	901331	937457
PocGH01_09_v2	937571	947164
PocGH01_09_v2	947272	955201
PocGH01_09_v2	955329	978992
PocGH01_09_v2	979102	998564
PocGH01_09_v2	998671	1001229
PocGH01_09_v2	1001341	1017714
PocGH01_09_v2	1017821	1038196
PocGH01_09_v2	1038306	1071979
PocGH01_09_v2	1072089	1108231
PocGH01_09_v2	1108319	1112094
PocGH01_09_v2	1112183	1143975

PocGH01_09_v2	1144089	1157715
PocGH01_09_v2	1157812	1178952
PocGH01_09_v2	1179063	1197518
PocGH01_09_v2	1197628	1199783
PocGH01_09_v2	1199901	1273596
PocGH01_09_v2	1273705	1299738
PocGH01_09_v2	1299848	1334348
PocGH01_09_v2	1334443	1353668
PocGH01_09_v2	1353787	1364430
PocGH01_09_v2	1364527	1444993
PocGH01_09_v2	1446102	1462667
PocGH01_09_v2	1462760	1496207
PocGH01_09_v2	1496323	1499704
PocGH01_09_v2	1499816	1521819
PocGH01_09_v2	1521928	1599818
PocGH01_09_v2	1599922	1633788
PocGH01_09_v2	1633896	1638461
PocGH01_09_v2	1638579	1653357
PocGH01_09_v2	1653476	1674890
PocGH01_09_v2	1675000	1719600
PocGH01_09_v2	1719711	1728298
PocGH01_09_v2	1728407	1744179
PocGH01_09_v2	1744276	1745794
PocGH01_09_v2	1745902	1761072
PocGH01_09_v2	1761189	1778507
PocGH01_09_v2	1778615	1781147
PocGH01_09_v2	1781256	1804923
PocGH01_09_v2	1804983	1833880
PocGH01_09_v2	1833889	1865930
PocGH01_09_v2	1866077	1881439
PocGH01_10_v2	25	11016
PocGH01_10_v2	11124	42527
PocGH01_10_v2	42642	56413
PocGH01_10_v2	56522	71635
PocGH01_10_v2	71730	90236
PocGH01_10_v2	90340	101147
PocGH01_10_v2	101244	126972
PocGH01_10_v2	127080	131636
PocGH01_10_v2	131720	149814
PocGH01_10_v2	149933	158524
PocGH01_10_v2	158632	168558
PocGH01_10_v2	168677	201144
PocGH01_10_v2	201254	246395

PocGH01_10_v2	246519	254046
PocGH01_10_v2	254154	266536
PocGH01_10_v2	266645	271989
PocGH01_10_v2	272095	304229
PocGH01_10_v2	304337	339384
PocGH01_10_v2	339633	388561
PocGH01_10_v2	388673	394458
PocGH01_10_v2	394578	447908
PocGH01_10_v2	448012	459008
PocGH01_10_v2	459110	475871
PocGH01_10_v2	475920	480114
PocGH01_10_v2	480223	499426
PocGH01_10_v2	499544	553587
PocGH01_10_v2	553705	568444
PocGH01_10_v2	568552	572216
PocGH01_10_v2	572332	577950
PocGH01_10_v2	578046	589531
PocGH01_10_v2	589639	615662
PocGH01_10_v2	615772	631551
PocGH01_10_v2	631663	641761
PocGH01_10_v2	641870	647363
PocGH01_10_v2	647471	664794
PocGH01_10_v2	664902	715265
PocGH01_10_v2	715374	769065
PocGH01_10_v2	769173	806589
PocGH01_10_v2	806698	815529
PocGH01_10_v2	815640	825181
PocGH01_10_v2	825289	826967
PocGH01_10_v2	827076	840733
PocGH01_10_v2	840841	864357
PocGH01_10_v2	864467	908456
PocGH01_10_v2	908565	915376
PocGH01_10_v2	915485	953706
PocGH01_10_v2	953832	981572
PocGH01_10_v2	981696	1040874
PocGH01_10_v2	1040980	1092762
PocGH01_10_v2	1092872	1162663
PocGH01_10_v2	1162773	1165872
PocGH01_10_v2	1165977	1226634
PocGH01_10_v2	1226744	1260523
PocGH01_10_v2	1260632	1271708
PocGH01_10_v2	1271822	1293476
PocGH01_10_v2	1293536	1299758

PocGH01_11_v2	0	57850
PocGH01_11_v2	57951	108934
PocGH01_11_v2	109007	131438
PocGH01_11_v2	131557	140771
PocGH01_11_v2	140879	157647
PocGH01_11_v2	157762	209490
PocGH01_11_v2	209598	234775
PocGH01_11_v2	234885	250308
PocGH01_11_v2	250417	322411
PocGH01_11_v2	322540	333248
PocGH01_11_v2	333253	348582
PocGH01_11_v2	348691	351426
PocGH01_11_v2	351530	364826
PocGH01_11_v2	364935	377084
PocGH01_11_v2	377193	389601
PocGH01_11_v2	389709	397691
PocGH01_11_v2	397799	428414
PocGH01_11_v2	428522	454522
PocGH01_11_v2	454630	458685
PocGH01_11_v2	458793	471244
PocGH01_11_v2	471350	490486
PocGH01_11_v2	490609	494112
PocGH01_11_v2	494221	544245
PocGH01_11_v2	544353	549018
PocGH01_11_v2	549101	560797
PocGH01_11_v2	560892	572983
PocGH01_11_v2	573086	596158
PocGH01_11_v2	596277	621055
PocGH01_11_v2	621147	623549
PocGH01_11_v2	623659	713672
PocGH01_11_v2	713784	746114
PocGH01_11_v2	746223	799145
PocGH01_11_v2	799251	858079
PocGH01_11_v2	858174	859703
PocGH01_11_v2	859817	864503
PocGH01_11_v2	864612	886679
PocGH01_11_v2	886793	899926
PocGH01_11_v2	900048	904644
PocGH01_11_v2	904733	908325
PocGH01_11_v2	908445	969659
PocGH01_11_v2	969677	989931
PocGH01_11_v2	990041	1001965
PocGH01_11_v2	1002971	1022489

PocGH01_11_v2	1022598	1030125
PocGH01_11_v2	1030233	1049494
PocGH01_11_v2	1049604	1080045
PocGH01_11_v2	1080155	1084821
PocGH01_11_v2	1084931	1093219
PocGH01_11_v2	1093327	1106479
PocGH01_11_v2	1106589	1116412
PocGH01_11_v2	1116520	1127623
PocGH01_11_v2	1127732	1132535
PocGH01_11_v2	1132645	1149090
PocGH01_11_v2	1149327	1176465
PocGH01_11_v2	1176487	1177819
PocGH01_11_v2	1177924	1195131
PocGH01_11_v2	1195246	1231665
PocGH01_11_v2	1231772	1268596
PocGH01_11_v2	1268696	1286366
PocGH01_11_v2	1286477	1293439
PocGH01_11_v2	1293553	1332834
PocGH01_11_v2	1332943	1364980
PocGH01_11_v2	1365084	1386051
PocGH01_11_v2	1386160	1412395
PocGH01_11_v2	1412503	1417888
PocGH01_11_v2	1417996	1432677
PocGH01_11_v2	1432765	1468056
PocGH01_11_v2	1468164	1478831
PocGH01_11_v2	1478939	1483147
PocGH01_11_v2	1483242	1486770
PocGH01_11_v2	1486879	1527828
PocGH01_11_v2	1527923	1549806
PocGH01_11_v2	1549915	1646702
PocGH01_11_v2	1646790	1653331
PocGH01_11_v2	1653433	1656751
PocGH01_11_v2	1656877	1659651
PocGH01_11_v2	1659760	1751368
PocGH01_11_v2	1751476	1770744
PocGH01_11_v2	1770848	1784908
PocGH01_11_v2	1785020	1802329
PocGH01_11_v2	1802437	1892419
PocGH01_11_v2	1892485	1909907
PocGH01_12_v2	0	66153
PocGH01_12_v2	66261	76250
PocGH01_12_v2	76359	87612
PocGH01_12_v2	87723	127691

PocGH01_12_v2	127799	159426
PocGH01_12_v2	159533	181567
PocGH01_12_v2	181676	199641
PocGH01_12_v2	199751	217419
PocGH01_12_v2	217528	223885
PocGH01_12_v2	223993	248280
PocGH01_12_v2	248389	276646
PocGH01_12_v2	276754	281210
PocGH01_12_v2	281312	288986
PocGH01_12_v2	289090	301087
PocGH01_12_v2	301178	305826
PocGH01_12_v2	305941	330822
PocGH01_12_v2	330920	366386
PocGH01_12_v2	366498	429442
PocGH01_12_v2	429551	434906
PocGH01_12_v2	435014	461644
PocGH01_12_v2	461752	500056
PocGH01_12_v2	500165	510577
PocGH01_12_v2	510696	514188
PocGH01_12_v2	514297	522197
PocGH01_12_v2	522305	569347
PocGH01_12_v2	569455	583581
PocGH01_12_v2	583678	633611
PocGH01_12_v2	633729	671848
PocGH01_12_v2	671958	714171
PocGH01_12_v2	714289	737119
PocGH01_12_v2	737217	755954
PocGH01_12_v2	756063	763261
PocGH01_12_v2	763375	823568
PocGH01_12_v2	823678	859427
PocGH01_12_v2	859539	870427
PocGH01_12_v2	870638	876066
PocGH01_12_v2	876176	883843
PocGH01_12_v2	883960	887122
PocGH01_12_v2	887230	896766
PocGH01_12_v2	896875	901773
PocGH01_12_v2	901893	919169
PocGH01_12_v2	919280	967762
PocGH01_12_v2	968012	993807
PocGH01_12_v2	993916	1009262
PocGH01_12_v2	1009370	1012106
PocGH01_12_v2	1012116	1015936
PocGH01_12_v2	1016045	1040490

PocGH01_12_v2	1040600	1062106
PocGH01_12_v2	1062218	1096280
PocGH01_12_v2	1096390	1105509
PocGH01_12_v2	1105619	1125858
PocGH01_12_v2	1125966	1148893
PocGH01_12_v2	1149005	1170695
PocGH01_12_v2	1170803	1173812
PocGH01_12_v2	1173909	1175812
PocGH01_12_v2	1175909	1182512
PocGH01_12_v2	1182621	1199315
PocGH01_12_v2	1199416	1200612
PocGH01_12_v2	1200721	1210880
PocGH01_12_v2	1210988	1220940
PocGH01_12_v2	1221043	1229902
PocGH01_12_v2	1230011	1232984
PocGH01_12_v2	1233090	1246139
PocGH01_12_v2	1246250	1249805
PocGH01_12_v2	1249817	1279237
PocGH01_12_v2	1279345	1295587
PocGH01_12_v2	1295697	1378853
PocGH01_12_v2	1378963	1411488
PocGH01_12_v2	1411597	1432971
PocGH01_12_v2	1433077	1437837
PocGH01_12_v2	1437955	1466175
PocGH01_12_v2	1466255	1477922
PocGH01_12_v2	1478032	1543169
PocGH01_12_v2	1543287	1549601
PocGH01_12_v2	1549711	1575519
PocGH01_12_v2	1575629	1604834
PocGH01_12_v2	1604942	1628510
PocGH01_12_v2	1628614	1675097
PocGH01_12_v2	1675192	1683823
PocGH01_12_v2	1683923	1688717
PocGH01_12_v2	1688827	1750272
PocGH01_12_v2	1750392	1764575
PocGH01_12_v2	1764685	1831935
PocGH01_12_v2	1832043	1847553
PocGH01_12_v2	1847661	1874582
PocGH01_12_v2	1874692	1881613
PocGH01_12_v2	1881714	1883136
PocGH01_12_v2	1883140	1894469
PocGH01_12_v2	1894578	1960651
PocGH01_12_v2	1960759	2012770

