Prevalence of Artemisinin Resistance Associated Single Nucleotide Polymorphisms in Kenyan Isolates

Elizabeth M. Glenn
University of New Mexico - Main Campus

Follow this and additional works at: https://digitalrepository.unm.edu/biol_etds

Recommended Citation

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Biology ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.
Elizabeth Glenn
Candidate

Biology
Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Dr. Eric Toolson, Chair

Dr. Douglas Perkins

Dr. Dorothy Scholl
Prevalence of Artemisinin Resistance Associated Single Nucleotide Polymorphisms in Kenyan Isolates

BY

Elizabeth Glenn
B.S. Biology

THESIS
Submitted in Partial Fulfillment of the Requirements for the Degree of
Masters of Science
Biology

The University of New Mexico
Albuquerque, New Mexico

May, 2018
Malaria has had an unspeakable toll on human economy and wellbeing. Every year, there are approximately 212 million malaria infections caused by Plasmodium falciparum leading to 429,000 deaths, the vast majority of which occur in sub-Saharan Africa where the parasite is endemic. Decades of widespread use, and probable misuse, of antimalarial drugs has created extraordinary selective pressure on the parasite and made many formerly effective drugs ineffective. Currently, artemisinin and artemisinin derivatives are the most effective treatment available but resistance to dihydroartemisinin (DHA), the active ingredient, is emerging. Continued monitoring for genetic mutations linked with resistance is vital to global health. My project tested for the prevalence of artemisinin resistant SNPs using a simple and inexpensive TaqMan Assay™. Of the five SNPs examined, I found evidence of only one mutant, the Y630F SNP, within the pfkelch13 gene. Though this information is valuable, there were limitations to the study. Samples should be genotyped to identify other mutations and monoclonal vs. polyclonal infections. Samples collected at other time points, particularly prior to 2006, should also be run.
# TABLE OF CONTENTS

## CHAPTER 1 BACKGROUND
- Parasite characteristics .......................................................... 1
- Disease characteristics ............................................................ 3
- Methods of containment ......................................................... 5

## CHAPTER 2 INTRODUCTION
- Tracking spread of resistance using microsatellite markers .......... 7
- History of drug resistance .......................................................... 7
- Mechanism of resistance ........................................................... 9
- Current prevalence of DHA resistance .................................... 11

## CHAPTER 3 METHODS
- Protocol for concurrent extraction ......................................... 15
- TaqMan assay ........................................................................ 16

## CHAPTER 4 RESULTS........................................................................ 17

## CHAPTER 5 DISCUSSION
- Significance of results ............................................................. 20
- Limitations of study ............................................................... 21
- Future direction .................................................................... 21

## APPENDICES

### APPENDIX A  EXTRACTION PROTOCOL ................................. 23

### APPENDIX B  LIST OF PRIMER/PROBE PAIRS .................... 26

## REFERENCES .............................................................................. 27
Chapter 1

Background

There are few diseases more persistent, virulent and well adapted to infect humans than malaria. The tremendous impact that this parasitic disease has had on the human population is hard to exaggerate. The World Health Organization (WHO) estimates every year approximately 212,000,000 people are infected with malaria. Of these cases, there are approximately 429,000 deaths annually, the vast majority of which occur in sub-Saharan Africa where the parasite is endemic. It is a leading cause of death in children under the age of five [3]. Although there have been enormous improvements in diagnosis and treatment in the last fifty years, the disease remains a leading cause of morbidity and mortality in the developing world [9]. Malaria is a disease caused by a protozoan parasite in the genus *Plasmodium*. There are four species of *Plasmodium* known to cause malaria in Homo sapiens; *P. falciparum, P. vivax, P. malariae, P. ovale*. *P. falciparum* is responsible for the majority of infections globally [3]. *P. knowlesi*, a fifth species of malaria known to infect monkeys, is increasingly causing infection in humans in Southeast Asia (SEA) [1, 2]. Although length of life stages and morphology varies between species, all share a complex life cycle [4].

*Plasmodium* parasites are spread from host to host by mosquito vectors belonging to the genus *Anopheles* [3-5]. Parasites, in their sexual form known as gametocytes, are diploid (2n) and are taken up in a blood meal by a female *Anopheles* mosquito. During the sporogenic cycle, male and female gametocytes sexually reproduce in the mosquito midgut. The resulting oocysts eventually burst, releasing asexual sporozoites. When the mosquito takes another blood meal the sporozoites are released into the host’s blood stream through the salivary glands. The
sporozoites travel to the liver where they remain for approximately two weeks and mature into schizonts that then rupture releasing merozoites into the blood stream. This is known as the blood stage of the parasite. The merozoites invade erythrocytes and become trophozoites, or ring-stage parasites, so named for their ring-like appearance in blood smears. While inside red blood cells, the trophozoites digest the hemoglobin protein for the essential amino acids needed for growth and reproduction. The heme released by hemoglobin catabolism is toxic. The parasite packages this heme into a crystalline structure called hemozoin, also known as malaria pigment because of its distinct staining under light microscopy [3,5,6]. Trophozoites begin reproducing asexually, maturing into schizonts that rupture the parasitic and erythrocytic cell membranes, releasing more merozoites into the blood stream [10-12]. In a phenomenon not well understood [7], a small proportion of trophozoites will mature into sexually reproducing male and female gametocytes that are taken up in a blood meal and the cycle begins again (Fig. 1).

![Diagram of the malaria life cycle](image-url)
The brief periods in which parasites are circulating in the bloodstream as sporozoites and merozoites are the only times they are exposed directly to the immune system. By invading hepatocytes in the liver stage and erythrocytes in the blood stage, *Plasmodium* evades host immune defense.

