ASSOCIATION OF GENETIC VARIANTS IN *IL7* AND *CSF2* WITH *PLASMODIUM FALCIPARUM* MALARIA OUTCOMES IN CHILDREN ATTENDING SIAYA COUNTY REFERAL HOSPITAL

BY

LILY KISIA ELAZIA

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN MEDICAL IMMUNOLOGY

SCHOOL OF PUBLIC HEALTH AND COMMUNITY DEVELOPMENT

MASENO UNIVERSITY

DECLARATION

This thesis is my original work and has not been presented for an award of a degree in any other University

Signature…… Date……

Lily Kisia Elazia, MSc
(PHD/PH/00014/2014)

DECLARATION BY SUPERVISORS

This thesis has been submitted for examination with our approval as University Supervisors

Signature…… Date……

Kiprotich Chelimo, PhD
Department of Biomedical Sciences and Technology,
School of Public Health and Community Development,
Maseno University, Maseno
Kenya

Signature…… Date……

John M. Ong'echa, PhD
Center for Global Health Research,
Kenya Medical Research Institute, Kisumu
Kenya

ACKNOWLEDGEMENT

This work was conducted as part of larger investigations on immunogenetics of severe malarial anemia by the University of New Mexico (UNM)/Kenya Medical Research Institute (KEMRI)/Maseno University (MU). My utmost appreciation is directed to Prof. Douglas J. Perkins (Director, Global Health and Research, UNM, USA) for enabling me to work in his laboratories, both in Kenya (KEMRI-Kisumu, Siaya County and Referral Hospital and MU) and the USA (UNM-Albuquerque), where he provided a conducive environment for me to carry out the assays for this thesis. I am grateful to my thesis supervisors -Dr. Kiprotich Chelimo (MU) and Dr. Michael O. Ong'echa (Center for Global Health and Research-KEMRI, Kisumu) for the guidance they provided in all the stages of this PhD course. I thank my mentors Prof. Collins Ouma (MU), Dr. Evans Raballah (Masinde Muliro University of Science and Technology), Dr. Samuel Anyona (MU) and Dr. Prakasha Kempaiah (UNM, USA) for their contribution towards the development of this thesis. I wish to sincerely thank all the staff that were involved with this work for their cooperation and support. I appreciate the parents/guardians of the study participants and the children that participated in this study for providing the necessary samples for the experimental analysis.

DEDICATION

I dedicate this work to my beloved parents, all my siblings, my good friends, my greatest friend Dianne Hart Lee, and last but not least my two beloved children and their father.

ABSTRACT

Immune naïve children residing in endemic *Plasmodium falciparum* (*P. falciparum*) transmission regions experience repeated episodes of malaria infection that may lead to bone marrow suppression and subsequently severe malaria anemia [SMA, hemoglobin (Hb)<5g/dL, with any density parasitemia] that is responsible for the high morbidity and mortality in these regions. Candidate gene approach studies have shown that genetic variation in cytokine genes play a key role in influencing malaria infection outcomes by altering production of immune inflammatory mediators. The current study determined the association of genetic variants in *IL7* (72194T>C and -2440A>G) and *CSF2* (-7032G>A, and 64544T>C) with *P. falciparum* malaria infection outcomes [susceptibility to inefficient erythropoiesis (IE) and SMA; rate of malaria infections and SMA episodes; and risk of all-cause mortality] in falciparum parasitemic children (aged <36 months) attending Siaya County Referral Hospital (SCRH). The study designs were cross-sectional case control to achieve objective 1 (n=883) and longitudinal case control to achieve objectives 2-4 (n=1654). All the children (n=2537) that participated in the current study had been enrolled in a parent study (2003-2012) that evaluated genetic basis of SMA in children attending SCRH. Genomic DNA was extracted from buccal swabs using the Gentra System and genotyped using a TaqMan 5' allelic discrimination Assay. Cross-sectional data was analyzed using SPSS (Version 24.0) while longitudinal data was analyzed using R (version 3.1.4). The Kolmogorov–Smirnov test was used to determine normality of continuous variables. For normally distributed variables, ANOVA and Student's t-test were used to compare means. If normality was violated Mann-Whitney-*U* test was used for pairwise comparisons of medians. Chi square (χ^2) and Fisher's exact test were used to find differences in proportions of genetic variants. Binomial logistic regression was used to determine associations of genetic variants in *IL7* with susceptibility to IE and SMA. Poisson regression was used to determine the influence of genetic variants in *CSF2* on the rate of in malaria infections and SMA episodes. Cox proportional hazard model was used to determine the relationship between genetic variants in *CSF2* and the risk of malaria infections, SMA episodes and all-cause mortality. *P* values ≤0.05 were considered significant. The results showed that carriage of *IL7*72194 TC genotype was associated with enhanced susceptibility to IE (Odds ratio (OR) =1.90; 95% (confidence interval (CI) 1.09–3.30; *P*=0.02) as was homozygous CC genotype (OR 5.14; 95%CI=1.20–21.99; *P*=0.03). Individuals with the *IL7* CA haplotype had an increased risk of IE (OR=1.90; 95%CI=1.10–3.30; *P*=0.02), whereas TA haplotype carriers were protected against IE (OR=0.24; 95%CI=0.06–1.21; *P*=0.05). Further, a decreased incidence rate ratio (IRR) for malaria was conferred by inheritance of the *CSF2*64544 TC genotype (*P*=0.02) and *CSF2* AC/GC diplotype (*P*<0.01). An increased IRR for malaria was observed in carriers of the *CSF2* AT/GC diplotype (*P*=0.02), while the inheritance of the *CSF2* AT haplotype increased the IRR for SMA (*P*=0.01). Decreased hazard rates (HR) for malaria were observed among *CSF2* AC haplotype carriers (*P*<0.01) while inheritance of the *CSF2*-7032 GA genotype increased the HR for all-cause mortality (*P*=0.03). Cumulative results showed that variation in *IL7* gene is associated with erythropoietic responses in children with falciparum malaria while variation in *CSF2* gene influences susceptibility to malaria, SMA, and all-cause mortality, suggesting that genetic variations in *IL7* and *CSF2* are potential modulators and or predictors of malaria infection outcomes in young children residing in malaria endemic areas. The findings highlight novel immuno-genetic associations with malaria infection, IE, SMA and all-cause mortality that have potential to guide immunotherapy and vaccine design for severe malaria in young children.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

AIC	Akaike's information criterion
AIDS	Acquired immune deficiency syndrome
ANOVA	Analysis of variance
ARV	Antiretrovirals
CDC	Center for Disease Control
CI	Confidence interval
CBC	Complete blood count
CD	Cluster of differentiation
C/EBP	CCAAT/enhancer-binding protein beta
c-Ets	c-E-twenty-six transformation-specific-
AP-2	Activating enhancer binding protein
COX	Cyclooxygenase
CSF2	Colony stimulating factor 2
DNA	Deoxyribonucleic acid
Elk	Early gene expression factor
ER	Estrogen receptor
GATA	GATA binding protein
G-CSF	Granulocyte colony stimulating factor
GM-CSF	Granulocyte macrophage colony stimulating factor
G6PD	Glucose-6-phosphate dehydrogenase

GWAS	Genome wide association studies
Hb	Hemoglobin
Hb-AA	Normal hemoglobin
Hb-AS	Sickle cell trait
Hb-SS	Sickle cell disease
Hct	Hematocrit
HDP	High density parasitemia
HIV	Human immunodeficiency virus
HR	Hazard ratio
HWE	Hardy Weinberger equilibrium
IE	Inefficient erythropoiesis
IL-	Interleukin
IFN	Interferon gamma
IRF	Interferon regulatory factor
IRR	Incidence rate ratio
KEMRI	Kenya Medical Research Institute
LAIR	Leukocyte associated immunoglobulin like receptor
LD	Linkage disequilibrium
MAF	Minor allele frequency
MCHC	Mean corpuscular hemoglobin concentration
MCV	Mean corpuscular volume

MIP	Macrophage inflammatory protein
MOH	Ministry of Health
MPS	Malaria parasites
MU	Maseno University
NFAT	Nuclear factor of activated T cells
NF1-X3	Nuclear factor 1-X3
NO	Nitric oxide
OR	Odds ratio
PCR	Polymerase chain reaction
RANTES	Regulated upon activation, normal T cell expressed and secreted
RBC	Red blood cell
RDW	Red cell distribution width
RPI	Reticulocyte production index
SCGF	Stem cell growth factor
SCRH	Siaya County Referral Hospital
SDH	Siaya District Hospital
SMA	Severe malarial anemia
SNPs	Single nucleotide polymorphisms
TFBS	Transcription factor binding sites
TLR	Toll-like receptor
TNF	Tumor necrosis factor

UM	Uncomplicated malaria
UNM	University of New Mexico
USA	United States of America
USF	Upstream transcription factor
WBC	White blood cell
WHO	World Health Organization
ZEB	Zinc finger E box-binding homeobox 1

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1. Background Information

Malaria is the leading cause of morbidity and mortality worldwide with an estimated record of 247 million cases in 2021 *(WHO 2022).* A large proportion (95%) of these cases occurred in the African region, and over 99.8% of the total cases were attributed to *Plasmodium falciparum* (*P. falciparum*) infections *(WHO 2022).* Globally, there were ~619,000 malaria-related deaths reported, 76% of which were children (especially those <5 years of age) residing in the World Health Organization (WHO) African region *(WHO 2022).* Kenya experienced an 88% overall reduction in the prevalence of malaria infection since 1990 (21.2%) to 2015 (2.6%) as a result of increased optimized treatment and vector control *(Macharia et al. 2018).* However, 9 counties in Kenya have remained highly endemic for malaria, with 7 of them around Lake Victoria i.e. Migori, HomaBay, Kisumu, Siaya, Kakamega, Vihiga, and Busia, all having a malaria prevalence of ≥30% *(Macharia et al. 2018).*

Siaya County is one of the regions with the highest malaria endemicity in Kenya and more than 99% of all malaria infections are caused by *P. falciparum* with transmission modeling based on parasitic screening showing that the area has predicted rates of over 40% parasite prevalence year round *(Hollowell et al. 2023).* Slide positivity rates among children (0–14 years) visiting health clinics with a history of or documented fever were found to be 52.5% (2007 and 2012), giving an indication of high malaria incidence in the region *(Khagayi et al. 2017).* Although data indicates that major progress has been made in the region regarding reducing malaria mortality in children, malaria still remains the most prevalent cause of morbidity and mortality in children under five years of age in Siaya County (33.2%) followed

by acute respiratory infections (26.7%), HIV/AIDS (18.6%) and diarrheal disease with younger children bearing the greatest burden *(Amek et al. 2014; Desai et al. 2014; Hollowell et al. 2023).*

Children residing in endemic *P. falciparum* transmission regions such as Siaya County typically experience repeated episodes of malaria prior to developing naturally-acquired immunity predisposing them to malaria complications including bone marrow suppression and the resultant anemia *(Helleberg et al. 2005).* Severe malaria anemia [SMA, hemoglobin (Hb) <5.0 g/dL, with any density parasitemia] accounts for the greatest worldwide proportion of malaria-associated morbidity and mortality *(WHO 2022).* In Kenya, SMA is also the most common complication of severe malaria in children <15 years of age in a number of counties neighboring Siaya, i.e. Kakamega, Vihiga, Busia and Kisumu, with a prevalence of 45.2% and case fatality rates at 5.7% between 2014-2018, followed by severe respiratory distress at 19.7%, *(Akech et al. 2020).* Earlier studies in Siaya, had indicated that SMA had a 26% prevalence rate in children aged 1–5 years of age and is almost exclusively (i.e., 89% of the cases) in children <3 years of age *(Obonyo et al. 2007).* Severe malaria anemia contributed 52.7% of all malaria-related deaths with a case fatality rate of 22.8% and resulted in significant in-hospital morbidity and mortality in the Siaya community *(Obonyo et al. 2007).*

Although SMA etiology includes hemolysis and splenic clearance of infected and uninfected erythrocytes, abnormal erythropoiesis (ineffective erythropoiesis and dyserythropoiesis) has been observed in malaria patients and could potentially be instrumental in anemia development *(White 2018).* Ineffective erythropoiesis is described as an abnormal proportion of erythroblasts in the bone marrow in comparison to steady-state erythropoiesis where the normal ratio between erythroblasts is 1 for Pro-E; 2 for Baso-E1; 4 for Baso-E2; 8 for Poly-E and 16 for Ortho-E, leading to a decreased number of red cells while dyserythropoiesis refers

to qualitative abnormality of erythropoiesis with morphological defects of erythroblasts, which may finally be responsible for a decreased erythrocytes production *(Dumarchey et al. 2022).* Ineffective erythropoiesis is associated with an inhibition of the erythroid cell proliferation and cellular division which has been observed in children infected with *P. falciparum*, where there were decreased numbers of cells in G1 and a decreased S/G2 ratio occurred in Baso-E, an observation that was also reported in Poly-E, suggesting an accumulation in the G2 phase *(Wickramasinghe et al. 1982).* Dyserythropoiesis has also been seen in microscopic observations of bone marrow aspirates from patients infected with *P. falciparum* revealing erythroid hyperplasia and ultra-structural nuclear abnormalities such as multi-nuclearity, nuclear fragmentation, inter-nuclear bridges and irregular nuclear shapes *(Baro et al. 2017; Wickramasinghe et al. 1989).*

One of drivers of an abnormal erythropoiesis in malarial anemia is the altered innate immune inflammatory response to malaria antigens such as hemozoin/malaria pigment that is released with merozoites during rupture of parasitized-red blood cells *(Shio et al. 2010).* Hemozoin stimulates the secretion of inflammatory mediators such as cytokines, chemokines, and effector molecules such as nitric oxide (NO) by monocytes and macrophages which are altered in some patients and drive abnormal erythropoiesis *(Perkins et al. 2011).* For instance, tumor necrosis factor (TNF)-α and interferon (IFN)-γ appear to play a protective role during the early stages of malaria infection through their ability to stimulate monocyte/macrophage activation and aid in controlling parasitemia but their excessive release along with NO enhances malarial anemia by contributing to abnormal erythropoiesis and erythrophagocytosis *(Shio et al. 2010).* Tumor necrosis factor (TNF) -α significantly decreases human erythroid progenitor cell proliferation while IFN-γ inhibits the proliferation and the differentiation of erythroblasts *(Felli et al. 2005).* In general, several studies show

evidence that inflammatory cytokines are critical in initiating the process of abnormal erythropoiesis.