PocGH01_12_v2	2012870	2057792
PocGH01_12_v2	2057900	2070686
PocGH01_12_v2	2070794	2082430
PocGH01_12_v2	2082534	2187492
PocGH01_12_v2	2187502	2216356
PocGH01_12_v2	2216781	2225252
PocGH01_12_v2	2225297	2286613
PocGH01_12_v2	2286668	2306843
PocGH01_12_v2	2306921	2319929
PocGH01_12_v2	2320038	2326034
PocGH01_12_v2	2326135	2330855
PocGH01_12_v2	2330950	2362348
PocGH01_12_v2	2362462	2372886
PocGH01_12_v2	2372972	2393415
PocGH01_12_v2	2393521	2415500
PocGH01_12_v2	2416491	2425987
PocGH01_12_v2	2426096	2428652
PocGH01_12_v2	2428771	2445793
PocGH01_12_v2	2445901	2467581
PocGH01_12_v2	2467690	2517924
PocGH01_12_v2	2518919	2605916
PocGH01_12_v2	2606032	2652210
PocGH01_12_v2	2652312	2710158
PocGH01_12_v2	2710266	2771558
PocGH01_12_v2	2772664	2843522
PocGH01_12_v2	2843582	2945718
PocGH01_13_v2	0	96235
PocGH01_13_v2	96342	158145
PocGH01_13_v2	158252	180670
PocGH01_13_v2	180780	192093
PocGH01_13_v2	192201	213688
PocGH01_13_v2	213780	230004
PocGH01_13_v2	230112	243796
PocGH01_13_v2	244630	249361
PocGH01_13_v2	249485	309004
PocGH01_13_v2	309125	342107
PocGH01_13_v2	342225	360572
PocGH01_13_v2	360681	399543
PocGH01_13_v2	399662	439683
PocGH01_13_v2	439801	441457
PocGH01_13_v2	441566	445794
PocGH01_13_v2	445903	496228
PocGH01_13_v2	496339	531749

PocGH01_13_v2	531860	562354
PocGH01_13_v2	562463	568748
PocGH01_13_v2	568840	611438
PocGH01_13_v2	611546	620587
PocGH01_13_v2	620634	651202
PocGH01_13_v2	651310	697737
PocGH01_13_v2	697833	700518
PocGH01_13_v2	700633	716564
PocGH01_13_v2	716672	842123
PocGH01_13_v2	842240	863887
PocGH01_13_v2	863996	889526
PocGH01_13_v2	889635	898458
PocGH01_13_v2	898567	1008246
PocGH01_13_v2	1008354	1015068
PocGH01_13_v2	1015176	1040085
PocGH01_13_v2	1040209	1056590
PocGH01_13_v2	1056698	1093296
PocGH01_13_v2	1093391	1106429
PocGH01_13_v2	1106539	1136334
PocGH01_13_v2	1136438	1143632
PocGH01_13_v2	1143752	1184972
PocGH01_13_v2	1185076	1191020
PocGH01_13_v2	1191133	1227992
PocGH01_13_v2	1228101	1238568
PocGH01_13_v2	1238660	1261050
PocGH01_13_v2	1261162	1279722
PocGH01_13_v2	1279814	1280899
PocGH01_13_v2	1281007	1365731
PocGH01_13_v2	1365841	1395427
PocGH01_13_v2	1395537	1396559
PocGH01_13_v2	1396669	1518446
PocGH01_13_v2	1518554	1571486
PocGH01_13_v2	1571598	1578939
PocGH01_13_v2	1579048	1591581
PocGH01_13_v2	1591689	1609733
PocGH01_13_v2	1609833	1614856
PocGH01_13_v2	1614965	1617624
PocGH01_13_v2	1617749	1679769
PocGH01_13_v2	1679796	1687046
PocGH01_13_v2	1687140	1718557
PocGH01_13_v2	1718575	1747920
PocGH01_13_v2	1748037	1768073
PocGH01_13_v2	1768177	1783478

PocGH01_13_v2	1783588	1784941
PocGH01_13_v2	1785059	1799730
PocGH01_13_v2	1799828	1840950
PocGH01_13_v2	1841063	1848174
PocGH01_13_v2	1848281	1855475
PocGH01_13_v2	1855585	1915529
PocGH01_13_v2	1915638	1928710
PocGH01_13_v2	1930279	1945897
PocGH01_13_v2	1946007	2005957
PocGH01_13_v2	2006067	2056254
PocGH01_13_v2	2056349	2060994
PocGH01_13_v2	2061103	2078504
PocGH01_13_v2	2078564	2107900
PocGH01_14_v2	0	31739
PocGH01_14_v2	31754	56103
PocGH01_14_v2	56213	125187
PocGH01_14_v2	125299	152799
PocGH01_14_v2	152806	177842
PocGH01_14_v2	177950	194629
PocGH01_14_v2	194693	201705
PocGH01_14_v2	201816	281139
PocGH01_14_v2	281231	321962
PocGH01_14_v2	322004	328380
PocGH01_14_v2	328480	338099
PocGH01_14_v2	338101	379767
PocGH01_14_v2	379875	413898
PocGH01_14_v2	413995	459439
PocGH01_14_v2	459526	465695
PocGH01_14_v2	465812	499992
PocGH01_14_v2	500102	534967
PocGH01_14_v2	535064	541867
PocGH01_14_v2	541979	557007
PocGH01_14_v2	557058	582552
PocGH01_14_v2	582660	601275
PocGH01_14_v2	601385	613987
PocGH01_14_v2	614092	641359
PocGH01_14_v2	641469	674643
PocGH01_14_v2	674752	743415
PocGH01_14_v2	743540	750090
PocGH01_14_v2	750199	753765
PocGH01_14_v2	753878	766635
PocGH01_14_v2	766743	770070
PocGH01_14_v2	770188	775241

PocGH01_14_v2	775351	793315
PocGH01_14_v2	793429	829315
PocGH01_14_v2	829423	869047
PocGH01_14_v2	869952	873330
PocGH01_14_v2	873345	890399
PocGH01_14_v2	890509	900584
PocGH01_14_v2	900692	923503
PocGH01_14_v2	923612	933008
PocGH01_14_v2	933116	936968
PocGH01_14_v2	937037	938550
PocGH01_14_v2	938658	955860
PocGH01_14_v2	955968	973448
PocGH01_14_v2	973566	986791
PocGH01_14_v2	986888	1002687
PocGH01_14_v2	1002797	1018959
PocGH01_14_v2	1019069	1042126
PocGH01_14_v2	1042483	1044109
PocGH01_14_v2	1044225	1047410
PocGH01_14_v2	1048440	1068033
PocGH01_14_v2	1068143	1076551
PocGH01_14_v2	1076659	1101228
PocGH01_14_v2	1101335	1116566
PocGH01_14_v2	1116679	1129122
PocGH01_14_v2	1129217	1134537
PocGH01_14_v2	1134651	1221344
PocGH01_14_v2	1221449	1224492
PocGH01_14_v2	1224601	1254552
PocGH01_14_v2	1254660	1288127
PocGH01_14_v2	1288242	1303007
PocGH01_14_v2	1303115	1341620
PocGH01_14_v2	1341729	1359119
PocGH01_14_v2	1359216	1373607
PocGH01_14_v2	1373715	1376242
PocGH01_14_v2	1376350	1385794
PocGH01_14_v2	1385902	1392252
PocGH01_14_v2	1392358	1405044
PocGH01_14_v2	1405144	1410728
PocGH01_14_v2	1410836	1414084
PocGH01_14_v2	1414139	1415618
PocGH01_14_v2	1417749	1499355
PocGH01_14_v2	1499465	1506548
PocGH01_14_v2	1506658	1523863
PocGH01_14_v2	1523972	1529442

PocGH01_14_v2	1529552	1532591
PocGH01_14_v2	1532717	1583114
PocGH01_14_v2	1583228	1609349
PocGH01_14_v2	1609457	1618296
PocGH01_14_v2	1618406	1624721
PocGH01_14_v2	1624839	1670104
PocGH01_14_v2	1670199	1683489
PocGH01_14_v2	1683507	1749654
PocGH01_14_v2	1749762	1753835
PocGH01_14_v2	1753943	1764958
PocGH01_14_v2	1765054	1778274
PocGH01_14_v2	1778354	1793080
PocGH01_14_v2	1793191	1806281
PocGH01_14_v2	1806390	1826302
PocGH01_14_v2	1826364	1864714
PocGH01_14_v2	1864825	1877200
PocGH01_14_v2	1877303	1940612
PocGH01_14_v2	1940717	1948820
PocGH01_14_v2	1948928	1969022
PocGH01_14_v2	1969139	1978389
PocGH01_14_v2	1978497	1993961
PocGH01_14_v2	1994069	2016728
PocGH01_14_v2	2016836	2019539
PocGH01_14_v2	2019574	2034388
PocGH01_14_v2	2034498	2060342
PocGH01_14_v2	2060456	2067408
PocGH01_14_v2	2067517	2075696
PocGH01_14_v2	2075804	2080616
PocGH01_14_v2	2080638	2084570
PocGH01_14_v2	2084678	2098591
PocGH01_14_v2	2098709	2127793
PocGH01_14_v2	2127907	2131468
PocGH01_14_v2	2131563	2207490
PocGH01_14_v2	2207576	2213871
PocGH01_14_v2	2213979	2225083
PocGH01_14_v2	2225179	2251253
PocGH01_14_v2	2251371	2259121
PocGH01_14_v2	2259210	2261176
PocGH01_14_v2	2261285	2292848
PocGH01_14_v2	2292956	2299807
PocGH01_14_v2	2299916	2307267
PocGH01_14_v2	2307357	2331087
PocGH01_14_v2	2331196	2336078

PocGH01_14_v2	2336187	2340866
PocGH01_14_v2	2340974	2351359
PocGH01_14_v2	2351462	2355986
PocGH01_14_v2	2356094	2460148
PocGH01_14_v2	2460258	2466188
PocGH01_14_v2	2466296	2467992
PocGH01_14_v2	2468116	2492107
PocGH01_14_v2	2492224	2496777
PocGH01_14_v2	2496886	2500673
PocGH01_14_v2	2500797	2503903
PocGH01_14_v2	2504012	2538774
PocGH01_14_v2	2538888	2544820
PocGH01_14_v2	2544838	2547653
PocGH01_14_v2	2547761	2598645
PocGH01_14_v2	2598674	2599795
PocGH01_14_v2	2599905	2663027
PocGH01_14_v2	2663127	2664949
PocGH01_14_v2	2665051	2671501
PocGH01_14_v2	2671610	2695036
PocGH01_14_v2	2695145	2707348
PocGH01_14_v2	2707447	2725162
PocGH01_14_v2	2725269	2728008
PocGH01_14_v2	2728114	2740584
PocGH01_14_v2	2740703	2745759
PocGH01_14_v2	2745878	2854760
PocGH01_API_v2	270	1860
PocGH01_API_v2	1946	3473
PocGH01_API_v2	3501	9017
PocGH01_API_v2	11309	12633
PocGH01_API_v2	13389	15035
PocGH01_API_v2	15587	17247
PocGH01_API_v2	21338	23208

Supplemental Table 2: Alternative calls for Mixed(all) infections

broad_id	pf_reads	po_reads	pm_reads	pv_reads	total	Alternative Diagnosis	PCR
NWD942096	97726	202	979	63	98970	Pf+	Pf+

NWD620594	2327	37229	146	115	39817	Po+	Mixed (Pf+ and Po+)
NWD197132	114611	6698	436	86	121831	Pf+	Mixed (Pf+ and Po+)
NWD301738	29091	105	152	87	29435	Pf+	Pf+
NWD616442	62546	153	501	53	63253	Pf+	Pf+
NWD594495	74335	410	5383	79	80207	Pf+	Mixed (Pf+ and Pm+)
NWD562912	24554	79	2484	52	27169	Mixed (Pf and Pm)	Mixed (Pf+ and Pm+)
NWD386040	143828	338	584	78	144828	Pf+	Pf+
NWD896265	43719	224	34210	76	78229	Mixed (Pf and Pm)	Mixed (Pf+ and Pm+)
NWD959504	2035274	4548	8208	1013	2049043	Pf+	Pf+
NWD899632	259148	869	6741	99	266857	Pf+	Mixed (Pf + Pm)
NWD777492	108684	366	384	70	109504	Pf+	Pf+
NWD664694	40225	946	181	140	41492	Pf+	PF+
NWD994664	913728	2055	31002	480	947265	Pf+	Pf+
NWD275394	88073	236	296	50	88655	Pf+	Pf+
NWD230045	162730	389	533	101	163753	Pf+	Pf+
NWD383089	186888	468	729	100	188185	Pf+	Pf+
NWD846454	130553	373	24895	95	155916	Mixed (Pf and Pm)	Mixed (Pf+ and Pm+)
NWD366746	79398	196	286	57	79937	Pf+	Pf+
NWD141952	539173	1269	1879	312	542633	Pf+	Pf+
NWD363693	27440	315	66039	120	93914	Mixed (Pf and Pm)	Pf+

