POLYMORPHISMS IN THE IL-12 AND IL-12R GENE1S: ALTERING *PLASMODIUM FALCIPARUM* DISEASE OUTCOME

by

LaToya Michelle Strong

B.S., Biology Wofford College, 2006

Submitted to the Graduate Faculty of

The Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

LaToya Michelle Strong

It was defended on

July 27, 2009

and approved by

Amy L. Hartman, Ph.D
Research Instructor
Infectious Diseases and Microbiology
Graduate School of Public Health
University of Pittsburgh

Eleanor Feingold, Ph.D
Associate Professor
Human Genetics
Graduate School of Public Health
University of Pittsburgh

Thesis Advisor:
Jeremy Martinson, DPhil
Assistant Professor
Infectious Diseases and Microbiology
Graduate School of Public Health
University of Pittsburgh

POLYMORPHISMS IN THE IL-12 AND IL-12R GENES: ALTERING *PLASMODIUM FALCIPARUM* DISEASE OUTCOME

LaToya Michelle Strong, M.S.

University of Pittsburgh, 2009

Malaria is a major public health concern as greater than 40% of the world's population is at risk. Globally and annually, there are approximately 300-500 million incident cases a year resulting in between 1 and 2 million deaths; the majority of these deaths occur in children under the age of 5 and in pregnant women. There are several disease complications that can arise from malaria, two of which are high density parasitemia (HDP) and severe malaria anemia (SMA). Not everyone who gets malaria gets HDP or SMA, and the underlying reason for this is unknown, however, research has shown that innate immune mediators, including Interleukin-12 (IL-12), play an important role. Currently there is no vaccine for malaria and drug resistance is a major issue. This study is of public health significance because it can give insight into the difference between those individuals who progress to severe disease complications and those that do not; potentially giving rise to novel drug and vaccine development. The severity and occurrences of these complications vary by age, region, and level of malaria endemicity. Previous studies have indicated a role for not only circulating levels of IL-12, but also for polymorphisms in the IL-12 and Interleukin-12 Receptor (IL-12R) genes. To gain a better understanding of the role that polymorphisms in the IL-12 and IL-12R genes may have we conducted a case-control study to compare phenotypic data to genotypic data. We investigated four Single Nucleotide Polymorphisms (SNPs) - rs2243113, rs2243140, rs383483 and rs429774 - in the IL-12 and IL-12R genes to determine if a correlation existed between disease status and

genotype. We found that for all four SNPs, there was not a correlation between disease status and genotype. We also investigated the distribution of these four SNPs across populations with varying malaria endemicity. We also found that in all of the populations except for the Kenyan population there is a higher frequency of homozygous wildtype alleles for both rs2243113 and rs2243140 and a higher frequency of heterozygotes for rs383483.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

“Piglet sidled up to Pooh from behind. “Pooh,” he whispered.

“Yes, Piglet?”

“Nothing,” said Piglet, taking Pooh’s paw, “I just wanted to be sure of you”

-Winnie the Pooh

"Here's to the crazy ones, the misfits, the rebels, the troublemakers, the round pegs in a square hole, the ones who see things differently. They're not fond of rules, and they have no respect for the status quo. You can quote them, disagree with them, glorify or vilify them. About the only thing you can't do is ignore them, because they change things. They push the human race forward, and while some may see them as the crazy ones, we see genius, because the people who are crazy enough to think they can change the world, and are the ones who'll do it."

- Apple Computer TV Ad

"Be daring, be different, be impractical, be anything that will assert integrity of purpose and imaginative vision against the play-it-safers, the creatures of the commonplace, the slaves of the ordinary."

- Sir Cecil Beaton

1.0 INTRODUCTION

1.1 DESCRIPTION OF THE PROBLEM

Globally, there are 300 – 500 million incident cases of malaria each year resulting in approximately 1 million deaths. Most of these deaths occur in children under the age of 5 and pregnant women [1]. Malaria is a disease of impoverished countries as well as tropical countries, particularly sub-Saharan Africa [2, 3] There are several complications that can arise following a malaria infection. These complications include, but are not limited to, cerebral malaria (CM), severe malaria anemia (SMA) and high density parasitemia (HDP) [4]. The clinical manifestations of these complications vary by age, region, and level of endemicity [5].

Currently there is no vaccine for malaria, and drug resistance to anti-malarials has become a global problem [6]. Drugs that were at the forefront of malaria treatment, such as Chloroquine and, increasingly, sulphadoxine-pyrimethamine (SP), have become ineffective as single drug treatments because of resistance. The World Health Organization (WHO) now recommends artemisinin-based combination therapy (ACT) as a first line of treatment for malaria. Although it is very efficacious, it is too costly for some countries. [7].

Many studies have illustrated the importance of cytokines such as interleukin (IL)-12, and interferon-gamma (IFN-γ) in the immune response to malaria infection. These cytokines are initially produced by innate immune cells such as dendritic cells (DCs) and macrophages. Studies looking at these cytokines have shown that individuals with higher levels of IL-12, IFN-γ, and other innate immune mediators have a strong correlation with lower rates of severe disease complications [3, 8]. In addition to looking at plasma levels or circulating levels of cytokines, previous studies have demonstrated that host genetics plays a major role in the development of malaria disease outcomes. It has been shown that polymorphisms in cytokine genes can be determinants of malaria disease outcome [9, 10].

This study also addresses the role that polymorphisms in immune mediators may have in malaria disease outcome. This study like the ones before it addresses an important issue; what is the difference between individuals who develop complications and those who are able to fight off the disease. As more understanding is gained about the pathogenesis of malaria and the role that host genetics plays, more options are gained for drug development.

1.2 MALARIA

Malaria is caused by single celled protozoan parasites of the Plasmodium genus [11]. There are four Plasmodium species, *malariae, ovale, vivax* and *falciparum. Plasmodium ovale* is found in Western Africa and is the least severe. Similar to *P. ovale, P. malariae* has low frequency of infection; however, it is distributed worldwide. *P. vivax* is the most widespread species, but it is rarely fatal. *P. falciparum* can be attributed to most of the mortality associated with malaria [12].

This thesis will focus on *P. falciparum.* Malaria is a vector-borne disease transmitted by the bite of a female anopheline mosquito. There are three stages to the parasite life cycle: the mosquito stage, the human liver stage, and the human blood stage [12]. The bite of an infected female mosquito injects sporozoites into the host. The sporozoites are then quickly transported through the blood to the liver beginning the liver stage of the life cycle. In the liver the sporozoites mature and undergo asexual replication, known as exo-erythrocytic schizogony, forming schizonts. [11] The schizonts eventually burst releasing thousands of merozoites, which immediately invade host red blood cells [13]. This marks the beginning of the human blood stage. This stage is extremely important because it is where the clinical manifestations of the disease appear. The merozoites mature and undergo asexual reproduction to form erythrocytic schizonts, each containing daughter merozoites. These merozoites can do one of two things. They can re-infect red blood cells forming a cyclical blood stage infection or differentiate into male and female gametocytes which are taken up by a mosquito. [11, 13] Once the gametocytes are taken up the beginning of the third stage of the life cycle, the mosquito stage begins. In the mosquito the gametocytes give rise to gametes which subsequently develop into sporozoites [13]. The three stages of the malaria life cycle can be seen in Figure 1.

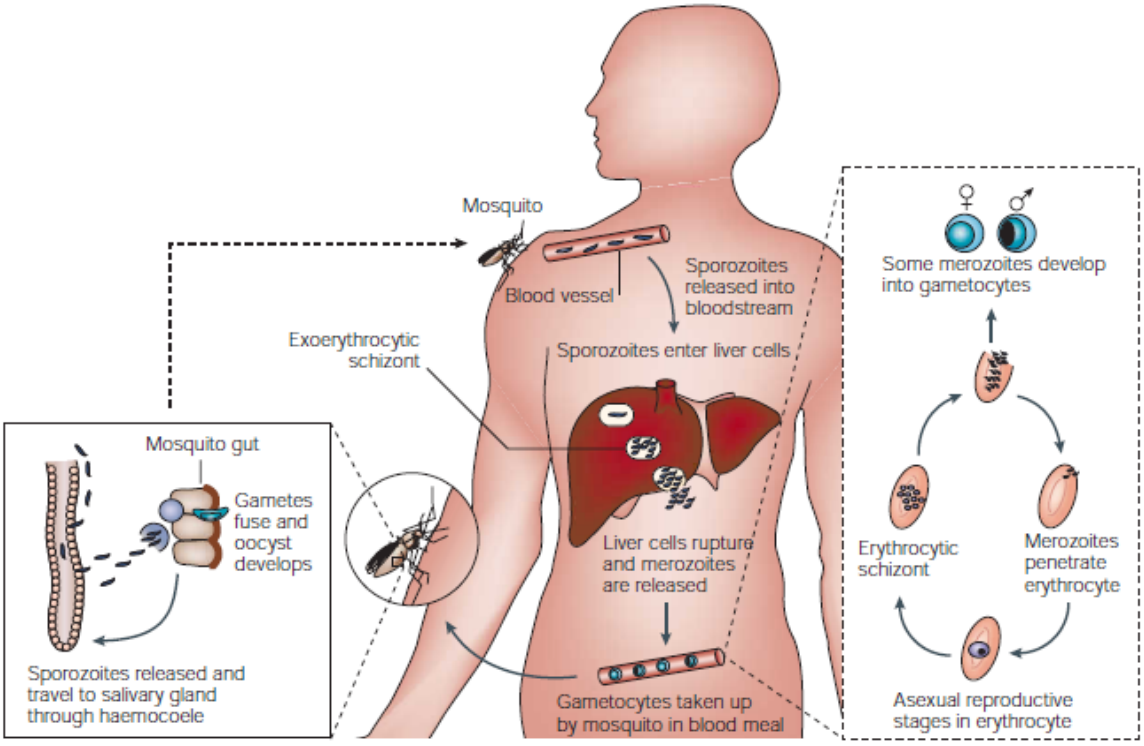

Figure 1. Malaria Life Cycle [11].

Copyright Permissions Granted

A characteristic of infection that occurs in *P. falciparum* malaria, but not in the other Plasmodium species is sequestration. Sequestration is the accumulation of large numbers of parasitized erythrocytes in various organs [14]. Infected erythrocytes (IE) containing young forms of the parasite are able to circulate freely, but on the contrary, IEs infected with more mature forms of the parasite are bound to endothelial cells and are subsequently sequestered away from peripheral circulation [13].

1.2.1 MALARIA DISEASE COMPLICATIONS

A malaria infection can lead to many disease complications, two of which are High density parasitemia (HDP) and Severe malaria anemia (SMA) [3]. Parasitemia is measured by the number of parasites (P) per liter of blood (P/L) and malaria anemia is expressed as grams of hemoglobin per dL of blood (g/dL). A table representing the hemoglobin and parasitemia levels for each respective disease outcome that was used in this study can be seen in Table 1 and Table 2, respectively. SMA and HDP are important because these complications are seen most often in children under the age of 5 and pregnant women. These groups also have the highest malaria mortality rate [15].

Although the underlying cause of these outcomes are not currently known, research has shown that the production of innate inflammatory mediators plays an important role [11], [16]. Intereukin-12 (IL-12), a major pro-inflammatory cytokine, is important in the immune response to malaria is IL-12 exerts immunoregulatory effects on T-cells and natural killer (NK) cells as well as inducing a Type-1 T Helper (T_H1) response [9].

Table 1. Classification for Malaria Anemia Clinical Disease Status

Classifications of Malaria Anemia Status	
Hemoglobin Level (g/dL of blood)	**Clinical Disease Status**
≥ 11	Non-Anemic (NA)
8.0 – 10.9	Mild Malaria Anemia (milMA)
6.0 – 7.9	Moderate Malaria Anemia (modMA)

Table 1: Continued

Hemoglobin Level (g/dL of blood)	Clinical Disease Status
<6	Severe Malaria Anemia (SMA)

Table 2. Classification of Parasitemia Clinical Disease Status

Classifications of Parasitemia Status	
Parasitemia Level (P/L of blood)	**Clinical Disease Status**
0	Aparasitemic Controls (APs)
< 10,000	Low Density Parasitemia (LDP)
≥ 10,000	High Density Parasitemia (HDP)

1.3 INTERLEUKIN-12

Cytokines are small proteins released by immune cells upon activation. Upon release, cytokines induce responses when bound to their receptors. Cytokines can act in an autocrine, paracrine or endocrine manner, meaning they can affect the behavior of the cell that released it, an adjacent cell, or a distant cell, respectively [17]. Interleukin-12 or IL-12p70 (IL-12) is a 70 kDa, heterodimeric, proinflammatory, cytokine produced by phagocytic cells. IL-12 is made up of a 35 kDa light chain and a 40 kDa heavy chain termed p35 and p40, respectively. The p35 subunit is also referred to as IL-12A or IL-12α and the p40 subunit is also referred to as IL-12B or IL-12β. The p35 and p40 subunits are linked by two disulfide bonds [18]. The genes encoding the

p35 and p40 subunit are located on different chromosomes and protein expression is therefore regulated independently. The p35 subunit is found on chromosome 5q31-33 and p40 is found on 3p12-3q13.2. [19] Both the p35 and p40 subunit must be expressed coordinately in the same cell to form the bioactive IL-12p70 [18]. IL-12α or p35 is expressed ubiquitously and constitutively at low levels and the p40 subunit is limited to expression in phagocytic cells. IL-12α serves as the rate limiting factor in the expression of IL-12 because IL-12βis always made in excess of IL-12α. IL-12α is regulated both transcriptionally and translationally, whereas IL-12β is only regulated transcriptionally [20].

IL-12 is produced by antigen presenting cells (APCs) and phagocytic cells, which include monocytes/macrophages, dendritic cells (DCs), neutrophils, and to a lesser extent, B-cells [20] [18] [21]. The major IL-12 responders are natural killer (NK) cells and activated T-cells. The effects that IL-12 has on both these cells are to induce proliferation and IFN-γ production, increase cytotoxicity activity, and polarize undifferentiated T cells to a T_H1 response [20]. A picture depicting the sources, responders and functions of IL-12 can be seen in Figure 2. In addition to the immunoregulatory effects that IL-12 has on T cells and NK cells, IL-12 has several other effects on the immune response. Not only does IL-12 induce the production of IFN-γ by T-cells and NK-cells, IL-12 itself also produces IFN-γ. A positive feedback loop exists between IL-12 and IFN-γ in which IL-12 secretes IFN-γ, which in turn primes monocytes and peripheral mononuclear blood cells (PMNCs), which in turn produces more IL-12. [20] A schematic of the overall biology of IL-12 can be seen in Figure 3.

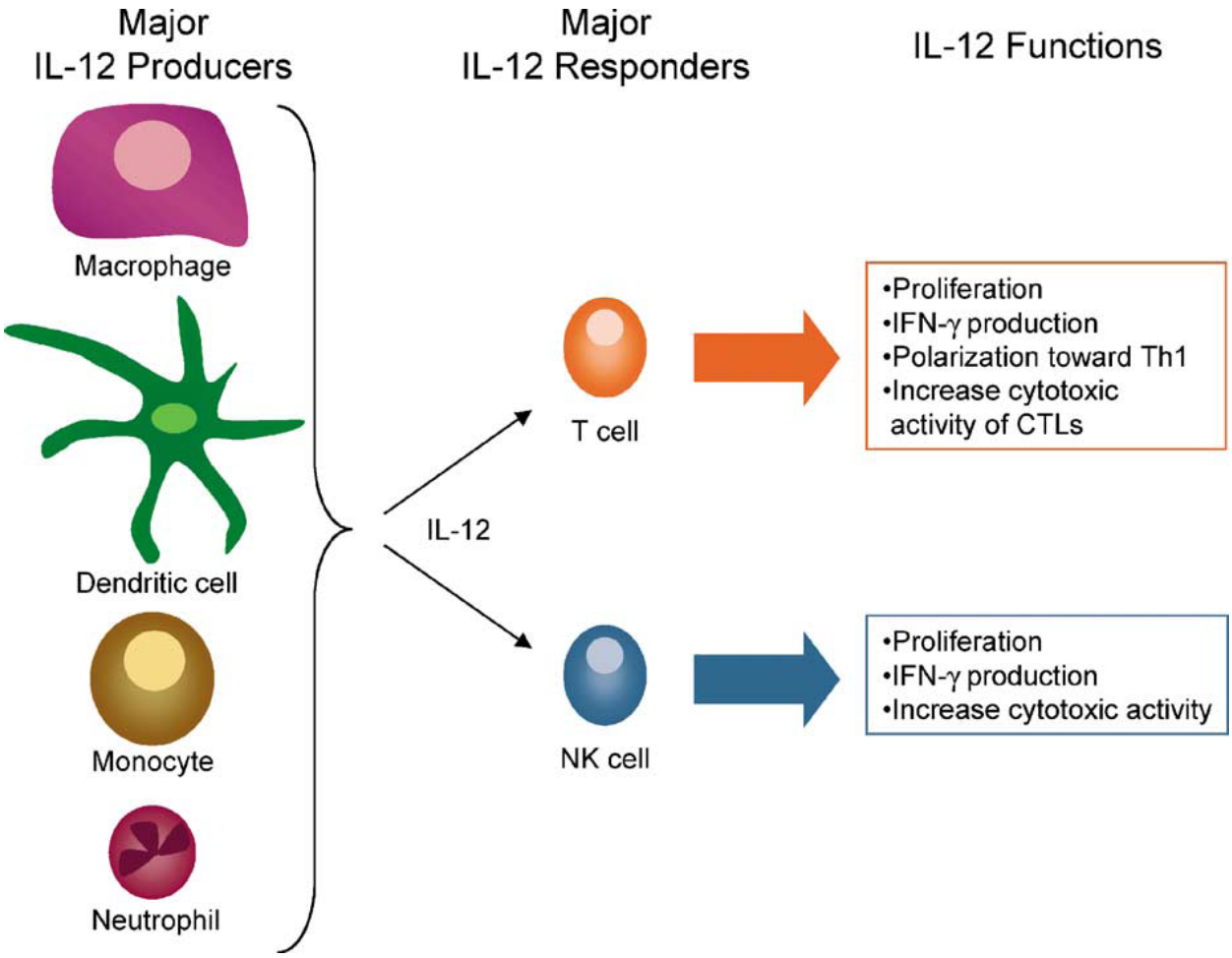

Figure 2. IL-12 producers, responders and functions [20].