Candidate gene approach studies have shown that genetic variation in immune response genes (i.e., cytokines, chemokines, growth factors and effector molecules) alters the production of inflammatory mediators that influences the malaria infection outcomes including risk of malaria, SMA episodes and all case mortality *(Achieng et al. 2019; Ouma et al. 2012).* For instance, functional variation in stem cell growth factor (SCGF, a hematopoietic growth factors that promote erythroid and myeloid colony development) promoter is associated with elevated SCGF production, enhanced erythropoiesis, and protection against the development of SMA *(Ouma et al. 2010).* Additionally, genetic variation in the Fc gamma receptor (Fcγ) Fcγ11A, Fcγ111A and Fcγ111B conditions susceptibility to SMA whereas Fcγ11A and Fcγ111B predicts the risk of repeated malaria episodes, SMA and mortality through a mechanism that involves functional changes in IFN-γ production *(Munde et al. 2017;Ouma et al. 2012).* Furthermore, cyclooxygenase-2 *(COX-2)* promoter variants are associated with the longitudinal risk of malaria, SMA, and all-cause mortality among children *(Anyona et al. 2020).* These data among others shows that cytokine genes condition *P. falciparum* malaria infection outcomes.

Although the candidate gene approach which employs the targeted investigation of a particular gene/gene pathway can be a viable means of uncovering the molecular mechanisms that influence malaria infection outcomes, unbiased approaches, such as global genomic characterization can identify genetic variants associated with complex human phenotypes *(Kim et al. 2010).* Therefore, before the current study was conducted, a pilot Genome-wide Association Study (GWAS) was conducted to identify gene/gene pathways associated with SMA *(Achieng et al. 2019).* The current study hypothesized that variation in immune

response genes that influenced susceptibility to SMA during the GWAS would also influence other malaria infection outcomes including rate and risk of malaria episodes; erythropoiesis efficiency; and all-cause mortality through shared mechanisms that influence malaria disease progression. The GWAS *(Achieng et al. 2019)* was performed in children (n=144, aged 3–36 months attending the former Siaya District Hospital (SDH) selected from a parasitemic cohort (n=1,220) excluding children with covariates that could affect hemoglobin levels. Children were stratified into 'polarized extremes' of non-SMA [(Hb=8.0–10.9g/dL; avg. Hb=9.8), n=74] and SMA; [(Hb<5.0 g/dL; avg. Hb= 4.1), n=70]. Single nucleotide polymorphisms (SNPs) in immune response genes including leukocyte associated immunoglobulin like receptor (*LAIR*)-1 and 2; interleukin *(IL)-7;* and colony stimulating factor (*CSF)-2;* were identified that were associated with enhanced risk of SMA ($P \leq 0.05$). In the *IL7* haploblock, SNPs 72194T>C (rs2583759) and -2440A>G (rs7007634) were identified whereas 64544T>C (rs246835) was found in *CSF2* haploblock. One SNP *CSF2* 7032G>A (rs168681) was not identified from the Immunochip® but was selected from online databases (dbSNP; http://www.ncbi.nlm.nih.gov/SNP, HapMap; http://www.hapmap.org/) because it was a promoter SNP that fitted the SNP selection criteria of the current study.

Interleukin-(IL)-7 is a hematopoietic growth factor which modulates both T- and B-cell development and T-cell homeostasis *(Fry and Mackall 2002).* Interleukin-7 induces differentiation of pluripotent hematopoietic stem cells into lymphoid progenitor cells, promotes proliferation, differentiation, and survival of naïve and matures T cells, and stimulates the production of myelopoietic factors [e.g., IL-3 and granulocyte monocyte – colony stimulating factor (GM-CSF)]. A role for *IL7* in malaria is suggested by studies in murine models in which malarial anemia was associated with suppression of erythropoietic-related cytokines [i.e., granulocyte colony stimulating factor (G-CSF), GM-CSF, and IL-7], and elevation of IL-10 and TNF-α *(Xu et al. 2013).* Previous findings in human studies

showed that levels of circulating IL-7 are suppressed in children with SMA *(Ong'echa et al. 2011).* Since IL-7 induces erythropoiesis, reduced levels in children with SMA may contribute to delayed and/or inappropriate erythroid development *(Aiello et al. 2007).* Interleukin (IL)-7 also mediates cell-mediated immune responses to human pathogens, and induces tissue destruction in inflammatory diseases such as rheumatoid arthritis and multiple sclerosis *(Churchman et al. 2014; Kreft et al. 2012; Sieling et al. 1995).* Interleukin (IL)-7 induced tissue destruction may arise from its ability to promote T cell-dependent activation of monocyte/macrophages which, in turn, release a number of cytokines (i.e., IL-1, IL-3, IL-6, IL-8, GM-CSF, macrophage inflammatory protein (MIP)-1β, TNF-α and IFN-γ) that exacerbate the local inflammatory response leading to tissue damage (Winer et al. 2022). In human malaria, overproduction of proinflammatory cytokines such as TNF-α, IFN-γ, IL-6, and IL-1 also contributes to anemia *(Angulo and Fresno 2002).*

Human *IL7* gene is located on chromosome 8q12-13 and contains 6 exons distributed over more than 33-kbp *(Lupton et al. 1990).* In non-African populations, variation in the *IL7* locus has been shown to influence susceptibility to inflammatory diseases, including multiple sclerosis and osteoarthritis *(Ghavimi et al. 2014; Zhang et al. 2016; Zuvich et al. 2010).* However, functional studies that would have elaborated how the genetic variation influences disease development remain largely undone. In African population, a targeted genetic-association study in pregnant women from Mozambique revealed that *IL7* was strongly associated with susceptibility to placental malaria *(Sikora et al. 2011).* Placental malarial pathology involves increased T helper 1 responses resulting in increased recruitment of macrophages to intervillous spaces of the placenta and increased concentrations of TNF-α and IFN-γ to destroy sequestered parasitized erythrocytes. However the excessive Th1 responses enhance oxidative stress and apoptotic cell death of the placenta leading to poor pregnancy outcomes *(Sharma and Shukla 2017).* These studies showing significant

associations between *IL7* and susceptibility to inflammatory diseases including malaria formed the basis to investigate the associations of genetic variants in *IL7* (72194T>C, rs2583759 and -2440A>G, rs7007634) on susceptibility to inefficient erythropoiesis and SMA in children attending Siaya County Referral Hospital (SCRH) located in a *P.falciparum* malaria endemic region of Kenya.

Granulocyte monocyte-colony stimulating factor (GM-CSF) is a hematopoietic growth factor that facilitates the differentiation of progenitor cells into the lymphoid, myeloid, and erythroid lineages *(Barreda et al. 2016).* Granulocyte monocyte-colony stimulating factor has been shown to promote growth and differentiation of leucocytes, and enhances the release of other cytokines, that mediate host immune responses *(Barreda et al. 2016).* The importance of GM-CSF in host immune response to inflammatory diseases has been reported in multiple sclerosis, tuberculosis, allergic disease and obesity *(Lee et al. 2020).* In part, the mechanism through which GM-CSF contributes to pathology during inflammatory diseases is through induction of an enhanced inflammatory state in the tissues *(Lee et al. 2020).* Granulocyte monocyte-colony stimulating factor (GM-CSF) has not been shown to have a direct role in malaria pathology but it has been shown to induce the synthesis of TNF-α and IL-1 *(Ananth et al. 2003).* Studies utilizing murine models have reported a reduction in the levels of erythropoietic-related cytokines, including GM-CSF (Miyashita *et al.* 2005); a negative correlation between GM-CSF concentrations and enhanced pathology in malarial anemia; and elevated levels of GM-CSF in lethal malaria *(Owhashi et al. 2013).* In the context of human malaria, the toll-like receptors (TLR) 7/8 stimulated production of GM-CSF was elevated in cord blood cells of infants with evidence of past placental malaria, suggesting a profound effect on the fetal immune system, with the differential alternations in innate immune responses predicting the risk of malaria during the first year of life *(Natama et al. 2018).* Moreover, elevated serum levels of GM-CSF have been reported in cases of severe

P. falciparum malaria *(Ringwald et al. 1991).* Previous investigations identified elevated levels of GM-CSF in children with SMA compared to those with non-SMA, and elevated GM-CSF levels in children with *P. falciparum* and HIV-1 co-infection relative to children with malaria alone *(Davenport et al. 2012; Ong'echa et al. 2011).*

The gene that encodes GM-CSF is colony-stimulating factor 2 (*CSF2*), which is located on the human chromosome at 5q23-31 *(van Leeuwen et al. 1989),* spans ~2.5kb in length, and encompasses 4 exons and 3 introns *(Miyatake et al. 1985).* Polymorphic variability in the *CSF2* gene has been associated with disease outcomes in asthma, ankylosing spondylitis, cancers and trachomatis *(Burkhardt et al. 2012; Chen et al. 2017; Cotterchio et al. 2015; Hardikar et al. 2015; Johnson et al. 2011; Natividad et al. 2009)* while other studies observed no relationships with diseases outcomes in acute respiratory distress and smoke induced chronic obstructive pulmonary disease *(Brown et al. 2015; He et al. 2008).* Although these studies did not explore the molecular mechanisms that link the genetic variations with disease pathogenesis, GM-CSF could be mediating inflammatory induced pathology during these diseases. For instance in asthma GM-CSF stimulates macrophages and neutrophils *(Burkhardt et al. 2012).* Release of GM-CSF by infected epithelial cells in *Chlamydia trachomatis* infection may mediate the influx and activation of inflammatory cells at the site of infection leading to trachomatous scarrying *(Natividad et al. 2009).* These investigations showing significant associations between *CSF2* and susceptibility to different inflammatory diseases including malaria formed the basis to investigate the influence of two SNPs flanking the *CSF2* (-7032 G>A, rs168681 and 64544 T>C, rs246835) on the rate and risk of malaria infections and SMA episodes; and all-cause mortality in children attending Siaya County Referral Hospital (SCRH) located in a *P.falciparum* malaria endemic region of Kenya.