NWD702221	126951	273	395	65	127684	Pf+	Pf+
NWD767250	182814	593	662	97	184166	Pf+	Pf+
NWD436146	15172	249	41612	71	57104	Mixed (Pf and Pm)	Mixed (Pf+ and Pm+)
NWD334060	406729	942	1406	185	409262	Pf+	Pf+
NWD603327	147332	362	553	92	148339	Pf+	Pf+
NWD276915	853131	1877	2906	434	858348	Pf+	Pf+
NWD452630	107054	221	379	61	107715	Pf+	Pf+
NWD901807	165714	673	5352	66	171805	Pf+	Pf+
NWD928158	516	486	165499	206	166707	Pm+	Pm+
NWD356682	337792	1294	4276	178	343540	Pf+	Pf+
NWD898494	411	239270	668	235	240584	Po+	Po+
NWD345308	36899	219	46714	70	83902	Mixed (Pf and Pm)	Mixed (Pf+ and Pm+)
NWD908043	630756	1424	2135	301	634616	Pf+	Pf+
NWD115137	802083	1664	2626	364	806737	Pf+	Pf+
NWD986644	90345	236	339	60	90980	Pf+	not tested
NWD265886	126014	285	2202	64	128565	Pf+	not tested
NWD251063	101981	239	351	58	102629	Pf+	not tested
NWD887797	136501	291	508	50	137350	Pf+	not tested
NWD857534	183288	390	651	109	184438	Pf+	not tested
NWD492274	115008	260	402	54	115724	Pf+	not tested
NWD104090	498172	1166	1788	326	501452	Pf+	not tested

NWD756394	278525	601	885	165	280176	Pf+	not tested
NWD355358	165627	351	588	82	166648	Pf+	not tested

Chapter 3 Appendix

Supplemental Table 1: Exact nucleotide position of literature-back variation in important antimalarial resistance genes

pfdhps				
Wt amino acid	**Codon**	**Mut amino acid**	**Codon**	**Position**
S	T	A*/F+	**G*/T+**	549681
	C		C*/**T+**	549682
	T		T*/T+	549683
A	G	G	G	549684
	C		**G**	549685
	T		T	549686
K	A	E	**G**	549993
	A		A	549994
	A		A	549995
A	G	G	G	550116
	C		**G**	550117
	G		G	550118
A	G	S	**T**	550212
	C		C	550213
	C		C	550214
pfdhfr				
Wt amino acid	**Codon**	**Mut amino acid**	**Codon**	**Position**
N	A	I	A	748238
	A		**T**	748239
	T		T	748240
C	T	R	**C**	748262
	G		G	748263
	T		T	748264
S	A	N*/T+	A*/A+	748409
	G		**A***/**C+**	748410
	C		C*/C+	748411
I	A	L	**C**	748577
	T		T	748578
	A		A	748579

pfmdr1				
Wt amino acid	**Codon**	**Mut amino acid**	**Codon**	**Position**
N	A	Y	**T**	958145
	A		A	958146
	T		T	958147
Y	T	F	T	958439
	A		**T**	958440
	T		T	958441
S	A	C	**T**	960989
	G		G	960990
	T		T	960991
N	A	D	**G**	961013
	A		A	961014
	T		T	961015
D	G	Y	**T**	961625
	A		A	961626
	T		T	961627
pfcrt				
Wt amino acid	**Codon**	**Mut amino acid**	**Codon**	**Position**
C	T	S	T	403612
	G		**C**	403613
	T		T	403614
V	G	*		403615
	T			403616
	A			403617
M	A	I	A	403618
	T		T	403619
	G		**A**	403620
N	A	E	**G**	403621
	A		A	403622
	T		**G/A**	403623
K	A	T	A	403624
	A		**C**	403625
	A		A	403626
pfk13				
Wt amino acid	**Codon**	**Mut amino acid**	**Codon**	**Position**
F	T	I	A	1725662
	T		T	1725661

	T		T	1725660
M	A	I	A	1725572
	T		T	1725571
	G		**T**	1725570
Y	T	H	**C**	1725521
	A		A	1725520
	C		C	1725519
R	A	T	A	1725383
	G		**C**	1725382
	A		A	1725381
I	A	T	A	1725371
	T		**C**	1725370
	T		T	1725369
P	C	L	C	1725341
	C		**T**	1725340
	G		G	1725339

Supplemental Table 2: Total unmapped reads (avg,range) available for query broken down by population

Country	Unmapped Reads (Avg)	Unmapped Reads (Range)
Ghana	4,547,026	(1,012,360-65,293,224)
Ethiopia	18,634,438	(758,702-158,646,922)
Guinea-Bissau	7,408,687	(1,871,744-17,346,606)

Botswanna	10,745,922	(1,131,964-588,201,938)
Tanzania	6,899,562	(340,734-141,996,086)
Kenya	74,644,620	(1,360,994-842,171,966)
Cameroon	172,816,097	(1,672,016-541,892,254)

Supplemental Table 3: Alternative calls for Mixed(all) infections

country	pf_reads	po_reads	pm_reads	pv_reads	Alternative Diagnosis
Cameroon	1771	1438	500958	1094	Pm+
Cameroon	25176	264	69604	227	Mixed (Pf and Pm)
Cameroon	3022	1917	699621	1009	Pm+
Cameroon	1170	209	69613	89	Pm+
Cameroon	134	103	40594	76	Pm+

Cameroon	27672	34524	241	80	Mixed (Pf and Po)
Cameroon	63461	205	19658	72	Mixed (Pf and Pm)
Cameroon	2159	779	261963	321	Pm+
Cameroon	277	145	37141	87	Pm+
Cameroon	21256	174	18587	64	Mixed (Pf and Pm)
Cameroon	3525	334	73117	235	Pm+
Cameroon	229927	568	856	140	Pf+
Cameroon	171919	409	5864	147	Pf+
Cameroon	482	376	123997	277	Pm+
Cameroon	1606	57	19542	51	Pm+
Cameroon	111643	269	403	70	Pf+
Cameroon	134660	537	920	187	Pf+
Ghana	97726	202	979	63	Pf+
Ghana	2327	37229	146	115	Po+
Ghana	114611	6698	436	86	Pf+
Ghana	29091	105	152	87	Pf+
Ghana	62546	153	501	53	Pf+
Ghana	74335	410	5383	79	Pf+
Ghana	24554	79	2484	52	Mixed (Pf and Pm)
Ghana	143828	338	584	78	Pf+
Ghana	43719	224	34210	76	Mixed (Pf and Pm)
Ghana	2035274	4548	8208	1013	Pf+
Ghana	259148	869	6741	99	Pf+
Ghana	108684	366	384	70	Pf+
Ghana	40225	946	181	140	Pf+
Ghana	913728	2055	31002	480	Pf+
Ghana	88073	236	296	50	Pf+
Ghana	162730	389	533	101	Pf+
Ghana	186888	468	729	100	Pf+
Ghana	130553	373	24895	95	Mixed (Pf and Pm)
Ghana	79398	196	286	57	Pf+
Ghana	539173	1269	1879	312	Pf+
Ghana	27440	315	66039	120	Mixed (Pf and Pm)
Ghana	126951	273	395	65	Pf+
Ghana	182814	593	662	97	Pf+
Ghana	15172	249	41612	71	Mixed (Pf and Pm)
Ghana	406729	942	1406	185	Pf+
Ghana	147332	362	553	92	Pf+
Ghana	853131	1877	2906	434	Pf+

Ghana	107054	221	379	61	Pf+
Ghana	165714	673	5352	66	Pf+
Ghana	516	486	165499	206	Pm+
Ghana	337792	1294	4276	178	Pf+
Ghana	411	239270	668	235	Po+
Ghana	36899	219	46714	70	Mixed (Pf and Pm)
Ghana	630756	1424	2135	301	Pf+
Ghana	802083	1664	2626	364	Pf+
Ghana	90345	236	339	60	Pf+
Ghana	126014	285	2202	64	Pf+
Ghana	101981	239	351	58	Pf+
Ghana	136501	291	508	50	Pf+
Ghana	183288	390	651	109	Pf+
Ghana	115008	260	402	54	Pf+
Ghana	498172	1166	1788	326	Pf+
Ghana	278525	601	885	165	Pf+
Ghana	165627	351	588	82	Pf+
Kenya	201126	458	872	132	Pf+
Kenya	164776	339	532	67	Pf+
Kenya	83734	220	576	56	Pf+
Kenya	262950	2260	605249	880	Mixed (Pf and Pm)
Kenya	15120	141	30892	53	Mixed (Pf and Pm)
Kenya	46783	42254	14909787	23475	Pm+
Kenya	12341	471	153759	329	Pm+
Kenya	849461	1794	2818	458	Pf+
Kenya	81917	596	2028	51	Pf+
Kenya	377496	853	1399	162	Pf+
Kenya	175867	589	808	133	Pf+
Kenya	474231	1082	1805	275	Pf+
Kenya	50574	256	46107	80	Mixed (Pf and Pm)
Kenya	215467	478	764	136	Pf+
Tanzania	157418	472	791	112	Pf+
Tanzania	853881	1895	2839	441	Pf+

Supplemental Table 4: Cameroon Breakdown of coverage across all antimalarial resistance genes at positions of historical interest

Cameroon

Gene	Codon	Wild type amino acid	Mutant amino acid	Individuals with Mutations	Individuals with Coverage	Read Coverage in pooled samples
sulfadoxine-pyrimethamine						
pf-dhps	436	S	A	7	13	43
	437	A	G	13	13	44
	540	K	E	0	17	46
	581	A	G	2	11	41
	613	A	S	2	13	40
pf-dhfr	51	N	I	12	12	32
	59	C	R	9	10	31
	108	S	N	16	16	51
	164	I	L	0	16	50
chloroquine						
pf-mdr1	86	N	Y	6	12	41
	184	Y	F	11	20	47
	1034	S	C	0	13	45
	1042	N	D	0	12	43
	1246	D	Y	0	17	55
pf-crt	72	C	S	0	10	19
	73	V	^	0	10	19
	74	M	I	1	9	16
	75	N	E	0	8	16
	76	K	T	0	8	16
artemisinin						
pf-k13	446	F	I	0	19	72
	476	M	I	0	20	62
	493	Y	H	0	14	48
	539	R	T	0	15	58
	543	I	T	0	15	59
	553	P	L	0	16	63

Supplemental Table 5: Ethiopia Breakdown of coverage across all antimalarial resistance genes at positions of historical interest

Ethiopia

Gene	Codon	Wild type amino acid	Mutant amino acid	Individuals with Mutations	Individuals with Coverage	Read Coverage in pooled samples
sulfadoxine-pyrimethamine						
pf-dhps	436	S	A	0	2	3
	437	A	G	1	2	3
	540	K	E	0	1	1
	581	A	G	0	2	2
	613	A	S	0	1	1
pf-dhfr	51	N	I	1	1	1
	59	C	R	0	1	1
	108	S	N	1	1	1
	164	I	L	0	1	1
chloroquine						
pf-mdr1	86	N	Y	1	2	3
	184	Y	F	1	1	1
	1034	S	C	0	2	3
	1042	N	D	0	2	2
	1246	D	Y	1	3	3
pf-crt	72	C	S	0	1	5
	73	V	^	0	1	5
	74	M	I	0	0	0
	75	N	E	0	0	0
	76	K	T	0	0	0
artemisinin						
pf-k13	446	F	I	0	3	3
	476	M	I	0	2	2
	493	Y	H	0	2	3
	539	R	T	0	3	6
	543	I	T	0	3	5
	553	P	L	0	2	5

Supplemental Table 6: Ghana Breakdown of coverage across all antimalarial resistance genes at positions of historical interest