The blood stage is the stage in which clinical characteristics of disease manifest. This stage involves destruction of red blood cells, clumping, or rosetting, of infected red blood cells, the release of host and parasitic cell components upon lysis and immune response of the host [13-17]. Symptoms of blood stage infection can range widely, as can disease outcome. Mild symptoms include symptoms such as fever, chills, nausea and joint or muscle ache. Severe malaria is defined as the presence of systemic dysfunction including coma, severe anemia, severe hypoglycemia, metabolic acidosis and organ failure in the presence of asexual parasitemia and no other identifiable cause [13, 21]. Severe disease can be attributed to two general mechanisms; factors that affect microcirculation and immune hyperactivity/dysfunction in response to parasite or host-derived products released during red blood cell lysis [15, 20]. Although several symptoms can be present, there are two classic manifestations of severe disease; cerebral malaria (CM) and severe malarial anemia (SMA).

Cerebral malaria is characterized by coma, evaluated with the Glasgow Coma Scale, in the presence of any asexual parasitemia. CM is not well understood but is influenced by three factors [22]. Rosetting is among the most important factors [15,22]. Rosetting is the clustering of infected red blood cells (iRBCs) to form small clumps. These iRBCs can also adhere to non-infected RBCs creating clumps of both infected and non-infected cells that can adhere to the vascular endothelium of the microvasculature, constricting or impeding blood flow [18]. This sequestration likely evolved to prevent iRBCs being removed from the blood stream, as rosettes
remain in the microvasculature and are not filtered out by the spleen. Clumps can travel to and lodge in the microvasculature of the brain, inhibiting blood supply to tissues [22]. Another important contributing factor to CM is acidosis. Metabolic acidosis is a process in which the pH of the blood begins to acidify because the kidneys are not properly filtering acid out of the blood and are not producing bicarbonate at a sufficient rate. Hyperactive immune response is the final component to CM. There is a notable increase in circulating pro-inflammatory cytokines, particularly TNF-α [23]. Increased levels of TNF-α cause increased expression of ICAM-1, an intercellular adhesion protein that is expressed on the surface endothelial cells and promotes the migration of lymphocytes out of the bloodstream into tissues. ICAM-1 is also a key protein in the adhesion of rosettes to the endothelium [23]. These factors, and likely many others including health of the host and characteristics of the parasite itself, contribute to generally poor outcome. Even if a patient survives, CM can still cause significant cognitive impairment.

Another form of severe malaria is severe malarial anemia (SMA). SMA is defined as a hemoglobin count less than 5g/dL ([Hb] < 5g/dL) in the presence of any parasitemia and is most common in children under 3 years of age living in endemic or high transmission regions [21, 27]. SMA is caused by a simultaneous destruction of erythrocytes and failure to replace them. Rosetting within the microvasculature of the body removes infected and uninfected RBCs from circulation. Other iRBCs are phagocytized by macrophages, monocytes and neutrophils and many more are destroyed by schizonts during the release of merozoites. The hemozoin acquired by phagocytosis of iRBCs, and free hemozoin released by iRBC lysis, stimulates an increase in the Type I inflammatory immune response in the host characterized by the release of IFNγ, IL-12 and TNF-α [26-28]. This response causes increased inflammation, activation of macrophages and peripheral blood monocytes and IL-12 promotes increased erythropoiesis. However, there is
evidence that circulating levels of IL-10 and TNF-α are greatly increased in children with SMA. Both IL-10 and TNF-α are potent IL-12 suppressors. There is strong positive correlation between the ratio of circulating IL-10 and TNF-α to IL-12 and severity of anemia in children [27]. In endemic areas with high rates of exposure, SMA overwhelmingly manifests in children under five years of age [26, 27]. However, in areas of low or unstable rates of exposure, severe disease is more evenly distributed among children and adults [21]. Although many host and parasite genetic factors have been identified, there are no reliable predictors of disease outcome. Therefore, transmission control is our best tool for preventing disease.

Vector control is a key component of public health. This includes seemingly simple measures such as the use of insecticide and bed nets. Compliance, however, has proven difficult in developing countries with limited funds, governmental corruption and isolated rural communities [30]. Another method is the development of a vaccine. Vaccines work by introducing antigens specific to an infectious agent into an immune naïve host to generate an adaptive immune response and the production of antibodies. If the host is again exposed to that infectious agent, the agent is recognized and bound by the circulating antibodies. The host is able to mount an adaptive immune response much more quickly, thereby eliminating the agent’s ability to infect the host and cause disease. The key to this system is the use of stable and reliable antigens that are expressed by the infectious agent. Plasmodium spp. have evolved methods to constantly change the character of the antigens exposed to the host immune system. A good example of this is Plasmodium falciparum Membrane Protein 1 (PfEMP1). This is a protein that the parasite manufactures and exports to the membrane of infected RBCs and is the primary factor mediating resetting [31]. Ostensibly, this would be a prime vaccine target.
However expression of $var$ genes, the genes coding for the PfEMP1 proteins, are situated near parasite chromosome telomeres and are subject to high degrees of mutation and rearrangement [33] and the parasite can change its expression patterns during different life stages [34]. This high variance in PfEMP1 expression has stymied vaccine development [32].