1.2. Statement of the Problem

Despite significant progress in combating malaria in recent years the burden of severe disease and death due to *Plasmodium* infections remains a global public health concern *(WHO, 2022).* Only a fraction of infected people develops severe clinical syndromes motivating a longstanding search for genetic determinants of malaria severity *(Penha-Goncalves 2019).* Although strong genetic effects have been repeatedly ascribed to mutations and allelic variants of proteins expressed in red blood cells *(Mohandas and An 2012),* evidence provides a number of inflammatory response genes that have been repeatedly associated with malaria infection outcomes including, TNF, nitric oxide synthase (NOS)-2, IFNγR1, TLR, CD16 and CD40LG *(Penha-Goncalves 2019).* Scrutiny of genetic effects of the individual variants corroborates a pathogenesis model where pro-inflammatory genetic variants acting in early stages of malaria infection contribute to resolve infection but at later stages confer increased risk to organ damage driven by tissue inflammation *(Penha-Goncalves 2019).* In line with this model, several studies have demonstrated that variation in *IL7* and *CSF2* mediates tissue destruction during inflammatory diseases such as human pathogens, multiple sclerosis and rheumatoid arthritis *(Churchman et al. 2014; Kreft et al. 2012; Lee et al. 2020).* Further, genetic variation in *IL7* and *CSF2* has been shown to modulate susceptibility to placental malaria, ankylosing spondylitis, pancreatic cancer, placental malaria, multiple sclerosis and osteoarthritis *(Chen et al. 2017; Cotterchio et al. 2015; Ghavimi et al. 2014; Sikora et al. 2011; Zhang et al. 2016; Zuvich et al. 2010)* but their influence on malaria infection outcomes in children is still unknown. Therefore, the current study determined genetic associations of SNPs in *IL7* (72194T>C and -2440A>G) with susceptibility to inefficient erythropoiesis and SMA in children attending Siaya County Referral Hospital (SCRH) located in western Kenya, a region highly endemic for *P. falciparum* infections. In addition,

the influence of SNPs in *CSF2* (-7032 G>A, and 64544 T>C) on rate and risk of malaria infections and SMA episodes; and all-cause mortality was also determined.

1.3. Study Objectives

1.3.1. General objective

To investigate the association of genetic variants in *IL7* (72194 T>C and -2440 A>G) and *CSF2* (-7032 G>A, and 64544 T>C) with *Plasmodium falciparum* infection outcomes in children attending Siaya County Referral Hospital

1.3.2. Specific objectives

1. To determine the association of genetic variants in *IL7* (72194T>C and -2440A>G) with inefficient erythropoiesis and SMA
2. To establish the influence of genetic variants in *CSF2* (-7032G>A and 64544T>C) on the rate of malaria infections and SMA episodes
3. To determine relationship between genetic variants in *CSF2* (-7032G>A and 64544T>C) and the risk of malaria infections and SMA episodes
4. To establish the relationship between genetic variants in *CSF2* (-7032G>A and 64544T>C) and the risk of all-cause mortality

1.3.3. Null Hypotheses

1. Genetic variants in *IL7* are not associated with inefficient erythropoiesis and SMA
2. Genetic variants in *CSF2* do not influence the rate of malaria infections and SMA episodes

3. Genetic variants in *CSF2* do not relate to the risk of malaria infections and SMA episodes

4. Genetic variants in *CSF2* do not relate to the risk of all-cause mortality

1.4. Significance of the Study

Analysis of the effects of genetic variation in inflammatory response genes on malaria infection outcomes is a key step in malaria research to motivate experimental investigations of the underlying molecular mechanisms involved in disease pathogenesis. This knowledge will be critical to identify adjuvant therapies to prevent severe malaria related fatalities and to reduce the economic burden associated with treatment of severe malaria in endemic regions which are also among the poorest regions of the world. The findings of the current study highlight novel immuno-genetic associations of *IL7* and *CSF2* with the rate and risk of malaria infections and SMA episodes; inefficient erythropoiesis and all-cause mortality in children residing in Siaya County, western Kenya. The genetic correlates of protection observed included carriage of *CSF2*64544TC genotype and *CSF2* AC/GC diplotype associated with decreased rates of falciparum malaria infections and *CSF2*AC haplotype associated with a decreased risk of falciparum malaria infections. These protective genetic variants have potential to guide the development and testing of therapeutic alternatives, including vaccines for severe malaria outcomes which are suitable for children under five years of age, especially when further investigated in the context of the exact molecular mechanisms they engage in order to advance or ameliorate malaria disease. Furthermore, the genetic variants associated with poor outcomes included carriage of *IL7*72194TC and *IL7*72194CC genotypes and *IL7*CA haplotype associated with increased susceptibility to inefficient erythropoiesis. Carriage of *CSF2*AT/GC diplotype increased the rates of malaria episodes; presence of *CSF2*AT haplotype increased the rate of SMA episodes, whereas carriage of *CSF2*GA genotype increased the risk of all-cause mortality. These variants

associated with adverse outcomes may act as genetic predictors or markers that may guide management of children from birth until when they are five years old since this is the duration when they are immune naïve and thus most vulnerable to *P. falciparum* infections and subsequent severe malaria anemia.

CHAPTER TWO

LITERATURE REVIEW

2.1. Malaria Etiology and Epidemiology

Malaria in humans is caused by protozoan parasites of the genus *Plasmodium* that are transmitted by the female *Anopheles* mosquitoes with the major species of *Plasmodium* that infect humans being *P. falciparum, P. vivax, P. ovale, P. knowlesi* and *P. Malariae (Rossati et al. 2016).* Malaria is an endemic disease that is more prevalent in tropical countries because of the prevailing climatic conditions, especially temperature and rainfall, which are also the major drivers of malaria transmission *(Ryan et al. 2010).* The disease patterns of malaria transmission and clinical outcomes also vary as a result of variation in the malaria parasite and mosquitoes vectors, ecological conditions and socio-economic factors *(Rossati et al. 2016).* Uncomplicated malaria infection presents symptoms that are non-specific including fever, chills, myalgia, headache, anorexia, and cough, gastrointestinal symptoms, respiratory symptoms, and jaundice *(Ashley et al. 2018).* Failure of antimalarials or host mediated immunity to control the primary infection may lead to life-threatening complications which manifest as hyperparasitemia, hypoglycemia, hyperlactatemia, renal impairment, metabolic acidosis, cerebral malaria, respiratory distress and SMA *(Ashley et al. 2018).*

Malaria is the leading cause of morbidity and mortality worldwide with an estimated record of 247 million cases in 2021 *(WHO 2022).* A large proportion (95%) of these cases occurred in the African region, and over 99.8% of the total cases were attributed to *Plasmodium falciparum* (*P. falciparum*) infections *(WHO 2022).* Globally, there were ~619,000 malaria-related deaths reported, 76% of the of which were children (especially those <5 years of age) residing in the World Health Organization (WHO) African region *(WHO 2022).* Kenya experienced an 88% overall reduction in the prevalence of malaria infection since 1990

(21.2%) to 2015 (2.6%) as a result of increased optimized treatment and vector control *(Macharia et al. 2018).* However, 9 counties in Kenya have remained highly endemic for malaria, with 7 of them around Lake Victoria i.e. Migori, HomaBay, Kisumu, Siaya, Kakamega, Vihiga, and Busia, all having a malaria prevalence of ≥30% *(Macharia et al. 2018).*

Siaya County is one of the regions with the highest malaria endemicity in Kenya and more than 99% of all malaria infections are caused by *P. falciparum* with transmission modeling based on parasitic screening showing that the area has predicted rates of over 40% parasite prevalence year round *(Hollowell et al. 2023).* Slide positivity rates among children (0–14 years) visiting health clinics with a history of documented fever were found to be 52.5% (2007 and 2012), giving an indication of high malaria incidence in the region *(Khagayi et al. 2017).* Although data indicates that major progress has been made in the region regarding reducing malaria mortality in children, malaria still remains the most prevalent cause of morbidity and mortality in children <5 years of age in Siaya County (33.2%) followed by acute respiratory infections (26.7%), HIV/AIDS (18.6%) and diarrheal disease with younger children bearing the greatest burden *(Amek et al. 2014; Desai et al. 2014; Hollowell et al. 2023).*

Children residing in endemic *P. falciparum* transmission regions such as Siaya County typically experience repeated episodes of malaria prior to developing naturally-acquired immunity predisposing them to malaria complications including bone marrow suppression and the resultant anemia *(Helleberg et al. 2005).* Severe malaria anemia [SMA, hemoglobin (Hb) <5.0 g/dL, with any density parasitemia] accounts for the greatest worldwide proportion of malaria-associated morbidity and mortality *(WHO 2022).* In Kenya, SMA is also the most common complication of severe malaria in children <15 years of age in a number of counties

neighboring Siaya, i.e. Kakamega, Vihiga, Busia and Kisumu, with a prevalence of 45.2% and case fatality rates at 5.7% between 2014-2018, followed by severe respiratory distress at 19.7%, *(Akech et al. 2020).* Earlier studies in Siaya, had indicated that SMA had a 26% prevalence rate in children aged 1–5 years of age and is almost exclusively (i.e., 89% of the cases) in children <3 years of age *(Obonyo et al. 2007).* Severe malaria anemia contributed 52.7% of all malaria-related deaths with a case fatality rate of 22.8% and resulted in significant in-hospital morbidity and mortality in the Siaya community *(Obonyo et al. 2007).*

2.2. Abnormal Erythropoiesis and Severe Malaria Anemia

The pathological mechanisms that lead to SMA development include lysis of infected and uninfected red blood cells (RBCs), dyserythropoiesis and bone marrow suppression *(White 2018).* Loss of infected erythrocytes occurs through parasite maturation and subsequent hemolysis as well as through erythrophagocytosis in the spleen and possibly the liver while clearance of uninfected erythrocytes is due to activation of splenic macrophages and to extrinsic and intrinsic changes to the red blood cells that enhance their recognition and phagocytosis *(Jakeman et al. 1999; Lamikanra et al. 2007; Waitumbi et al. 2000).* Although the above mechanisms contribute to a reduction in hemoglobin levels in childhood malaria, one of the primary mechanisms responsible for low hemoglobin levels in children with SMA is abnormal erythropoiesis (ineffective erythropoiesis and dyserythropoiesis), features that have been observed in malaria patients *(White 2018).* Ineffective erythropoiesis is described as an abnormal proportion of erythroblasts in the bone marrow in comparison to steady-state erythropoiesis where the normal ratio between erythroblasts is 1 for Pro-E; 2 for Baso-E1; 4 for Baso-E2; 8 for Poly-E and 16 for Ortho-E, leading to a decreased number of red cells while dyserythropoiesis refers to qualitative abnormality of erythropoiesis with morphological defects of erythroblasts, which may finally be responsible for a decreased

erythrocytes production *(Dumarchey et al. 2022).* Ineffective erythropoiesis is associated with an inhibition of the erythroid cell proliferation and cellular division for instance in children infected with *P. falciparum*, there were decreased numbers of cells in G1 and a decreased S/G2 ratio occurred in Baso-E which was also reported in Poly-E, suggesting an accumulation in the G2 phase *(Wickramasinghe et al. 1982).* Dyserythropoiesis has also been seen in microscopic observations of bone marrow aspirates from patients infected with *P. falciparum* revealing erythroid hyperplasia and ultra-structural nuclear abnormalities such as multi-nuclearity, nuclear fragmentation, inter-nuclear bridges and irregular nuclear shapes *(Baro et al. 2017; Wickramasinghe et al. 1989).*

2.3. Role of cytokines in Inducing Abnormal Erythropoiesis and Severe Malaria Anemia

Although the precise mechanisms responsible for abnormal erythropoiesis in malaria anemia remain obscure, inappropriate balance in innate inflammatory mediators can induce bone marrow suppression *(Perkins et al. 2011).* An innate inflammatory response to malaria antigens such as hemozoin/malaria pigment (that is released with merozoites during rupture of parasitized-RBCs) *(Shio et al. 2010)* stimulates the host to secrete inflammatory mediators such as cytokines, chemokines, and effector molecules such as NO by monocytes and macrophages to control parasitemia but the responses may be altered in some patients thus contributing to abnormal erythropoiesis *(Perkins et al. 2011).* For instance, TNF-α and IFN-γ appear to play a protective role during the early stages of malaria infection through their ability to stimulate monocyte/macrophage activation and aid in controlling parasitemia but their excessive release along with NO enhances malarial anemia by contributing to abnormal erythropoiesis and erythrophagocytosis *(Shio et al. 2010).* TNF-α significantly decreases human erythroid progenitor cell proliferation while IFN-γ inhibits the proliferation and the differentiation of erythroblasts *(Felli et al. 2005).* Nitric oxide inhibits the proliferation and

erythroid differentiation of human primary erythroblasts *(Cokic and Schechter 2008).* IL-10 down regulates TNF-α leading to the stimulation of erythropoiesis and upon *Plasmodium* infection, a reduced level of IL-10 may result in erythropoiesis repression *(Othoro et al. 1999).* Interleukin-12 (IL-12) is a direct stimulator of erythropoiesis increasing the production of burst forming U-erythrocytes (BFU-Es) and colony forming U-erythrocytes (CFU-Es), leading to an enhanced erythrocyte count *(Dybedal et al.1995).* Children with falciparum malaria have suppressed circulating IL-12 which is associated with enhanced malarial anemia *(Luty et al. 2000).* Evidence indicates that chemokines such as macrophage inflammatory proteins (MIP) and regulated on activation, normal T-cell expressed and secreted (RANTES) also suppress erythropoietic responses *(Awandare et al. 2009;Were et al. 2006).*