Copyrights Permissions Granted

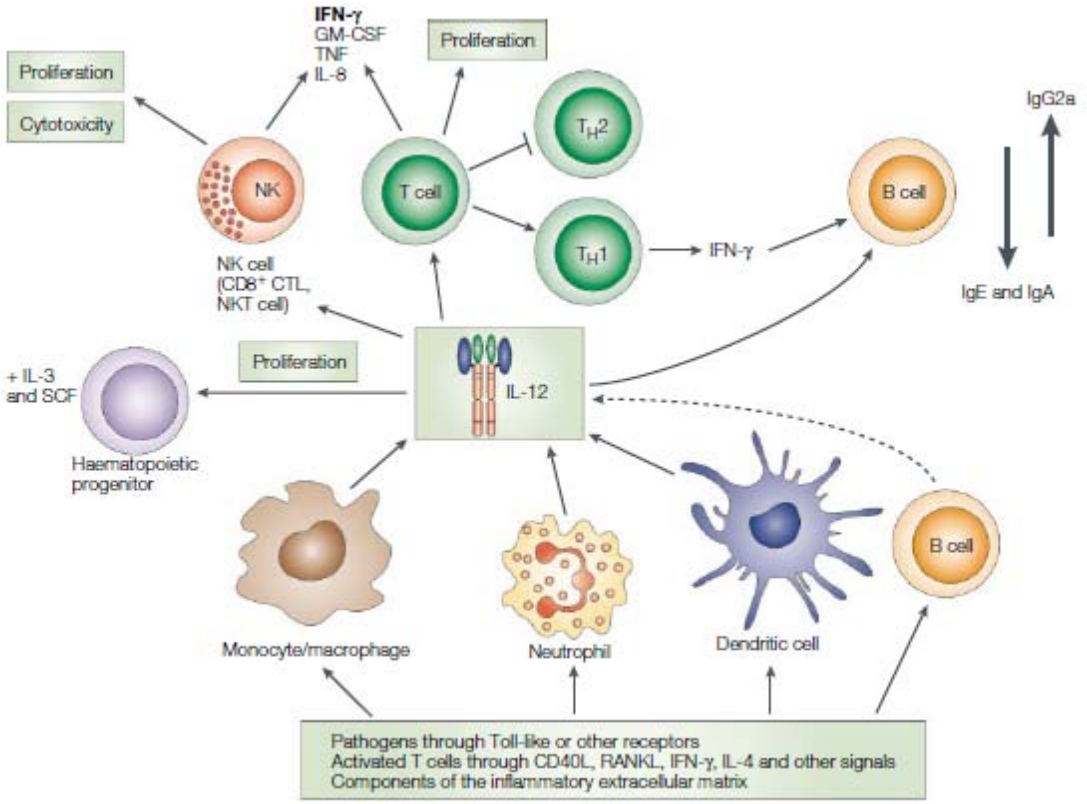

Figure 3. Overall Biology of IL-12 [18].

Copyright Permissions Granted.

The IL-12 receptor (IL-12R) is composed of two chains: IL12-Rβ1 (β1) and IL12-R (β2). These subunits activate the Janus Kinase – Signal Transducer and Activator (JAK-STAT) pathway [18]. The β1 and β2 subunits are found on separate chromosomes, β1on 19p13.1 and B2 on 1p31.2, and are therefore independently regulated. β1, which is predominately used for ligand binding, interacts with the p40 subunit of IL-12. Unlike β1, β2 has tyrosine residues and is therefore used as the signal transducing chain of the receptor. β2 interacts with the p35 chain of IL-12. [20] Co-expression of both subunits is required for high affinity IL-12 binding sites. Both the B1 and B2 subunits are type-I transmembrane glycoproteins with molecular weights of 100

kDa and 130 kDa, respectively [22]. The IL-12 receptors are found on NK cells and activated T cells.

1.4 TYPE 1 T HELPER CELL RESPONSE

Type-I T Helper (T_H1) cells are important cells in the immune response to malaria. Upon activation, naïve $CD4^+$ T cells can differentiate into either T_H1 or T_H2 cells. T_H1 cells leads to cell-mediated immunity and T_H2 leads to humoral immunity [23]. Humoral immune responses are mediated by antibody production by B-cells and cell-mediated immunity is mediated by antigen specific T-cells [17]. Specific cytokine profiles present at the beginning stages of $CD4^+$ T Cells are determining factors for polarization towards a T_H1 or T_H2 response. IL-4, IL-5 and IL-6 are thought to be the T_H2 polarizing cytokines and IL-2 and IL-12 are thought to be the T_H1 polarizing cytokines [24]. IL-12Rβ2 is thought to be the determinant of T_H phenotype because both the β1 and β2 receptors are present on $T_H1$cells but β2 expression is lost in T_H2 cells [20], [24]. A T_H1 immune response presupposes an effective fight against intracellular pathogens [23] and is therefore important in the body's fight against malaria. A T_H1 response is characterized by the production of interferon-gamma (IFN-γ), interleukin (IL)-2 and tumor necrosis factor-beta (TNF-β). TNF-β and IFN-λ both recruit and activate inflammatory cytokines [25]. A schematic of factors involved in the T_H1 differentiation can be seen in Figure 4.

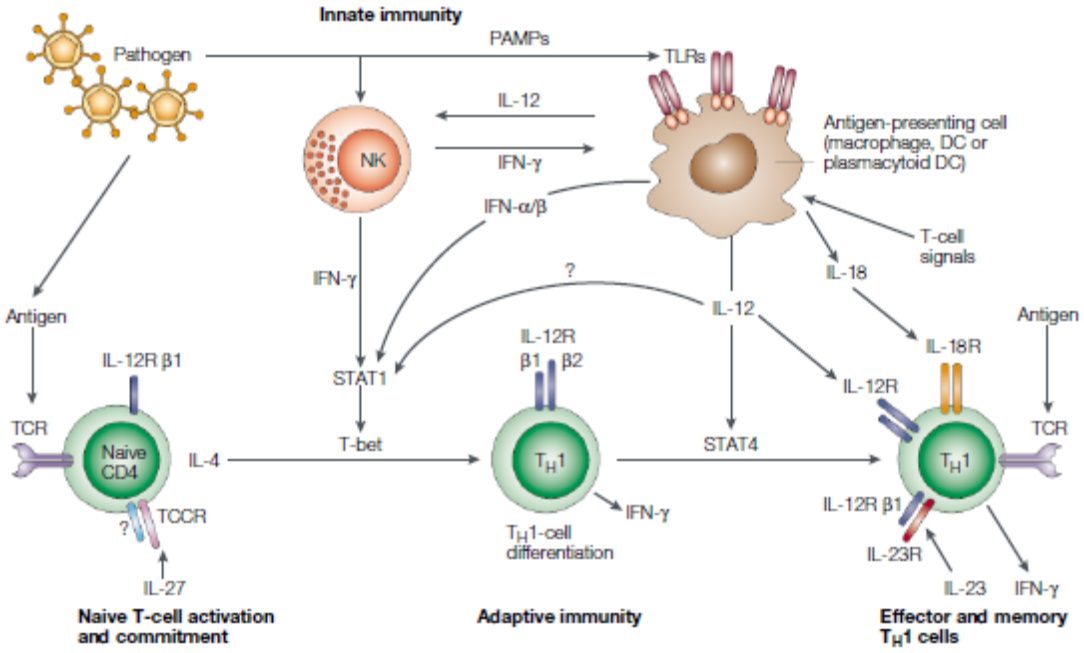

Figure 4. T_H1 differentiation [18].

Copyright Permissions Granted.

1.5 SINGLE NUCLEOTIDE POLYPORPHISMS

Single nucleotide polymorphisms (SNPs) are single base pair changes in the genome that occur every 300-1000 bases. SNPs account for 90% of the genetic variation that exists in humans. In order for a single base pair mutation to be considered a SNP, it must occur in at least 1% of the population. SNPs are important for several reasons. They can affect how the body responds to a specific drug, they can affect how the body responds to an infection, and SNPs can be markers for diseases. A single SNP can also cause a disease such as in the case of Sickle Cell Anemia.

1.6 PREVIOUS STUDIES

Many studies have been done to show the significance of IL-12 and a T_H1 response in the immune response to malaria. Malaguarnera *et al.* (2002) did a study in which they measured the plasma level of IL-12 in a group of 73 children and correlated that level with the severity of the disease. They found that IL-12 levels were significantly elevated in children with less severe disease. They also found that the level of IL-12 was significantly lower in patients with more severe disease. Patients with lower levels of IL-12 showed more severe anemia, lower platelet count, and higher levels of parasite density. [3]

Moreover, studies have also been done to show the influence that SNPs in the IL-12 gene may have in malaria disease outcome. McGuire *et al.* showed that homozygotes for a variant in the tumor necrosis factor-α (TNF-α) had an increased risk for developing CM [10]. Although this study did not focus on IL-12 it was important because it was the first to show a relationship between a cytokine gene polymorphism and a malaria disease outcome. The study that specifically influenced our study hypothesis used 75 tagging SNPs that covered the p35, p40, β1 and β2 genes for IL-12 and IL-12R, respectively. They found a correlation between susceptibility to SMA and SNPS rs2243140 and rs2243133. They also found a correlation between a protection from SMA and SNPs rs383483 and rs429774 as well as a protection against HDP in rs383483. This study is the basis for why we chose to use these SNPs in this study.

2.0 MATERIALS AND METHODS

2.1 STUDY OVERVIEW

For this study, we wanted to analyze four single nucleotide polymorphisms (SNPs), two in the gene encoding the p35 subunit of IL-12 and two in the gene encoding the β_1 subunit of the IL-12 receptor. The specific SNP and its location, effect, and nucleotide change can be seen in Table 3 below. We wanted to investigate a possible correlation between these SNPs and malaria disease outcome. The disease outcomes that we were investigating were severe malaria anemia (SMA) and High density parasitemia (HDP). To do this, we conducted a blinded case-control study; phenotype data was revealed only after samples were genotyped. We used fluorescent polarization (FP) to genotype the samples. After all samples were genotyped we then did the Hardy-Weinberg Equilibrium and chi-squared statistical tests to see if the overall genotype distribution was out of equilibrium. We then stratified our data by disease status to see if the genotype distribution was in Hardy-Weinberg equilibrium.

We then wanted to investigate differences existing in the distribution of the genotypes of the SNPs throughout different regions of the world. In addition to our Kenyan pediatric

population, we also looked at a group from Bogoin Village in Central African Republic, two regions of Papua New Guinea, and four tribes of Madagascar. Furthermore, we also took information from the International HapMap Project to compare our samples with. The populations used from HapMap are the Yoruba in Ibadan, Nigeria (YRI), the Japanese in Tokyo, Japan (JPT), the Han Chinese in Beijing, China (CHB), and the CEPH Collection [(Utah Residents with Northern and Western European Ancestry)CEU] [26].

Table 3. Description of SNPs used in the Study

SNP	Nucleotide Change	Gene	Location	Associated Effect	Associated Disease Outcome
rs2243113	A/G	IL-12A	Promotor	Susceptibility	SMA
rs2243140	T/C	IL-12A	3' UTR	Susceptibility	SMA
rs383483	G/A	IL-12Rβ1	Intron	Protective	HDP
rs383483	G/A	IL-12Rβ1	Intron	Protective	SMA
rs429774	T/C	IL-12Rβ1	Intron	Protective	SMA

UTR – Untranslated Region; SMA – Severe Malaria Anemia; HDP – High Density Parasitemia

2.2 HYPOTHESIS

For the SNPs associated with disease susceptibility, rs2243113 and rs2243140, we expected to see a higher frequency of the homozygous mutant genotype and a deviation from the Hardy-Weinberg Equilibrium (HWE) in individuals with the more severe disease complication.

For the SNPs associated with disease protection, rs383483 and rs429774, we expected to see a higher frequency of the homozygous mutant genotype and a deviation from the HWE in individuals with the less severe disease complications.

For the haplotype data, we expected to see higher frequency of at least one haplotype that is more prevalent in the non-SMA groups and not in the SMA disease group or vice versa.

Geographically, we expected to see higher rates of homozygous mutant protective genotypes in a deviation from HWE in populations with higher rates of malaria transmission.

2.3 STUDY POPULATIONS

2.3.1 KENYAN POPULATION

Children less than three years of age (n=650) who presented at the hospital with acute malaria or for routine immunizations were recruited to the study. All participants were from the Luo ethnic group. Children with detectable *P. falciparum* parasitemia were categorized into two groups depending on their anemia status: non-SMA (Hb $\geq$ 6.0 g/dL, n=309) and SMA (Hb < 6.0 g/dL, n=210). The definition of SMA was based on the distribution of anemia determined by > 14,000

longitudinal Hb measurements in age and gender matched children from the same geographic area. Aparasitemic controls (APs, n=124) were children with *P. falciparum* parasitemia negative blood smears, free from fever or other related malaria symptoms in the two weeks prior to enrollment. Exclusion criteria were co-infection with other *Plasmodium* species, prior hospitalizations, a blood transfusion for any reasons, and CM [Awandare, unpublished data]. It should be noted that the patient status of SMA, non-SMA, and APs were not known by us prior to data collection. Although the total sample size is 650, the sample size used to analyze the data of each SNP was much less because of missing data, no data collection, or problems with genotype calling.

2.3.2 CENTRAL AFRICAN REPUBLIC, MADAGASCAR, PAPUA NEW GUINEA

We wanted to investigate the geographic distribution of the four SNPs under investigation across several populations. The populations used for this part of the study included sample groups from Bogoin Village in Central African Republic, two regions of Papua New Guinea (Kavieng and Madang) and four tribes of Madagascar (Betsileo, Bezanozano, Merina and Sihanaka). These samples were taken as part of an anthropological study; therefore information on malaria disease status is unknown. The total sample size for each population can be seen in Table 4.

Table 4. Sample Size For Each Population

Total Sample Sizes For Each Population	
Population	**Sample Size**
Betsileo	42

Table 4: Continued

Population	Sample Size
Merina	42
Bezanozano	17
Sihanaka	19
Bogoin Village	52
Kavieng	31
Madang	65

2.3.3 HAPMAP DATABASE

HAPMAP is an online database of common genetic variants that occur in human beings. Furthermore, we also took information from the International HapMap Project. The populations used from HapMap are the Yoruba in Ibadan, Nigeria (YRI), the Japanese in Tokyo, Japan (JPT), the Han Chinese in Beijing, China (CHB), and the CEPH Collection [(Utah Residents with Northern and Western European Ancestry)CEU] [26].