Ghana

Gene	Codon	Wild type amino acid	Mutant amino acid	Individuals with Mutations	Individuals with Coverage	Read Coverage in pooled samples
sulfadoxine-pyrimethamine						
pf-dhps	436	S	A	41	51	134
	437	A	G	43	52	135
	540	K	E	0	37	172
	581	A	G	0	37	160
	613	A	S	6	35	149
pf-dhfr	51	N	I	21	48	111
	59	C	R	32	51	114
	108	S	N	40	54	127
	164	I	L	0	66	158
chloroquine						
pf-mdr1	86	N	Y	19	75	189
	184	Y	F	45	73	166
	1034	S	C	0	71	162
	1042	N	D	0	68	163
	1246	D	Y	3	64	184
pf-crt	72	C	S	0	51	88
	73	V	^	0	50	82
	74	M	I	1	34	51
	75	N	E	0	34	50
	76	K	T	0	34	48
artemisinin						
pf-k13	446	F	I	0	69	160
	476	M	I	0	68	152
	493	Y	H	0	67	138
	539	R	T	0	76	162
	543	I	T	0	78	167
	553	P	L	1	74	164

Supplemental Table 7: Guinea-Bissau Breakdown of coverage across all antimalarial resistance genes at positions of historical interest

Guinea-Bissau

Gene	Codon	Wild type amino acid	Mutant amino acid	Individuals with Mutations	Individuals with Coverage	Read Coverage in pooled samples
sulfadoxine-pyrimethamine						
pf-dhps	436	S	A	0	0	0
	437	A	G	0	0	0
	540	K	E	0	0	0
	581	A	G	0	0	0
	613	A	S	0	0	0
pf-dhfr	51	N	I	0	0	0
	59	C	R	0	0	0
	108	S	N	0	0	0
	164	I	L	0	0	0
chloroquine						
pf-mdr1	86	N	Y	0	0	0
	184	Y	F	0	0	0
	1034	S	C	0	1	1
	1042	N	D	0	1	1
	1246	D	Y	1	1	1
pf-crt	72	C	S	0	0	0
	73	V	^	0	0	0
	74	M	I	0	0	0
	75	N	E	0	0	0
	76	K	T	0	0	0
artemisinin						
pf-k13	446	F	I	0	0	0
	476	M	I	0	0	0
	493	Y	H	0	0	0
	539	R	T	0	0	0
	543	I	T	0	0	0
	553	P	L	0	0	0

Supplemental Table 8: Kenya Breakdown of coverage across all antimalarial resistance genes at positions of historical interest

Kenya

Gene	Codon	Wild type amino acid	Mutant amino acid	Individuals with Mutations	Individuals with Coverage	Read Coverage in pooled samples
sulfadoxine-pyrimethamine						
pf-dhps	436	S	A	1	11	89
	437	A	G	10	11	94
	540	K	E	12	12	70
	581	A	G	0	10	84
	613	A	S	0	9	96
pf-dhfr	51	N	I	15	15	69
	59	C	R	12	15	67
	108	S	N	19	19	89
	164	I	L	0	17	75
chloroquine						
pf-mdr1	86	N	Y	8	13	78
	184	Y	F	8	14	63
	1034	S	C	0	20	77
	1042	N	D	0	21	86
	1246	D	Y	8	17	88
pf-crt	72	C	S	0	13	55
	73	V	^	0	13	52
	74	M	I	1	3	4
	75	N	E	0	2	4
	76	K	T	1	3	5
artemisinin						
pf-k13	446	F	I	0	15	133
	476	M	I	0	19	109
	493	Y	H	0	19	101
	539	R	T	1	21	93
	543	I	T	0	21	94
	553	P	L	0	18	87

Supplemental Table 9: Tanzania Breakdown of coverage across all antimalarial resistance genes at positions of historical interest

Tanzania

Gene	Codon	Wild type amino acid	Mutant amino acid	Individuals with Mutations	Individuals with Coverage	Read Coverage in pooled samples
sulfadoxine-pyrimethamine						
pf-dhps	436	S	A	0	4	12
	437	A	G	4	4	12
	540	K	E	2	2	24
	581	A	G	0	4	28
	613	A	S	0	4	20
pf-dhfr	51	N	I	3	3	6
	59	C	R	2	3	4
	108	S	N	1	1	3
	164	I	L	0	2	7
chloroquine						
pf-mdr1	86	N	Y	1	4	12
	184	Y	F	1	3	9
	1034	S	C	0	3	13
	1042	N	D	0	4	13
	1246	D	Y	0	6	10
pf-crt	72	C	S	0	2	4
	73	V	^	0	1	3
	74	M	I	0	1	1
	75	N	E	0	1	1
	76	K	T	0	1	1
artemisinin						
pf-k13	446	F	I	0	3	8
	476	M	I	0	3	6
	493	Y	H	0	3	6
	539	R	T	0	2	12
	543	I	T	0	3	12

	553	P	L	0	3	12

Supplemental Table 10: Systemic sequencing error found in *pfcrt* across populations

Pf3D7_07_v3 403612	Pf3D7_07_v3 403613	Pf3D7_07_v3 403614
T	G	T
<*>	<*>	<*>
PL:DP:AD	PL:DP:AD	PL:DP:AD
0,3,40:1:1,0	0,3,40:1:1,0	0,3,40:1:1,0
0,3,40:1:1,0	0,3,40:1:1,0	0,3,40:1:1,0

72

Pf3D7_07_v3 403615	Pf3D7_07_v3 403616	Pf3D7_07_v3 403617
G	T	A
<*>	<*>	<*>
PL:DP:AD	PL:DP:AD	PL:DP:AD
0,3,40:1:1,0	0,3,40:1:1,0	0,3,40:1:1,0
0,3,40:1:1,0	0,3,40:1:1,0	0,3,40:1:1,0

73

Pf3D7_07_v3 403618	Pf3D7_07_v3 403619	Pf3D7_07_v3 403620
A	T	G
<*>	<*>	<*>
PL:DP:AD	PL:DP:AD	PL:DP:AD
0,3,40:1:1,0	0,3,40:1:1,0	0,3,40,3,40,40:1:1,0
0,3,40:1:1,0	0,0,0:0:0,0	0,3,30,3,30,30:1:1,0

74

Pf3D7_07_v3 403621	Pf3D7_07_v3 403622	Pf3D7_07_v3 403623
A	A	T
<*>	<*>	<*>
PL:DP:AD	PL:DP:AD	PL:DP:AD

0,3,40:1:1,0	0,3,40:1:1,0	0,3,40:1:1,0
0,3,23:1:1,0	0,0,0:0:0,0	0,0,0:0:0,0

75

Pf3D7_07_v3 403624	Pf3D7_07_v3 403625	Pf3D7_07_v3 403626
A	A	A
<*>	<*>	<*>
PL:DP:AD	PL:DP:AD	PL:DP:AD
0,3,40:1:1,0	0,3,40:1:1,0	0,3,40:1:1,0
0,0,0:0:0,0	0,0,0:0:0,0	0,3,21:1:1,0

76

Normal coverage is represented by colored in cells, the repetitive systemic error is represented by white cells.

Supplemental Table 11: Variation in antimalarial resistance genes that are not supported by previous literature which eliminated the individual

Individuals not included in estimates

Mutation Found	Gene	Position	Read Coverage	Country	Synonymous
TTC (F)	*pfk13*	446	4	Cameroon	Y
TTA (L)	*pfcrt*	73	1	Cameroon	N
ATT (I)	*pfcrt*	75	1	Cameroon	N
AAT (N)	*pfcrt*	76	1	Cameroon	N
TTC (F)	*pfk13*	446	2	Ghana	Y
TGT(C)	*pfk13*	446	1	Ghana	N
AGG (R)	*pfk13*	476	1	Ghana	N
CAT (H)	*pfmdr1*	1042	3	Ghana	N
AAA (K)	*pfmdr1*	1042	1	Ghana	N
TTA (L)	*pfcrt*	73	1	Ghana	N
ATT (I)	*pfcrt*	75	1	Ghana	N

AAT (N)	*pfcrt*	76	1	Ghana	N
ATT (I)	*pfmdr1*	1034	1	Ghana	N
TTC (F)	*pfk13*	446	5	Kenya	Y
TGT(C)	*pfk13*	446	1	Kenya	N
CCA (P)	*pfk13*	553	5	Kenya	Y
TAA (STOP)	*pfcrt*	75	1	Kenya	N
GGT (V)	*pfcrt*	72	1	Tanzania	N
GAA (I)	*pfcrt*	73	1	Tanzania	N

Supplemental Table 12: Variation in antimalarial resistance genes that are not supported by previous literature which eliminated the read

Reads not included in estimates

Mutation Found	Gene	Position	Read Coverage	Country	Synonymous
TCG (S)	*pfk13*	553	1	Cameroon	N
GAC (D)	*pfdhps*	581	1	Cameroon	N
GAT (D)	*pfmdr1*	86	1	Cameroon	N
TAG (STOP)	*pfk13*	493	1	Ghana	N
AGT (S)	*pfk13*	539	1	Ghana	N
ATA (1)	*pfk13*	543	2	Ghana	Y
AAA (K)	*pfmdr1*	1042	1	Ghana	N
TAA (STOP)	*pfmdr1*	86	1	Ghana	N
AAA (K)	*pfdhfr*	164	1	Ghana	N
ACT (T)	*pfk13*	476	1	Kenya	N
GTC (V)	*pfk13*	543	1	Kenya	N
GCG (A)	*pfdhps*	613	1	Kenya	Y
ATT (I)	*pfdhps*	613	1	Kenya	N
CGC (R)	*pfdhps*	581	1	Kenya	N
TGA (STOP)	*pfdhps*	540	1	Kenya	N
TAT (Y)	*pfdhps*	437	1	Kenya	N
TGT (C)	*pfdhps*	436	1	Kenya	N
AGT (S)	*pfdhfr*	51	1	Kenya	N
CTA (L)	*pfdhfr*	59	1	Kenya	N
AAA (K)	*pfdhfr*	108	1	Kenya	N
TTA (L)	*pfcrt*	72	1	Kenya	N

TTT (F)	*pfcrt*	73	1	Kenya	N
AGT (S)	*pfcrt*	72	1	Kenya	Y
CAT (H)	*pfmdr1*	1042	1	Kenya	N
CAT (H)	*pfmdr1*	1246	1	Kenya	N
GAT (D)	*pfmdr1*	86	1	Tanzania	N
ACC (T)	*pfdhps*	613	1	Tanzania	N

Citations

1. Neghina, R., Iacobiciu, I., Neghina, A.M., and Marincu, I. (2010). Malaria, a Journey in Time: In Search of the Lost Myths and Forgotten Stories. The American Journal of the Medical Sciences 340, 492-498.
2. Schlagenhauf, P. (2004). Malaria: from prehistory to present. Infectious disease clinics of North America 18, 189-205, table of contents.
3. Cox, F.E.G. (2010). History of the discovery of the malaria parasites and their vectors. Parasites & Vectors 3, 5.
4. Meibalan, E., and Marti, M. (2017). Biology of Malaria Transmission. Cold Spring Harb Perspect Med 7.
5. Phillips, M.A., Burrows, J.N., Manyando, C., van Huijsduijnen, R.H., Van Voorhis, W.C., and Wells, T.N.C. (2017). Malaria. Nature Reviews Disease Primers 3, 17050.
6. White, N.J. (1996). The treatment of malaria. The New England journal of medicine 335, 800-806.
7. Fitri, L.E., Widaningrum, T., Endharti, A.T., Prabowo, M.H., Winaris, N., and Nugraha, R.Y.B. (2022). Malaria diagnostic update: From conventional to advanced method. J Clin Lab Anal 36, e24314.
8. Zimmerman, P.A., and Howes, R.E. (2015). Malaria diagnosis for malaria elimination. Current opinion in infectious diseases 28, 446-454.
9. White, N.J. (2022). Severe malaria. Malaria Journal 21, 284.
10. Bartoloni, A., and Zammarchi, L. (2012). Clinical aspects of uncomplicated and severe malaria. Mediterranean journal of hematology and infectious diseases 4, e2012026.
11. Organization, W.H. (2010). Guidelines for the treatment of malaria.(Geneva, Switzerland: WHO Press).
12. Moxon, C.A., Gibbins, M.P., McGuinness, D., Jr., D.A.M., and Marti, M. (2020). New Insights into Malaria Pathogenesis. 15, 315-343.
13. Wahlgren, M., Goel, S., and Akhouri, R.R. (2017). Variant surface antigens of Plasmodium falciparum and their roles in severe malaria. Nature Reviews Microbiology 15, 479-491.
14. Rahimi, B.A., Thakkinstian, A., White, N.J., Sirivichayakul, C., Dondorp, A.M., and Chokejindachai, W. (2014). Severe vivax malaria: a systematic review and meta-analysis of clinical studies since 1900. Malaria Journal 13, 481.
15. Trape, J.F. (2001). The public health impact of chloroquine resistance in Africa. The American journal of tropical medicine and hygiene 64, 12-17.
16. Ministry of Health, G. (2009). Anti-malaria drug policy for Ghana. In. (Ghana, Ministry of Health.
17. Global Partnership to Roll Back, M., and United Nations Children's, F. (2005). World malaria report : 2005. In. (Geneva, World Health Organization.
18. R, O. (2000). Beta-Thalassemia. In GeneReviews® E.D. Adam MP, Mirzaa GM, et al., editors., ed. (Seattle (WA), University of Washington, Seattle.
19. Kariuki, S.N., and Williams, T.N. (2020). Human genetics and malaria resistance. Human genetics 139, 801-811.
20. Bender, M.A., and Carlberg, K. (1993). Sickle Cell Disease. In GeneReviews(®), M.P. Adam, D.B. Everman, G.M. Mirzaa, R.A. Pagon, S.E. Wallace, L.J.H. Bean, K.W. Gripp, andA. Amemiya, eds. (Seattle (WA), University of Washington, Seattle