2.4. Genetic Variation in Cytokine genes (*IL7* and *CSF2*) and Malaria Infection Outcomes

2.4.1. Genetic variation in cytokine genes and malaria infection outcomes

Candidate gene approach studies have shown that genetic variation in immune response genes (i.e., cytokines, chemokines, growth factors and effector molecules) alters the production of inflammatory mediators that influences the malaria infection outcomes including risk of malaria, SMA episodes and all cause mortality *(Achieng et al. 2019; Ouma et al. 2012).* For instance, functional variation in stem cell growth factor (SCGF), a hematopoietic growth factors that promote erythroid and myeloid colony development) promoter is associated with elevated SCGF production, enhanced erythropoiesis, and protection against the development of SMA *(Ouma et al. 2010).* Additionally, genetic variation in the Fc gamma receptor (Fcγ) Fcγ11A, Fcγ111A and Fcγ111B conditions susceptibility to SMA whereas Fcγ11A and Fcγ111B predicts the risk of repeated malaria episodes, SMA and mortality through a mechanism that involves functional changes in IFN-γ

production *(Munde et al. 2017;Ouma et al. 2012).* Interleukin (IL)-18 promoter haplotypes that conditioned elevated IL-18 gene products during acute infection were associated with increased risk of SMA through overproduction of IL-18 while carriage of a rare haplotype significantly increased the risk of childhood mortality *(Anyona et al. 2011).* Other reports demonstrate that IL-10 promoter variants condition susceptibility to SMA and functional changes in IL-10, TNF-α and IL-12 production in children with malaria (Ouma et al. 2008)**.** Findings of a more recent study using Integrated OMICS platforms demonstrated that Leukocyte associated immunoglobulin like receptor 1 (*LAIR1)* variants are associated with altered susceptibility to malaria, SMA, and all-cause mortality over a 36-month follow-up period where 'protective' *LAIR1* variants were associated with elevated *LAIR1* transcript expression, while variants that enhanced susceptibility to SMA had reduced mRNA levels *(Achieng et al. 2019).* Infection with *P. falciparum* had earlier been shown to suppress cyclooxygenase-2 (*COX-2)* gene products and reduced prostaglandin E_2 (PGE_2) levels resulting in increased severity of malarial anemia *(Anyona et al. 2012)* while more recently additional investigations revealed that *COX-2* promoter variants are associated with the longitudinal risk of malaria, SMA, and all-cause mortality among children *(Anyona et al. 2020).* Recently, it has also been demonstrated that complement component-3 *(C3)* mutations altered the longitudinal risk of pediatric malaria and SMA *(Raballah et al. 2022).*

Although the candidate gene approach which employs the targeted investigation of a particular gene/gene pathway can be a viable means of uncovering the molecular mechanisms that influence malaria infection outcomes, unbiased approaches, such as global genomic characterization can identify genetic variants associated with complex human phenotypes *(Kim et al. 2010).* Therefore, before the current study was conducted, a pilot Genome-wide Association Study (GWAS) was conducted to identify gene/gene pathways associated with SMA *(Achieng et al. 2019).* The current study hypothesized that variation in immune

response genes that influenced susceptibility to SMA during the GWAS would also influence other malaria infection outcomes including rate and risk of malaria episodes; erythropoiesis efficiency; and all-cause mortality through shared mechanisms that influence malaria disease progression.

The GWAS *(Achieng et al. 2019)* was performed using the Illumina® Infinium® HD Super Assay in conjunction with Illumina's® Human Omni2.5-8v1 BeadChip (with N2.45M markers), and high-throughput genotyping with the Human Immunochip (coated with N196K markers, Illumina®, CA, USA) in children [n=144, aged <36 months attending the former Siaya District Hospital (SDH)] selected from a parasitemic cohort (n=1220) excluding children with covariates that could affect hemoglobin levels i.e. hemoglobinopathies [glucose-6-phosphate dehydrogenase deficiency (G-6-PD), sickle cell and $\alpha^{3.7}$thallasemia deletions] and co-infections (hookworms, bacteria and HIV-1) *(Otieno et al. 2006; Perkins et al. 2011; Were et al. 2011).* Children were stratified into 'polarized extremes' of non-SMA [(Hb=8.0–10.9g/dL; avg. Hb=9.8), n=74] and SMA; [(Hb<5.0 g/dL; avg. Hb= 4.1), n=70]. Immunochip® data were analyzed using logistic regression analysis with an additive mode of inheritance. Single nucleotide polymorphisms (SNPs) in immune response genes (including *LAIR 1 and 2, IL7* and *CSF2)* were identified that were associated with enhanced risk of SMA ($P \leq 0.05$). In the *IL7* haploblock, SNPs (72194T>C, rs2583759) and (-2440A>G, rs7007634) were identified whereas (64544T>C, rs246835) was found in *CSF2* haploblock. One SNP *CSF2* (7032G>A, -rs168681) was not identified from the Immunochip® but was selected from online databases (dbSNP; http://www.ncbi.nlm.nih.gov/SNP, HapMap; http://www.hapmap.org/) because it was a promoter SNP that fitted the SNP selection criteria of the current study.

Generally the SNP selection criteria for the current study involved selecting SNPs that (i) contained high rankings on the *P*-value ($P \leq 0.05$) in the immunochip® results (ii) had MAFs >10% in African populations [identified from online databases; dbSNP; http://www.ncbi.nlm.nih.gov/SNP and HapMap; http://www.hapmap.org/)] so as to increase the power to detect rare variants associated with malaria infection outcomes (iii) had significant association with pathogenesis of other inflammatory diseases. Specifically, in one study, the SNP (rs2583759) of *IL7* was a risk factor for osteoarthritis in the Chinese Han population *(Zhang et al. 2016)* with other studies reporting that higher IL-7 levels are produced by cartilage tissue in patients with osteoarthritis that could be responsible for bone loss *(Hartgring et al. 2009)* (iv) they contained variation that could impart functional changes based on *in-silico* analysis (ALGENE PROMO or TRANSFAC). Functional changes may influence malaria infection outcomes by altering gene expression and hence immune profiles *(Achieng et al. 2019).* It was determined that the simulated presence of the major allele A in the promoter variant of (*IL7*-2440A>G, rs7007634, interferon regulatory transcription factor (IRF-3), c-E-twenty-six transformation-specific-2(c-Ets-2), activating enhancer binding protein 2 alpha (AP-2 Alpha), and Nuclear Factor of Activated T Cells 1 (NFAT1) bind while in the simulated presence of the minor G allele, Early gene expression factor (Elk-1) and CCAAT/enhancer-binding protein beta (C/EBP beta) bind. An A to G transition at this locus results in loss of the NFAT1 binding site. NFAT1 is a transcription factor that induces expression of cytokines in T cells (e.g., IL-2, IL-3, IL-4, TNF-α and GM-CSF) *(Hogan 2017).* Likewise for *CSF2* 7032G>A, rs168681 locus, presence of the wild type G allele creates transcription factor binding sites (TFBS) for estrogen receptor α (ER-α) and zinc finger E box-binding homeobox 1 (ZEB1) while transition to the minor A allele ablates such binding and creating a TFBS for nuclear factor I-X3 (NF1-X3) which has been implicated in tumor growth where it controls the migration of differentiating astrocytes as well as

migration and invasion of glioma cells *(Singh et al. 2011).* SNP *CSF2*64544T>C, rs246835 is part of a potential distal enhancer which a switch of T allele to C allele results in the loss of transcription factor binding sites (TFBS) for upstream transcription factor 2 (USF2) and creation of TFBS for GATA binding protein-1 (GATA-1. Various distal enhancers have been shown to modulate transcription of human target genes by forming cognate enhancer-promoter loops with the target gene promoter site *(Croft et al. 2018; Dahan et al. 2021; Koay et al. 2021).* GATA-1 is master transcription factor in erythropoiesis that regulates at the transcriptional level of all aspects of erythroid maturation and function, as revealed by gene knockout studies in mice and by genome-wide occupancies in erythroid cells *(Gutierrez et al. 2020).*

2.4.2. Genetic variation in *IL7* and malaria infection outcomes

Interleukin (IL)-7 is a hematopoietic growth factor which modulates both T- and B-cell development and T-cell homeostasis whose effects of IL-7 are mediated through binding to both the IL-7 receptor α and common cytokine γ chain which is shared with members of the IL-2 cytokine family (i.e., IL-2, IL-4, IL-9, IL-15, and IL-21) *(Chen et al. 2021).* Production of IL-7 is primarily from non-hematopoietic cells such as stromal cells in the bone marrow and thymus, dendritic cells, hepatocytes, keratinocytes, and epithelial cells *(Winer et al. 2022).* Interleukin-7 induces differentiation of pluripotent hematopoietic stem cells into lymphoid progenitor cells, promotes proliferation, differentiation, and survival of naïve and matures T cells, and stimulates the production of myelopoietic factors [e.g., IL-3 and GM-CSF] *(Chen et al. 2021).* A role for *IL7* in malaria is suggested by studies in murine models in which malarial anemia was associated with suppression of erythropoietic-related cytokines [i.e., GCSF, GM-CSF, IL-7 and IL-17], and elevation of IL-10 and TNF-α *(Xu et al. 2013).* Previous findings in human studies showed that levels of circulating IL-7 are suppressed in

children with SMA *(Ong'echa et al. 2011).* Since IL-7 induces erythropoiesis, reduced levels in children with SMA may contribute to delayed and/or inappropriate erythroid development *(Aiello et al. 2007).*

Interleukin-7 also mediates cell-mediated immune responses to human pathogens, and induces tissue destruction in inflammatory diseases such as multiple sclerosis and rheumatoid arthritis *(Churchman et al. 2014; Kreft et al. 2012; Sieling et al. 1995).* In rheumatoid arthritis several synovial cell types produce IL-7 which induces IFN, IL-17, TNF and macrophage inflammatory protein (MIP)-1α and induces bone loss by stimulating osteoclastogenesis *(Hartgring et al. 2006).* In multiple sclerosis (MS), inflammation plays a major role in the onset and propagation of the disease and there are high serum IL-7 levels that may signify a T helper 1 driven form of MS and may predict outcome in these patients *(Lee et al. 2020).* In human malaria, overproduction of proinflammatory cytokines such as TNF-α, IFN-γ, IL-6, and IL-1 also contributes to pathology *(Angulo and Fresno 2002).* For instance, cerebral malaria pathology is characterized by sequestration of malaria infected erythrocytes in the vascular endothelium inducing a local inflammatory response that induces tissue damage *(Dunst et al. 2017).*

Human *IL7* gene is located on chromosome 8q12-13 and contains 6 exons distributed over more than 33-kb *(Winer et al. 2022).* In non-African populations, variation in the *IL7* locus has been shown to influence susceptibility to inflammatory diseases, including multiple sclerosis and rheumatoid arthritis *(Churchman et al. 2014; Kreft et al. 2012).* In African population, a targeted genetic-association study in pregnant women from Mozambique revealed that *IL7* was strongly associated with susceptibility to placental malaria *(Sikora et al. 2011).* During placental malarial infection, there is an increased recruitment of macrophages to intervillous spaces of placenta and increased concentrations of TNF-α and

IFN-γ to counteract the sequestered infected erythrocytes leading to oxidative stress and apoptotic cell death of the placental cells resulting in poor pregnancy outcomes *(Sharma and Shukla 2017).* However, functional studies that would have elaborated how the genetic variation in *IL7* influences inflammatory diseases and placental malaria development remain largely undone. These studies showing significant associations between *IL7* and susceptibility to inflammatory diseases including malaria formed the basis to investigate the associations of genetic variants in *IL7* (72194T>C, rs2583759 and -2440A>G, rs7007634) on susceptibility to SMA and inefficient erythropoiesis in children attending Siaya County Referral Hospital (SCRH) located in a *P.falciparum* malaria endemic region of Kenya.