2.4 DNA SAMPLE QUANTIFCATION

To quantify the concentration of DNA in each sample we used the PicoGreen assay, which is a dye based assay used to quantify double stranded (ds) DNA. To quantify our DNA we used the Invitrogen Quant-iT™ PicoGreen® Kit. First, 99 µL of 1X TE buffer was added to each well of a 96-well microtitre plate followed by adding one 1 µL of the DNA sample. We then constructed standard curves by serial dilutions of lambda (λ) DNA stock. λ DNA was prepared by diluting the stock solution 1:50 in TE buffer, making enough for three replicate reads.

Table 5. Lambda DNA dilution

(λ) DNA (µL)	**TE buffer (µL)**	**Final Volume (µL)**
12	588	600

We then added the dilutions shown in Table 5 in triplicate to a 96-well microtitre plate. The average of the three readings was the standard to which our DNA samples were compared. The TE buffer was added first and then the λ DNA was added. The following table depicts the serial dilutions (Table 6). These dilutions were used to create the standard curve.

Table 6. λ DNA Serial Dilutions Used to Make Standard Curve

Volume of λ DNA (µL)	**Volume of TE Buffer (µL)**	**Final Volume (µL)**
100	0	1 µg/mL

Table 6: Continued

Volume of λ DNA (µL)	Volume of TE Buffer (µL)	Final Volume (µL)
50	50	500 ng/mL
25	75	250 ng/mL
10	90	100 ng/mL
5	95	50 ng/mL
1	99	10 ng/mL
1	99	1 ng/mL**
0	100	0 ng/ml

** To make the 1 ng/ml dilution, diluted the diluted stock of λ DNA an additional 1:10 with TE buffer. Then added 1µL of the newly diluted stock to the three wells that corresponded to the 1ng/ml well.

Table 7. Volumes Added to Make Pico Green Solution

Pico Green (µL)	TE buffer (mL)	Final Volume (mL)
162.5	32.3375	32.5

Fifth, 100 µL of the Pico Green solution was added to each well of the microtitre plate containing the λ DNA and the microtitre plate containing the sample DNA. After adding the PicoGreen mix, both plates were covered with foil and fluorescence was read on an HBJ BioSystems machine.

2.5 FLUORESCENT POLARIZATION

2.5.1 FLUORESCENT POLARIZATION OVERVIEW

Fluorescent polarization (FP) is a template-directed, primer extension assay that uses fluorophore labeled dideoxy-nucleoside triphosphates (ddNTPs), a primer (referred to as internal primer or IP) and a DNA polymerase. The IP is annealed one base pair upstream of the polymorphic site and in the presence of DNA polymerase and fluorophore labeled ddNTPs, the IP is extended by one base that is complementary to the mutation site. The alleles present in the target DNA can be inferred by the ddNTP that is incorporated. FP is based on the idea that when a fluorescent molecule is excited by plane-polarized light, the molecule itself will also emit polarized light. The FP of a molecule is proportional to the molecule's rotational relaxation time, which is the time it takes to rotate through an angle of 68.5º. The rotational relaxation time is directly proportional to the viscosity of the solvent, absolute temperature, and molecular weight of what of the fluorescent molecule. If viscosity and temperature are held constant then the FP is directly proportional to molecular weight. If the molecule is small it will have lower molecular weight and will therefore tumble faster and emission of light will be depolarized. Likewise, if the molecule is large it will have a higher molecular weight and will therefore tumble slower and light emission will remain polarized. FP is expressed as the ratio of fluorescent detected in the vertical and horizontal axes for each dye and as a result it is independent of fluorescent intensity. [27, 28]. Fluorescence polarization was measured using LJL Biosystems Analyst HT. Samples were plotted on a graph based on their mP value for each dye and their genotype will be determined based on combinations of mP values. Samples with a high mP for only one dye were

genotyped as homozygous for that dye's allele, while samples with high mP values for both dyes were heterozygous for the SNP. FP is done in three steps: Polymerase Chain Reaction (PCR) Amplification of DNA, Degradation and Single Base Extension [27-29].

2.5.2 PCR AMPLIFICATION OF DNA

Polymerase Chain Reaction (PCR) is an enzymatic dependent *invitro* method for amplifying specific segments of DNA [30]. PCR is performed in three steps: (1) denaturation of the double stranded DNA template into single strands; (2) the annealing of primers to each of the single strands for new strand synthesis; and (3) extension of the new DNA strands from the primers [31]. The reagents needed to carry out a PCR reaction are a DNA polymerase, Magnesium Chloride ($MgCl_2$), 10x buffer without $MgCl_2$, primers, deoxynucleotides (dNTPs) and sterile deionized water (SDW). *Taq* Polymerase, $MgCl_2$ and the 10x buffer were all purchased from Sigma. The sequence of the PCR primers used for each SNP can be seen in Table 8. The dNTPs were purchased from New England BioLabs. The deoxynucleotide solution set contained four separate solutions (one for each dNTP) at 0.25 ml of 100 mM ultrapure nucleotide. A dNTP mix was made by adding 20 µL of each dNTP into a 0.2 µL tube. For all plates the master mix was first made and then 10 µL of the master mix was aliquoted into each individual well. The DNA was then added to each well using an eight channel pipette. For each well of every plate, 2.5 µL of DNA was added. The master mix and DNA were added in this order to eliminate the possibility of contamination. The reaction mixture used for each 96 well plate and specific SNP can be seen in tables 9-14. The reaction was run in a Genemate 96 well 0.2 mL PCR low profile

flat plate purchased from ISC Bioexpress. The specific PCR program that was used for each plate is as follows:

95º C for 2 minutes

95º C for 30 seconds

63º C for 30 seconds

The temperature was increased by 3 º/s until 72 º C

72 º C for 30 seconds

Steps 2-5 were repeated 13 times.

95 º C for 30 seconds

56º C for 30 seconds

72º C for 30 seconds

Steps 6-8 were repeated 21 times

Table 8. Sequences of PCR Primers

PCR Primer Sequences		
SNP	**Forward Primer**	**Reverse Primer**
rs2243113	CCCACCCTGTGTTCTCCATACTG	AACAAGCAGGCATTAGACCGC
rs2243140	TGAGTGGCAGGTGAGCAAGTATTAC	GGGCAAAGATAGTTTCGGGATG
rs383483	TGGAGCACTGGTTTTGGAGTCC	GGGAGGGGATAGTAACTTTTGGGTC
rs429774	TGCTCCACTTGCTCCAGGAACTTC	TGCGGACCTCAGGATGATGATAC

Table 9. PCR Reaction Mixture for rs2243113 for plate 1.

PLATE 1			
REAGENT	**PER CELL VOLUME (µL)**		**MASTER MIX VOLUME (µL)**
10x buffer	1.10		110
MgCl2	1.25		125
dNTPs	0.05		5
PRIMERS	0.10	x 100	10
TAQ	0.15		15
SDW	7.35		735

Table 10. PCR Reaction Mixture for rs2243113 plates 2-8

PLATE 2-8			
REAGENT	**PER CELL VOLUME (µL)**		**MASTER MIX VOLUME (µL)**
10x buffer	1.10		110
MgCl2	1.25		125
dNTPs	0.05		5
PRIMERS	0.10	x 100	10
TAQ	0.10		10
SDW	7.40		740

Table 11. PCR Reaction Mixture for rs2243140, plate 1

PLATE 1			
REAGENT	**PER CELL VOLUME (µL)**		**MASTER MIX VOLUME (µL)**
10x buffer	1.10		110
MgCl2	1.25		125
dNTPs	0.05		5
PRIMERS	0.10	x 100	10
TAQ	0.15		15
SDW	7.35		735

Table 12. PCR Reaction Mixture for rs2243140, plates 2-8

PLATES 2-8			
REAGENT	**PER CELL VOLUME (µL)**		**MASTER MIX VOLUME (µL)**
10x buffer	1.10		110
MgCl2	1.25		125
dNTPs	0.05		5
PRIMERS	0.10	x 100	10
TAQ	0.10		15
SDW	7.40		740

Table 13. PCR Mixture for rs429774, plate 1

PLATE 1			
REAGENT	**PER CELL VOLUME (µL)**		**MASTER MIX VOLUME (µL)**
10x buffer	1.10		110
MgCl2	1.25		125
dNTPs	0.05		5
PRIMERS	0.10	x 100	10
TAQ	0.15		15
SDW	7.35		735

Table 14. PCR Mixture for rs429774, plates 2-8

PLATES 2-8			
REAGENT	**PER CELL VOLUME (µL)**		**MASTER MIX VOLUME (µL)**
10x buffer	1.10		110
MgCl2	1.25		125
dNTPs	0.05	x 100	5
PRIMERS	0.10		10
TAQ	0.10		15
SDW	7.40		740

2.5.3 GEL ELECTROPHORESES

After the PCR product, we ran 1 µL of the PCR product on a gel to verify that the DNA amplification was successful. We used 1 µL of 2x dye mixed with 1 µL of our PCR product to run on the gel. Our agarose gels were made by heating 5 g of agarose into 250 mL of TE buffer. After all of the agarose was dissolved into the buffer, the mixture was allowed to cool. Then 12.5 µL of ethidium bromide (EtBr) was added to the mixture. The mixture was poured into a loading tray and allowed to solidify for 30 minutes. A solution of 300 µL of TE buffer and 15 µL of EtBr was added to the loading tray. The gel was then added into the loading tray. Using an 8 channel pipette, the DNA was added to the gel. The gel ran at 100 volts for 30 minutes.

2.5.4 DEGRADATION

The degradation step of the FP is used to degrade excess dNTPs and unincorporated primers. 1 µL of rapid alkaline phosphatase (Roche), 0.05 µL of EXOI and 0.1 µL of 10x rap buffer was added to each individual well. A master mix was made and a multi-dispensing pipette was used to dispense 10 µL of the master mix into each well. The volumes used for each reagent can be seen in table The reaction was run in a genemate 0.2 mL 96 well low-profile PCR plate (purchased from ISC BioExpress) for total of 75 minutes. The first 60 minutes was run at 37º C and the last 15 minutes was run at 85ºC.

Table 15. Reaction Volumes Used for Degradation Step.

REAGANT	PER CELL VOLUME (µL)		MASTER MIX VOLUME (µL)
RAPID ALKALINE PHOSPHATASE	0.1	x 100 (for the master mix)	100
10X BUFFER	0.1		100
EXO1	0.05		5
SDW	7.95		795
TOTAL	10		1000

2.5.5 SINGLE BASE EXTENSION

The single base extension step was done following the degradation step. A reaction using the forward internal primer and a separate reaction using the reverse internal primer were carried out. Primer sequences can be seen in Table 16. To avoid contamination, the reaction mixture was first made and then added to a 96 well, skirted, black twin tec PCR plate with yellow/black wells, purchased from Eppendorf, using a multi dispenser pipette. 5 µL was added to each individual well. 10 µL of the degraded-PCR product was added to the respective well using an 8 channel pipette. The program used for all plates is as follows:

94º C for 1 minute

94º C for 10 seconds

52º C for 30 seconds

Repeat steps 2 and 3 40 times

72º C for 10 seconds

Table 16. Forward and Reverse Internal Primer Sequences

FP Internal Primer Sequences		
SNP	**Internal Primer (Forward)**	**Internal Primer (Reverse)**
rs2243113	AAGACTCCCCACTACTCTCAGCAC	TGCCATACTAAGGAGTGTACTTTC
rs2243140	ATGCTCCTTATAAGAATCTAAGTAA	CTGTTCCACCTCGTATTATCAGGC
rs383483	GCTAATGCGTAACCCTTGTCCAGC	CCAAGTCTTTTTATTGGGTGCACA
rs429774	GGCACCAGTCACTTAACCATCAT	TACACAGTTGGAGCTCAATAAGC

Table 17. Single Base Extension Reaction Mixture

REAGENT	**VOLUME PER CELL (µl)**		**MASTER MIX VOLUME (µL)**
THERMO SEQUENASE	0.05	x 100 (for master mix)	5
10x BUFFER	0.5		50
DYE LABLED ddNTPs	0.025		2.5
INTERNAL PRIMER	0.5		5
SDW	3.925		392.5

3.0 STATISTICAL METHODS

3.1 HARDY-WEINBERG PRINCIPLE

The Hardy-Weinberg Principle (HWP) asserts that under specific conditions the allele and genotype frequencies of a particular population remain constant when certain assumptions are met. This equilibrium is maintained from generation to generation unless those assumptions about the population are disturbed. The main assumptions made under the HWP are:

1. Diploid organisms
2. Sexual reproduction
3. Non-overlapping generations
4. Random mating
5. Large population size
6. Equal allele frequencies in the sexes
7. No mutation
8. No selection

With these assumptions, the genotype frequencies for a biallelic gene can be calculated using the Hardy-Weinberg Equation [32].

3.2 HARDY-WEINBERG EQUATION

The Hardy-Weinberg Equilibrium (HWE) is a theoretical method which describes and predicts genotype and allele frequency in a given population. HWE enables the actual genetic make up of a population to be compared to what is expected to be seen under the Hardy-Weinberg Principle. The Hardy-Weinberg Principle asserts that under specific conditions the allele and genotype frequency of a particular gene remain constant. The Hardy-Weinberg equation takes the following the form:

$$p^2 + 2pq + q^2 = 1$$

where p^2 is homozygous for one allele (AA), 2pq is heterozygous for two different alleles (Aa) and q^2 is homozygous for the second allele (aa) [32]. The total number of alleles must be calculated in order to perform the HWE. These values are then used to compute the expected values. Once the observed alleles were obtained, these values were used to calculate the p^2, q^2 and 2pq values [32]. The observed and expected values can be placed into a contingency table (Table 18):

Table 18. Example of Chi Squared Contingency Table

ALLELE	OBSERVED	EXPECTED
AA	OAA	EAA
GG	OGG	EGG
GA	OGA	EGA
TOTAL	OAA + OGG + OGA	EAA + EGG + EGA

3.3 STATISTICAL TESTS

3.3.1 CHI-SQUARED TEST OF INDEPENDENCE

For the overall data, the chi-squared (χ^2) test of independence was used to determine if the expected values were significantly different from the observed values. The χ^2 test statistic is calculated by the following formula:

$$\chi^2 = \sum\left(\frac{(o-e)^2}{e}\right)$$

The null (H_0) and alternative (H_1) hypotheses were set as the following:

H_0: There is no significant difference between observed and expected values

H_a: There is a significant difference between observed and expected values

We set our alpha as (α) = 0.05 and our degrees of freedom (df) as one. Under the $\chi 2$ distribution, the $\chi 2$ critical value for α = 0.05 and one degree of freedom is 3.94. Any statistic found that is equal or less than 3.94 would lead to a failure to reject H_0. Any statistic found that is greater than 3.94 would result in rejecting H_0.

3.3.2 FAST EXACT HARDY-WEINBERG TEST FOR SNPS

For the geographical distribution data and the stratified Kenya data, the sample sizes were small (< 5) so instead of using the $\chi 2$ test statistic, a fast exact Hardy-Weinberg Equilibrium test for SNPs (SNP-HWE) was performed using the R statistical software package version of SNP-HWE [33].

The null (H_0) and alternative (H_1) hypotheses were set as the following:

H_0: There is no significant difference between observed and expected values

H_a: There is a significant difference between observed and expected values

We set our p-value at 0.05; any value less than or equal to 0.05 we will reject H_0 and any value greater than 0.05 we will fail to reject H_0.

3.4 HAPLOTYPE CONSTRUCTION

A haplotype is a profile of SNPs; it is the combination of SNP alleles on a particular gene or region of a gene. Because individuals have 2 copies of each gene, there are also 2 haplotypes for each combination of SNPs. We created haplotypes to enhance our allelic information because sometimes a single SNP is not sufficient to explain an observed phenotype. We used the PHASE software package to create haplotypes of the SNPs that are found on the same gene [34]. We constructed haplotypes for rs2243113 and rs2243140 which are both found on IL-12A and haplotypes for rs383483 and rs429774, which are both found on IL-12Rβ1.

4.0 RESULTS

4.1 DESCRIPTION OF FP RESULTS

We analyzed 4 SNPs, two in the gene encoding the p35 subunit of IL-12 and 2 in the gene encoding the β1 receptor of the IL-12 receptor. We genotyped the DNA samples for these SNPs using Fluorescent Polarization (FP). With Fluorescent Polarization (FP), the data is clustered into four groups: homozygous AA, homozygous GG, heterozygous GA or blank. The mP values of the sample determine how that specific sample will be called. Samples that fall within the mP value range for homozygous AA will be called 'AA,' samples that fall within the mP value range of GG will be called 'GG' and samples that fall within the mP values range of blank will be called 'Blank'. We had samples that did not fall within those ranges and thus were unable to be called into any particular group. These data samples will be termed 'No Call'. The 'No Call' samples are the samples that lie within the yellow and red bars on the scatter plot. An example of scatter plot readout can be seen in Figure 5.