21. Rockett, K.A., Clarke, G.M., Fitzpatrick, K., Hubbart, C., Jeffreys, A.E., Rowlands, K., Craik, R., Jallow, M., Conway, D.J., Bojang, K.A., et al. (2014). Reappraisal of known malaria resistance loci in a large multicenter study. Nature Genetics 46, 1197-1204.
22. Torres, L.d.S., Okumura, J.V., Silva, D.G.H.d., and Bonini-Domingos, C.R. (2015). Hemoglobin D-Punjab: origin, distribution and laboratory diagnosis. Revista Brasileira de Hematologia e Hemoterapia 37, 120-126.
23. Sangare, A., Sanogo, I., Meite, M., Ambofo, Y., Abesopie, V., Segbena, A., and Tolo, A. (1992). [Hemoglobin O Arab in Ivory Coast and western Africa]. Medecine tropicale : revue du Corps de sante colonial 52, 163-167.
24. Azevedo, J.H.a.A.M. (2022). StatPearls [Internet]. In Alpha Thalassemia

(National Library of Medicine, Treasure Island (FL): StatPearls Publishing.
25. Ndila, C.M., Uyoga, S., Macharia, A.W., Nyutu, G., Peshu, N., Ojal, J., Shebe, M., Awuondo, K.O., Mturi, N., Tsofa, B., et al. (2018). Human candidate gene polymorphisms and risk of severe malaria in children in Kilifi, Kenya: a case-control association study. The Lancet Haematology 5, e333-e345.
26. Maheshwari, A., and Killeen, R.B. (2022). Duffy Blood Group System. In StatPearls. (Treasure Island (FL), StatPearls Publishing

27. Kasehagen, L.J., Mueller, I., Kiniboro, B., Bockarie, M.J., Reeder, J.C., Kazura, J.W., Kastens, W., McNamara, D.T., King, C.H., Whalen, C.C., et al. (2007). Reduced Plasmodium vivax erythrocyte infection in PNG Duffy-negative heterozygotes. PLoS One 2, e336.
28. World Health, O. (2021). World malaria report 2021.(Geneva: World Health Organization).
29. Organization, W.H. (2022). World malaria report 2022. In, G.M. Programme, ed. (Geneva, Switzerland, World Health Organization), p 296.
30. Naserrudin, N.A., Hassan, M.R., Jeffree, M.S., Culleton, R., Hod, R., and Ahmed, K. (2022). A systematic review of asymptomatic Plasmodium knowlesi infection: an emerging challenge involving an emerging infectious disease. Malaria Journal 21, 373.
31. Battle, K.E., Lucas, T.C.D., Nguyen, M., Howes, R.E., Nandi, A.K., Twohig, K.A., Pfeffer, D.A., Cameron, E., Rao, P.C., Casey, D., et al. (2019). Mapping the global endemicity and clinical burden of Plasmodium vivax, 2000–17: a spatial and temporal modelling study. The Lancet 394, 332-343.
32. Howes, R.E., Patil, A.P., Piel, F.B., Nyangiri, O.A., Kabaria, C.W., Gething, P.W., Zimmerman, P.A., Barnadas, C., Beall, C.M., Gebremedhin, A., et al. (2011). The global distribution of the Duffy blood group. Nature Communications 2, 266.
33. Zimmerman, P.A. (2017). Plasmodium vivax Infection in Duffy-Negative People in Africa. The American journal of tropical medicine and hygiene 97, 636-638.
34. Ménard, D., Barnadas, C., Bouchier, C., Henry-Halldin, C., Gray, L.R., Ratsimbasoa, A., Thonier, V., Carod, J.F., Domarle, O., Colin, Y., et al. (2010). Plasmodium vivax clinical malaria is commonly observed in Duffy-negative Malagasy people. Proceedings of the National Academy of Sciences of the United States of America 107, 5967-5971.
35. Gumbo, A., Topazian, H.M., Mwanza, A., Mitchell, C.L., Puerto-Meredith, S., Njiko, R., Kayange, M., Mwalilino, D., Mvula, B., Tegha, G., et al. (2022). Occurrence and Distribution of Nonfalciparum Malaria Parasite Species Among Adolescents and Adults in Malawi. The Journal of Infectious Diseases 225, 257-268.

36. Daniels, R.F., Deme, A.B., Gomis, J.F., Dieye, B., Durfee, K., Thwing, J.I., Fall, F.B., Ba, M., Ndiop, M., Badiane, A.S., et al. (2017). Evidence of non-Plasmodium falciparum malaria infection in Kédougou, Sénégal. Malaria Journal 16, 9.
37. Betson, M., Clifford, S., Stanton, M., Kabatereine, N.B., and Stothard, J.R. (2018). Emergence of Nonfalciparum Plasmodium Infection Despite Regular Artemisinin Combination Therapy in an 18-Month Longitudinal Study of Ugandan Children and Their Mothers. The Journal of Infectious Diseases 217, 1099-1109.
38. Lover, A.A., Baird, J.K., Gosling, R., and Price, R.N. (2018). Malaria Elimination: Time to Target All Species. The American journal of tropical medicine and hygiene 99, 17-23.
39. Amoah, L.E., Donu, D., Abuaku, B., Ahorlu, C., Arhinful, D., Afari, E., Malm, K., and Koram, K.A. (2019). Probing the composition of Plasmodium species contained in malaria infections in the Eastern region of Ghana. BMC public health 19, 1617.
40. Gimenez, A.M., Marques, R.F., Regiart, M., and Bargieri, D.Y. (2021). Diagnostic Methods for Non-Falciparum Malaria. Frontiers in cellular and infection microbiology 11, 681063.
41. Nchinda, T.C. (1998). Malaria: a reemerging disease in Africa. Emerging infectious diseases 4, 398-403.
42. MacCormack, C.P. (1984). Human ecology and behaviour in malaria control in tropical Africa. Bulletin of the World Health Organization 62 Suppl, 81-87.
43. Cibulskis, R.E., Alonso, P., Aponte, J., Aregawi, M., Barrette, A., Bergeron, L., Fergus, C.A., Knox, T., Lynch, M., Patouillard, E., et al. (2016). Malaria: Global progress 2000 – 2015 and future challenges. Infectious Diseases of Poverty 5, 61.
44. Mbacham, W.F., Ayong, L., Guewo-Fokeng, M., and Makoge, V. (2019). Current Situation of Malaria in Africa. In Malaria Control and Elimination, F. Ariey, F. Gay, andR. Ménard, eds. (New York, NY, Springer New York), pp 29-44.
45. (1997). The spirit of Dakar: a call for action on malaria. Nature 386, 541-541.
46. Roberts, L., and Enserink, M. (2007). Malaria. Did they really say ... eradication? Science 318, 1544-1545.
47. Programme, G.M. (2021). Global technical strategy for malaria 2016–2030: 2021 Update. In, W.H. Organization, ed. (Geneva, Switzerland, World Health Organization), p 40.
48. Cohen, J.M., Okumu, F., and Moonen, B. The fight against malaria: Diminishing gains and growing challenges. Science Translational Medicine 14, eabn3256.
49. Organization, W.H. (2016). Eliminating Malaria. In. (Geneva, Switzerland, WHO Press), p 28.
50. Nájera, J.A., González-Silva, M., and Alonso, P.L. (2011). Some lessons for the future from the Global Malaria Eradication Programme (1955-1969). PLoS Med 8, e1000412.
51. Fenner, F., Henderson, D.A., Arita, I., Jezek, Z., Ladnyi, I.D., and World Health, O. (1988). Smallpox and its eradication / F. Fenner ... [et al.]. In. (Geneva, World Health Organization.
52. Zavala, F. (2022). RTS,S: the first malaria vaccine. The Journal of clinical investigation 132.
53. Laurens, M.B. (2020). RTS,S/AS01 vaccine (Mosquirix™): an overview. Human vaccines & immunotherapeutics 16, 480-489.
54. Ashley, E.A., Pyae Phyo, A., and Woodrow, C.J. (2018). Malaria. Lancet 391, 1608-1621.
55. Nabarro, D.N., and Tayler, E.M. (1998). The "roll back malaria" campaign. Science 280, 2067-2068.
56. WHO Regional Office for Africa. (2004). A Strategic Framework for Malaria Prevention and Control During Pregnancy in the African Region. In. (Brazzaville, World Health Organization.
57. Shretta, R., Liu, J., Cotter, C., Cohen, J., Dolenz, C., Makomva, K., Newby, G., Ménard, D., Phillips, A., Tatarsky, A., et al. (2017). Malaria Elimination and Eradication. In. (The

International Bank for Reconstruction and Development / The World Bank, Washington (DC).
58. World Health, O. (2015). World malaria report 2015.(Geneva: World Health Organization).
59. Gething, P.W., Casey, D.C., Weiss, D.J., Bisanzio, D., Bhatt, S., Cameron, E., Battle, K.E., Dalrymple, U., Rozier, J., Rao, P.C., et al. (2016). Mapping Plasmodium falciparum Mortality in Africa between 1990 and 2015. 375, 2435-2445.
60. Lin, J.T., Juliano, J.J., and Wongsrichanalai, C. (2010). Drug-Resistant Malaria: The Era of ACT. Current Infectious Disease Reports 12, 165-173.
61. Hanboonkunupakarn, B., and White, N.J. (2022). Advances and roadblocks in the treatment of malaria. Br J Clin Pharmacol 88, 374-382.
62. Xie, S.C., Ralph, S.A., and Tilley, L. (2020). K13, the Cytostome, and Artemisinin Resistance. Trends in parasitology 36, 533-544.
63. Alker, A.P., Kazadi, W.M., Kutelemeni, A.K., Bloland, P.B., Tshefu, A.K., and Meshnick, S.R. (2008). dhfr and dhps genotype and sulfadoxine-pyrimethamine treatment failure in children with falciparum malaria in the Democratic Republic of Congo. Tropical medicine & international health : TM & IH 13, 1384-1391.
64. Chaturvedi, R., Chhibber-Goel, J., Verma, I., Gopinathan, S., Parvez, S., and Sharma, A. (2021). Geographical spread and structural basis of sulfadoxine-pyrimethamine drug-resistant malaria parasites. International Journal for Parasitology 51, 505-525.
65. Otienoburu, S.D., Suay, I., Garcia, S., Thomas, N.V., Srisutham, S., Björkman, A., and Humphreys, G.S. (2019). An online mapping database of molecular markers of drug resistance in Plasmodium falciparum: the ACT Partner Drug Molecular Surveyor. Malar J 18, 12.
66. Borges, S., Cravo, P., Creasey, A., Fawcett, R., Modrzynska, K., Rodrigues, L., Martinelli, A., and Hunt, P. (2011). Genomewide Scan Reveals Amplification of mdr1 as a Common Denominator of Resistance to Mefloquine, Lumefantrine, and Artemisinin in Plasmodium chabaudi Malaria Parasites. 55, 4858-4865.
67. Ndiath, M.O. (2019). Insecticides and Insecticide Resistance. In Malaria Control and Elimination, F. Ariey, F. Gay, andR. Ménard, eds. (New York, NY, Springer New York), pp 287-304.
68. Sougoufara, S., Doucouré, S., Backé Sembéne, P.M., Harry, M., and Sokhna, C. (2017). Challenges for malaria vector control in sub-Saharan Africa: Resistance and behavioral adaptations in Anopheles populations. Journal of vector borne diseases 54, 4-15.
69. Prusty, D., Gupta, N., Upadhyay, A., Dar, A., Naik, B., Kumar, N., and Prajapati, V.K. (2021). Asymptomatic malaria infection prevailing risks for human health and malaria elimination. Infection, Genetics and Evolution 93, 104987.
70. Lindblade, K.A., Steinhardt, L., Samuels, A., Kachur, S.P., and Slutsker, L. (2013). The silent threat: asymptomatic parasitemia and malaria transmission. Expert Review of Anti-infective Therapy 11, 623-639.
71. Heinemann, M., Phillips, R.O., Vinnemeier, C.D., Rolling, C.C., Tannich, E., and Rolling, T. (2020). High prevalence of asymptomatic malaria infections in adults, Ashanti Region, Ghana, 2018. Malaria Journal 19, 366.
72. Andolina, C., Rek, J.C., Briggs, J., Okoth, J., Musiime, A., Ramjith, J., Teyssier, N., Conrad, M., Nankabirwa, J.I., Lanke, K., et al. (2021). Sources of persistent malaria transmission in a setting with effective malaria control in eastern Uganda: a longitudinal, observational cohort study. The Lancet Infectious Diseases 21, 1568-1578.