2.4.3. Genetic variation in CSF2 and malaria infection outcomes

Granulocyte monocyte-colony stimulating factor (GM-CSF) is a hematopoietic growth factor that facilitates the differentiation of progenitor cells into the lymphoid, myeloid, and erythroid lineages and it is produced by an array of cell types including mast cells, B cells, activated T cells, fibroblasts, macrophages, vascular endothelial cells, and various oncogenic cells *(Barreda et al. 2004).* Granulocyte monocyte-colony stimulating factor has been shown to promote growth and differentiation of leucocytes, and enhances the release of other cytokines, that mediate host immune responses *(Becher et al. 2016).* The importance of GM-CSF in inflammatory diseases has been reported in multiple sclerosis, tuberculosis, allergic disease and obesity and its associated meta inflammation where it contributes to through induction of an enhanced inflammatory state in the tissues *(Lee et al. 2020).* For instance, T helper 17 secreted GM-CSF is the main cytokine contributing to encephalitogenicity within the central nervous system via activation of microglia which adopt an inflammatory phenotype and produce highly neurotoxic molecules such as TNF-α, IL-1 and IL-6 in multiple sclerosis *(Parajuli et al. 2012).* In tuberculosis infections, GM-CSF induces

macrophage TNF-α production with TNF-α being a critical mediator of granuloma formation *(Szeliga et al. 2008).* In a mouse model of asthma, allergen exposed epithelial cells secretes GM-CSF that promotes eosinophils survival and activates dendritic cells which produce the chemokine CCL17 that perpetuate allergic airway inflammation *(Ait Yahia Saliha et. al. 2014).* Obesity is now widely considered as a low-grade, chronic inflammatory disease that contributes to metabolic dysfunction, ectopic lipid deposition and insulin resistance and elevated GM-CSF levels can be detected in adipose tissue where it promotes adipose tissue inflammation *(Hotamisligil 2017).*

Granulocyte monocyte-colony stimulating factor (GM-CSF) has not been shown to have a direct role in malaria pathology but it has been shown to induce the synthesis of TNF-α and IL-1 *(Ananth et al. 2003).* Studies utilizing murine models have reported: (i) a reduction in the levels of erythropoietic-related cytokines, including GM-CSF (Miyashita *et al.* 2005) (ii) a negative correlation between GM-CSF concentrations and enhanced pathology in malarial anemia, and (iii) elevated levels of GM-CSF in lethal malaria *(Owhashi et al. 2013).* In the context of human malaria, the TLR 7/8 stimulated production of GM-CSF was elevated in cord blood cells of infants with evidence of past placental malaria, suggesting a profound effect on the fetal immune system, with the differential alternations in innate immune responses predicting the risk of malaria during the first year of life *(Natama et al. 2018).* Previous investigations identified elevated levels of GM-CSF in children with SMA compared to those with non-SMA, and elevated GM-CSF levels in children with *P. falciparum* and HIV-1 co-infection relative to children with malaria alone *(Davenport et al. 2012; Ong'echa et al. 2011).*

The gene that encodes GM-CSF is colony-stimulating factor 2 (*CSF2*) is located on the human chromosome at 5q23-31 and spans ~2.5kb in length, and encompasses 4 exons and 3

introns *(Miyatake et al. 1985).* Polymorphic variability in the *CSF2* gene has been associated with disease outcomes in asthma, ankylosing spondylitis, cancers and trachomatis *(Burkhardt et al. 2012; Chen et al. 2017; Cotterchio et al. 2015; Hardikar et al. 2015; Johnson et al. 2011; Natividad et al. 2009)* while other studies observed no relationships with diseases outcomes in acute respiratory distress and smoke induced chronic obstructive pulmonary disease *(Brown et al. 2015; He et al. 2008).* In asthma, GM-CSF could stimulate macrophages and neutrophils during inflammation and may be related to the etiology of asthma *(Burkhardt et al. 2012).* Release of GM-CSF by infected epithelial cells in *Chlamydia trachomatis* infection may mediate the influx and activation of inflammatory cells at the site of infection leading to trachomatous scarrying *(Natividad et al. 2009).* Largely these studies did not explore the molecular mechanisms that link the genetic variations with disease pathogenesis. These investigations showing significant associations between *CSF2* and susceptibility to different inflammatory diseases including malaria formed the basis to investigate the influence of two SNPs flanking the *CSF2* (-7032 G>A, rs168681 and 64544 T>C, rs246835) on the rate and risk of *P.falciparum* malaria infections, SMA episodes and all-cause mortality in children attending Siaya County Referral Hospital (SCRH) located in a malaria endemic region of western Kenya.

CHAPTER THREE

MATERIALS AND METHODS

3.1. Study Site

This study was conducted at Siaya County Referral Hospital (SCRH) located in Siaya County, western Kenya (Appendix 1.0) located in a region endemic for *P. falciparum* malaria. Siaya County is approximately 1520 km^2 in size and lies between Latitude 0º 26' to 0º 18' North and Longitude 33º 58' East and 34º 33' West *(Ong'echa et al. 2006).* The residents of Siaya County receive an average of 9.6 infective mosquito bites per annum *Anopheles gambiae* s.s., *Anopheles arabiensis,* and *Anopheles funestus* with *Anopheles gambiae s.s* being the most dominant species *(Amek et al. 2012).* Malaria transmission intensity in western Kenya is largely dependent on two rainfall seasons (April to August and November to January) annually *(Matsushita et al. 2019).* In Siaya, more than 99% of all malaria infections are caused by *P. falciparum* with transmission modeling based on parasitic screening showing that the area has predicted rates of over 40% parasite prevalence year round *(Hollowell et al. 2023).* Malaria still remains the most prevalent cause of morbidity and mortality in children <5 years of age residing in Siaya County (33.2%) followed by acute respiratory infections (26.7%), HIV/AIDS (18.6%) and diarrheal disease with younger children bearing the greatest burden *(Amek et al. 2014; Desai et al. 2014; Hollowell et al. 2023).* SMA is the most common complication of severe malaria in children in Siaya whereby the rate in children aged 1–5 years is 26% and it is almost exclusively (i.e., 89% of the cases) in children <3 years of age *(Obonyo et al. 2007).* SMA contributed 52.7% of all malaria-related deaths with a case fatality rate of 22.8% and resulted in significant in-hospital morbidity and mortality in the Siaya community *(Obonyo et al. 2007).* In four counties neighboring Siaya that are also located in the highest malaria endemic areas in Kenya, the

prevalence of SMA was 45.2% and case fatality rates at 5.7% (2014-2018), followed by severe respiratory distress at 19.7%, *(Akech et al. 2020).* Individuals inhabiting the study area are predominantly from the Luo ethnic group, a genetically homogeneous population suitable for conducting host-genetic investigations on factors that condition susceptibility to malaria disease outcomes *(Ong'echa et al. 2006).*

3.2. Study Design

All the children enrolled in the current study and their samples utilized in the current study were obtained from 2003 to 2012 during a parent study (SSC Protocol No. 696-*Genetic Basis of Severe Malarial Anemia*) conducted at Siaya District Hospital (SDH) (Appendix 2.0). The parent study received ethical approval from the KEMRI/National Ethics Review Committee (Appendix 2.0). Authorization for sample/data collection is found in (Appendix 2.0). The samples were archived in the form of genomic DNA (from buccal swabs) at the University of New Mexico/KEMRI Global Health Research Laboratories located in Kisumu County. All procedures of current study have been adopted from the parent study except for the genotyping experiments which were conducted by the current study. The current study utilized two types of study designs i.e. cross-sectional case-control and longitudinal case-control. The cross-sectional design examined the associations of genetic variants in *IL7* and susceptibility to inefficient erythropoiesis and SMA to complete objective 1. The cross-sectional design was used because the available samples to perform this objective were collected on Day 0 (at enrollment). The longitudinal design determined the relationship between genetic variants in *CSF2* and the rate and risk of malaria infections, SMA episodes and all-cause mortality to complete objectives 2, 3 and 4. The longitudinal design was used because the available samples to perform objectives 2-4 were collected on Day 0 (at enrollment) and over the 36 months follow-up period. The follow-up involved recording any

instances of malaria infections, SMA and mortality by any cause. The records were obtained at SDH during a day 14 scheduled visit (if febrile on day 0), during acute visits (due to febrile status of the participant), during quarterly scheduled visits (to check their health status) and during home visits (in cases where participants failed to report to the hospital on scheduled visits). The parent/legal guardian was also supposed to report cases of participant death.

3.3. Study Participants

The study participants of the current study were enrolled from the children who participated in the parent study. The procedure for enrollment in the parent study was as follows; Children presenting with suspected malaria infections and those reporting for routine immunization at SDH were approached for enrollment. The parent/legal guardian of the child consented after receiving an explanation of the study. Pre-test counseling for HIV was provided by a professionally trained counselor who also administered an interviewer administered questionnaire to collect relevant demographic and clinical information. For those who consented, heel and/or finger-prick blood samples (<300µL) were obtained from the children. The blood was used to determine the HIV-1 status because HIV enhances the severity of malarial anemia *(Otieno et al. 2006).* The blood was also used to determine the presence of malaria parasitemia (this enabled a malaria diagnosis to be made so as to offer needed treatment for malaria). It also enabled the inclusion of only *falciparum ssp.* positive children and to exclude those positive for other Plasmodium species. Children that satisfied all inclusion criteria and did not meet the exclusion criteria were enrolled in the parent study.

3.3.1. Inclusion Criteria

i. Children aged <36 months.

ii. Children positive for *P. falciparum* parasitemia.

iii. Aparasitemic controls.

iv. Parent/legal guardian able and willing to sign informed consent.

v. Parent/guardian able to schedule and keep study appointments.

vi. Distance to hospital ≤25 km.

3.3.2. Exclusion Criteria

i. Refusal by parent/legal guardian to participate in the study.

ii. Planned relocation during study period.

iii. Children presenting with non-*falciparum* parasite strains.

iv. Reported use of antimalarial therapy by the child in the two preceding weeks.

v. A cerebral malaria positive test, though it is rare in Siaya and its environs *(Akech et al. 2020; Ong'echa et al. 2006)*

vi. Presence of chronic illnesses such as HIV and AIDS related symptoms.

vii. Children who had prior hospitalization for any reason.

After enrollment into the study, based on hemoglobin concentrations and any density parasitemia, children targeted for the cross-sectional study were stratified into uncomplicated malaria (UM, Hb≥5.0g/dL) or SMA (Hb<5.0g/dL) groups according to the WHO definition of SMA *(WHO 2012).* Children targeted for the longitudinal study were stratified into aparasitemic controls (*P. falciparum* negative, Hb≥11.0 g/dL and reporting for routine vaccinations and were a febrile i.e. had no fever or diarrhea for the last 14 days), non-SMA (Hb≥5.0g/dL) and SMA (Hb<5.0g/dL). Prior to treatment interventions, venous blood (1-3 mL) was taken from the enrolled children by venipuncture. The blood was used for determining hematological measurements [complete blood counts (CBCs) that included red blood cell (RBC) indices, white blood cell (WBC) indices, hemoglobin (Hb) and hematocrit

(Hct)] and co-infection status (HIV-1 and bacteremia). Co-infection status for hookworm was also determined by obtaining stool samples. The blood was also used for genetic analyses to determine $\alpha^{3.7}$-thalassemia deletions, G-6-PD and sickle cell status. All the above investigations were done because co-infections and hemoglobinopathies have been shown to enhance malaria anemia severity in children *(Otieno et al. 2006; Were et al. 2011; Perkins et al. 2011)*. Febrile malaria parasitemic children then promptly received appropriate antimalarial treatment. All treatment interventions were given according to MOH-Kenya guidelines at that time which included the use of Coartem (Artemether and Lumefantrin combination) for uncomplicated malaria and intravenous quinine to treat severe clinical malaria. Supportive care including hematinics and blood transfusions were also provided for children with SMA. HIV-1 exposed children were placed on sulfamethoxazole-trimethoprim (Bactrim®). This continued until those >20 months of age became negative on the serological tests for HIV-1. Children that met the WHO criteria for receiving anti-retrovirals (ARVs) were referred to the SDH Patient Support Centre for onward monitoring and treatment. At the time of sample collection, none of the HIV-1-positive study participants had been started on antiretroviral treatment. Bacterial positive children were referred to SDH for appropriate antibiotics therapy upon identification of the specific pathogen(s) and any accompanying anti-microbial resistance patterns. The anti-helminthic treatment for hookworm infections was also administered at SDH.

3.3.3. Longitudinal Follow-ups

For the longitudinal study design, after enrollment (Day 0 of the parent study), children were scheduled for follow-up visits on day 14 (if they were febrile upon enrollment) which represents a well-visit. If the child was still malaria-positive, he/she was brought back for subsequent visits every 7 days until a parasite-free measure was determined. The children

also came back for quarterly visits (after every 3 months) that served as wellness visits to check their health status. Parents/legal guardians who failed to return for scheduled quarterly follow-up visits were traced by the study team at their residence to check the child's health status, including mortality. Each residence was identified by a Geographic Information Systems (GIS)/Global Positioning System (GPS) surveillance system. In addition, parents/legal guardians were asked to return to the hospital any time their child was febrile (acute visits). The number of acute visits (due to either malaria infection or SMA) were noted for each child and the time period between the visits since this was required to compute the data on the risk and rate of malaria infections and SMA episodes. Physical evaluations and laboratory tests required for comprehensive clinical management of the patients were performed at enrollment (Day 0), day 14, and each acute and quarterly visit including CBCs, malaria parasitemia measures, and evaluation of bacteremia were clinically indicated. For all acute and scheduled visits, children were managed according to the MOH-Kenya guidelines. Parents were also informed to report on any deaths of participants to the study clinicians.