Arguably the most important factor in disease control and outcome is the availability of effective drug treatment. Emergence of drug resistance in $P. falciparum$ has made many formerly effective drugs ineffective [36]. Decades of widespread use, and misuse, of antimalarial drugs creates extraordinary selective pressure on the parasite to develop resistance. Today, artemisinin and artemisinin derivatives are the most effective treatment [3]. Fears of global emergence of artemisinin resistance are increasing as resistance associated mutations spread.
Chapter 2
Introduction

The advent of genomic sequencing has helped scientists trace the development and spread of drug resistance. Until the use of sequencing technology, it was thought that resistance in malaria typically developed endemic regions [36]. Sequencing microsatellite markers flanking genes correlated with drug resistance has changed this idea. Microsatellites are tandem repeats of one to ten nucleotides that repeat 5-50 times [37]. The microsatellites used in resistance tracking are in non-coding regions upstream of genes correlated with drug resistance [36]. Because they are non-coding segments, unique mutations can accumulate. By sequencing unique microsatellites within resistant parasites, scientists can identify different strains and track the progression of spread globally. This process has shed light on the spread of resistant parasites for many drugs including chloroquine, pyrimethamine/sulfadoxine, mefloquine and most recently artemisinin.

Modern drug treatment for malaria began in the early 1940’s. Chloroquine (CQ) was the best drug treatment used globally for infection. It acts as a schizonticide by interfering with the polymerization of heme, creating buildup of this toxic intermediate and causing parasite death [38]. Resistance to CQ was first recognized on the Thai/Cambodia border in the late 1950’s. By 1959, resistant parasites had spread throughout Thailand and by 1962 had spread through neighboring countries Cambodia, Vietnam and Malaysia. CQ resistance first arrived in eastern Africa in 1978. Over the following decade, resistant parasites spread steadily though the continent and became a significant health concern in western Africa by the early 1990’s [39]. Mutations within the \textit{P. falciparum chloroquine resistance transporter (pfcrt)} gene are largely responsible for CQ resistance [40]. Several non-synonymous single nucleotide polymorphisms
(SNPs) are directly correlated with CQ resistance, the most prominent being the K76T SNP [38, 40]. This mutation occurs at the 76th amino acid position in the polypeptide sequence and changes the amino acid from a lysine to a threonine (K→T). It is in the first transmembrane domain of the protein and is thought to either decrease transportation of CQ into, or increase transportation of CQ out of, the food vacuole. Microsatellite sequencing upstream of the \textit{pfcr}t gene has identified the origins and route of spread globally. The resistance genotype that originated on the Thai/Cambodia border in the 1950’s is named CVIET, naming the amino acid sequence 72-76 with the underlined portion indicating the mutant haplotype. Microsatellite sequencing reveals this is the predominant mutation in SE Asia and Africa indicating that resistance is able to spread globally. CQ resistance was so strong that it was replaced in the 1990’s.

\textbf{Fig. 2}

Map of the origin and migration of CQ resistant \textit{P. falciparum} strains built by identifying microsatellite haplotypes upstream of the \textit{pfcr}t gene.

Mita et al. 2009
Pyrimethamine was a drug first tested in SE Asia in the late 1950’s but was not used widely until the 1960’s. It was used as a monotherapy in prophylaxis and eradication efforts in endemic regions within SE Asia. Resistance was first reported on the Thai/Cambodia border in the 1960’s. In the early 1990’s pyrimethamine/sulfadoxine (SP) replaced CQ as the primary drug treatment for uncomplicated malaria [41]. However in the mid 1990’s resistance to SP became so prevalent that it was replaced by Mefloquine (MQ). Mefloquin was first used in Thailand in the late 1984 as a monotherapy [43]. Despite strict control of its use, mefloquine resistant became prevalent by the early 1990’s. SE Asia adopted a strategy of combination therapy and paired mefloquine with artesunate, an artemisinin derivative. By 2010 more than 80% of endemic countries adopted Artemisinin Combination Therapy (ACT) as primary treatment for uncomplicated malaria [43, 44]. Today, the World Health Organization recommends artemether-lumefantrine as the primary treatment for uncomplicated malaria [3]. Artemisinin and its derivatives are the best treatment for malaria currently available and are most effective at treating early ring-stage parasites [3]. The active ingredient is dihydroartemisinin (DHA). DHA is thought to inhibit a kinase vital in the pathway for hemoglobin transport [45]. Phosphatidylinositol-3-kinase (PfPI3K) is a parasite-derived enzyme exported through the parasite endoplasmic reticulum (ER) to the host erythrocytic cytoplasm and is also localized near the parasite cell membrane and the food vacuole [46]. There, the precursor molecule phosphoinositide (PI) is phosphorylated by PfPI3K into phosphatidyl-3-phosphate (PI3P) [46]. PI3P is an important lipid signaling molecule used for transportation. Of particular importance is the role PI3P plays in the targeting of vesicles containing hemoglobin from the
erythrocytic cytoplasm to the food vacuole where the hemoglobin is digested [47]. Recent experiments demonstrated that reduction in the amount of PI3P available in the parasite cell led to an increase in the amount of hemoglobin in the parasitic cell [46]. Upon further investigation,

Diagram of a parasitized red blood cell (iRBC). Vesicles containing hemoglobin are targeted to the parasite food vacuole (FV) for digestion. PI3P is an important signaling molecule used to target vesicles from the erythrocytic cytoplasm to the parasitic FV.

Fig. 3

Vaid et al. 2017

Demonstration of effect of wortmannin, a PI3K inhibitor. **B.** number of Hb filled vesicles in parasites treated with DMSO vs. wortmannin. **C.** Transmission electron micrograph of parasites treated with DMSO vs. wortmannin.