73. Alves, F.P., Gil, L.H.S., Marrelli, M.T., Ribolla, P.E.M., Camargo, E.P., and Da Silva, L.H.P. (2005). Asymptomatic Carriers of Plasmodium spp. as Infection Source for Malaria Vector Mosquitoes in the Brazilian Amazon. Journal of Medical Entomology 42, 777-779.
74. Gonçalves, B.P., Kapulu, M.C., Sawa, P., Guelbéogo, W.M., Tiono, A.B., Grignard, L., Stone, W., Hellewell, J., Lanke, K., and Bastiaens, G.J.J.N.c. (2017). Examining the human infectious reservoir for Plasmodium falciparum malaria in areas of differing transmission intensity. 8, 1-11.
75. Akiyama, T., Pongvongsa, T., Phrommala, S., Taniguchi, T., Inamine, Y., Takeuchi, R., Watanabe, T., Nishimoto, F., Moji, K., Kano, S., et al. (2016). Asymptomatic malaria, growth status, and anaemia among children in Lao People's Democratic Republic: a cross-sectional study. Malaria Journal 15, 499.
76. Tamiru, A., Tolossa, T., Regasa, B., and Mosisa, G. (2022). Prevalence of asymptomatic malaria and associated factors in Ethiopia: Systematic review and meta-analysis. SAGE open medicine 10, 20503121221088085.
77. Nwali, M., Umeora, O., Ozumba, B., Onoh, R., Ezeonu, P., Agwu, U.J.I.J.o.D., and Sciences, M. (2014). Anemia among unbooked parturients with asymptomatic malaria parasitemia at a tertiary institution southeast Nigeria. 13, 54-57.
78. Sherrard-Smith, E., Hogan, A.B., Hamlet, A., Watson, O.J., Whittaker, C., Winskill, P., Ali, F., Mohammad, A.B., Uhomoibhi, P., Maikore, I., et al. (2020). The potential public health consequences of COVID-19 on malaria in Africa. Nature medicine 26, 1411-1416.
79. Weiss, D.J., Bertozzi-Villa, A., Rumisha, S.F., Amratia, P., Arambepola, R., Battle, K.E., Cameron, E., Chestnutt, E., Gibson, H.S., Harris, J., et al. (2021). Indirect effects of the COVID-19 pandemic on malaria intervention coverage, morbidity, and mortality in Africa: a geospatial modelling analysis. The Lancet Infectious diseases 21, 59-69.
80. Agrawal, P.K., Agrawal, C., and Blunden, G. (2022). Artemisia Extracts and Artemisinin-Based Antimalarials for COVID-19 Management: Could These Be Effective Antivirals for COVID-19 Treatment? Molecules (Basel, Switzerland) 27.
81. Cao, R., Hu, H., Li, Y., Wang, X., Xu, M., Liu, J., Zhang, H., Yan, Y., Zhao, L., Li, W., et al. (2020). Anti-SARS-CoV-2 Potential of Artemisinins In Vitro. ACS Infectious Diseases 6, 2524-2531.
82. Cruciani, M., Pati, I., Masiello, F., Malena, M., Pupella, S., and De Angelis, V. (2021). Ivermectin for Prophylaxis and Treatment of COVID-19: A Systematic Review and Meta-Analysis. Diagnostics (Basel, Switzerland) 11.
83. Ataba, E., Dorkenoo, A.M., Nguepou, C.T., Bakai, T., Tchadjobo, T., Kadzahlo, K.D., Yakpa, K., and Atcha-Oubou, T. (2022). Potential Emergence of Plasmodium Resistance to Artemisinin Induced by the Use of Artemisia annua for Malaria and COVID-19 Prevention in Sub-African Region. Acta parasitologica 67, 55-60.
84. Wollman, A.J., Nudd, R., Hedlund, E.G., and Leake, M.C. (2015). From Animaculum to single molecules: 300 years of the light microscope. Open biology 5, 150019.
85. Barcia, J.J. (2007). The Giemsa stain: its history and applications. International journal of surgical pathology 15, 292-296.
86. Wambani, J., and Okoth, P. (2022). Impact of Malaria Diagnostic Technologies on the Disease Burden in the Sub-Saharan Africa. Journal of tropical medicine 2022, 7324281.
87. Martiáñez-Vendrell, X., Skjefte, M., Sikka, R., and Gupta, H. (2022). Factors Affecting the Performance of HRP2-Based Malaria Rapid Diagnostic Tests. 7, 265.
88. Mpina, M., Stabler, T.C., Schindler, T., Raso, J., Deal, A., Acuche Pupu, L., Nyakarungu, E., del Carmen Ovono Davis, M., Urbano, V., Mtoro, A., et al. (2022). Diagnostic performance and comparison of ultrasensitive and conventional rapid diagnostic test, thick blood

smear and quantitative PCR for detection of low-density Plasmodium falciparum infections during a controlled human malaria infection study in Equatorial Guinea. Malaria Journal 21, 99.
89. Alkan, M.L. (2020). The Importance of Submicroscopic Diagnosis of Malaria. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 71, 175-176.
90. Katrak, S., Nayebare, P., Rek, J., Arinaitwe, E., Nankabirwa, J.I., Kamya, M., Dorsey, G., Rosenthal, P.J., and Greenhouse, B. (2018). Clinical consequences of submicroscopic malaria parasitaemia in Uganda. Malar J 17, 67.
91. Mahittikorn, A., Masangkay, F.R., Kotepui, K.U., De Jesus Milanez, G., and Kotepui, M. (2021). Comparative performance of PCR using DNA extracted from dried blood spots and whole blood samples for malaria diagnosis: a meta-analysis. Sci Rep 11, 4845.
92. Tedrow, R.E., Ratovonjato, J., Walker, E.D., Ratsimbasoa, A.C., and Zimmerman, P.A. (2019). A Novel Assay for Simultaneous Assessment of Mammalian Host Blood, Mosquito Species, and Plasmodium spp. in the Medically Important Anopheles Mosquitoes of Madagascar. The American journal of tropical medicine and hygiene 100, 544-551.
93. McNamara, D.T., Thomson, J.M., Kasehagen, L.J., and Zimmerman, P.A. (2004). Development of a multiplex PCR-ligase detection reaction assay for diagnosis of infection by the four parasite species causing malaria in humans. Journal of clinical microbiology 42, 2403-2410.
94. Yerlikaya, S., Campillo, A., and Gonzalez, I.J. (2018). A Systematic Review: Performance of Rapid Diagnostic Tests for the Detection of Plasmodium knowlesi, Plasmodium malariae, and Plasmodium ovale Monoinfections in Human Blood. J Infect Dis 218, 265-276.
95. Tanizaki, R., Kato, Y., Iwagami, M., Kutsuna, S., Ujiie, M., Takeshita, N., Hayakawa, K., Kanagawa, S., Kano, S., and Ohmagari, N. (2014). Performance of Rapid Diagnostic Tests for Plasmodium ovale Malaria in Japanese Travellers. Tropical medicine and health 42, 149-153.
96. Oriero, E.C., Amenga-Etego, L., Ishengoma, D.S., and Amambua-Ngwa, A. (2021). Plasmodium malariae, current knowledge and future research opportunities on a neglected malaria parasite species. Critical reviews in microbiology 47, 44-56.
97. Beshir, K.B., Grignard, L., Hajissa, K., Mohammed, A., Nurhussein, A.M., Ishengoma, D.S., Lubis, I.N.D., Drakeley, C.J., and Sutherland, C.J. (2020). Emergence of Undetectable Malaria Parasites: A Threat under the Radar amid the COVID-19 Pandemic? The American journal of tropical medicine and hygiene 103, 558-560.
98. Volkman, S.K., Sabeti, P.C., DeCaprio, D., Neafsey, D.E., Schaffner, S.F., Milner, D.A., Jr., Daily, J.P., Sarr, O., Ndiaye, D., Ndir, O., et al. (2007). A genome-wide map of diversity in Plasmodium falciparum. Nat Genet 39, 113-119.
99. Carlton, J.M., Adams, J.H., Silva, J.C., Bidwell, S.L., Lorenzi, H., Caler, E., Crabtree, J., Angiuoli, S.V., Merino, E.F., Amedeo, P., et al. (2008). Comparative genomics of the neglected human malaria parasite Plasmodium vivax. Nature 455, 757-763.
100. Färnert, A., Rooth, I., Svensson, Å., Snounou, G., and Björkman, A. (1999). Complexity of Plasmodium falciparum Infections Is Consistent over Time and Protects against Clinical Disease in Tanzanian Children. The Journal of Infectious Diseases 179, 989-995.
101. Mayengue, P.I., Kouhounina Batsimba, D., Niama, R.F., Ibara Ottia, R., Malonga-Massanga, A., Fila-Fila, G.P.U., Ahombo, G., Kobawila, S.C., and Parra, H.J. (2020). Variation of prevalence of malaria, parasite density and the multiplicity of Plasmodium falciparum infection throughout the year at three different health centers in Brazzaville, Republic of Congo. BMC Infectious Diseases 20, 190.