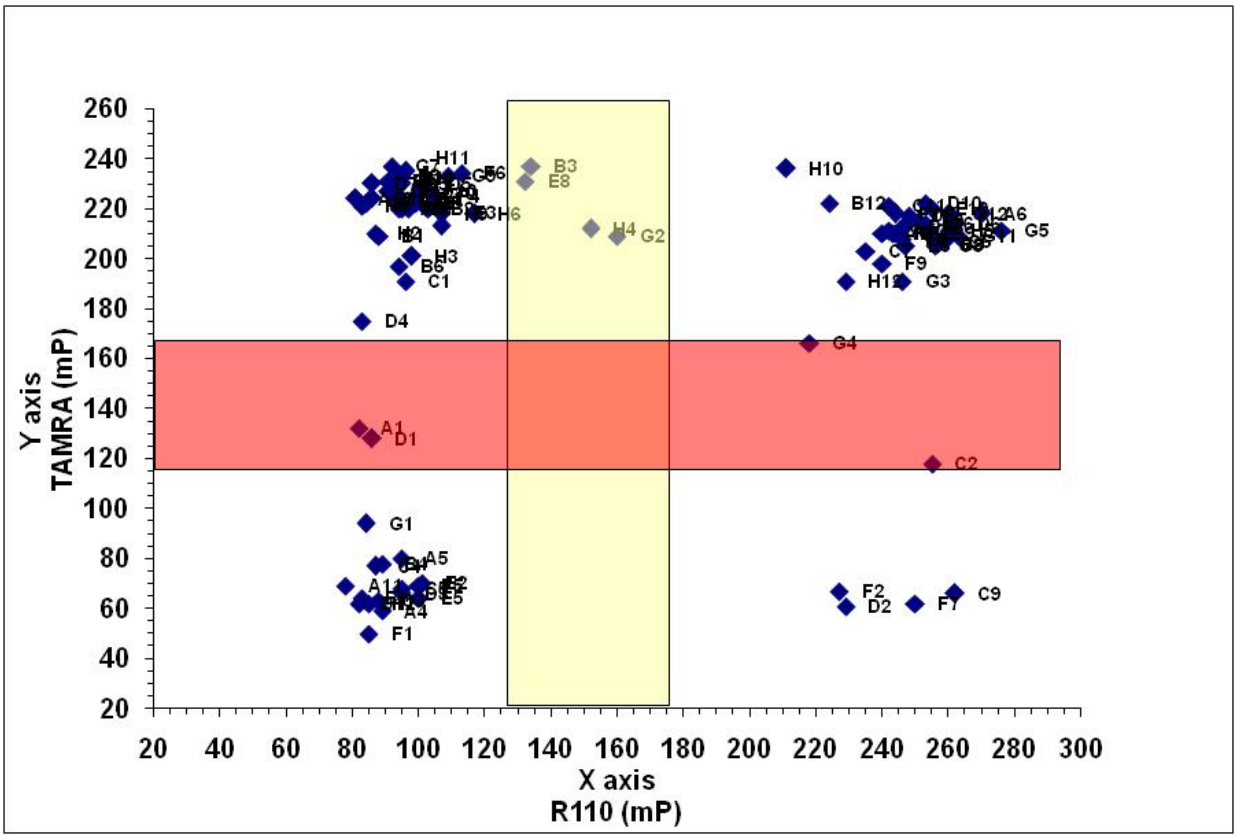

Figure 5. Scatterplot readout of FP data.

From the FP results we calculated the frequency of each genotype for each SNP. To do this the results from each 96 well plate added together. From this data, we compiled two tables: an overall genotype frequency and an informative genotype frequency. The overall genotype frequency table includes all the samples even if they were 'Blank' or 'No Call'. The Informative Genotype Frequency Table includes only those samples that were called 'AA', 'GG' or 'GA'. The total frequency of each genotype for each particular SNP can be seen Tables 19-26.

Table 19. Overall Genotype Frequency for rs2243113.

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
AA	95	14.1
GA	271	40.3
GG	9	1.3
BLANK	97	14.4
NO CALL	200	29.8
TOTAL	672	100

Table 20. Informative Genotype Frequencies for rs2243113

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
AA	95	25.3
GA	271	72.3
GG	9	2.4
TOTAL	375	100

Table 21. Overall Genotype Frequency for rs2243140

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
AA	126	26.3
GA	38	7.9
GG	141	29.4
BLANK	100	20.8

Table 21: Continued

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
NO CALL	75	15.6
TOTAL	480	100

Table 22. Informative Genotype Frequency for rs2243140

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
AA	126	41.3
GA	38	12.5
GG	141	46.2
TOTAL	305	100

Table 23. Overall Genotype Frequency for rs383483

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
TT	49	6
TC	23	12.8
CC	96	25
BLANK	126	32.8
NO CALL	87	22.7
TOTAL	672	100

Table 24. Informative Genotype Frequency for rs383483

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
TT	49	29.2
TC	23	13.7
CC	96	57.1
TOTAL	168	100

Table 25. Overall Frequency for rs429774

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
AA	256	38.3
GA	81	12.1
GG	55	8.2
BLANK	173	25.9
NO CALL	103	15.4
TOTAL	668	100

Table 26. Informative Genotype Information for rs429774

GENOTYPE	FREQUENCY	PERCENTAGE OF SAMPLES (%)
TT	256	65
TC	81	21.1
CC	55	14
TOTAL	394	100

4.2 HARDY-WEINBERG AND CHI-SQUARED TEST

4.2.1 OVERALL HWE AND CHI-SQUARED TEST RESULTS

We used the Hardy-Weinberg Equilibrium (HWE) to compare the genotypic makeup of our samples to what we would expect to see under the Hardy-Weinberg Principle. We calculated the frequency of A, G, C or T alleles for each specific SNP. This information can be observed in Tables 27-30.

Table 27. Allele Frequency For rs22431113

ALLELE	OBSERVED FREQUENCY
A	451
G	289
TOTAL	740

Table 28. Allele Frequency For rs2243140

ALLELE	OBSERVED FREQUENCY
A	290
G	320
TOTAL	610

Table 29. Allele Frequency for rs383483

ALLELE	OBSERVED FREQUENCY
T	121
C	215
TOTAL	336

Table 30. Allele Frequency for rs22431113

ALLELE	OBSERVED FREQUENCY
A	595
G	193
TOTAL	788

We used the Hardy-Weinberg equation ($p^2 + 2pq + q^2$) in conjunction with our allele frequencies and observed number of genotype frequencies to calculate the values for $p^{2,}$ 2pq, and q^2. We then used our p^2, q^2 and 2pq values to calculate our expected observations. To do this, we multiplied our observed total population by our p^2, q^2 and 2pq values. This data can be observed in the Tables 31-34.

Table 31. Observed and Expected Genotype Frequency for rs2243113

GENOTYPE	OBSERVED (O)	EXPECTED (E)	(O-E)	$(O-E)^2$	$(O-E)^2 / E$
AA	95	139.125	-44.125	1947	13.99
GG	9	57.375	-48.375	2340	40.79
GA	271	178.5	92.5	8556	47.93
TOTAL	375	375			

Table 32. Observed and Expected Genotype Frequency for rs2243140

GENOTYPE	OBSERVED (O)	EXPECTED (E)	(O-E)	$(O-E)^2$	$(O-E)^2 / E$
AA	126	68.93	57.07	3256	47.25
GG	141	83.875	57.13	3263	38.91
GA	38	152.195	-114.20	13030	85.68
TOTAL	305	305			

Table 33. Observed and Expected Genotype Frequency for rs383483

GENOTYPE	OBSERVED (O)	EXPECTED (E)	(O-E)	$(O-E)^2$	$(O-E)^2 / E$
AA	49	21.84	27.16	737	33.77
GG	96	68.712	27.29	744.63	10.84
GA	23	77.448	-54.45	2964	38.27
TOTAL	168	168			

Table 34. Observed and Expected Genotype Frequency for rs429774

GENOTYPE	OBSERVED (O)	EXPECTED (E)	(O-E)	$(O-E)^2$	$(O-E)^2$ / E
AA	256	224.58	31.42	987	4.40
GG	55	23.64	31.36	983	41.60
GA	83	145.78	-62.78	3941	27.03
TOTAL	394	394			

We then calculated our χ^2 value by using the χ^2 equation ($\chi^2 = \sum\left(\frac{(o-e)^2}{e}\right)$). The values for these can be found in Table 35.

Table 35. Calculated χ^2 values for each SNP.

SNP	χ2 Value
rs2243113	102.71
rs2243140	171.84
rs383483	82.88
rs429774	73.3

All of our calculated χ^2 values are greater than 3.94 so we would reject H_0 thus concluding that there is a significant difference between our observed and expected values. We then used the HWE to determine if the difference that exists between our observed and expected values could be accounted for by the specific clinical statuses.

4.2.2 HWE AND SNP_HWE RESULTS BY DISEASE STATUS

Observed and expected values for each SNP by the specific clinical disease status can be seen in Tables 36-40.

SNP rs2243113 is a change from an A → G and rs2243140 is a change from a T → C and they are both been associated with a susceptibility to SMA. Although we saw a significant difference between the observed and expected values within each clinical status we did not see a significant difference between each status. Looking at the graphical representations (Figures 6 and 7) we can see that each disease group follows the same trend. If this SNP was a playing a role in susceptibility to SMA, we would only expect to see a higher frequency of the homozygous mutant genotype and a deviation from HWE in the SMA disease group. We can conclude that for our particular population rs2243113 and drs2243140 do not display a susceptibility to SMA.

SNP rs383483 has been shown to have a protective effect against SMA and HDP and rs429774 has been shown to have a protective effect for SMA. SNP rs383483 is a change from a G → A and rs429774 is a change from a C → T; rs383483 has been associated with a protection against SMA and HDP and rs429774 has been associated with a protection against SMA. For rs429774 and rs383483 (HDP protection) we observed a significant difference between the observed and expected values within each disease group, but between each group there is not a difference (Figures 8 – 10). If this SNP was playing a protective role, we would expect to see a higher frequency of the homozygous mutant genotype and a deviation from HWE in the non-SMA disease groups. We can conclude that for our population, SNP rs383483 does not play a protective role against SMA and HDP and SNP rs429774 does display a protection against SMA.

Table 36. Observed and Expected Values and Corresponding p-values for Anemia Status by Genotype – rs2243113

rs2243113 Anemia Status by Genotype			
Non-Anemic			
Genotype	**Observed**	**Expected**	**p-value**
AA	8	13	0.0005
AG	26	16	
GG	0	5	
Mild Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
AA	34	45	4.86e-8
AG	67	45	
GG	0	11	
Moderate Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
AA	23	31	2.93e-6
AG	49	17	
GG	0	8	
Severe Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
AA	18	24	0.002
AG	41	29	
GG	3	9	

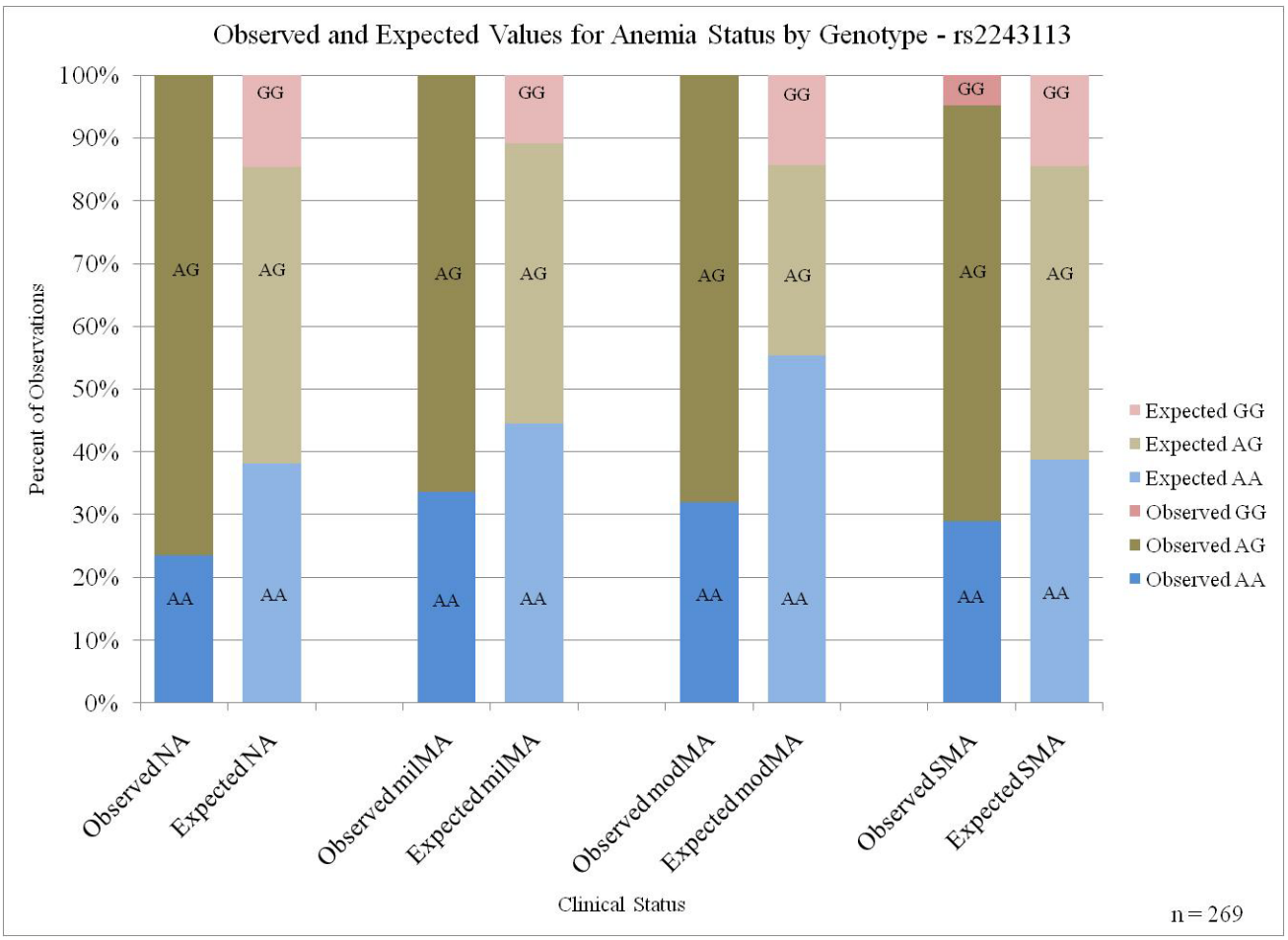

Figure 6. Observed and Expected Values for Anemia Status by Genotype – rs2243113

Table 37. Observed and Expected Values and Corresponding p-values for Anemia Status by Genotype –rs22431140

rs2243113 Anemia Status by Genotype			
Non-Anemic			
Genotype	**Observed**	**Expected**	**p-value**
TT	10	4	1.08e-6
TC	2	14	
CC	19	13	
Mild Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
TT	41	22	6.43e-15
TC	13	50	
CC	47	28	
Moderate Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
TT	27	13	4.78e-13
TC	8	35	
CC	37	23	
Severe Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
TT	31	20	1.91e-8
TC	11	33	
CC	25	14	