Vaid et al. 2017
it was found that the number of hemoglobin containing vesicles within the parasitic cell was greatly increased when PI3K was chemically inhibited with wortmannin. This indicates that PI3P is vital in targeting hemoglobin containing vesicles to the food vacuole [46, 48]. DHA is thought primarily to inhibit PI3K thereby decreasing levels of PI3P and ultimately starving the parasite of vital amino acids needed for growth and reproduction. DHA is the best drug treatment available. Yet there is evidence of spread of resistance that could cause a global increase in mortality.

Resistance to artemisinin was first reported in western Cambodia in 2006 [50]. Over the next several years the pfkelch13 gene on chromosome 13 was identified as consistently mutated in artemisinin resistant P. falciparum [51]. Pfkelch13 is a gene that codes for a propeller protein consisting of 6 blade-like beta sheath subunits connected by alpha helices. The exact function of this kelch13 protein is unknown, however the mammalian orthologue acts as a substrate adapter for E3 ubiquitin ligase [49]. Once attached to the substrate the E3 ubiquitin ligase recruits E2 ubiquitin conjugating enzyme that has been loaded with ubiquitin protein. The ubiquitin protein is transferred to the substrate thereby tagging the substrate for degradation [52]. Pfkelch13 protein is likely a substrate adapter between E3 and PI3K to assist in marking the enzyme for degradation. Studies have demonstrated a significantly increased concentration of PI3K enzyme in artemisinin resistant parasites [49]. Mutations within the pfkelch13 gene lead to structural differences in the protein significant enough to decrease its affinity for PI3K binding therefore less PI3K is ubiquitinated, increasing levels of the kinase, increasing levels of PI3P and increasing the number of hemoglobin filled vesicles that reach the food vacuole. This mechanism completely negates the effect of DHA on PI3K.
Fig. 5

Map of SNPs within PfK13. SNPs tended to be found in one of two groups or clusters. **A.** Cluster I is yellow, Cluster II in red. SNPs outside of a cluster are green. N and C terminals for the protein are indicated. **B.** Network of mutant SNPs.

Singh et al. 2016

Through extensive genotyping of samples taken from Cambodia, several single nucleotide polymorphisms (SNPs) were identified within *phkelch13*. However, there was a single SNP that was correlated with greatest DHA survivability [54, 55]. The mutation, C580Y, is a missense mutation that changes the amino acid codon from *tgt* to *tat*. This changes the resulting amino acid in the primary structure from a cysteine (polar side chain) to a tyrosine (neutral side chain), changing the tertiary structure of the Kelch 13 protein. Resistance to DHA was measured in both wildtype and mutant parasites using the Ring Survival Assay (RSA 0-3). Researchers found survival time was on average 7.19 hours compared to about 3.3 hours in wildtype samples [51]. Through genome sequencing they also found the C580Y SNP to be present in 85% of all mutant parasites isolated from 2011-2012 [51]. The spread of this single
mutation in such a short period shocked the research community and triggered a wave of intense investigation worldwide.

Significant energy has gone into identifying mutant SNPs worldwide since the identification of C580Y in Cambodia. Of particular concern is the spread of this and other SE Asian SNPs to sub-Saharan Africa, where up to 90% of malaria infections occur [57]. In 2013 and 2014, clinical samples were collected from 12 sub-Saharan countries and the parasitic DNA

Table 1

<table>
<thead>
<tr>
<th>Country (Region)</th>
<th>Samples, No.</th>
<th>Geen Position</th>
<th>Reference as (id)</th>
<th>Mutant as (id)</th>
<th>Samples With Mutant Allele, No.</th>
<th>Presence, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>470</td>
<td>Tvr (iso)</td>
<td>Tvr (iso)</td>
<td>1</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>486</td>
<td>Gly (iso)</td>
<td>Gly (iso)</td>
<td>1</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>587</td>
<td>Asp (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>607</td>
<td>Gly (iso)</td>
<td>Gly (iso)</td>
<td>1</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>555</td>
<td>Gly (iso)</td>
<td>Gly (iso)</td>
<td>1</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>DPR, Vietnam (n = 23)</td>
<td>450</td>
<td>Tyr (iso)</td>
<td>Tyr (iso)</td>
<td>1</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>510</td>
<td>Ala (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>DPR, Vietnam, Cambodia, West Africa (n = 22)</td>
<td>No SNPs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gabon (Libreville; n = 20)</td>
<td>471</td>
<td>Arg (iso)</td>
<td>Arg (iso)</td>
<td>1</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>578</td>
<td>Ala (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>586</td>
<td>Leu (iso)</td>
<td>Leu (iso)</td>
<td>1</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>Ghana (K prayer; n = 87)</td>
<td>468</td>
<td>Cys (iso)</td>
<td>Cys (iso)</td>
<td>2</td>
<td>2.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>505</td>
<td>Val (iso)</td>
<td>Val (iso)</td>
<td>3</td>
<td>3.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>510</td>
<td>Leu (iso)</td>
<td>Leu (iso)</td>
<td>1</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td>Ghana (Kawange; n = 90)</td>
<td>459</td>
<td>Ser (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>565</td>
<td>Cys (iso)</td>
<td>Cys (iso)</td>
<td>1</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>469</td>
<td>Cys (iso)</td>
<td>Cys (iso)</td>
<td>1</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>402</td>
<td>Tyr (iso)</td>
<td>Tyr (iso)</td>
<td>1</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>603</td>
<td>Thr (iso)</td>
<td>Thr (iso)</td>
<td>1</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>576</td>
<td>Ala (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>Kenya (Kinawa; n = 169)</td>
<td>476</td>
<td>Tvr (iso)</td>
<td>Tvr (iso)</td>
<td>1</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>603</td>
<td>Gly (iso)</td>
<td>Gly (iso)</td>
<td>1</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>505</td>
<td>Asp (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>576</td>
<td>Ala (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Madagascar (Berembo; n = 69)</td>
<td>450</td>
<td>Tyr (iso)</td>
<td>Tyr (iso)</td>
<td>1</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>603</td>
<td>Ser (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>505</td>
<td>Gly (iso)</td>
<td>Gly (iso)</td>
<td>1</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>576</td>
<td>Ala (iso)</td>
<td>Ser (iso)</td>
<td>1</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table with distribution and frequency of SNPs from 12 sub-Saharan countries.