102. Escalante, A.A., and Pacheco, M.A. (2019). Malaria Molecular Epidemiology: An Evolutionary Genetics Perspective. 7, 7.4.25.
103. Tchekounou, C., Zida, A., Zongo, C., Soulama, I., Sawadogo, P.M., Guiguemde, K.T., Sangaré, I., Guiguemde, R.T., and Traore, Y. (2022). Antimalarial drugs resistance genes of Plasmodium falciparum: a review. Annals of parasitology 68, 215-225.
104. Shambhu, S., Koundal, D., Das, P., Hoang, V.T., Tran-Trung, K., and Turabieh, H. (2022). Computational Methods for Automated Analysis of Malaria Parasite Using Blood Smear Images: Recent Advances. Computational intelligence and neuroscience 2022, 3626726.
105. Nate, Z., Gill, A.A.S., Chauhan, R., and Karpoormath, R. (2022). Recent progress in electrochemical sensors for detection and quantification of malaria. Analytical biochemistry 643, 114592.
106. Jameela, T., Athotha, K., Singh, N., Gunjan, V.K., and Kahali, S. (2022). Deep Learning and Transfer Learning for Malaria Detection. Computational intelligence and neuroscience 2022, 2221728.
107. Ikerionwu, C., Ugwuishiwu, C., Okpala, I., James, I., Okoronkwo, M., Nnadi, C., Orji, U., Ebem, D., and Ike, A. (2022). Application of machine and deep learning algorithms in optical microscopic detection of Plasmodium: A malaria diagnostic tool for the future. Photodiagnosis and photodynamic therapy 40, 103198.
108. Meienberg, J., Bruggmann, R., Oexle, K., and Matyas, G. (2016). Clinical sequencing: is WGS the better WES? Human genetics 135, 359-362.
109. Kozarewa, I., Ning, Z., Quail, M.A., Sanders, M.J., Berriman, M., and Turner, D.J. (2009). Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. Nature methods 6, 291-295.
110. Taliun, D., Harris, D.N., Kessler, M.D., Carlson, J., Szpiech, Z.A., Torres, R., Taliun, S.A.G., Corvelo, A., Gogarten, S.M., Kang, H.M., et al. (2021). Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature 590, 290-299.
111. Sirugo, G., Williams, S.M., and Tishkoff, S.A. (2019). The Missing Diversity in Human Genetic Studies. Cell 177, 26-31.
112. Jooma, S., Hahn, M.J., Hindorff, L.A., and Bonham, V.L. (2019). Defining and Achieving Health Equity in Genomic Medicine. Ethnicity & disease 29, 173-178.
113. Wonkam, A. (2021). Sequence three million genomes across Africa. Nature 590, 209-211.
114. Gurdasani, D., Carstensen, T., Tekola-Ayele, F., Pagani, L., Tachmazidou, I., Hatzikotoulas, K., Karthikeyan, S., Iles, L., Pollard, M.O., Choudhury, A., et al. (2015). The African Genome Variation Project shapes medical genetics in Africa. Nature 517, 327-332.
115. Tessema, S.K., Raman, J., Duffy, C.W., Ishengoma, D.S., Amambua-Ngwa, A., and Greenhouse, B. (2019). Applying next-generation sequencing to track falciparum malaria in sub-Saharan Africa. Malar J 18, 268.
116. Inzaule, S.C., Tessema, S.K., Kebede, Y., Ogwell Ouma, A.E., and Nkengasong, J.N. (2021). Genomic-informed pathogen surveillance in Africa: opportunities and challenges. The Lancet Infectious diseases 21, e281-e289.
117. Stockdale, J.E., Liu, P., and Colijn, C. (2022). The potential of genomics for infectious disease forecasting. Nature Microbiology 7, 1736-1743.
118. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics 43, 491-498.
119. Vijayvargiya, P., Jeraldo, P.R., Thoendel, M.J., Greenwood-Quaintance, K.E., Esquer Garrigos, Z., Sohail, M.R., Chia, N., Pritt, B.S., and Patel, R. (2019). Application of

metagenomic shotgun sequencing to detect vector-borne pathogens in clinical blood samples. PLoS One 14, e0222915.
120. Ramesh, A., Nakielny, S., Hsu, J., Kyohere, M., Byaruhanga, O., de Bourcy, C., Egger, R., Dimitrov, B., Juan, Y.F., Sheu, J., et al. (2019). Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. PLoS One 14, e0218318.
121. McFadden, G.I. (2011). The apicoplast. Protoplasma 248, 641-650.
122. Crawford, N.G., Kelly, D.E., Hansen, M.E.B., Beltrame, M.H., Fan, S., Bowman, S.L., Jewett, E., Ranciaro, A., Thompson, S., Lo, Y., et al. (2017). Loci associated with skin pigmentation identified in African populations. 358, eaan8433.
123. Tishkoff, S.A., Gonder, M.K., Henn, B.M., Mortensen, H., Knight, A., Gignoux, C., Fernandopulle, N., Lema, G., Nyambo, T.B., Ramakrishnan, U., et al. (2007). History of Click-Speaking Populations of Africa Inferred from mtDNA and Y Chromosome Genetic Variation. Molecular Biology and Evolution 24, 2180-2195.
124. Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.-M., Doumbo, O., et al. (2009). The Genetic Structure and History of Africans and African Americans. 324, 1035-1044.
125. Tishkoff, S.A., Reed, F.A., Ranciaro, A., Voight, B.F., Babbitt, C.C., Silverman, J.S., Powell, K., Mortensen, H.M., Hirbo, J.B., Osman, M., et al. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. Nature Genetics 39, 31-40.
126. National Heart, L., and Blood Institute. (2020). The NHLBI BioData Catalyst. In, N.I.o. Health, ed. (Zenodo.
127. Kodaman, N., Aldrich, M.C., Sobota, R., Asselbergs, F.W., Poku, K.A., Brown, N.J., Moore, J.H., and Williams, S.M. (2016). Cardiovascular Disease Risk Factors in Ghana during the Rural-to-Urban Transition: A Cross-Sectional Study. PLoS One 11, e0162753.
128. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.
129. Boratyn, G.M., Thierry-Mieg, J., Thierry-Mieg, D., Busby, B., and Madden, T.L. (2019). Magic-BLAST, an accurate RNA-seq aligner for long and short reads. BMC bioinformatics 20, 405.
130. Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., and Ray, S.C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. Journal of virology 73, 152-160.
131. Sievers, F., and Higgins, D.G. (2021). The Clustal Omega Multiple Alignment Package. Methods in molecular biology (Clifton, NJ) 2231, 3-16.
132. Darling, A.C., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome research 14, 1394-1403.
133. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome research 19, 1639-1645.
134. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods 9, 357-359.
135. (2001). PlasmoDB: An integrative database of the Plasmodium falciparum genome. Tools for accessing and analyzing finished and unfinished sequence data. The Plasmodium Genome Database Collaborative. Nucleic acids research 29, 66-69.
136. Team, R.C. (2020). R: A language and environment for statistical computing. In. (Vienna, Austria, R Foundation for Statistical Computing.

137. Manske, M., Miotto, O., Campino, S., Auburn, S., Almagro-Garcia, J., Maslen, G., O'Brien, J., Djimde, A., Doumbo, O., Zongo, I., et al. (2012). Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. Nature 487, 375-379.
138. Brown, C.A., Pappoe-Ashong, P.J., Duah, N., Ghansah, A., Asmah, H., Afari, E., and Koram, K.A. (2021). High frequency of the Duffy-negative genotype and absence of Plasmodium vivax infections in Ghana. Malar J 20, 99.
139. Mueller, I., Zimmerman, P.A., and Reeder, J.C. (2007). Plasmodium malariae and Plasmodium ovale--the "bashful" malaria parasites. Trends in parasitology 23, 278-283.
140. Gunalan, K., Niangaly, A., Thera, M.A., Doumbo, O.K., and Miller, L.H. (2018). Plasmodium vivax Infections of Duffy-Negative Erythrocytes: Historically Undetected or a Recent Adaptation? Trends in parasitology 34, 420-429.
141. Lo, E., Russo, G., Pestana, K., Kepple, D., Abagero, B.R., Dongho, G.B.D., Gunalan, K., Miller, L.H., Hamid, M.M.A., Yewhalaw, D., et al. (2021). Contrasting epidemiology and genetic variation of Plasmodium vivax infecting Duffy-negative individuals across Africa. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases 108, 63-71.
142. Mensah, B.A., Aydemir, O., Myers-Hansen, J.L., Opoku, M., Hathaway, N.J., Marsh, P.W., Anto, F., Bailey, J., Abuaku, B., and Ghansah, A. (2020). Antimalarial Drug Resistance Profiling of Plasmodium falciparum Infections in Ghana Using Molecular Inversion Probes and Next-Generation Sequencing. 64, e01423-01419.
143. Apinjoh, T.O., Ntui, V.N., Chi, H.F., Moyeh, M.N., Toussi, C.T., Mayaba, J.M., Tangi, L.N., Kwi, P.N., Anchang-Kimbi, J.K., Dionne-Odom, J., et al. (2022). Intermittent preventive treatment with Sulphadoxine-Pyrimethamine (IPTp-SP) is associated with protection against sub-microscopic P. falciparum infection in pregnant women during the low transmission dry season in southwestern Cameroon: A Semi - longitudinal study. PLoS One 17, e0275370.
144. Duah, N.O., Quashie, N.B., Abuaku, B.K., Sebeny, P.J., Kronmann, K.C., and Koram, K.A. (2012). Surveillance of molecular markers of Plasmodium falciparum resistance to sulphadoxine-pyrimethamine 5 years after the change of malaria treatment policy in Ghana. The American journal of tropical medicine and hygiene 87, 996-1003.
145. Ndiaye, T., Sy, M., Gaye, A., and Ndiaye, D. (2019). Genetic polymorphism of Merozoite Surface Protein 1 (msp1) and 2 (msp2) genes and multiplicity of Plasmodium falciparum infection across various endemic areas in Senegal. African health sciences 19, 2446-2456.
146. Pinkevych, M., Petravic, J., Bereczky, S., Rooth, I., Färnert, A., and Davenport, M.P. (2015). Understanding the relationship between Plasmodium falciparum growth rate and multiplicity of infection. J Infect Dis 211, 1121-1127.
147. Kyabayinze, D.J., Karamagi, C., Kiggundu, M., Kamya, M.R., Wabwire-Mangen, F., Kironde, F., and Talisuna, A. (2008). Multiplicity of Plasmodium falciparum infection predicts antimalarial treatment outcome in Ugandan children. African health sciences 8, 200-205.
148. Singh, A., Bhandari, S., Das, A., and Bharti, P.K. (2021). Asymptomatic low-density Plasmodium falciparum infections: A challenge in malaria elimination in India. J Infect Public Health 14, 1600-1602.
149. Aguilar, J.B., and Gutierrez, J.B. (2020). An Epidemiological Model of Malaria Accounting for Asymptomatic Carriers. Bulletin of mathematical biology 82, 42.

150. Cheaveau, J., Mogollon, D.C., Mohon, M.A.N., Golassa, L., Yewhalaw, D., and Pillai, D.R. (2019). Asymptomatic malaria in the clinical and public health context. Expert Rev Anti Infect Ther 17, 997-1010.
151. Okyere, B., Owusu-Ofori, A., Ansong, D., Buxton, R., Benson, S., Osei-Akoto, A., Owiredu, E.W., Adjei, C., Xorse Amuzu, E., Marfo Boaheng, J., et al. (2020). Point prevalence of asymptomatic Plasmodium infection and the comparison of microscopy, rapid diagnostic test and nested PCR for the diagnosis of asymptomatic malaria among children under 5 years in Ghana. PLoS One 15, e0232874.
152. Høgh, B. (1996). Clinical and parasitological studies on immunity to Plasmodium falciparum malaria in children. Scandinavian journal of infectious diseases Supplementum 102, 1-53.
153. Ramirez, A.H., Sulieman, L., Schlueter, D.J., Halvorson, A., Qian, J., Ratsimbazafy, F., Loperena, R., Mayo, K., Basford, M., Deflaux, N., et al. (2022). The All of Us Research Program: Data quality, utility, and diversity. Patterns (New York, NY) 3, 100570.
154. Hay, S.I., and Snow, R.W. (2006). The malaria Atlas Project: developing global maps of malaria risk. PLoS Med 3, e473.
155. H, W. (2016). ggplot2: Elegant Graphics for Data Analysis.(Springer-Verlag New York).
156. Areas, G.A. (2012). GADM database of Global Administrative Areas, version 2.0. In. (
157. Programme, G.M. (2012). World malaria report 2012. In. (WHO Headquarters (HQ)), p 195.
158. Programme, G.M. (2016). World malaria report 2016. In, WHO, ed., p 186.
159. (!!! INVALID CITATION !!! {}).
160. Mehlotra, R.K., Mattera, G., Bockarie, M.J., Maguire, J.D., Baird, J.K., Sharma, Y.D., Alifrangis, M., Dorsey, G., Rosenthal, P.J., Fryauff, D.J., et al. (2008). Discordant patterns of genetic variation at two chloroquine resistance loci in worldwide populations of the malaria parasite Plasmodium falciparum. Antimicrobial agents and chemotherapy 52, 2212-2222.
161. DaRe, J.T., Mehlotra, R.K., Michon, P., Mueller, I., Reeder, J., Sharma, Y.D., Stoneking, M., and Zimmerman, P.A. (2007). Microsatellite polymorphism within pfcrt provides evidence of continuing evolution of chloroquine-resistant alleles in Papua New Guinea. Malar J 6, 34.
162. Lewontin, R.C. (1988). On measures of gametic disequilibrium. Genetics 120, 849-852.
163. Devlin, B., and Risch, N. (1995). A Comparison of Linkage Disequilibrium Measures for Fine-Scale Mapping. Genomics 29, 311-322.
164. Awine, T., Malm, K., Bart-Plange, C., and Silal, S.P. (2017). Towards malaria control and elimination in Ghana: challenges and decision making tools to guide planning. Global health action 10, 1381471.
165. Quin, J.E., Bujila, I., Chérif, M., Sanou, G.S., Qu, Y., Vafa Homann, M., Rolicka, A., Sirima, S.B., O'Connell, M.A., Lennartsson, A., et al. (2017). Major transcriptional changes observed in the Fulani, an ethnic group less susceptible to malaria. eLife 6.
166. Rojas-Pirela, M., Medina, L., Rojas, M.V., Liempi, A.I., Castillo, C., Pérez-Pérez, E., Guerrero-Muñoz, J., Araneda, S., and Kemmerling, U. (2021). Congenital Transmission of Apicomplexan Parasites: A Review. Frontiers in microbiology 12, 751648.
167. Hagenlocher, M., and Castro, M.C. (2015). Mapping malaria risk and vulnerability in the United Republic of Tanzania: a spatial explicit model. Population Health Metrics 13, 2.
168. Mulugeta, A., Assefa, A., Eshetie, A., Asmare, B., Birhanie, M., and Gelaw, Y. (2022). Six-year trend analysis of malaria prevalence at University of Gondar Specialized Referral Hospital, Northwest Ethiopia, from 2014 to 2019. Scientific Reports 12, 1411.