3.4. Sample Size Determination

All the children who met the inclusion criteria of the parent study participated in the current study. The children were enrolled in the parent study under two cohorts (Cohort 1:2003-2007, n=883 and Cohort 2: 2009-2012, n=1,654). The Cohort 1 children (n=883) participated to achieve objective 1 through the cross-sectional design. To study the influence of genetic variants on susceptibility to SMA, a total of 883 children participated with the case group being children with SMA (n=165) while the control group were children with uncomplicated malaria (UM, n=718). Additionally, to study efficiency of erythropoietic response, a total of 883 children participated with the case group being children with RPI<2 (n=741) and the control were children with RPI>2 (n=142). To achieve objective 2, 3 and 4 through the

longitudinal design, all the Cohort 2 children (n=1,654) participated in the current study. They were grouped as; aparasitemic controls, age and gender matched (n=335); non-SMA (n=1029); and SMA (n=290). The ratio of controls to cases in both study designs was about 4:1 which is consistent with the prevalence of SMA (cases) at SDH during the enrollment period *(Obonyo et al. 2007)* which is also consistent with acceptable limits for maximum ratio of controls to cases recommended in case control studies.

3.5. Laboratory Procedures

3.5.1. Sample Collection

Blood collection was conducted by trained phlebotomists employed by the parent study. Heel/finger prick blood <300µL was removed by pricking the site with a small lancet after the area had been cleaned with an alcohol swab. A dried blood spot (DBS) from the pricked area was also obtained on filter papers. Venipuncture blood (1-3mL) was also collected aseptically in EDTA-containing vacutainer tubes. The stool samples (~20mg) were collected in dry, clean, leak proof plastic containers and preserved in 10% formalin and polyvinyl-alcohol. One volume of the stool specimen was added to three volumes of the preservative and mixed well and sealed with parafilm.

3.5.2. Malaria Diagnosis and Parasite Density Determination

Thick and thin blood films were used to determine *P. falciparum* parasitemia in peripheral circulation. Blood films were prepared from venous blood, stained with 3% Giemsa, and examined under oil immersion by microscopy for malaria parasites. The number and species of asexual *Plasmodium* parasites were determined per 300 leukocytes, and the parasite density was calculated based on the total leukocyte count/µL of blood for each individual as follows; Parasites/µL of blood=(No. of parasites counted × No. of WBCs/µL)/ No. of

WBCs counted (WHO: Malaria microscopy standard operating procedures). The different species of *Plasmodium* were differentiated on thick blood films based on their unique features *(WHO, 2022*). *Plasmodium falciparum* malaria was identified by observing high grade parasitemia, doubly parasitized cells and crescent-shaped gametocytes and appliqué forms. *P. Ovale* malaria infections were characterized by reticulocyte infections, Schuffner's dots, oval shaped RBCs and feathering. *Plasmodium malariae* was identified by its three unique characteristics; senescent RBC infection; band-like trophozoites and rosette forms.

3.5.3. Cerebral Malaria Diagnosis

Cerebral malaria is a manifestation of severe malaria though rare in the Siaya Community *(Akech et al. 2020; Ong'echa et al. 2006).* It is defined by the WHO as a clinical syndrome characterized by coma with the presence of asexual forms of *P. falciparum* in peripheral blood, and exclusion of other factors that could cause unconsciousness such as other infections or hypoglycemia *(WHO 2022).*

3.5.4. Complete Blood Counts

Complete blood counts (CBCs) included measurement of WBCs, RBCs, Hb, Hct, platelet count. RBC indices measured included mean corpuscular volume (MCV), mean corpuscular hemoglobin concentration (MCHC) and red cell distribution width (RDW). All the measurements were done using a Beckman Coulter AcT diff2 (Beckman-Coulter Corporation, Brea, CA, USA).

3.5.5. Reticulocyte Production Index

The reticulocyte count was performed manually by preparing smears form peripheral blood mixed with 2-3 drops of new methylene blue solution. The smear was then incubated for 15 minutes and counterstained with Wright's stain. Reticulocytes were identified by a dark blue

reticulum or network in the cells. Reticulocyte counts were performed by counting 1000 RBCs and then expressed as a percentage of RBCs. The reticulocyte count was used to calculate the reticulocyte production index (RPI) that was used to measure the efficiency of the erythropoietic response. The RPI was determined using the formula; RPI=reticulocyte index (RI)/maturation factor (MF), where RI=reticulocyte count [%] x Hct/0.36) and MF=b+(m)(x), where b=1, m=0.05, and x=(average normal population Hct - patient's Hct) An RPI<2.0 indicated that the reticulocyte count had not risen appropriately in response to the degree of anemia and vice versa *(Anyona et al. 2012; Novelli et al. 2010; Were et al. 2006)*

3.5.6. Co-infections Determination

All children were tested for HIV-1, bacterial and hookworm infections by the parent study. The bacterial cultures were prepared as follows: approximately 1.0 mL of venipuncture blood was collected aseptically into sterile pediatric isolator microbial tubes (Wampole Laboratories, Princeton, NJ) for bacterial cultures. Blood samples were inoculated directly onto chocolate agar plates and incubated for 18hrs at 37°C in 5% carbon IV oxide (CO_2), followed by subculture for 18-24hrs in an inverted position. If no growth was obtained, subcultures were incubated for an additional 4 days. Plates were inspected daily for signs of microbial growth. Bacterial colonies were identified by Gram staining, colonial characteristics and appearances, and biochemical tests (API biochemical galleries, bioMerieux, Louvres, France) and/or agglutination serology were used to confirm the presence of suspected blood-borne bacterial pathogens.

Exposure to HIV-1 was determined using venipuncture blood by two rapid serological antibody tests: Unigold (Trinity Biotech, Bray, and County Wicklow, Ireland) and Determine (Abbott Laboratories, Chicago, Illinois, USA). A confirmatory test for HIV-1 infection was

performed using molecular methods because antibody testing can be inaccurate for testing infants and breast feeding children because they could acquire circulating anti-HIV antibodies in-utero or via breast milk. The molecular methods involved RNA extraction, reverse transcription (RT), and nested PCR. For RT and primary PCR, the primers were GP40F1 (forward; 5′TCTTAGGAGCAGCAGGAAGCACTATGGG; nucleotides 7789 to 7816 based on HXB2 [GenBank accession no. K03455]) and GP41R1 (reverse; 5′AACGACAAAGGTGAGTATCCCTGCCTAA; nucleotides 8347 to 8374). Viral RNA was extracted from plasma by using the QIAamp viral RNA kit according to the manufacturer's protocol (Qiagen, Valencia, Calif.). Briefly, 200 μl of plasma was mixed with 800 μl of lysis buffer. After 10-min incubation, 800 μl of 100% ethanol was added to the lysate. The mixture was filtered through a column by centrifugation. After being washed with buffer, the RNA was eluted from the column by adding 50 μl of RNase-free water. For negative controls, RNA from normal human plasma was also extracted. Three to 10 μl of the RNA extract was used to synthesize cDNA with primer GP41R1 (20 μM) and the GeneAmp RNA PCR kit following the manufacturer's protocol (Perkin-Elmer Cetus, Norwalk, Conn.). The 20-μl cDNA reaction mixture was then added to a PCR mixture containing 50 μM GP40F1 and 30 μM GP41R1, 1× GeneAmp PCR buffer II, 1.25 mM $MgCl_2$, and 2.5U of Ampli*Taq* DNA polymerase (Perkin-Elmer Cetus, Foster City, Calif.) and was brought to a final volume of 100 μl with sterile distilled water. After initial denaturation at 94°C for 2 min, 35 cycles of PCR were performed in the GeneAmp 9600 thermocycler (Perkin-Elmer Cetus). Each cycle consisted of denaturation at 94°C for 30 s, annealing at 50°C for 30 s, and extension at 72°C for 60 s, with a final extension at 72°C for 5 min. For nested PCR, the primers were GP46F2 (forward; 5′ACAATTATTGTCTGGTATAGTGCAACAGCA; nucleotides 7850 to 7879) and GP47R2 (reverse; 5′TTAAACCTATCAAGCCTCCTACTATCATTA; nucleotides 8281 to 8310). The nested

PCR procedure was as follows; 5 µl of the primary PCR product was added to a 100-µl PCR mixture containing reagents similar to those in the primary PCR, except that the primers were replaced by 25 µM each GP46F2 and GP47R2. The PCR mixtures were subjected to 35 cycles under the same conditions as the primary PCR. After PCR, the nested PCR products were electrophoresed in 1.5% agarose gels along with a 100-bp ladder (Gibco, Grand Island, N.Y.) and visualized under UV light after ethidium bromide staining.

To diagnose a hookworm infection, the stool sample was concentrated using the ethyl acetate sedimentation technique by adding 10 mL of 10% formalin to the sediment and mixing thoroughly with wooden applicator sticks. Then 4 mL of ethyl acetates added to the mixture then shaken vigorously in an inverted position for 30s then centrifuged at 500×g for 10 min. The sample was examined microscopically to determine if there were hookworm eggs.

3.5.7. Hemoglobinopathies

Sickle-cell trait, $\alpha^{3.7}$-thalassemia deletions and G-6-PD were determined in all the children. The $\alpha^{3.7}$-thalassemia deletion variants were determined by PCR. Genomic DNA was extracted from 200 µl of peripheral whole blood using the QIA-amp DNA Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. Extracted DNA concentration was determined by measuring absorbance at 260 nm and 280 nm using a spectrophotometer (SPECORD Plus, Analytik, Jena AG, Jena, Germany), with A260/A280 ratios of between 1.6 and 1.8 accepted as of suitable quality. DNA was diluted with 1× PCR buffer to 5ng/µl and stored at -20°C until analysis. Polymerase chain reaction (PCR) was carried out in 25-µL reactions containing 0.75 µmol/l deoxynucleotide triphosphates (dNTPs), 1.25 units of AmpliTaq Gold polymerase, Dimethyl sulfoxide (DMSO), 200 ng genomic DNA, in 1× GeneAmp PCR buffer provided (Perkin-Elmer) by the manufacturer. The relative amounts of primers for $\alpha^{3.7}$-thalassemia deletion variants and reagents were

optimized with the use of AmpliTaq Gold polymerase. Three primer pairs were used to detect the -α (3.7) deletion mutant: forward primer (common) 3'AAGTCCACCCCTTCCT-TCCTCACC5' (Z84721 positions 32755-32778), reverse primer (mutant) 3'TC-CATCCCCTCCTCCCGCCCCTGCCT-TTTC5' (Z84721 positions 38492-38521), and reverse primer (normal) 3'AT-GAGAGAAATGTTCTGGCACCTG- CACTTG5' (Z84721 positions 34942-34971), generating products of 1963 and 2217 bp for -α (3.7) and the normal control, respectively. Primers were newly designed or modified from published primers. PCRs were performed in 25 μl containing 0.75 M betaine, 5% DMSO, 200 μmol of dNTPs, 1.25 U AmpliTaq Gold® DNA polymerase, 1× GeneAmp® buffer (Applied Biosystems, Foster City, CA), 0.2 μmol of each primer, and 5 μl of genomic DNA. Thermocycling (performed in a T100tm Thermo Cycler; Bio-Rad) conditions were as follows: 15 minutes at 95°C; 35 cycles of 95°C for 1 minute, 65°C for 1 minute, and 72°C for 2.5 minutes; with a final step of 72°C for 10 minutes. Amplicons were analyzed by 1.5% agarose gel-electrophoresis and visualized using ethidium bromide staining under UV illumination.

Glucose-6-phosphate dehydrogenase (G-6-PD) deficiency was determined by a fluorescent spot test (FST) (Trinity Biotech Plc., Bray, Ireland). A total of 5 μL of whole blood was added to 100 μL G-6-PDH substrate solution. A first aliquot was immediately spotted onto filter paper and the remaining solution incubated at room temperature. Second and third drops of blood-substrate mixture were blotted after 5 and 10 min of incubation. The filter papers were air dried for approximately 30 min and subsequently read under UV light (340nm). A sample with normal enzyme activity showed moderate to strong fluorescence after 5 min, and strong fluorescence after 10 min. A sample with intermediate G6PD activity showed no or weak fluorescence after 5 min and moderate fluorescence after 10 min, while a deficient sample had very little or no fluorescence after 10 min. The ST test was performed along with normal, intermediate and deficient controls (Catalogue numbers G6888, G5029, and G5888,

respectively, Trinity Biotech, Ireland). All results were read by two independent observers; when the interpretation of the readers was discordant, a third reader, blinded to the previous results was called in.