Kemau et al. 2015

13
was genotyped [56]. Kenya had the highest number of missense mutations at 3 out of 5. Only three mutant SNPs were found in more than one country; the missense A578S mutation and two silent mutations. The A578S SNP was found in the highest frequency in Kenya at 2.7%. In Ghana, the V566I SNP was present at a prevalence of 3.2%. The mutants A569T and Y630F were unique to Kenya [56]. There was no prevalence of the C580Y mutant in any of the 1212 samples run.
Chapter 3

Methods

For this project I developed a method of concurrent extraction of both DNA and RNA from the same clinical sample. To accomplish this, I began by using TRIzol™ Reagent to lyse the red blood cell and the parasitic cell membranes. Following TRIzol™ Reagent protocol published by the manufacturer Invitrogen™, I then used phenol-chloroform phase separation column to separate DNA, RNA and protein phases. Finally, I attempted to wash the DNA (bottom or organic phase) with sodium citrate and precipitate with ethanol. This method was extremely problematic. The RNA (top or aqueous phase) was consistently a clear or pale pink liquid that produced good quality RNA. However, the organic phase was consistently extremely viscous and sticky, to the point where I could not pipette or transfer the sample. I observed the same stickiness in both parasitemic and aparasitemic samples indicating a component of the RBC was reacting with the TRIzol™ Reagent and forming the viscous substance that could not be purified. I retooled the protocol to include the use of saponin prior to the TRIzol™ Reagent treatment. Saponin is a detergent that forms pores in cell membranes. Treating the clinical sample with 1% saponin and centrifuging the sample at moderate speed pelleted some of the component reacting with the TRIzol™ Reagent. Using the supernatant, I then treated it with TRIzol™ Reagent and performed the phenol-chloroform separation. In subsequent samples, the organic layer was still somewhat viscous but I was able to precipitate the DNA with ethanol and then purify it using the QIAamp DNA Blood Mini Kit produced by Qiagen™. The RNA was precipitated with ethanol and purified using the RNA Prep Kit produced by Omega™. In such a way I was able to reliably extract moderate quality DNA and RNA from the same clinical sample (See Appendix A for protocol).
After I purified the DNA samples, I ran each sample in a genomificaiton reaction using the Genomifi Kit produced by GE Healthcare™ to increase the quantity of DNA. The genomification reaction works by adding the genomificaiton mix, DNA polymerase enzyme and random hexamer primers to the sample and running the reaction at 30 degrees Celsius for 90 minutes. Once that was complete I diluted the genomified DNA samples 1:10 in PCR grade water. The final step was using the TaqMan Assay™ produced by Applied Biosystems™. I designed primers to amplify specific areas of interest within \textit{pfkelch13} and probes for both wildtype and mutant SNPs. I chose five primer/probe sets targeting five SNPs based on those previously identified; A569T and Y630F unique to Kenya, V566I that is highly prevalent in Ghana, A578S found in 5 African countries and the critical C580Y SNP identified in Cambodia that to date has not been found in Africa (See Appendix B for primer/probe pairs). The probes used in the assay are approximately 20 nucleotides in length and are highly specific meaning that, although they only differ by one nucleotide, the wildtype probe will not bind to the mutant SNP and vice versa. The TaqMan™ enzyme binds the template strand/primer double stranded DNA and begins to elongate the strand. When a probe binds to the respective target, the TaqMan™ enzyme releases the fluorophore from its quencher during elongation allowing the fluorophore to fluoresce. The computer then reads this signal as mutant or wildtype. This qPCR method is highly specific and allows me to directly quantify the ratio of wildtype to mutant alleles in each clinical sample.
Chapter 4

Results

One hundred and eleven (n=111) non-SMA samples were used for this project. All samples were collected in Siaya, western Kenya. Siaya is a rural town approximately 46 miles northwest of Kisumu, the provincial capital. This is a holoendemic malaria region where the parasite circulates year around, though peak occurs during the rainy season. Artemisinin Combination Therapy was officially adopted by the Kenyan government as first-line treatment for uncomplicated malaria in 2006 [58]. Our Cohort 2 study began in 2009, three years after the formal adoption of ACTs, and ended in 2011. For this project, samples were used from each year of the Cohort 2 study and selected based on parameters including presence of parasitemia, hemoglobin > 5g/dL and no hemoglobinopathies or coinfections such as bacteremia or HIV-1. I isolated parasitic genomic DNA and RNA from each sample. I genomified the DNA and ran each sample in the TaqMan™ Assay with five SNPs located in the PfKelch13 gene.