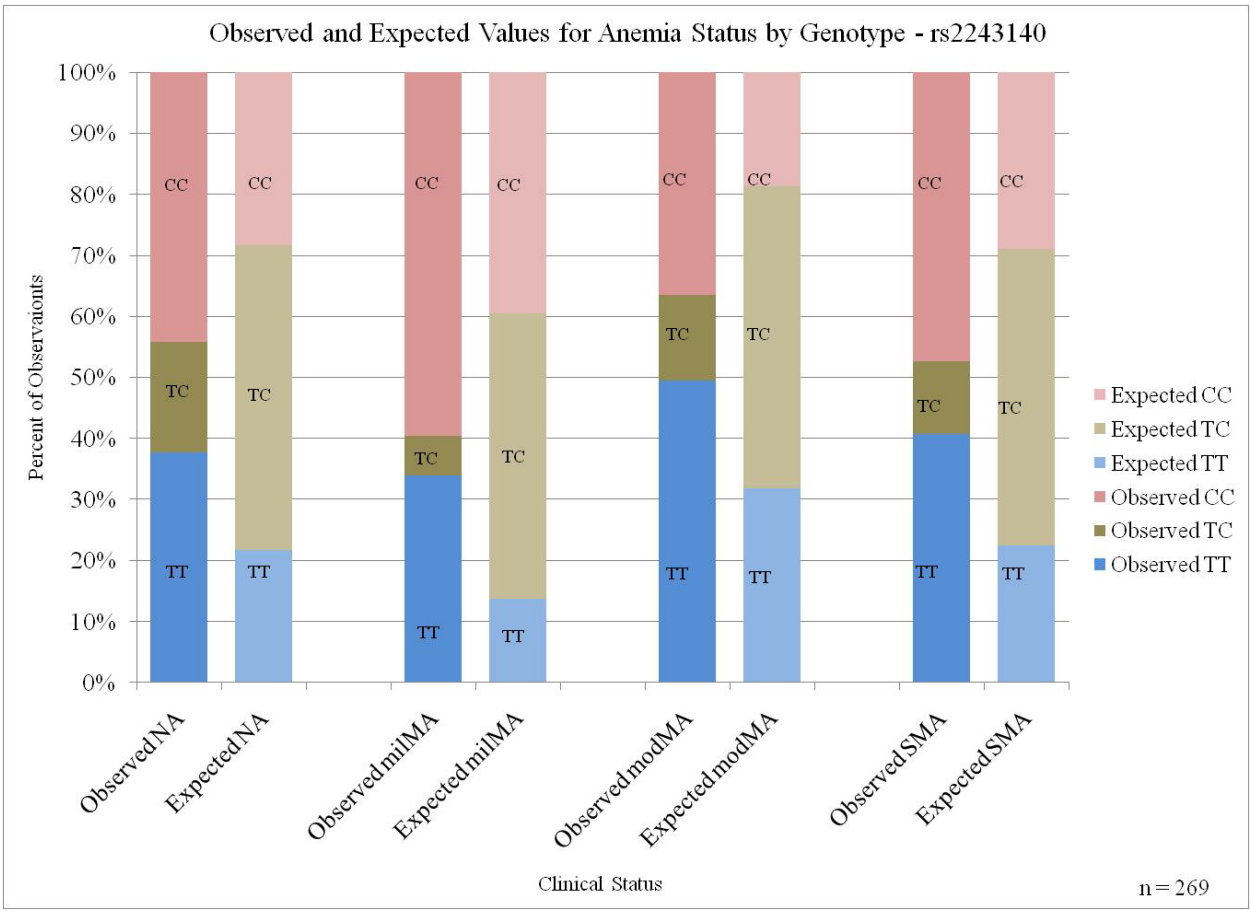

Figure 7. Observed and Expected Values for Anemia Status by Genotype – rs2243140

Table 38. Observed and Expected Values and Corresponding p-values for Anemia Status by Genotype –rs383483

rs383483 Anemia Status by Genotype			
Non-Anemic			
Genotype	**Observed**	**Expected**	**p-value**
GG	9	6	0.0006
GA	3	11	
AA	10	5	

Table 38: Continued

Mild Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
GG	29	20	1.35e-6
GA	9	26	
AA	17	8	
Moderate Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
GG	27	19	3.65e −7
GA	5	20	
AA	13	5	
Severe Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
GG	25	22	0.002
GA	5	12	
AA	5	2	

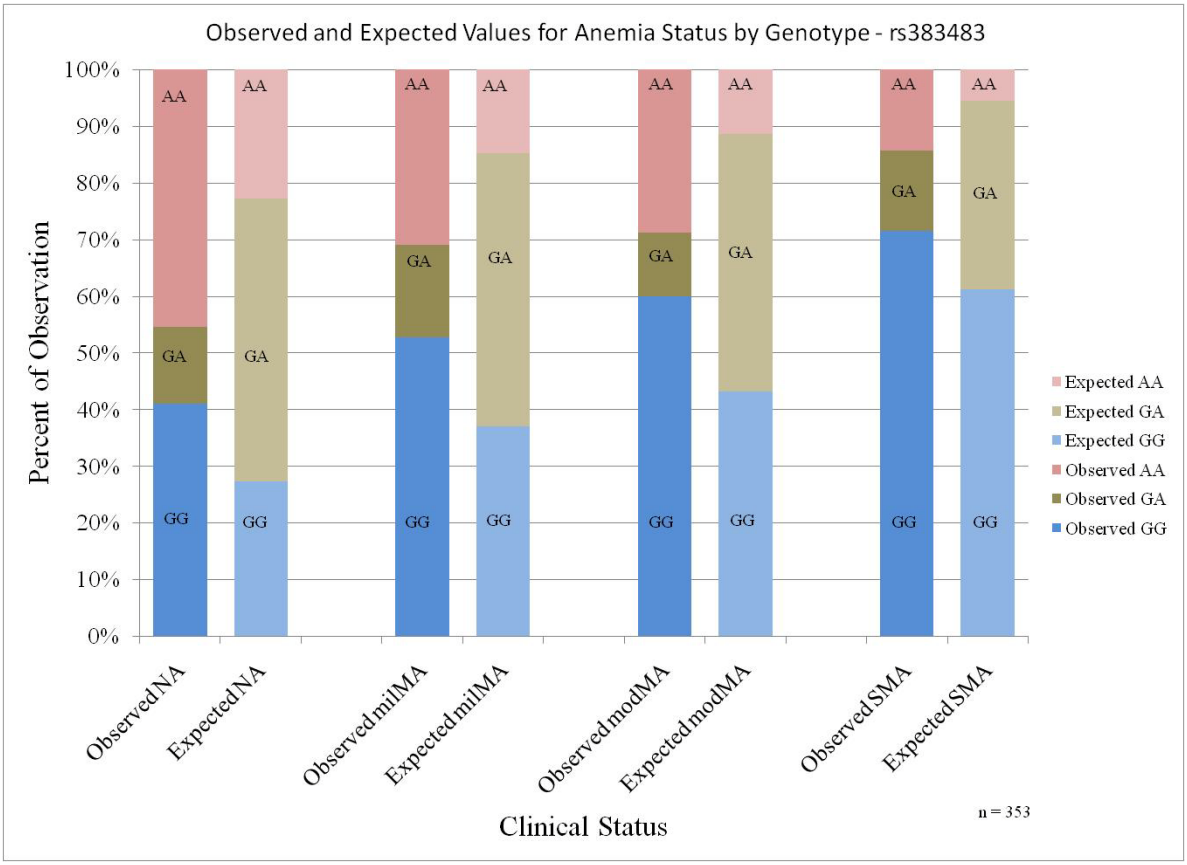

Figure 8. Observed and Expected Values for Anemia Status by Genotype – rs383483

Table 39. Observed and Expected Values and Corresponding p-values for Parasitemia Status by Genotype - rs383483

rs383483 Anemia Status by Genotype			
Aparasitemic			
Genotype	**Observed**	**Expected**	**p-value**
GG	18	11	3.15e −6
GA	7	17	
AA	10	6	

Table 39: Continued

Low Density Parasitemia			
Genotype	**Observed**	**Expected**	**p-value**
GG	48	34	9.3e −12
GA	9	36	
AA	23	9	
High Density Parasitemia			
Genotype	**Observed**	**Expected**	**p-value**
GG	24	19	0.0006
GA	8	17	
AA	9	4	

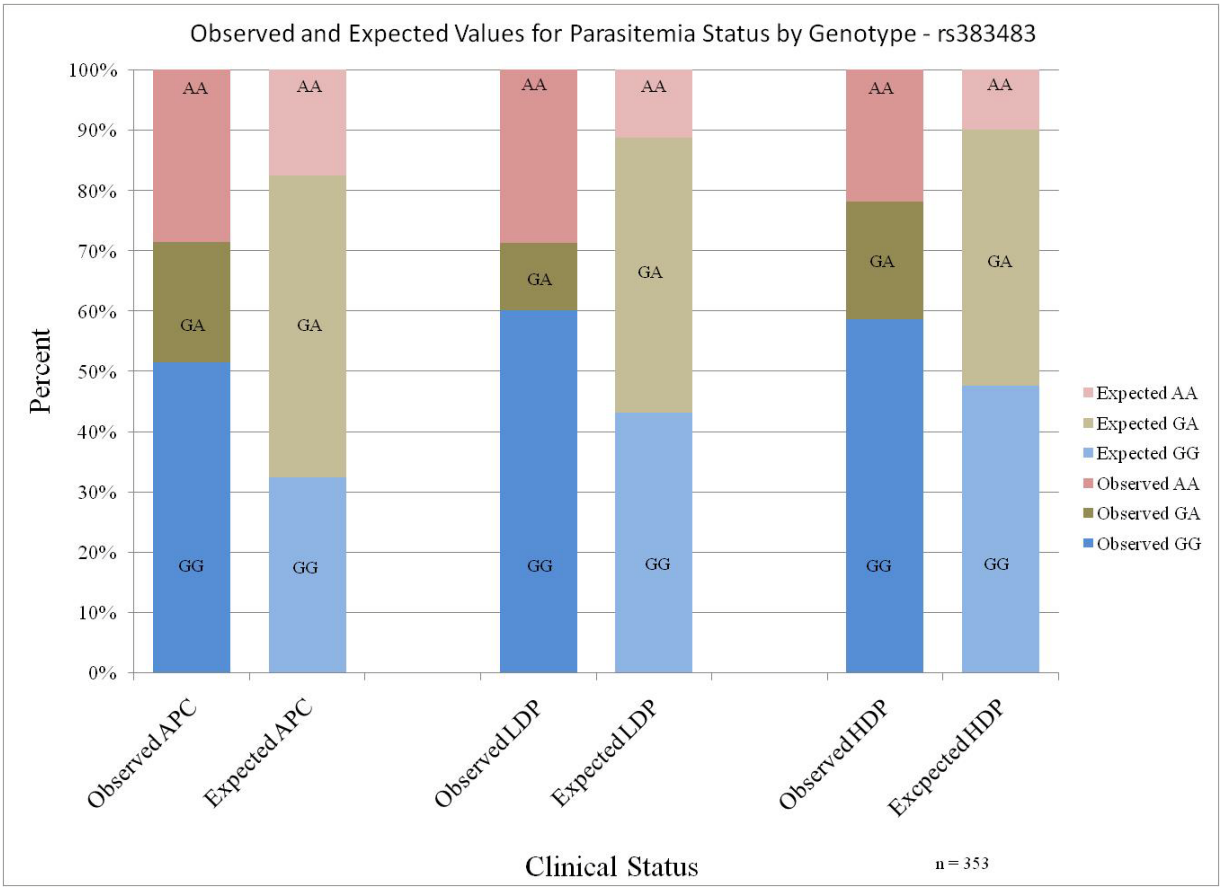

Figure 9. Observed and Expected Values for Parasitemia Status by Genotype – rs383483

Table 40. Observed and Expected Values and Corresponding p-values for Anemia Status by Genotype

rs429774 Anemia Status by Genotype			
Non-Anemic			
Genotype	**Observed**	**Expected**	**p-value**
CC	10	4	0.0002
CT	10	21	
TT	32	26	

Table 40: Continued

Mild Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
CC	19	9	3.62e-6
CT	25	45	
TT	69	59	
Moderate Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
CC	12	4	2.08e-5
CT	17	32	
TT	66	58	
Severe Malaria Anemia			
Genotype	**Observed**	**Expected**	**p-value**
CC	15	6	5.28e-6
CT	18	36	
TT	60	51	

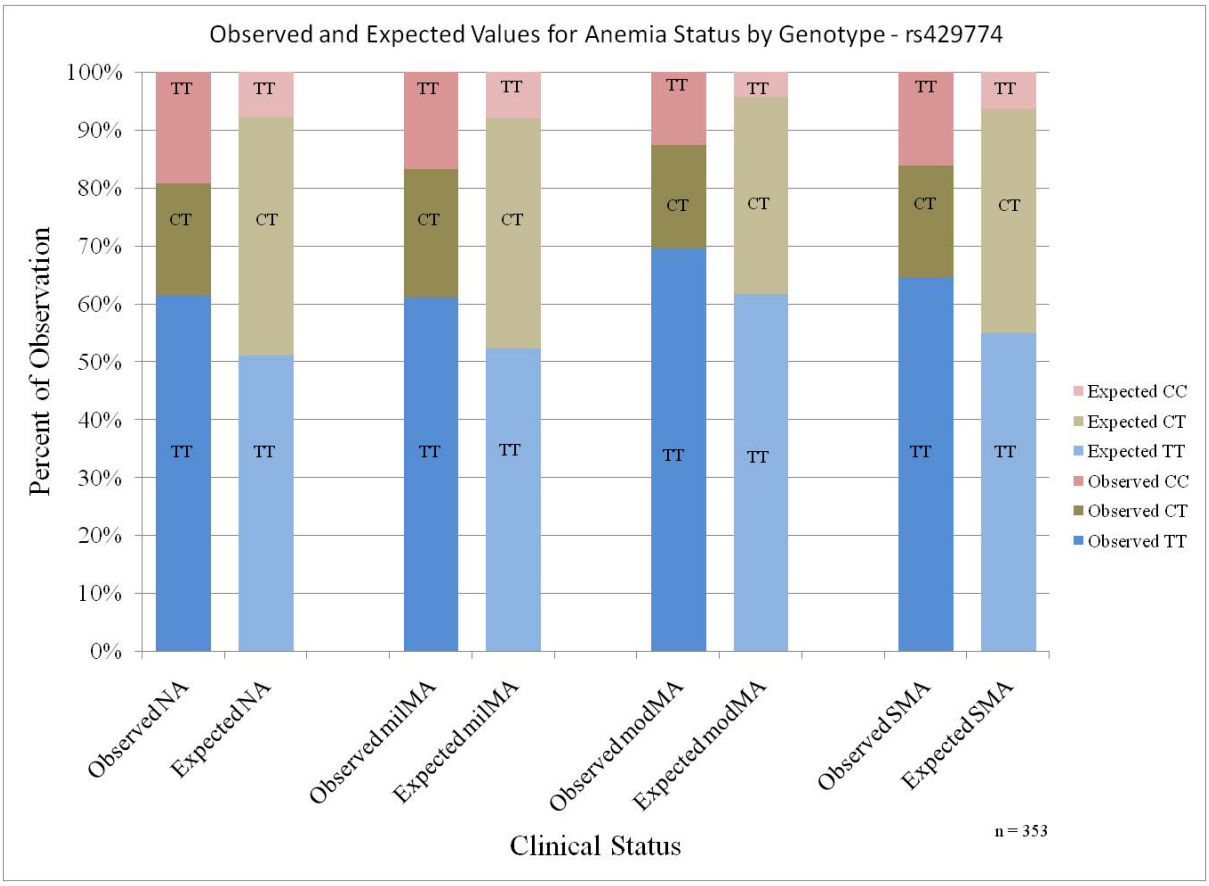

Figure 10. Observed and Expected Values for Anemia Status by Genotype – rs429774

4.3 HAPLOTYPE CONSTRUCTION

A haplotype is a combination of SNP alleles on a particular gene or region of a gene. Because individuals have two copies of each gene, there are also two haplotypes for each combination of SNPs. We created haplotypes to enhance our allelic information because sometimes a single SNP is not sufficient to explain a particular phenotype. Data for each haplotype can be seen in Tables 40 – 43 and graphical representations can be seen immediately following (Figures 11-14). The haplotypes for IL-12 consist of SNPs rs2243113 and rs2243140, both of which have been associated with a susceptibility to SMA. The haplotypes for IL-12Rβ1 consists of SNPs rs383483 and rs429774, both of which have been associated with a protection against SMA. Although rs383483 has also been associated with a protection against HDP, we only looked SMA because rs429774 has not been shown to be protective for HDP. For all of the haplotypes, there was not one that was either more prevalent in the less severe anemia status as compared to the SMA group or more prevalent in the SMA disease groups compared to the less severe disease group. Because the haplotype data did not reveal any results that differed or help to explain the results of our Hardy-Weinberg results, we decided to not pursue any further testing.