Sickle-cell status was determined using the alkaline cellulose acetate electrophoresis on Titan III plates (Helena BioSciences, Sunderland, UK). The plates were soaked in Supre-Heme ® Buffer for 5 minutes. 100 mL of Supre-Heme® buffer was poured into each of the outer sections of the Zip Zone® Chamber. Two chamber wicks were wetted in the buffer and one draped over each support bridge and the chamber was then covered to prevent buffer evaporation. The patient samples were prepared by adding whole blood to Hemolysate Reagent in ratio 1:3, mixed well and allowed to stand for 5 min. The patient hemolysates (5%) or controls were placed into the wells of the plates. The plate was then placed on the elecrophoresis and a weight (glass slide) was placed on the plate to ensure contact with the wicks. The plate was then electrophoresed for 25 min at 350 volts. The plates were then removed from the electrophoresis chamber and stained in Ponceau S for 5 min. the hemoglobin types were scored using the Hemo AFSC control. Determination of sickle cell status and G-6-PD deficiency provided important clinical information that was communicated back to study participants by a genetic counselor employed by the parent study.

3.5.8. Genotyping of SNPs in *IL7* and *CSF2* genes

Genotyping was performed by the current study on the study participants' samples to determine the frequency of the selected *IL7* and *CSF2* SNPs in the study groups.

3.5.8.1. Selection of SNPs in *IL7* gene

The pilot GWAS study *(Achieng et al. 2019)* identified 32 SNPs encompassing the *IL7* haploblock (10.8kb). Out of the 32 SNPs, two SNPs were selected (72194T>C, rs2583759)

and (-2440A>G, rs7007634) for the current study because (i) they contained high rankings on the *P*-value ($P \leq 0.05$) in the immunochip® results (ii) MAFs were >10% in African populations identified from online databases (dbSNP; http://www.ncbi.nlm.nih.gov/SNP, HapMap; http://www.hapmap.org/) to increase the power to detect rare variants associated with malaria disease (iii) Significant association with pathogenesis of other inflammatory diseases. For instance, SNP 72194T>C, (rs2583759) is associated with osteoarthritis in the Chinese Han population where haplotype AGAC, AGAT and GGGC of rs2583764, rs2583760, rs6993386 and rs2583759 were risk factors for osteoarthritis with other studies showing that higher IL-7 levels are produced by cartilage tissue in patients with osteoarthritis that could be responsible for bone loss *(Hartgring et al. 2006)* and (iv) they contained variation that could impart functional changes based on *in-silico* analysis (ALGENE PROMO). Functional changes may influence gene expression, immune mediator profiles and hence malaria infection outcomes *(Achieng et al. 2019).* In the simulated presence of the major allele A in the promoter variant *IL7* -2440 A>G (rs7007634), interferon regulatory transcription factor (IRF-3), c-E-twenty-six transformation-specific-2(c-Ets-2), activating enhancer binding protein 2 alpha (AP-2 alpha), and Nuclear Factor of Activated T Cells 1 (NFAT1) bind while in the simulated presence of the minor G allele, Early gene expression factor (Elk-1) and CCAAT/enhancer-binding protein beta (C/EBP beta) bind. An A to G transition at *IL7* (-2440A>G, rs7007634) locus results in loss of the NFAT1 binding site. NFAT1 is a transcription factor that induces expression of cytokines in T cells (e.g., IL-2, IL-3, IL-4, TNF-α and GM-CSF) *(Hogan 2017).* HaploView (version 4.2, Broad Institute, Cambridge, MA, USA) revealed a weak magnitude of linkage disequilibrium (LD) for the two *IL7* SNPs investigated ($D' = 0.11$; $r^2 < 0.01$).

3.5.8.2. Selection of SNPs in *CSF2* gene

The pilot GWAS study *(Achieng et al. 2019)* also identified an intergenic SNP 64544T>C (rs246835) within the *CSF2* haploblock that was associated with increased risk to SMA (OR =2.77; 95% CI =0.90-8.49; *P*=0.05). The SNP 64544T>C (rs246835) was selected because (i) it is part of a potential distal enhancer with the loss of transcription factor binding sites (TFBS) for upstream transcription factor 2 (USF2) and creation of TFBS for GATA binding protein-1 (GATA-1) by switch of T allele to C allele based on *in silico* analysis (TRANSFAC). Various distal enhancers have been shown to modulate transcription of human target genes by forming cognate enhancer-promoter loops with the target gene promoter site *(Croft et al. 2018; Dahan et al. 2021; Koay et al. 2021).* GATA-1 is master transcription factor in erythropoiesis *(Gutierrez et al. 2020)* (ii) it had a MAF (>10%) in African populations. The promoter SNP -7032G>A (rs168681) was selected (though not in the Immunochip®) because it had desirable characteristics for further exploration, including: (i) a MAF of (>10%) in African populations (ii) potential functional effects, where the wild type G allele creates TFBS for estrogen receptor α (ER-α) and zinc finger E box-binding homeobox 1 (ZEB1), and transition to the minor A allele ablates such binding and creates a TFBS for nuclear factor I-X3 (NF-X3) which has been implicated in tumor growth where it controls the migration of differentiating astrocytes as well as migration and invasion of glioma cells *(Singh et al. 2011).* HaploView (version 4.2, Broad Institute, Cambridge, MA, USA) revealed a weak LD (D′=0.41, r^2=0.01) for the two *CSF2* SNPs investigated.

3.5.8.3. Genotyping Procedure

Genetic variants rs168681 (*CSF3-36*32G>A), rs246835 (*CSF2*64544T>C), rs2583759 (*IL7*72194C>T) and rs7007634 (*IL7*-2440A>G) were genotyped using the TaqMan 5′ allelic discrimination Assay-By-Design (assay IDs C_3285157_20, C_2397167_10, C_9168017_10

and C_30930344_10 respectively). Genomic DNA (gDNA) was extracted from buccal swabs using the MasterAmpTM Buccal swab DNA extraction kit (Epicentre Biotechnologies, Madison, WI). The NanoDrop™ 2000 Spectrophotometer was used to estimate DNA concentration by measuring the absorbance at 260nm, adjusting the A_{260} measurement for turbidity (measured by absorbance at 230nm), multiplying by the dilution factor, and using the relationship that an A_{260} of 1.0 = 50µg/mL pure dsDNA. The gDNA with concentration of >50µg/mL was considered for amplification. Similarly, purity of the extracted DNA was determined by measuring absorbance at 260nm divided by the reading at 280nm. Extracted DNA with A_{260}/A_{280} ratio of 1.7–2.0 was considered pure enough for amplification. Extracted gDNA was initially amplified using Genomiphi DNA Amplification Kit (Amershan Biosciences SV Corp, CA, USA). This involved the use of Phi29 DNA polymerase, in combination with random-sequence hexamer primers to amplify DNA in an isothermal process. Genomiphied DNA was then used as template to genotype the selected *IL7* and *CSF2* SNPs, using predesigned assays outlined above via quantitative PCR on the StepOnePlus™ Real-Time PCR System (ABI) (Applied Biosystems, Foster City, CA, USA). The PCR was performed in a total reaction volume of 10µL with the following amplification cycles: initial denaturation (60ºC for 30s and 95°C for 10 min) followed by 40 cycles of (95ºC for 15s and 60ºC for 1 min.) and a final extension (60ºC for 30s). Thereafter, the genotype of each individual was determined using allele-specific fluorescence on the StepOnePlusTM Real-Time PCR Systems. StepOne™ Software Version 2.3 was used for allelic discrimination (Applied Biosystems, Foster City, CA, USA). Genotyping was performed to determine frequency of the select SNPs under investigation in the study groups.

3.6. Statistical Analyses

Statistical analyses for objective 1 were performed using SPSS (version 19.0). The Kolmogorov–Smirnov test was used to test for normality of continuous variables, and since normality was violated, Mann-Whitney-*U* test was used to compare median values between the clinical groups. Differences between proportions and comparisons of genotype and haplotypic frequencies of *IL7* polymorphisms between the uncomplicated malaria (UM) and SMA groups were performed using the chi-square ($\chi 2$) tests. The relationship between *IL7* genotypes/haplotypes with erythropoiesis efficiency or SMA were determined using bivariate logistic regression in a model that controlled for potential confounding variables [age, gender, HIV-1 and bacteremia status, hemoglobinopathies (sickle-cell traits, G-6-PD and $\alpha^{3.7}$-thalassemia variants)]. Statistical significance was set at $P \leq 0.05$. Haplotypes were constructed (using Haploview version 4.2, Broad Institute, Cambridge, MA, USA) since they can predict associations with disease outcomes that may not be revealed by analysis of individual SNPs *(Perkins et al. 2013).*

Statistical analyses for Objectives 2, 3 and 4 were performed using R version 3.1.4 *(Team 2018).* The Kolmogorov–Smirnov test was used to test for normality of continuous variables and all were normally distributed. One-way analysis of variance (ANOVA) and Students t-test were used to compare means of metric variables. The distribution of categorical variables was compared between the groups (aparasitemic, non-SMA, and SMA) using a chi-square test and a Fisher's exact test was used if the cell count was <20. Haplotypes and diplotypes were constructed to investigate the effects of the genetic architecture determined by the *CSF3-36*32G>A (rs168168) and *CSF2* 64544 T>C (rs246835) loci. The top-level was the individual diplotypes, consisting of 2 of the 4 possible haplotypes (GT, GC, AT, and AC). At the marginal levels, the diplotypes gave rise to the (1-locus) genotypes at *CSF2* -7032 G>A

(GG, GA, AA) and *CSF2* 64544T>C (TT, TC, CC). For all generalized linear and hazard models performed, the effects of the genetic architecture were coded as 0/1 variables that reflected: (i) per locus contributions determined by the mutant homozygotes and heterozygotes (indicating additive and dominant effects), (ii) haplotype contributions (reflecting epistatic interactions), and (iii) diplotype contributions (reflecting positional effects). The respective wild types were subsumed as a baseline factor. Hence, the following 0/1 variables were coded: (i) genotypes GA, AA for *CSF3-36*32G>A and TC, CC for *CSF2*64544T>C, (ii) haplotypes GC, AT, AC, and (iii) diplotypes GT/GC, GT/AT, GT/AC, GC/GC, GC/AC, AT/AT, AT/AC, AC/AC. The differences in the frequency distributions between the aparasitemic, non-SMA, and SMA groups were tested using the generalized Fisher's exact test with simulated *P*-values. Exact tests for deviations from Hardy-Weinberg equilibrium (HWE) at the *CSF3-36*32G>A (rs168681) and the *CSF2*64544T>C (rs246835) loci were performed using the HPlus software (version 2.5). Genetic variants for *CSF2*, age, sex, cohort, hemoglobinopathies (sickle-cell traits, G6PD and $\alpha^{3.7}$-thalassemia deletions), and co-infections (HIV-1 and bacteremia) were included in the models as covariates. Cohort was kept as a categorical covariate to account for potential changing malaria incidence across time. For objective 2 analyses, the effect of covariates on the rate of malaria and SMA episodes was investigated using generalized linear poisson regression. For objective 3 analyses, Cox proportional hazard model was used to investigate the influence of covariates on the risk of malaria and SMA episodes For objective 4 analyses, Cox proportional hazard model was used to investigate the influence of covariates on the risk of all-cause mortality.

3.7. Ethical Considerations

Ethical approval of the parent study was obtained from the KEMRI/National Ethics Review Committee (Appendix 2.0). Parents/legal guardians of children participating in the study

provided written informed consent in their language of choice (Dholuo, Swahili or English). They also received pre- and post-test counselling for HIV&AIDS. The HIV test result files were password protected and computer access restricted to the supervisor of the laboratory and the designated data entry staff. Febrile children at enrollment, acute visits or scheduled visits were managed according to the Ministry of Health-Kenya guidelines. Laboratory samples were bar-coded to add to the confidentially.

CHAPTER FOUR

RESULTS

4.1. Specific Objective 1: To Determine the Association of Genetic Variants in IL7 (72194 T>C and -2440 A>G) with Inefficient Erythropoiesis and SMA

This section (4.1) presents the demographic, clinical and laboratory characteristics of the study participants for objective 1; distribution of *IL7* genotypes and haplotypes in the clinical groups; and the outcome of associations of *IL7* genetic variants with susceptibility to inefficient erythropoiesis and SMA to complete objective 1.