Table 2

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Total Identified</th>
<th>Wildtype</th>
<th>Mutant</th>
<th>Heterozygous</th>
</tr>
</thead>
<tbody>
<tr>
<td>566 SNP</td>
<td>73</td>
<td>42</td>
<td>42</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>569 SNP</td>
<td>73</td>
<td>61</td>
<td>61</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>578 SNP</td>
<td>73</td>
<td>56</td>
<td>56</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>580 SNP</td>
<td>73</td>
<td>58</td>
<td>58</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>630 SNP</td>
<td>73</td>
<td>58</td>
<td>7</td>
<td>40</td>
<td>11</td>
</tr>
</tbody>
</table>
Table 3

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Total Identified</th>
<th>Wildtype</th>
<th>Mutant</th>
<th>Heterozygous</th>
</tr>
</thead>
<tbody>
<tr>
<td>566 SNP</td>
<td>26</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>569 SNP</td>
<td>26</td>
<td>23</td>
<td>23</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>578 SNP</td>
<td>26</td>
<td>14</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>580 SNP</td>
<td>26</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>630 SNP</td>
<td>26</td>
<td>13</td>
<td>2</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Total Identified</th>
<th>Wildtype</th>
<th>Mutant</th>
<th>Heterozygous</th>
</tr>
</thead>
<tbody>
<tr>
<td>566 SNP</td>
<td>12</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>569 SNP</td>
<td>12</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>578 SNP</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>580 SNP</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>630 SNP</td>
<td>12</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Of the five SNPs I tested, V566I, A569T, A578S and C580Y were wildtype for all samples for all three years. However, Y630F SNP showed interesting variance. In 2009, 68.9% of samples were mutant at the 630 position and 19.9% were heterozygous. In 2010 the trend was similar with 53.8% of samples mutant and 30.7% heterozygous. Of the five samples identified in 2011, 3 (60%) were mutant and 2 (40%) were heterozygous.
I compared the trends between the three years using a Chi-Square analysis. The mutant and heterozygous frequencies for Y630F SNP between the three years was not significant, with a p-value = 0.8368. I believe this is due to the small number of samples (n=5) from 2011 that were identified. Importantly, the C580Y SNP, so strongly correlated with ACT resistance in SE Asia, was 100% wildtype for all three years. These results are consistent with other studies that have not found evidence of SE Asian SNPs on the African continent [55, 56].
Chapter 5
Discussion

The vast majority of malaria infections occur in sub-Saharan Africa with high incidence in holoendemic regions including western Kenya. Because of high transmission rates, any mutation that confers an advantage in a parasite will quickly spread in the region, as we have seen with the C580Y SNP in Cambodia. To date, no one has examined clinical samples from Siaya for resistant mutations using the TaqMan™ Assay. Using this assay is advantageous because it is efficient and highly accurate when testing for a small number of SNPs in a large number of samples [59]. It is also relatively inexpensive and easy to use. This method could contribute to real-time monitoring of resistance in circulating parasites.

I found presence of only one mutant SNP, Y630F. In 2009, 68.9% of samples were mutant at the 630 position and 19.9% were heterozygous. All other SNPs were 100% wildtype across all three years. This is not consistent with previous findings. Kemau et al. found in 2014 the A578S mutant to have the highest prevalence in Kenya. It is worth noting that their samples were taken in Kisumu, an urban developed area whereas Siaya is a rural town. Mosquito populations and human demographics may be worth investigating. It is also important to note that Kemau et al. took samples from both children and adults, whereas all of my samples were taken from children less than 36 months. The majority of symptomatic infections occur in children under five years old. Drug pressure is higher in this age group than in adults as children are given drug therapy more frequently. This is a likely driver of higher proportion of mutants in my samples. Consist with recent research however [60], I found no evidence of the critical mutant C580Y SNP.
While my results are valuable, the project did have limitations. Wildtype/mutant SNP probes can only be designed based on SNPs already identified. Therefore, it can only be used to quantify known mutants; novel mutant SNPs will not be detected and may interfere with probe binding. SNP identification was inconsistent between SNPs and between samples; SNPs that were identified in one sample were not consistently identified in other samples. This could indicate the presence of novel SNPs within the probe sequence. Difference of a single nucleotide would prohibit probes from binding. The pfkelch13 gene from my samples should be genotyped to identify any novel SNPs that might be present.

Clonality of infection is impossible to determine with the TaqMan™ Assay. In holoendimic areas where the parasite is being transmitted year round, it is possible for a patient to be infected with more than one clonal strain of *P. falciparum* simultaneously. The presence of wildtype and mutant Y630F (heterozygous) in a single sample could indicate a polyclonal infection. The only way to identify a polyclonal infection is to sequence microsatellite markers. The Y630F heterozygous samples should be genotyped and microsatellite regions should be compared to establish clonality.

Sample size should also be expanded to include samples taken outside of the Cohort 2 timeframe (2009-2011). It would be especially useful to include samples taken prior to 2006, the year in which ACTs were officially adopted in Kenya. Another direction would be to test samples from patients who developed severe malarial anemia (SMA). This would examine if there is any correlation between the presence mutant SNPs and disease severity.

Finally, there is evidence that other gene mutations may affect resistance. The cysteine protease falcipain-2 is involved in hemoglobin hydrolysis. Inhibition of the protein or knockout of the gene decreases the efficacy of artemisinin. *In vitro* parasites selected for resistance
developed a nonsense mutation in the *fp2* gene. This implies that decreasing hemoglobin digestion decreases the deleterious effect of DHA [59]. In addition, there is emerging evidence of *pfkelch13* wildtype parasites exhibiting resistance to artemisinin. A study done in 2015 in Dakar, Senegal, followed nine patients who were parasitemic three days after being given intravenous artesunate followed by oral artemether-lumefantrine. Parasites collected from all nine patients were wildtype for the *pfkelch13* gene. This indicates K13 mutation is not the only predictor of resistance and there are other factors that have yet to be identified [60].