Table 41. IL-12 Haplotypes for SNP Profile 1

	Haplotype			
Disease Status	**AC**	**AT**	**GC**	**GT**
Non-Anemic	0	11	17	7
Mild Malaria Anemia	4	9	59	30

Table 41: Continued

Disease Status	AC	AT	GC	GT
Moderate Malaria Anemia	6	8	36	21
Severe Malaria Anemia	7	23	29	20

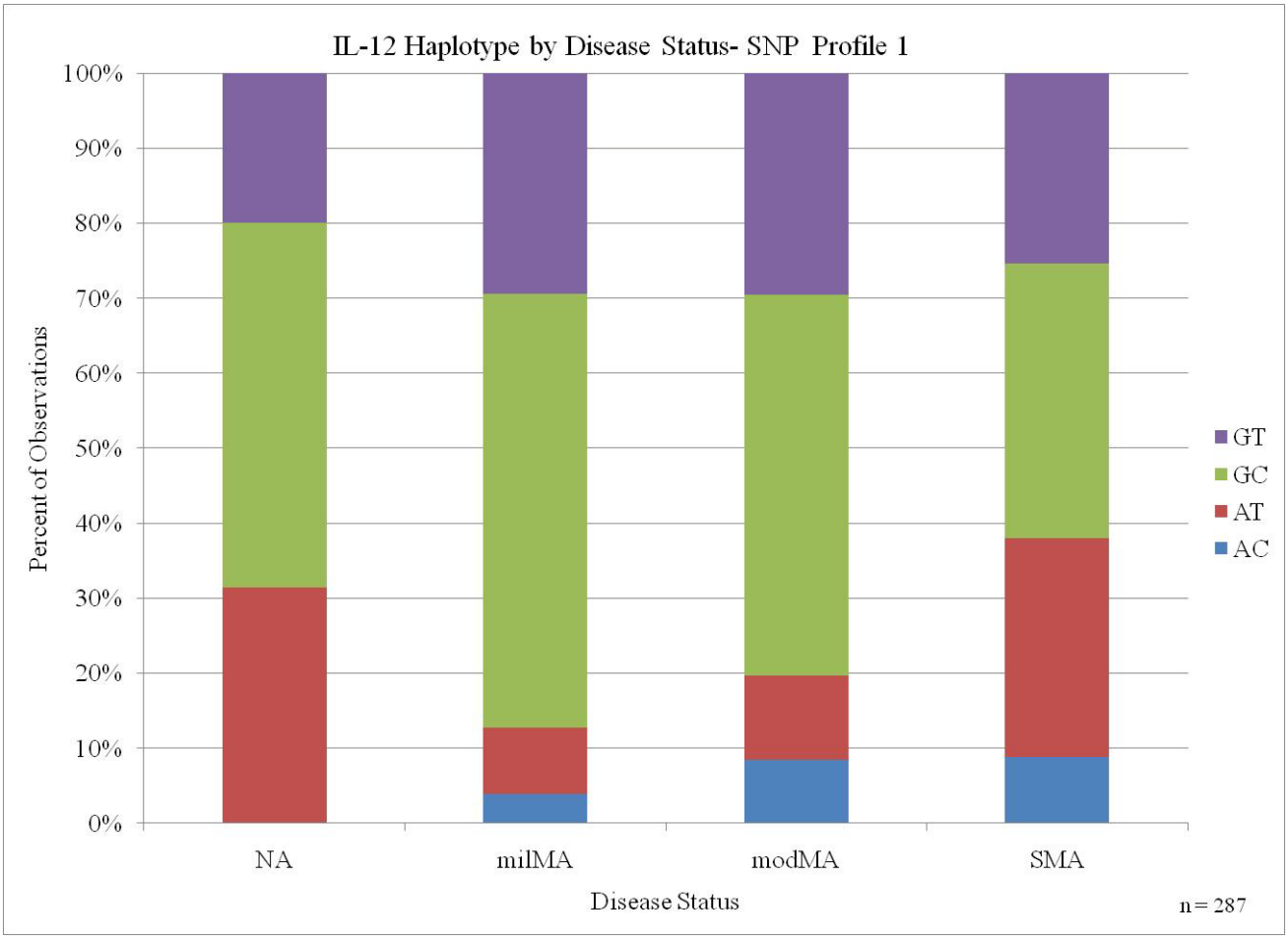

Figure 11. IL-12 SNP Profile 1 Haplotypes

Table 42. IL-12 Haplotypes for SNP Profile 2

	Haplotype			
Disease Status	**AC**	**AT**	**GC**	**GT**
Non-Anemic	27	8	0	0
Mild Malaria Anemia	76	26	0	0
Moderate Malaria Anemia	54	15	1	1
Severe Malaria Anemia	61	17	1	0

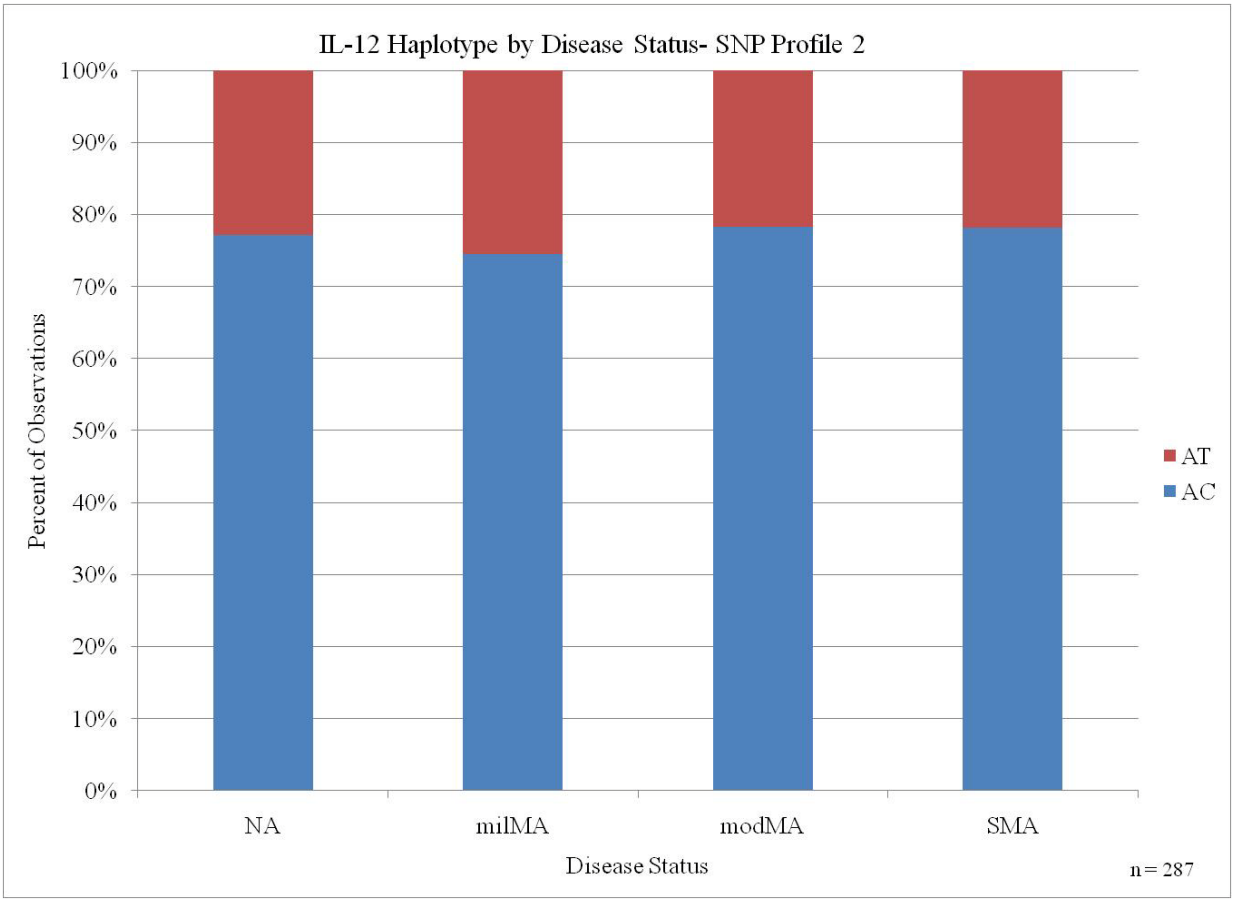

Figure 12. IL-12 SNP Profile 2 Haplotypes

Table 43. IL-12Rβ1 Haplotypes for SNP Profile 1

	Haplotype			
Disease Status	**AC**	**AT**	**GC**	**GT**
Non-Anemic	0	5	4	4
Mild Malaria Anemia	4	16	5	18
Moderate Malaria Anemia	2	9	8	7
Severe Malaria Anemia	2	13	5	9

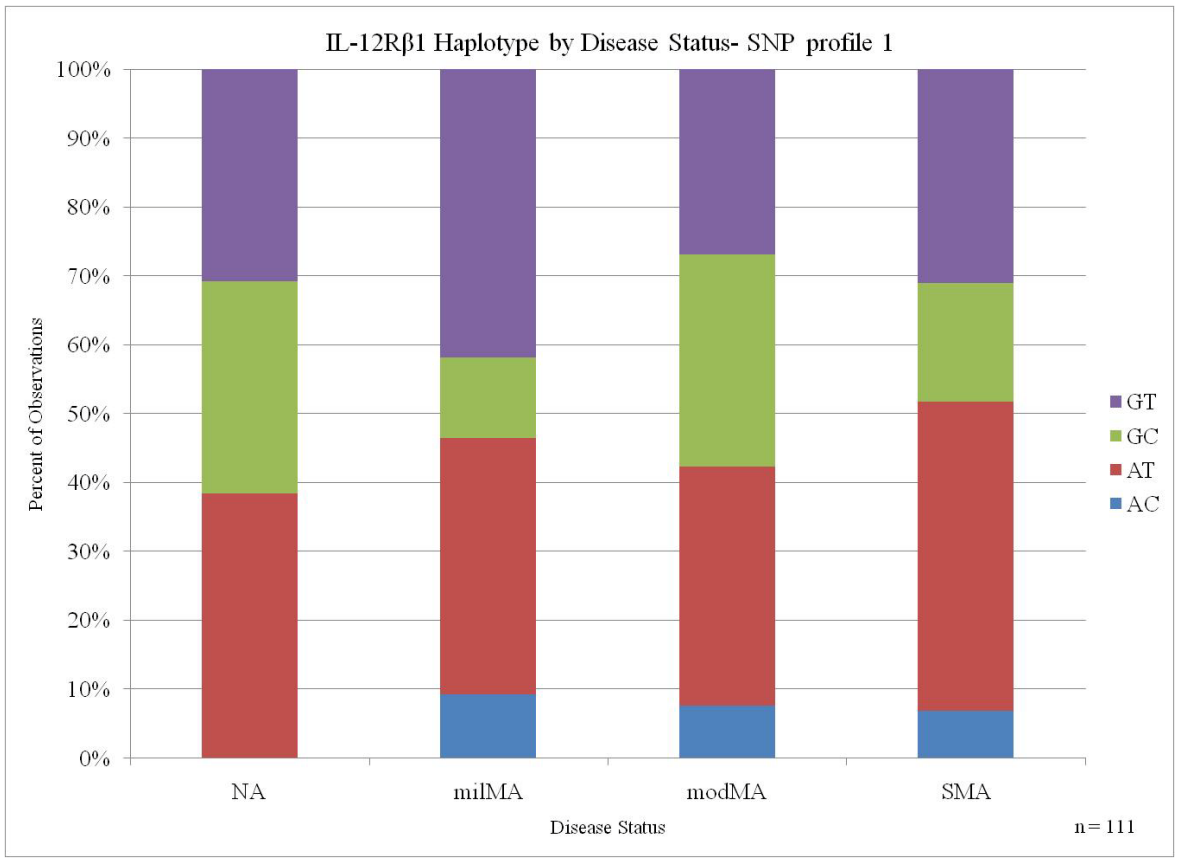

Figure 13. IL-12Rβ1 SNP Profile 1 Haplotypes

Table 44. IL-12Rβ1 Haplotypes for SNP Profile 1

	Haplotype			
Disease Status	**AC**	**AT**	**GC**	**GT**
Non-Anemic	0	2	2	9
Mild Malaria Anemia	1	15	4	23
Moderate Malaria Anemia	1	10	6	9
Severe Malaria Anemia	0	6	4	19

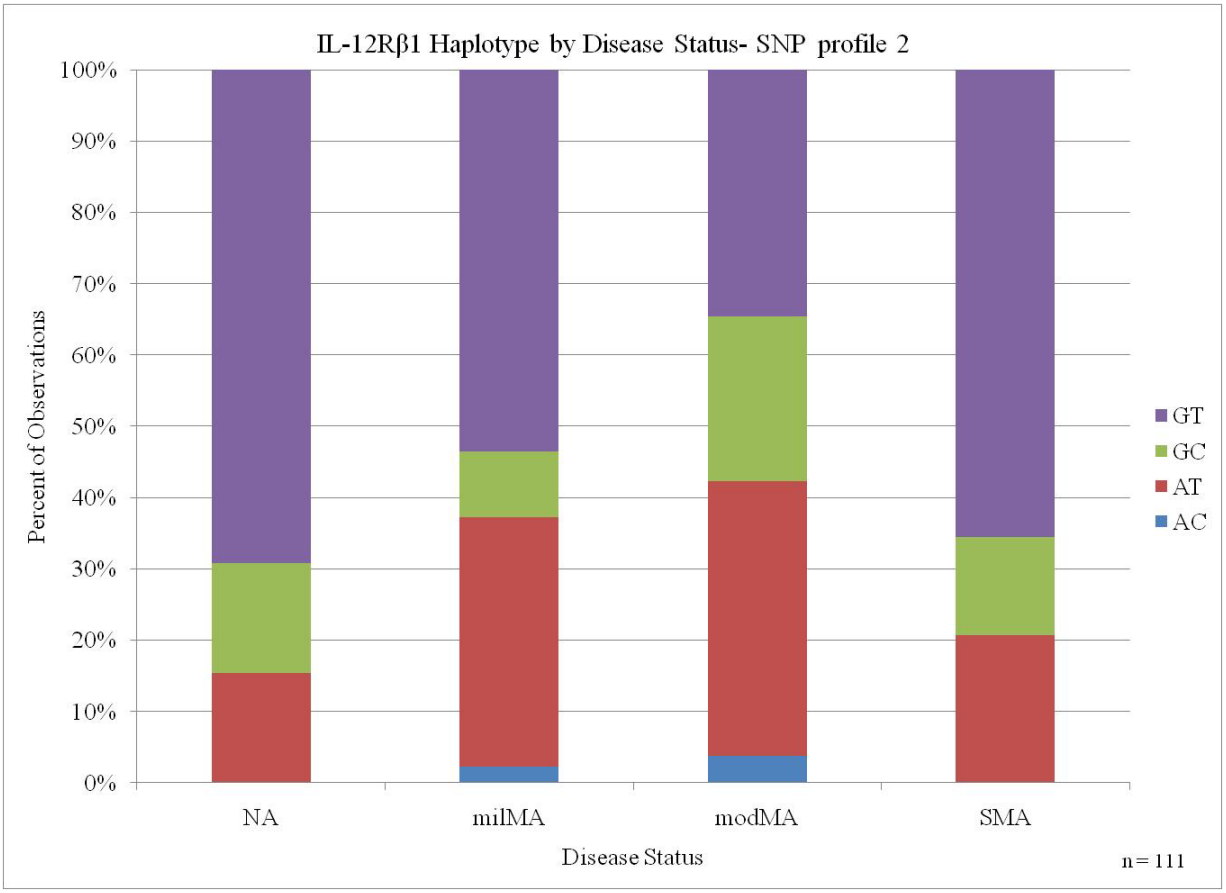

Figure 14. IL-12Rβ1 SNP Profile 2 Haplotypes

4.4 DISTRIBUTION OF SNPS ACROSS DIFFERENT GEOGRAPHICAL REGIONS

We investigated the distribution of SNPs rs2243113, rs2243140, rs383483 and rs429774 in different populations with varying malaria transmission rates. In addition to our Kenyan pediatric population, we also looked at the distribution of these SNPs in other populations. Our additional populations include samples from Bogoin Village, Central African Republic, the Kavieng and Madang regions of Papua New Guinea, and the Betsileo, Merina, Bezanozano and Sihanaka tribes of Madagascar. Furthermore, we also took information from the International HapMap Project. The populations used from HapMap are the Yoruba in Ibadan, Nigeria (YRI), the Japanese in Tokyo, Japan (JPT), the Han Chinese in Beijing, China (CHB), and the CEPH Collection [(Utah Residents with Northern and Western European Ancestry)CEU][26]. Data for the HapMap populations are not available for every SNP; therefore the HapMap populations will not be the same for every SNP. Data for rs2243113 does not have all of the populations due toissues with genotyping.