169. Ketema, T., Bacha, K., Getahun, K., Portillo, H.A.D., and Bassat, Q. (2021). Plasmodium vivax epidemiology in Ethiopia 2000-2020: A systematic review and meta-analysis. PLoS neglected tropical diseases 15, e0009781.
170. Motshoge, T., Ababio, G.K., Aleksenko, L., Read, J., Peloewetse, E., Loeto, M., Mosweunyane, T., Moakofhi, K., Ntebele, D.S., Chihanga, S., et al. (2016). Molecular evidence of high rates of asymptomatic P. vivax infection and very low P. falciparum malaria in Botswana. BMC Infectious Diseases 16, 520.
171. Motshoge, T., Haiyambo, D.H., Ayanful-Torgby, R., Aleksenko, L., Ntebela, D., Malleret, B., Rénia, L., Peloewetse, E., Paganotti, G.M., and Quaye, I.K. (2021). Recent Molecular Assessment of Plasmodium vivax and Plasmodium falciparum Asymptomatic Infections in Botswana. The American journal of tropical medicine and hygiene 104, 2159-2164.
172. National Malaria Control Programme, U.o.H.A.S., AGA Malaria Control Programme, World Health Organization and the INFORM Project. (2013). An Epidemiological Profile of Malaria and its Control in Ghana.
173. Asante, K.P., Zandoh, C., Dery, D.B., Brown, C., Adjei, G., Antwi-Dadzie, Y., Adjuik, M., Tchum, K., Dosoo, D., Amenga-Etego, S., et al. (2011). Malaria epidemiology in the Ahafo area of Ghana. Malaria Journal 10, 211.
174. Snow RW, O.J.M.I.J.D., Feachem RG, Makgoba MW, et al., editors. (2006). Disease and Mortality in Sub-Saharan Africa.(The International Bank for Reconstruction and Development / The World Bank).
175. World Health, O. (2008). World malaria report 2008. In. (Geneva, World Health Organization.
176. Ecker, A., Lehane, A.M., Clain, J., and Fidock, D.A. (2012). PfCRT and its role in antimalarial drug resistance. Trends in parasitology 28, 504-514.
177. Tuedom, A.G.B., Sarah-Matio, E.M., Moukoko, C.E.E., Feufack-Donfack, B.L., Maffo, C.N., Bayibeki, A.N., Awono-Ambene, H.P., Ayong, L., Berry, A., Abate, L., et al. (2021). Antimalarial drug resistance in the Central and Adamawa regions of Cameroon: Prevalence of mutations in P. falciparum crt, Pfmdr1, Pfdhfr and Pfdhps genes. PLoS One 16, e0256343.
178. Hemming-Schroeder, E., Umukoro, E., Lo, E., Fung, B., Tomás-Domingo, P., Zhou, G., Zhong, D., Dixit, A., Atieli, H., Githeko, A., et al. (2018). Impacts of Antimalarial Drugs on Plasmodium falciparum Drug Resistance Markers, Western Kenya, 2003-2015. The American journal of tropical medicine and hygiene 98, 692-699.
179. Frosch, A.E.P., Venkatesan, M., and Laufer, M.K. (2011). Patterns of chloroquine use and resistance in sub-Saharan Africa: a systematic review of household survey and molecular data. Malaria Journal 10, 116.
180. Owusu-Agyei, S., Asante, K.P., Adjuik, M., Adjei, G., Awini, E., Adams, M., Newton, S., Dosoo, D., Dery, D., Agyeman-Budu, A., et al. (2009). Epidemiology of malaria in the forest-savanna transitional zone of Ghana. Malar J 8, 220.
181. Apinjoh, T.O., Mugri, R.N., Miotto, O., Chi, H.F., Tata, R.B., Anchang-Kimbi, J.K., Fon, E.M., Tangoh, D.A., Nyingchu, R.V., Jacob, C., et al. (2017). Molecular markers for artemisinin and partner drug resistance in natural Plasmodium falciparum populations following increased insecticide treated net coverage along the slope of mount Cameroon: cross-sectional study. Infectious Diseases of Poverty 6, 136.
182. Asare, K.K., Boampong, J.N., Duah, N.O., Afoakwah, R., Sehgal, R., and Quashie, N.B. (2017). Synergism between Pfcrt and Pfmdr1 genes could account for the slow recovery of chloroquine sensitive Plasmodium falciparum strains in Ghana after chloroquine withdrawal. Journal of Infection and Public Health 10, 110-119.

183. Lumbala, C., Biéler, S., Kayembe, S., Makabuza, J., Ongarello, S., and Ndung'u, J.M. (2018). Prospective evaluation of a rapid diagnostic test for Trypanosoma brucei gambiense infection developed using recombinant antigens. PLoS neglected tropical diseases 12, e0006386.
184. Penn-Nicholson, A., Scriba, T.J., Hatherill, M., White, R.G., and Sumner, T. (2016). A novel blood test for tuberculosis prevention and treatment. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde 107, 4-5.
185. Chesnais, C.B., Vlaminck, J., Kunyu-Shako, B., Pion, S.D., Awaca-Uvon, N.P., Weil, G.J., Mumba, D., and Boussinesq, M. (2016). Measurement of Circulating Filarial Antigen Levels in Human Blood with a Point-of-Care Test Strip and a Portable Spectrodensitometer. The American journal of tropical medicine and hygiene 94, 1324-1329.
186. Yman, V., Wandell, G., Mutemi, D.D., Miglar, A., Asghar, M., Hammar, U., Karlsson, M., Lind, I., Nordfjell, C., Rooth, I., et al. (2019). Persistent transmission of Plasmodium malariae and Plasmodium ovale species in an area of declining Plasmodium falciparum transmission in eastern Tanzania. PLoS neglected tropical diseases 13, e0007414.
187. Cao, Y., Wang, W., Liu, Y., Cotter, C., Zhou, H., Zhu, G., Tang, J., Tang, F., Lu, F., Xu, S., et al. (2016). The increasing importance of Plasmodium ovale and Plasmodium malariae in a malaria elimination setting: an observational study of imported cases in Jiangsu Province, China, 2011-2014. Malar J 15, 459.
188. Hawadak, J., Dongang Nana, R.R., and Singh, V. (2021). Global trend of Plasmodium malariae and Plasmodium ovale spp. malaria infections in the last two decades (2000-2020): a systematic review and meta-analysis. Parasit Vectors 14, 297.
189. Ehrlich, H.Y., Jones, J., and Parikh, S. (2020). Molecular surveillance of antimalarial partner drug resistance in sub-Saharan Africa: a spatial-temporal evidence mapping study. The Lancet Microbe 1, e209-e217.
190. Ruybal-Pesántez, S., Tiedje, K.E., Tonkin-Hill, G., Rask, T.S., Kamya, M.R., Greenhouse, B., Dorsey, G., Duffy, M.F., and Day, K.P. (2017). Population genomics of virulence genes of Plasmodium falciparum in clinical isolates from Uganda. Scientific Reports 7, 11810.
191. Band, G., Leffler, E.M., Jallow, M., Sisay-Joof, F., Ndila, C.M., Macharia, A.W., Hubbart, C., Jeffreys, A.E., Rowlands, K., Nguyen, T., et al. (2022). Malaria protection due to sickle haemoglobin depends on parasite genotype. Nature 602, 106-111.
192. Buyon, L.E., Elsworth, B., and Duraisingh, M.T. (2021). The molecular basis of antimalarial drug resistance in Plasmodium vivax. International Journal for Parasitology: Drugs and Drug Resistance 16, 23-37.
193. Ferreira, M.U., Nobrega de Sousa, T., Rangel, G.W., Johansen, I.C., Corder, R.M., Ladeia-Andrade, S., and Gil, J.P. (2021). Monitoring Plasmodium vivax resistance to antimalarials: Persisting challenges and future directions. International Journal for Parasitology: Drugs and Drug Resistance 15, 9-24.
194. Mala, W., Wilairatana, P., Milanez, G.D.J., Masangkay, F.R., Kotepui, K.U., and Kotepui, M. (2022). Evidence of and deaths from malaria and severe pneumonia co-infections in malaria-endemic areas: a systematic review and meta-analysis. Scientific Reports 12, 17344.
195. Kamau, E., Yates, A., Maisiba, R., Singoei, V., Opot, B., Adeny, R., Arima, C.O., Otieno, V., Sumbi, C.S., Okoth, R.O., et al. (2021). Epidemiological and clinical implications of asymptomatic malaria and schistosomiasis co-infections in a rural community in western Kenya. BMC Infect Dis 21, 937.

196. Mandage, R., Kaur, C., Pramanik, A., Kumar, V., Kodan, P., Singh, A., Saha, S., Pandey, S., Wig, N., Pandey, R.M., et al. (2020). Association of Dengue Virus and Leptospira Co-Infections with Malaria Severity. Emerging infectious diseases 26, 1645-1653.
197. Potgieter, M., Bester, J., Kell, D.B., and Pretorius, E. (2015). The dormant blood microbiome in chronic, inflammatory diseases. FEMS microbiology reviews 39, 567-591.
198. Hidi, L., Kovács, G.I., Szabó, D., Makra, N., Pénzes, K., Juhász, J., Sótonyi, P., and Ostorházi, E. (2022). Human blood vessel microbiota in healthy adults based on common femoral arteries of brain-dead multi-organ donors. Frontiers in cellular and infection microbiology 12, 1056319.
199. Khan, I., Khan, I., Usman, M., Xiao Wei, Z., Ping, X., Khan, S., Khan, F., Jianye, Z., Zhiqiang, L., and Lizhe, A. (2022). Circulating microbiota and metabolites: Insights into cardiovascular diseases. J Clin Lab Anal 36, e24779.
200. Arbeeva, L., Azcarate-Peril, M.A., Cui, Y., Nelson, A.E., and Loeser, R.F. (2022). Association of plasma microbial composition with a leaky gut in obesity-related osteoarthritis: An exploratory study. Osteoarthritis and cartilage open 4, 100317.
201. Marquet, S. (2018). Overview of human genetic susceptibility to malaria: From parasitemia control to severe disease. Infection, Genetics and Evolution 66, 399-409.
202. Fugtagbi, G., Otu, P.S., Abdul-Rahman, M., Aidoo, E.K., Lo, A.C., Gyan, B.A., Afrane, Y.A., and Amoah, L.E. (2022). Association of TNF-Alpha, MBL2, NOS2, and G6PD with Malaria Outcomes in People in Southern Ghana. Genetics research 2022, 6686406.
203. Fitri, N., Na-Bangchang, K., Tjitra, E., Hutagalung, J., Sunarno, S., Dewi, R.M., Handayani, S., and Chaijaroenkul, W. (2023). Host susceptibility genes of asymptomatic malaria from South Central Timor, Eastern Indonesia. Parasitology Research 122, 61-75.