Effective drug therapy is the most important element of disease control especially in sub-Saharan Africa where 90% of malaria infections occurs. Tremendous increase in global travel has increased fears of rapid spread of drug resistance. Real-time monitoring of circulating parasites for known drug resistance mutations is vital to global health. Simple and relatively inexpensive methods, such as the use of the TaqMan™ Assay, are an important tool in the effort to monitor or even prevent the spread of resistance. Artemisinin combination therapies are the most effective treatment available. The spread of resistance could have devastating consequences and lead to the deaths of millions from this treatable disease.
Appendix A Extraction Protocol

RBC (P. falciparum) RNA/DNA extraction protocol with Saponin and Omega RNA Prep Kit or QIAamp blood mini kit

RBC (with Plasmodium falciparum) Pellet

1% Saponin Treatment

TRIzol/chloroform separation

PfRNA
Top (aqueous) phase

70% EtOH

Omega RNA Prep Kit

PfDNA
Bottom (organic) phase

100% EtOH

QIAamp Blood mini kit
Detailed procedure (suitable for pfRNA/pfDNA extraction from clinical RBC samples of malaria patients with pf or from cultures of P. f infected RBCs)

1. Take clinical RBC samples of malaria patients with pf out of -80°C freezer and thaw them at room temperature (RT).
   [If using cultured samples, collect the P. f infected RBCs (when parasitemia is >5%, and late trophozoites and early schizonts in predominate forms, high parasitemia is better) in 15 or 50mL falcon tubes → Spin down the parasitized RBCs at 3,000 RPM (RT) X 10 mins → Resuspend the resulting pellet in 10-20mL DPBS and repeat the centrifuge at 3,000 RPM (RT) X 5 mins→.Store the pellet at -20°C or processing on the same day (If not processing on the same day, thaw the sample at RT before proceeding to the next step)]

2. Add a 2:1 volume (RBC pellet:1% saponin) of 1% saponin (aliquots stored at -20°C works even better than freshly prepared 1% saponin when taken out for use. 300 mg saponin was dissolved into 30 mL of dd-H₂O to get 1% saponin solution) to RBC pellet in a tube and resuspend the pellet using a pipette. Incubate at RT for 10 mins.

3. Spin at 5000 RPM for 10 mins at 4°C.

4. Transfer supernatant to a new Eppendorf tube using pipet and add 1.5 volumes of TRIZol reagent, mix well, and then add 1/5 volume (to volume of TRIZol) of chloroform and mix well, (when there is clumps, let the tube sit at RT 10-30 min, and mix by inverting several times)

5. Resuspend pellet in the original tube in 500 µL of 1X DPBS and then spin at 6000 RPM for 3 min at 4°C→Discard the new supernatant and repeat wash one more time → Spin at 6000 RPM for 3 min at 4°C, and then resuspend the pellet in200 µL of 1X DPBS → add 1.5 volumes of TRIZol reagent, mix well, and then add 1/5 volume (to volume of TRIZol) of chloroform and mix well, (when there is clumps, let the tube sit at RT 10-30 min, and mix by inverting several times)

6. (Continuation from step 4 or step 5) Spin at 13,000 RPM for 15 minutes at 4 °C to separate layers. RNA is in the top (aqueous) phase, DNA is in the bottom (organic) phase.

RNA Isolation (continuation from step 6)

7. Transfer top (aqueous) layer containing RNA (without touching inter-phase or bottom phase) from the same RBC sample to a clean Eppendorf tube for immediate RNA isolation.

8. To this tube add equal volume of 70% ethanol and mix well.

9. Transfer contents of Eppendorf tube to Omega RNA prep column (700 µL at a time).

10. Centrifuge at 10,000 RPM (RT) for 1 minute. Discard flow through and repeat step 9 if necessary.

11. Add 500 µL of RNA Wash Buffer I. Centrifuge at 10,000 RPM (RT) for 30 seconds, discard flow through.

12. Add 500 µL RNA Wash Buffer II (supplemented with 100% ethanol), and centrifuge at 10,000 RPM for 1 minute, discard flow through.

13. Repeat step 12 once.

14. Centrifuge empty column at 13,000 RPM (RT) with the original collection tube for 2 minutes.

15. Transfer column to a clean microfuge tube and spin at 13,000 RPM (RT) for 2 minutes to assure no residue RNA Wash Buffer II left in the column (e.g., no fluid in clean microfuge tube. Otherwise, repeat this step and discard the tube with fluid).
16. Add 70 µL of DEPC water to the center of the column. Wait 1 minute at RT, and then centrifuge at 13,000 RPM (RT) for 2 minutes.
17. Discard column.
18. Measure RNA concentration using nanodrop, and then transfer 20 µL to a new tube and stored at -20 °C (for long-term storage, put the tube to-80 °C) for later use. To the remaining RNA product, add 80 µL of nuclease free water, 15 µL of 10X DNase I buffer, and 5 µL of DNase I (Note that both 10X DNase I buffer and DNase I enzyme are from NEB) and mix well, and then incubate at RT for 1 hour.
19. Then add 1.5 volume of TRizol to the tube and mix thoroughly.
20. Add 1/5 volume (to volume of TRizol) of chloroform and mix well, and then spin at 13,000 RPM X 15 min at 4° C.
21. Transfer top (aqueous) layer to a clean Eppendorf tube and add equal volume of 70% ethanol (discard the bottom organic phase).
22. Repeat steps 9-18 with new Omega RNA Prep column (only use 40 µL of DEPC water for RNA elution this time).
23. Quantify RNA using nanodrop, and store RNA samples at -80 °C or -20 °C for later use.