These populations vary in the rate of malaria transmission. With the exceptions of the JPT, CHB, CEPH, Merina and Betsileo populations, all of the populations are found in areas where malaria is endemic and transmission is stable (Figure 15). For the JPT population and Merina populations, malaria is not a risk. The malaria risk for the Betsileo population ranges

from no transmission to marginal transmission. We genotyped DNA samples and used the R version of SNP-HWE to determine if a significant difference exist between our genotyped data and what would be expected under the Hardy-Weinberg equilibrium. The p-value for each population for each SNP can be seen in Tables 45 – 48.

Country	Population	Stable	Marginal	No Risk
Kenya	W. Kenya			
Central African Republic	Bogoin Village			
Papua New Guinea	Kavieng			
	Madang			
Madagascar	Betsileo			
	Merina			
	Bezanozano			
	Sihanaka			
HapMap	YRI			
	JPT			
	CHB			
	CEU			
	pgaAA			
	pgaEUR			

Figure 15. Transmission rates for each population are indicated by a red block.

Table 45. Calculated p-values by Population – rs2243113

p-values by Population – rs2243113	
Population	**p-value**
Kenya	6.17e −28
Betsileo	0.048

Table 45: Continued

Population	p-value
Merina	0.269
Kavieng	0.0006
Madang	0.0006
YRI	0
AA	5.2e −13

Table 46. Calculated p-values by Population – rs2243140

p-values by Population – rs2243140	
Population	**p-value**
Kenya	3.83e −18
Betsileo	0.013
Merina	1.03e −8
Bezanozano	0.091
Sihanaka	0.014
Bogoin Village	0.001
Kavieng	2.32e −6
Madang	1.24e −14
HCB	7.00e −5
JPT	4.31e −7
YRI	6.62e −5

Table 46: Continued

Population	p-value
AA	0.100
EURO	0.062

Table 47. Calculated p-values by Population – rs383483

p-values by Population – rs383483	
Population	**p-value**
Kenya	0.039
Betsileo	0.616
Merina	0.0002
Bezanozano	1
Sihanaka	1
Bogoin Village	6.15e −9
Kavieng	2.32e −6
Madang	1.24e −14

Table 48. Calculated p-values by Population – rs429774

p-values by Population – rs429774	
Population	**p-value**
Kenya	3.79e −9
Betsileo	0.543

Table 48: Continued

Population	p-value
Merina	1.57e −5
Bezanozano	1
Sihanaka	0.009
Bogoin Village	3.28e −12
Kavieng	0.137
Madang	0.051
CEU	7.06e −9
HCB	2.83e −6
JPT	0
YRI	4.63e −16

For rs2243113, which has been associated with a susceptibility to malaria, the Merina population is the only population in which there is not a deviation from the Hardy-Weinberg Equilibrium (HWE). The Merina population is the only population that does not have a risk for malaria transmission, which suggests that there may be some selection for rs2243113 in the populations that are at risk for malaria. For rs2243113, all the populations excluding the Kenya population have a higher genotype frequency for the 'AA' genotype which is the homozygous wildtype genotype (Figure 16). For rs2243140, a deviation from HWE was found in all populations except the Bezanozano, AA and EURO. The deviation from HWE cannot be

explained by rates of transmission because the deviation exists in all transmissions rates. When compared to the Kenyan population all of the additional populations have a higher frequency of the ‘TT’ genotype which is the homozygous wildtype genotype (Figure 17). For rs383483, which has been associated with a protection against SMA and HDP, a deviation from HWE exists in all populations except for the Betsileo, Bezanozano and Sihananka. Similar to rs2243140, this deviation cannot be attributed to transmission rate because it also occurs in areas of varying transmission rate. The populations that did not have the deviations were in areas of no, marginal and stable transmission rates. When comparing the Kenyan population to the other populations, there is a higher frequency of the heterozygous genotype in all populations except for the Kenyan population. For rs429774, which has been associated with a protection against SMA, there is a deviation from HWE in all populations except the Betsileo, Bezanozano, Kavieng, Madang and JPT. Again, this deviation or lack of deviation can not be attributed to the rates of malaria transmission because of the occurrence in all levels of transmission. Unlike the previous three SNPs, a general trend does not exist for the additional populations when compared to the Kenyan population. However, when broken down by country, a trend does exist. For the Kavieng and Madang populations, which are both in Papua New Guinea, there appears to be an almost equal distribution of the genotypes. For the four tribes of Madagascar; the Betsileo, The Bezanozano, the Sihanaka and the Merina there is a higher frequency of the ‘TT’ genotype, which is the homozygous mutant genotype.

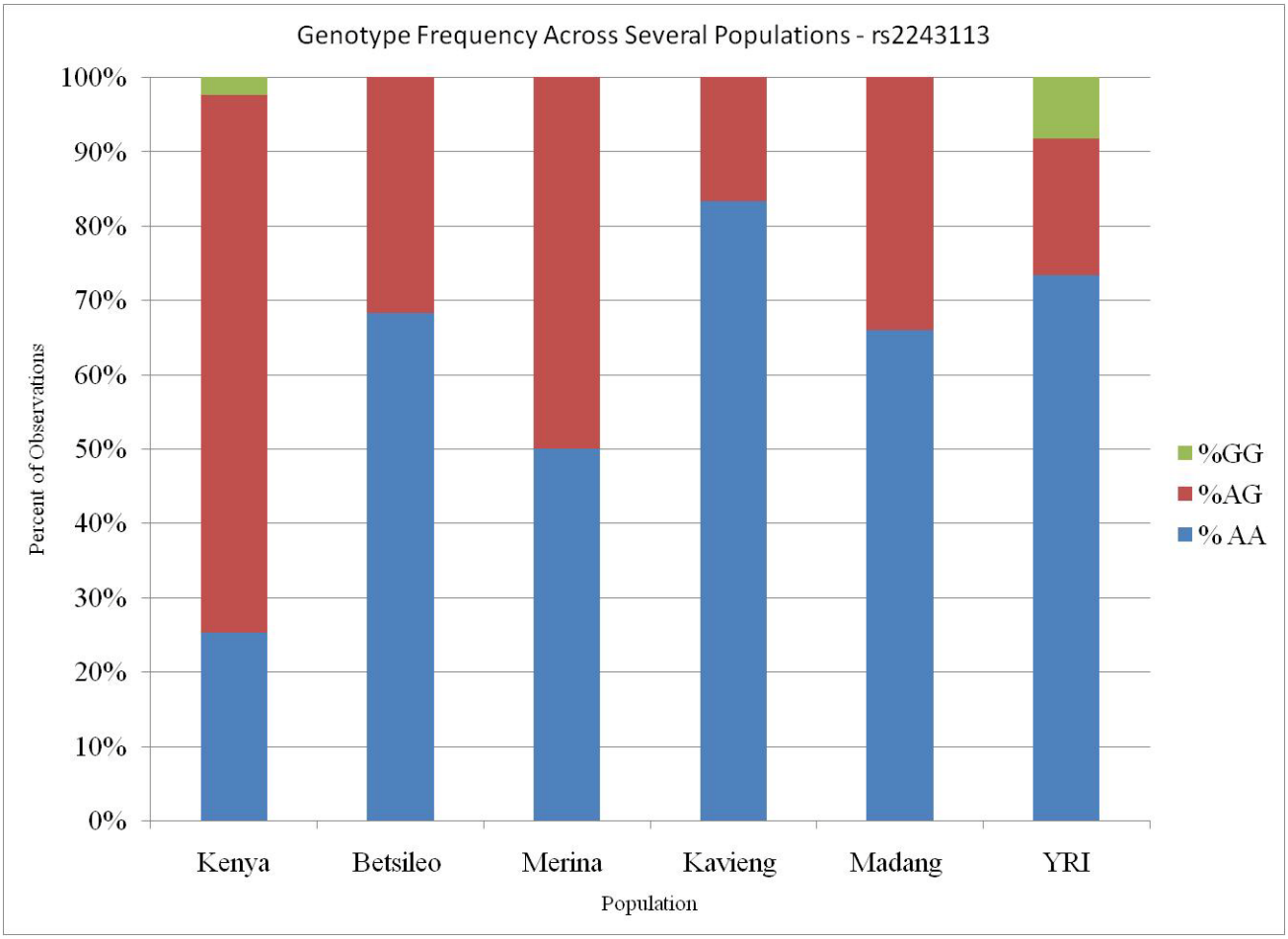

Figure 16. Distribution of rs2243113 Across Different Populations

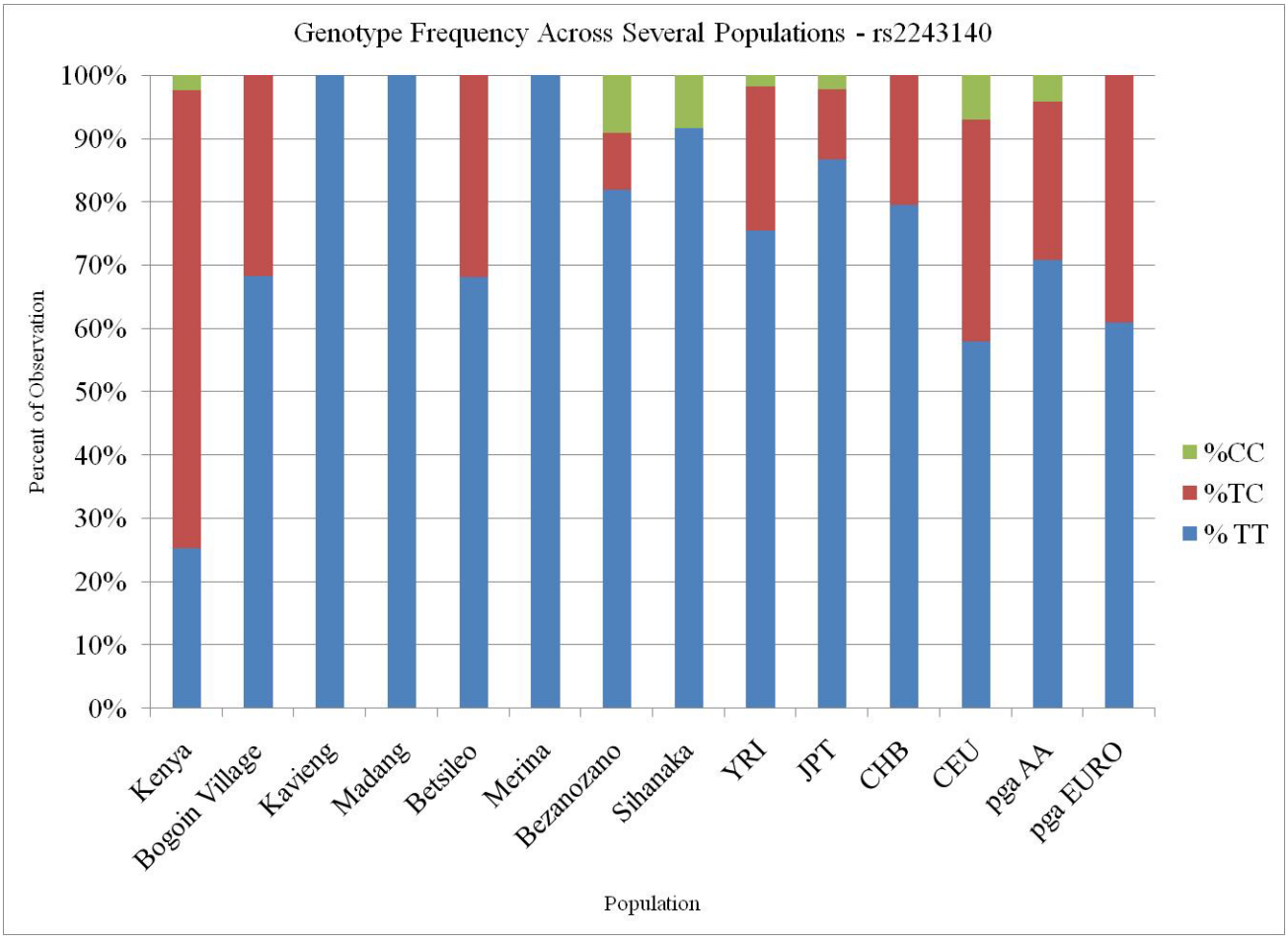

Figure 17. Distribution of rs2243140 Across Different Populations

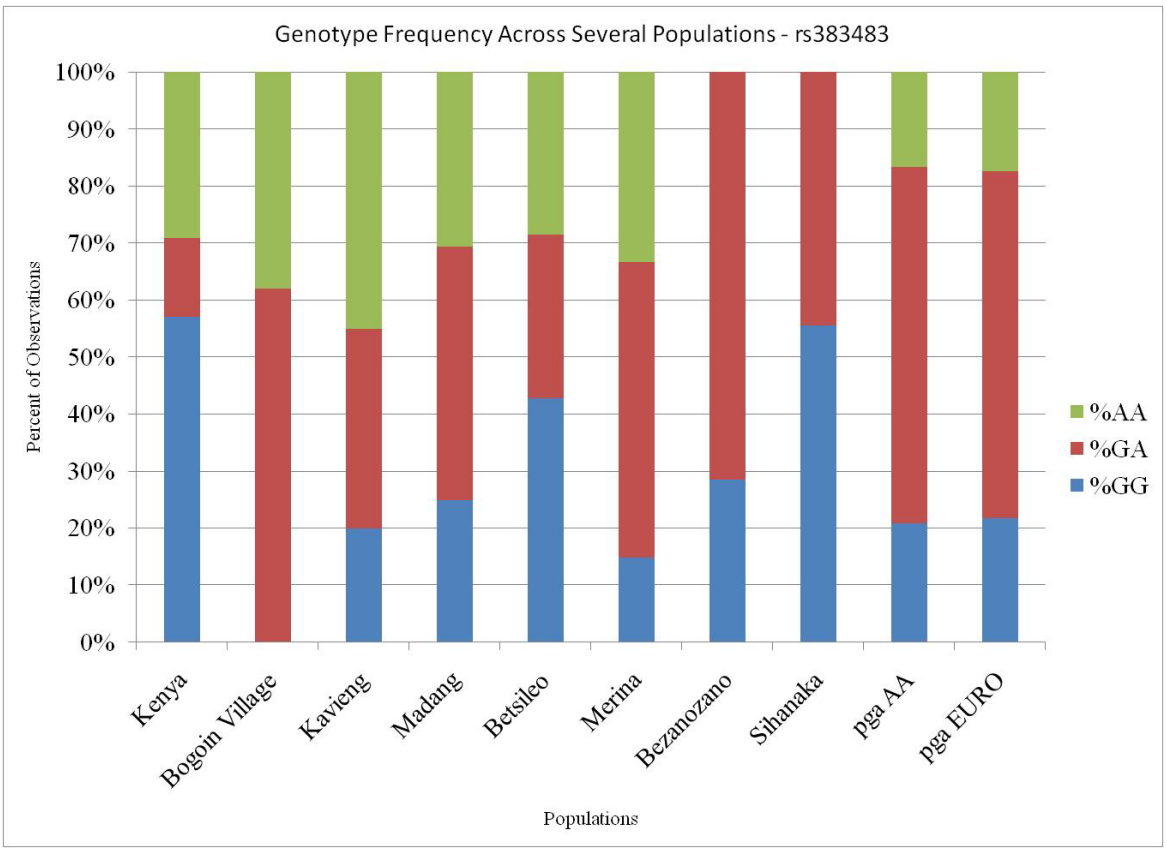

Figure 18. Distribution of rs383483 Across Different Populations

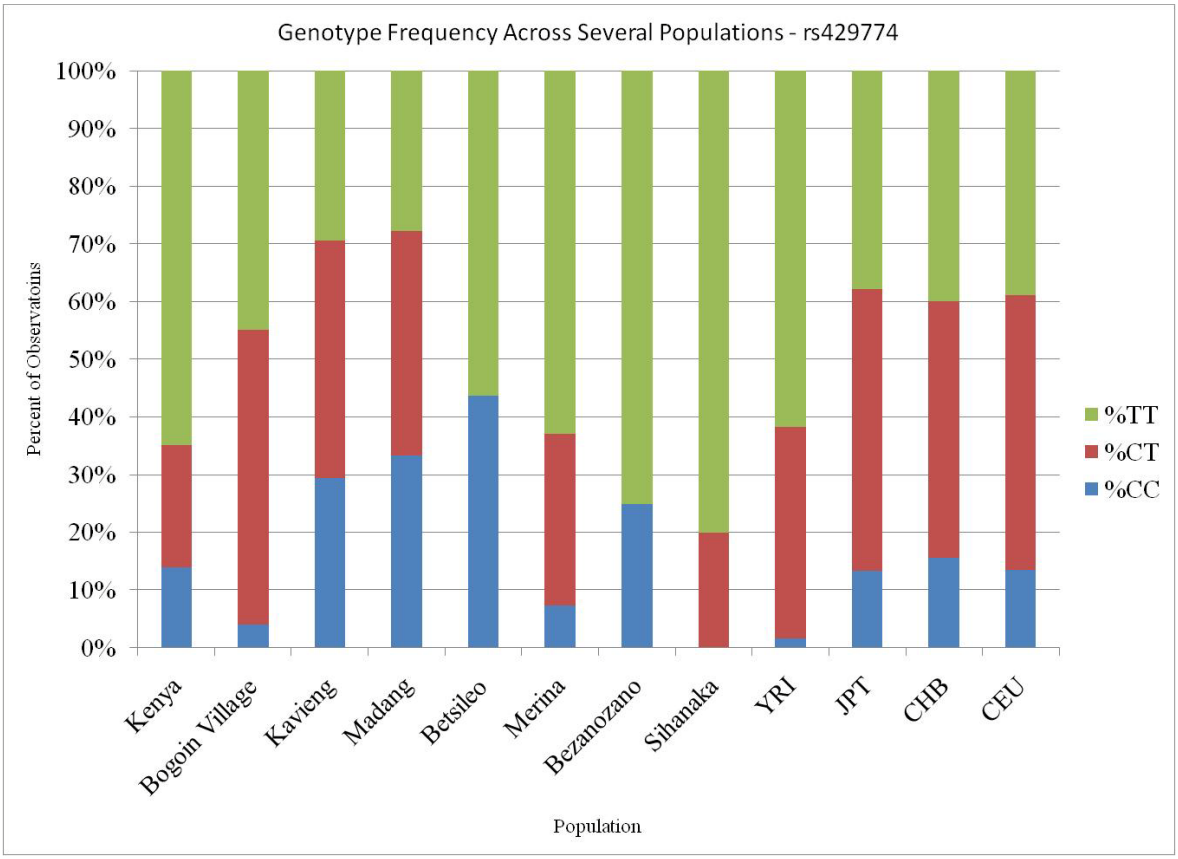

Figure 19. Distribution of rs429774 Across Different Populations

5.0 DISCUSSION

In the process of genotyping, significant difficulties were experienced; in some cases we lost more than 50% of our samples due to the inability to genotype. There are several possible explanations to account for this. To ensure that the fluorescent polarization (FP) assay was not the issue, we tried the TaqMan assay, which yielded either the same results or in most cases worse results (data not shown). The FP assay results can only be as good as the starting material. A lot of our samples were taken from children under the age of 3 who had malaria. Because of the complication of anemia that results from malaria, a small blood sample was taken, from which DNA was extracted. Due to the initial amount of extracted genomic DNA being so small, all of our samples are Wholoe-Genome Amplified. Because our samples have been stored for over 5 years, it is also possible that the DNA concentration may have altered due to evaporation. Unidentified SNPs located in the primer region or in the stretch of DNA that is being amplified for genotyping could also account for failed or ambiguous samples. If there is an unaccounted for SNP in a primer site (SNP-in-primer) and it has a high minor allele frequency, and it is in linkage disequilibrium with the target SNP or has an effect on the hybridization of the primers than the genotyping failure rate could increase [35].

Copy Number Variation (CNV) is a phenomenon that happens in chromosomal rearrangement and results in a deletion or duplication of stretches of DNA sequences. There are

several ways to check or account for these factors. One would be to sequence the samples to see if there are any additional mutations. A second method would be to perform an assay that checks for CNV such as multiplex amplifiable probe hybridization (MAPH) or multiplex ligation-dependent probe amplification (MLPA). It would also be plausible to compile a new set of samples and genotype the samples to see if the results are the same.

Geographically, we observed a higher frequency of the homozygous mutant genotype in all populations except the Kenyan population for SNPs rs2243113 and rs2243140. Our Kenyan population is made up of children whereas the other populations were all comprised of adults. Severe disease complications and death from malaria occur mostly in children. It is a possibility that in our additional populations we see more of the homozygous wildtype genotype for rs2243113 and rs2243140 because of selection bias. Specifically, the individuals chosen from the non-Kenyan populations were adults and it is possible that they survived the increased risks that are present during childhood. To control for this bias, a good study would be to sample a pediatric population and adult population from the same area of Kenya and see if the results hold true for the adult population.

Although it is known that these SNPs do not alter the amino acids in the protein products, it is unknown what effect these SNPs could be having on IL-12 or IL-12R. Gene expression assays could be done to assess whether there is an increase or decrease in gene expression. For SMA, it is possible that an increase in the expression of IL-12 could be the cause. For every infected red blood cell (iRBC) that is destroyed, ten uninfected RBCs are removed from circulation, a process in which macrophages play an important role [15]. If the levels of IL-12 are increased, then the amount of IFN-γ induced by IL-12 will also increase. An increased amount of IFN-γ will prime that much more monocytes thereby increasing the levels of

macrophages which help to contribute to the removal of uninfected RBCs. It has also been shown that IFN-γ plays a role in the mechanisms of bone marrow depression and dyserythropoiesis [36]. An increase in the IL-12 receptor would also have this same effect because it would work by the same mechanism of increasing the amount of IFN-γ and monocytes. A decrease in IL-12 could be the cause of HDP. If the expression of IL-12 is decreased, then the immune response to the malaria infection would be less than optimal. Decreased levels of IL-12 would also result in decreased levels of T-cells, NK-cells, macrophages and dendritic cells. This would result in a decreased removal of iRBCs and of parasites. A decrease in the IL-12 receptor would also display the same results. If there is a decrease in the receptor, the response to IL-12 would also decrease thereby creating the same less than optimal immune response.

The purpose of this study was to determine if there was a correlation between the disease status of either high density parasitemia (HDP) and/or severe malaria anemia (SMA) and SNPs rs2243113, 2243140 rs33483 and rs429774. We used the Hardy-Weinberg Equilibrium and either the Chi-Squared Test statistic or the R version of SNP-HWE to determine if the genotypic makeup of our population was what would be expected under the Hardy-Weinberg Principle. For all SNPs we found that there is a significant difference between our observed and expected values for the overall population. In addition, when the population was broken down by disease status, we found that a significant difference existed for our observed and expected values for all disease statuses. There was no difference between disease groups, which led us to conclude that SNPs rs2243113 and rs2243140 do not display a susceptibility to SMA. Likewise, we concluded that rs383483 does not display a protection against SMA or HDP and rs429774 does not display a protection against SMA. We also created haplotypes, which are combinations of SNP alleles on

a particular gene. However, our haplotypes did not increase our information because there was not any haplotype that was more prevalent in either the less severe disease complications as opposed to the more severe disease complications and vice versa. We found that for all four SNPs -rs2243113, rs224310, and rs429774 - there was not a correlation between disease status and genotype. We also investigated the distribution of these four SNPs across populations with varying malaria endemicity. We also found that in all of the populations except for the Kenyan population there is a higher frequency of homozygous wildtype alleles for both rs2243113 and rs2243140 and a higher frequency of heterozygotes for rs383483.

APPENDIX

GENOTYPE DISTRIBUTION ACROSS DIFFERENT POPULATIONS

The following charts are the 'percent of observations' by genotype for each population. It should be noted that the 'n' category in these charts does not represent the total sample size; it only represents the number of samples that were genotyped and able to be called the respective genotype.

Table 49. Distribution of rs2243113 Across Several Populations

		Genotype Frequency (%)		
Population	n	wt/wt (AA)	wt/mt (AG)	mt/mt (GG)
Western Kenya	294	32.3	65.6	2.0
Betsileo	22	68.18	31.18	0
Merina	22	50	50	0
Kavieng	24	83.33	16.67	0
Madang	47	65.96	24.04	0
YRI	60	73.3	18.3	8.3
AA	60	93.5	6.5	0

Table 50. Distribution of rs2243140 Across Several Populations

		Genotype Frequency (%)		
Population	n	wt/wt (TT)	wt/mt (TC)	mt/mt (CC)
Western Kenya	305	41.3	12.5	46.2
Betsileo	31	74.19	12.90	12.90
Merina	30	100	0	0
Bezanozano	11	81.82	9.09	9.09
Sihanaka	12	91.67	0	8.33
Bogoin Village	41	68.29	31.71	0
Kavieng	22	100	0	0
Madang	50	100	0	0
HCB	44	79.5	20.5	0
JPT	45	86.7	11.1	2.2
YRI	57	75.4	22.8	1.8
AA	24	70.8	25	4.2
EURO	23	60.9	39.1	0

Table 51. Distribution of rs383483 Across Several Populations

		Genotype Frequency (%)		
Population	n	wt/wt (GG)	wt/mt (GA)	mt/mt (AA)
Western Kenya	168	57.1	13.7	29.2

Table 51: Continued

		Genotype Frequency (%)		
Population	n	wt/wt (GG)	wt/mt (GA)	mt/mt (AA)
Betsileo	14	42.86	28.57	28.57
Merina	27	14.81	51.85	33.33
Bezanozano	7	28.57	71.43	0
Sihanaka	9	55.56	44.44	0
Bogoin Village	29	0	62.07	37.93
Kavieng	20	20	35	45
Madang	50	100	0	0

Table 52. Distribution of rs429774 Across Several Populations

		Genotype Frequency (%)		
Population	n	wt/wt (CC)	wt/mt (CT)	mt/mt (TT)
Western Kenya	394	14	21.1	65
Betsileo	16	43.75	0	56.25
Merina	27	7.41	29.63	62.96
Bezanozano	8	25	0	75
Sihanaka	10	0	20	80
Bogoin	49	4.08	51.02	44.90
Kavieng	17	29.41	41.18	29.41
Madang	36	33.33	38.89	27.78

Table 52: Continued

		Genotype Frequency (%)		
Population	n	wt/wt (CC)	wt/mt (CT)	mt/mt (TT)
CEU	60	13.6	47.5	39
HCB	45	15.6	44.4	40
JPT	45	13.3	48.9	37.8
YRI	60	1.7	36.7	61.7

BIBLIOGRAPHY

1. Teklehaimanot, A. and P. Mejia, *Malaria and poverty.* Ann N Y Acad Sci, 2008. **1136**: p. 32-7.
2. Breman, J.G., M.S. Alilio, and A. Mills, *Conquering the intolerable burden of malaria: what's new, what's needed: a summary.* Am J Trop Med Hyg, 2004. **71**(2 Suppl): p. 1-15.
3. Malaguarnera, L., et al., *Increased levels of interleukin-12 in Plasmodium falciparum malaria: correlation with the severity of disease.* Parasite Immunol, 2002. **24**(7): p. 387-9.
4. Mockenhaupt, F.P., et al., *Manifestation and outcome of severe malaria in children in northern Ghana.* Am J Trop Med Hyg, 2004. **71**(2): p. 167-72.
5. Marsh, K., et al., *Indicators of life-threatening malaria in African children.* N Engl J Med, 1995. **332**(21): p. 1399-404.
6. Ridley, R.G., *Medical need, scientific opportunity and the drive for antimalarial drugs.* Nature, 2002. **415**(6872): p. 686-93.
7. Pongtavornpinyo, W., et al., *Spread of anti-malarial drug resistance: mathematical model with implications for ACT drug policies.* Malar J, 2008. **7**: p. 229.
8. Luty, A.J., et al., *Interferon-gamma responses are associated with resistance to reinfection with Plasmodium falciparum in young African children.* J Infect Dis, 1999. **179**(4): p. 980-8.
9. Marquet, S., et al., *A functional promoter variant in IL12B predisposes to cerebral malaria.* Hum Mol Genet, 2008. **17**(14): p. 2190-5.
10. McGuire, W., et al., *Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria.* Nature, 1994. **371**(6497): p. 508-10.
11. Stevenson, M.M. and E.M. Riley, *Innate immunity to malaria.* Nat Rev Immunol, 2004. **4**(3): p. 169-80.
12. Tuteja, R., *Malaria - an overview.* FEBS J, 2007. **274**(18): p. 4670-9.
13. Kyes, S., P. Horrocks, and C. Newbold, *Antigenic variation at the infected red cell surface in malaria.* Annu Rev Microbiol, 2001. **55**: p. 673-707.
14. Beeson, J.G., et al., *Plasmodium falciparum isolates from infected pregnant women and children are associated with distinct adhesive and antigenic properties.* J Infect Dis, 1999. **180**(2): p. 464-72.
15. Lamikanra, A.A., et al., *Malarial anemia: of mice and men.* Blood, 2007. **110**(1): p. 18-28.
16. Keller, C.C., et al., *Acquisition of hemozoin by monocytes down-regulates interleukin-12 p40 (IL-12p40) transcripts and circulating IL-12p70 through an IL-10-dependent mechanism: in vivo and in vitro findings in severe malarial anemia.* Infect Immun, 2006. **74**(9): p. 5249-60.

17. Charles A. Janeway, J., *Activated macrophages secrete a range of cytokines that hafve a variety of local and distant effects.*, in *Immunobiology: The Immune System in Health and Disease*. E. Lawrence, Editor. 2005, Garland Science Publishing: New York. p. 76.
18. Trinchieri, G., *Interleukin-12 and the regulation of innate resistance and adaptive immunity.* Nat Rev Immunol, 2003. **3**(2): p. 133-46.
19. Stern, A.S., J. Magram, and D.H. Presky, *Interleukin-12 an integral cytokine in the immune response.* Life Sci, 1996. **58**(8): p. 639-54.
20. Watford, W.T., et al., *The biology of IL-12: coordinating innate and adaptive immune responses.* Cytokine Growth Factor Rev, 2003. **14**(5): p. 361-8.
21. Trinchieri, G., *Interleukin-12: a proinflammatory cytokine with immunoregulatory functions that bridge innate resistance and antigen-specific adaptive immunity.* Annu Rev Immunol, 1995. **13**: p. 251-76.
22. Holscher, C., *The power of combinatorial immunology: IL-12 and IL-12-related dimeric cytokines in infectious diseases.* Med Microbiol Immunol, 2004. **193**(1): p. 1-17.
23. Romagnani, S., *T-cell subsets (Th1 versus Th2).* Ann Allergy Asthma Immunol, 2000. **85**(1): p. 9-18; quiz 18, 21.
24. Szabo, S.J., et al., *Regulation of the interleukin (IL)-12R beta 2 subunit expression in developing T helper 1 (Th1) and Th2 cells.* J Exp Med, 1997. **185**(5): p. 817-24.
25. Abbas, A.K., K.M. Murphy, and A. Sher, *Functional diversity of helper T lymphocytes.* Nature, 1996. **383**(6603): p. 787-93.
26. *The International HapMap Project.* Nature, 2003. **426**(6968): p. 789-96.
27. Kwok, P.-Y., *SNP Genotyping With Flourescent Polarization Detection.* Human Mutation, 2002(19): p. 315-323.
28. Kwok, P.Y., *High-throughput genotyping with primer extension fluorescent polarization detection.* Curr Protoc Hum Genet, 2004. **Chapter 2**: p. Unit 2 11.
29. Chen, X., L. Levine, and P.Y. Kwok, *Fluorescence polarization in homogeneous nucleic acid analysis.* Genome Res, 1999. **9**(5): p. 492-8.
30. Kramer, M.F. and D.M. Coen, *The polymerase chain reaction.* Curr Protoc Protein Sci, 2002. **Appendix 4**: p. Appendix 4J.
31. Beverly C. Delidow, J.P.L., John J. Peluso, Bruce A. White, *Polymerase Chain Reaction*, in *PCR Protocols: Current Methods and Applications*, B.A. White, Editor. 1993, Humana Press Inc: Totowa, NJ. p. 1-29.
32. Hartl, D.L., *A Primer Of Population Genetics*. Third ed. 2000, Sunderland: Sinauer Associates, Inc.
33. Wigginton, J.E., D.J. Cutler, and G.R. Abecasis, *A note on exact tests of Hardy-Weinberg equilibrium.* Am J Hum Genet, 2005. **76**(5): p. 887-93.
34. Stephens, M., N.J. Smith, and P. Donnelly, *A new statistical method for haplotype reconstruction from population data.* Am J Hum Genet, 2001. **68**(4): p. 978-89.
35. Koboldt, D.C., R.D. Miller, and P.Y. Kwok, *Distribution of human SNPs and its effect on high-throughput genotyping.* Hum Mutat, 2006. **27**(3): p. 249-54.
36. Menendez, C., A.F. Fleming, and P.L. Alonso, *Malaria-related anaemia.* Parasitol Today, 2000. **16**(11): p. 469-76.