**DNA Isolation** (continuation from step 6)

24. Take the bottom (organic) phase from the same RBC sample by inserting a tip to the very bottom of the tube and aspirate the organic phase and transfer to a new Eppendorf tube (leave the aqueous phase and interphase inorganic phase not touched).
25. Add to the new Eppendorf tube 0.3 mL of 100% EtOH per 1 mL TRizol used for lysis of parasite in earlier step.
26. Cap the tube and mix by inverting several times.
27. Incubate 3 minutes at RT.
28. Carefully apply the mixture from step 27 to the QIAamp Mini spin column (in 2 mL collection tube) without wetting the rim. Close the cap and centrifuge at 8000 RPM (6000 x g) at RT for 1 min and discard filtrate (in case the mixture not fully pass through the QIAamp Mini spin column, spin the column for additional 1 min or up to 5 mins at RT at 13200 RPM, so that no mixture left in the column).
29. Carefully open the column and add 500 µL of Buffer AW1 without wetting the rim. Close the cap and centrifuge at 8000 RPM for 1 min. Discard filtrate.
30. Carefully open the column and add 500 µL of Buffer AW2 without wetting the rim. Close the cap and centrifuge at full speed (13,000 RPM) for 3 min.
31. Discard the filtrate and spin column again at 13,000 RPM for 1 min.
32. Transfer column to a clean microfuge tube and spin at 13,000 RPM (RT) for 1 min to assure no residue Buffer AW2 left in the column (e.g., no fluid in clean microfuge tube. Otherwise, repeat this step and discard the tube with fluid).
33. Add 40 µL of Buffer AE DEPC water to the center of the column.
34. Incubate at room temperature for 5 mins, then centrifuge at 13,000 RPM for 1 min to collect the elute. Determine DNA concentration using nanodrop, and store the elute at -20°C or -80°C freezer.
Appendix B Primer/Probe Pairs

**Plasmodium falciparum K13 Propeller gene (PF3D7_1343700)**

```
....GATTATAAGGCTTTATTTGAAAAGTTGCTGATATATGTGTTATGTTGTTATGTT
TCAAGTAATTTAAATACCATGGAAGAATATTTTGTTGTTACGTCCAAATAGTGAATT
TATTGTATGGGGGATA[ATGAGCTCTCTTATTATACCGAA]GTTGAAGCATAATGATCATCAG
CTATGAG[ACCCATGGGAAAAGGTTGCTCCTTGAAT][CCC]CTAG[ATCATCA]GATGCTGTT
TGATGGCTCTTCTATTATACCGAA
TGATGGCTCTTCTATTATACCGAA
TGATGGCTCTTCTATTATACCGAA
TGATGGCTCTTCTATTATACCGAA
TGATGGCTCTTCTATTATACCGAA
TGATGGCTCTTCTATTATACCGAA
TGATGGCTCTTCTATTATACCGAA
TGATGGCTCTTCTATTATACCGAA
```

**Codon 566:** GTA to ATA (Val to Ile)
**Codon 569:** GCA to ACA (Ala to Thr)
**Codon 578:** GCT to TCT (Ala to Ser)
**Codon 580:** TGT to TAT (Cys to Tyr)
**Codon 630:** TAC to TTC (Tyr to Phe)

<table>
<thead>
<tr>
<th>K13 Forward 566/580 (172bp)</th>
<th>TGATGGCTCTCTTATTATACCGAA</th>
</tr>
</thead>
<tbody>
<tr>
<td>K13 reverse 566/580</td>
<td>VIC-CTAGATCATCAGCTATGT-MGB</td>
</tr>
<tr>
<td>K13_566 Probe Mutant</td>
<td>VIC- AAGCATGGATAAGGCTTTT- MGB</td>
</tr>
<tr>
<td>K13_566 Probe Wild</td>
<td>FAM- AAGCATGGATAAGGCTTTT- MGB</td>
</tr>
<tr>
<td>K13_569 Probe Mutant</td>
<td>VIC- TAGAGGTGACCTTTGAATA- MGB</td>
</tr>
<tr>
<td>K13_569 Probe Wild</td>
<td>FAM- TAGAGGTGACCTTTGAATA- MGB</td>
</tr>
<tr>
<td>K13_578 Probe Mutant</td>
<td>VIC-CTAGATCATCAGCTATGT-MGB</td>
</tr>
<tr>
<td>K13_578 Probe Wild</td>
<td>FAM-CTAGATCATCAGCTATGT-MGB</td>
</tr>
<tr>
<td>K13_580 Probe Mutant</td>
<td>VIC-CTAGATCATCAGCTATGT-MGB</td>
</tr>
<tr>
<td>K13_580 Probe Wild</td>
<td>FAM- CTAGATCATCAGCTATGT-MGB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>K13 Forward 630 (121 bp)</th>
<th>GGGAAACAATTICCATATGCTCTTATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>K13 reverse 630</td>
<td>CATATTGICCAAGGAAATCAAATATGTG</td>
</tr>
<tr>
<td>K13_630 Probe Mutant</td>
<td>VIC- AGCAGCTTTTAATCTCTTA-MGB</td>
</tr>
<tr>
<td>K13_630 Probe Wild</td>
<td>FAM- AGCAGCTTTTAATCTCTTA-MGB</td>
</tr>
</tbody>
</table>
Reference

8. CDC website