4.1.1. Demographic, clinical and laboratory characteristics of the study participants

The demographic, clinical, and laboratory characteristics of the study participants are presented in Table 4.1. Gender was comparable between the two clinical categories (P=0.72). Children with UM/non-SMA and SMA differed by age, with the UM group being older than children with SMA (P=0.01). Axillary temperature differed between the two clinical groups (P<0.02). As expected based on the *a priori* grouping, Hb concentrations and RBC counts were lower in SMA patients in relation to the UM group (P<0.01 and P<0.01, respectively). Red cell distribution width (RDW), mean corpuscular hemoglobin concentration (MCHC), and reticulocyte production index (RPI) were comparable between the clinical groups (P=0.48, P=0.79 and P=0.42, respectively). Proportion of children with RPI<2.0 was marginally lower in children with SMA (P=0.07). This result is not expected however the etiology of SMA is complex and multi-factoral and thus other unprecedented factors could be at play *(Perkins et al. 2011).* Inefficient erythropoiesis has a number of etiologies some of which were tested and others not directly tested. Comparison of the UM and SMA groups also revealed that white blood cell (WBC) counts were marginally elevated

in children with SMA (P=0.08). Parasite density was significantly lower in children with SMA (P=0.03), while the proportions of children with high-density parasitemia (HDP: ≥10,000 parasites/μL) did not differ between the groups (P=0.11). The proportion of children with α-thalassemia ($-\alpha^{-3.7}/-\alpha^{-3.7}$), G-6-PD deficiency, HbAS were comparable between the groups (P=0.42, P=0.45, and P=0.62, respectively).

Table 4.1. Demographic, clinical and laboratory characteristics of the study participants

Characteristic	UM (n=718)	SMA (n=165)	*P*-value
Demographic parameters			
Sex: *Male,* [n (%)]	356 (49.6)	79 (47.7)	0.72[a]
Female, [n (%)]	362 (50.4)	86 (52.3)	
Age (months)	12.2 (7.3-18.4)	10.0 (6.3-16.7)	**0.01[b]**
Clinical parameters			
Axillary temperature (ºC)	37.6 (36.8-38.1)	38.0 (37.0-38.4)	**<0.02[b]**
Laboratory parameters			
Hemoglobin (g/dL)	7.8 (6.2-9.1)	4.3 (3.7-4.7)	**<0.01[b]**
RBC count ($\times 10^{12}$/µL)	3.7 (2.8-4.5)	2.9 (2.2-4.1)	**<0.01[b]**
RDW	20.3 (18.1-23.0)	20.7 (18.0-23.3)	0.48[b]
MCHC	30.7 (29.6-32.3)	30.6 (29.5-32.3)	0.79[b]
RPI	0.5 (0.3-1.1)	0.5 (0.29-0.98)	0.42[b]
RPI <2.0, [n (%)]	605 (89.5)	136 (94.4)	0.07[a]
WBC count ($\times 10^3$/µL)	11.9 (9.0-15.9)	12.7 (9.5-17.0)	0.08[b]
Parasite density (per µL)	30,424.60 (7163.1-91570.0)	24,180.00 (4605.7-71849.7)	**0.03[b]**
HDP(≥10000 parasites/µL) [n (%)]	506 (70.7)	106 (64.2)	0.11[a]
$-\alpha^{3.7}/-\alpha^{3}$ thalassaemia, [n (%)]	130 (21.0)	24 (17.1)	0.42[a]
G-6-PD deficiency, [n (%)]	25 (4.0)	5 (3.5)	0.45[a]
Sickle cell trait, [n (%)]	102 (14.9)	21 (14.1)	0.62[a]

[a]Data are median (interquartile range) or n (%); Parasitemic children were categorized into severe malarial anemia (SMA, Hb<5.0 g/dL) and UM (uncomplicated malaria, Hb ≥5.0 g/dL, both with any density parasitemia; RBC-red blood cell; RDW-red cell distribution width; MCHC- mean corpuscular hemoglobin concentration; WBC-white blood cell; HDP-high density parasitemia; G-6-PD, glucose-6-phosphate dehydrogenase deficiency; [a]Statistical significance was determined by Chi-square analysis; n=Number; [b]Statistical significance determined by Mann–Whitney U test and P-values ≤0.05 were considered significant and are indicated in bold .

4.1.2. Distribution of *IL7* genotypes and haplotypes in the clinical groups

The results are summarized in Table 4.2. Following genotyping, SNP calling was successful for the two *IL7* variants tested, with both genotypes present in all 883 study participants. The MAF for the study population is comparable to that from the Yoruba population (HapMap Project) and the African population (1000 Genome Project) (0.33 and 0.10, for IL7-72194T>C and *IL7*-2440A>G, respectively. The genotypic proportions for *IL7*-72194T>C (P=0.98), and -2440A>G (P=0.51) were comparable between UM and SMA groups contrary to the expectation which was that the proportions would be higher in SMA indicating that the signals from these loci may have been false due to the relatively small sample size in the GWAS. The Hardy Weinberger Equilibrium (HWE) analyses showed a departure for both the -72194T>C (P=0.83) and -2440A>G (P=0.81) SNPs in the overall population. There was departure from HWE at the *IL7*72194T>C variant in both the UM (P<0.01) and SMA groups (P<0.01). There was also a departure from HWE at the *IL7* -2440A>G variant in both the UM (P<0.01) and SMA groups (P<0.01). The distribution of haplotypes were comparable between UM and SMA groups except for TG haplotype where the proportion of carriers were higher in the UM relative to the SMA (P=0.03). The genotype proportions for *IL7*-72194T>C (P=0.63), and -2440A>G (P=0.44) were comparable between EE and IE groups contrary to the expectation which was that the proportions would be higher in IE. There was departure from HWE at the *IL7*72194T>C variant in both the EE (P<0.01) and EI groups (P<0.01). There was also a departure from HWE at the *IL7* -2440A>G variant in both the EE (P<0.01) and EI groups (P<0.01). The distribution of haplotypes was comparable between EE and EI groups.

Table 4.2. Distribution of *IL7* genotypes and haplotypes in the clinical groups

Genetic Variants	Total (n=883)	UM (n=718)	SMA (n=165)	*P* value	EE (n=142)	IE (n=741)	*P* value
Genotypes							
***IL7*2194T>C**							
TT	391(44.30)	318(44.30)	73 (44.20)	0.98	71(50.0)	320(43.18)	0.63
TC	399(45.20)	325(45.30)	74 (44.80)		50(35.21)	349(47.09)	
CC	93 (10.50)	75 (10.40)	18 (10.90)		21(14.78)	72(9.71)	
Allele Fr.(p/q)	0.77/0.33	0.63/0.27	0.69/0.31		0.64/0.26	0.61/0.39	
HWE *P* value	0.83	<0.01	<0.01		<0.01	<0.01	
***IL7*2440A>G**							
AA	733(83.00)	601(83.70)	132(80.00)	0.51	121(85.21)	612(82.59)	0.44
AG	142(16.10)	111(15.50)	31 (18.80)		18(12.67)	124(16.73)	
GG	8 (0.90)	6 (0.80)	2 (1.20)		3(2.11)	5(0.06)	
Allele Fr. (p/q)	0.90/0.10	0.73/0.27	0.71/0.29		0.89/0.11	0.74/0.26	
HWE *P*-value	0.81	<0.01	<0.01		<0.01	<0.01	
Haplotypes							
TA							
Non –carriers	99 (11.20)	79 (11.00)	20 (12.10)	0.68	21(14.78)	78(10.52)	0.32
Carriers	784(88.70)	639(89.00)	145(87.90)		121(85.21)	663(89.47)	
TG							
Non –carriers	821(93.00)	674(93.90)	147(89.10)	**0.03**	113(79.57)	708(95.54)	0.21
Carriers	62 (7.00)	44 (6.10)	18 (10.90)		29(20.42)	33(4.45)	
CA							
Non –carriers	468(53.10)	382(53.20)	86(52.10)	0.80	94(66.19)	374(50.47)	0.52
Carriers	415(46.90)	336(46.80)	79 (47.90)		48(33.80)	367(49.52)	
CG							
Non –carriers	792(89.70)	642(89.40)	150(90.90)	0.57	110(77.46)	682(92.03)	0.77
Carriers	91 (10.30)	76 (10.60)	15 (9.10)		32(22.53)	59(7.96)	

Data are presented as proportions [n, (%)] of IL7 genotypes (72194T>C and −2440A>G) and IL7 haplotypes (TA, TG, CA, and CG); Parasitemic children were categorized into severe malarial anemia (SMA, Hb<5.0g/dL) and uncomplicated malaria (UM, Hb≥5.0g/dL); Parasitemic children were also categorized into efficient erythropoiesis (EE) (Reticulocyte production index (RPI)≥2, n=142) and inefficient erythropoiesis (IE) (RPI<2.0, n=741); Statistical significance was determined by Chi-square analysis comparing UM versus SMA and EE versus IE; Allele Fr.-Allele frequency; P-major allele and q-minor allele; HWE Hardy–Weinberg Equilibrium; P-values ≤0.05 were considered significant and are indicated in bold; and n-Number.

4.1.3. The association of genetic variants in *IL7* (72194 T>C) and (-2440 A>G) with inefficient erythropoiesis and SMA

The association of genetic variants in *IL7* with inefficient erythropoiesis and SMA were performed using bivariate logistic regression controlling for potential confounding factors [age, gender, HIV-1 and bacteremia status, hemoglobinopathies (sickle-cell traits, G-6-PD and $\alpha^{3.7}$-thalassemia variants)]. Inefficient erythropoiesis and SMA were the dependent variables while *IL7* variants were the independent variables. The results are summarized in Table 4.3. Inheritance of the 72194TC (OR=1.90; 95%CI=1.09-3.30; *P*=0.02) and CC (OR=5.14; 95%CI=1.20-21.99; *P*=0.03) genotypes enhanced susceptibility to inefficient erythropoiesis. However, relative to -2440AA carriage, inheritance of the AG (OR=1.39; 95%CI=0.63-3.03; *P*=0.42) genotype did not significantly influence susceptibility to inefficient erythropoiesis. The sample size in the clinical groups for determining susceptibility to inefficient erythropoiesis in the GG group was too small for logistic regression analysis. Relative to 72194TT carriers, neither TC [Odds ratio (OR) =0.93; 95% confidence interval (CI) =0.61-1.42; *P*=0.73)] nor CC (OR=1.34; 95%CI=0.69-2.58; *P*=0.39) genotypes influenced susceptibility to SMA. Similarly, relative to -2440AA inheritance, AG (OR=1.24; 95%CI=0.74-2.09; *P*=0.41) and GG carriers (OR=0.80; 95%CI=0.09-7.13; *P*=0.84) did not alter susceptibility to SMA.

More so, carriage of the TA haplotype was associated with protection against inefficient erythropoiesis (OR=0.24; 95%CI=0.06-1.21; *P*=0.05), whereas individuals with the CA haplotype were at an increased risk of inefficient erythropoiesis (OR=1.90; 95%CI=1.10-3.30; *P*=0.02). Carriage of the TG (OR=1.17; 95%CI=0.40-3.44; *P*=0.78) and CG (OR=1.66; 95%CI=0.57-4.80; *P*=0.35) haplotypes were not associated with altered susceptibility to inefficient erythropoiesis. The results also showed that relative to non-haplotype carriers, TA

(OR=0.70; 95%CI=0.38-1.29; *P*=0.25), TG (OR=1.29; 95%CI=0.61-2.71; *P*=0.50), CA (OR=1.04; 95%CI=0.69-1.56; *P*=0.85), and CG (OR=0.94; 95%CI=0.48-1.85; *P*=0.86) haplotypic carriage had no impact on susceptibility to SMA.

Table 4.3. Association of genetic variants in *IL7* with malaria infection outcomes

Genetic variants	Inefficient erythropoiesis (n=714)				SMA (n=165)			
	N (%)	**OR**	**CI**	***P* value**	**N (%)**	**OR**	**CI**	***P* value**
Genotypes								
***IL7* 72194T>C**								
TT^{Ref}	320(43.18)				73(44.20)			
TC	349(47.09)	1.90	1.09-3.30	**0.02**	74(44.80)	0.93	0.61-1.42	0.73
CC	72(9.71)	5.14	1.20-21.99	**0.03**	18(10.90)	1.34	0.69-2.58	0.39
***IL7* -2440 A>G**								
AA^{Ref}	612(82.59)				132(80.00)			
AG	124(16.73)	1.38	0.63-3.03	0.42	31 (18.80)	1.24	0.74-2.09	0.41
GG	5(0.06)	-	-	-	2 (1.20)	0.8	0.09-7.13	0.84
Haplotypes								
TA	663(89.47)	0.24	0.06-1.21	**0.05**	145(87.90)	0.70	0.38-1.29	0.25
TG	33(4.45)	1.17	0.40-3.44	0.78	18(10.90)	1.29	0.61-2.71	0.50
CA	367(49.5)	1.90	1.10-3.30	**0.02**	79(47.90)	1.04	0.69-1.56	0.85
CG	59(7.96)	1.66	0.57-4.80	0.35	15 (9.10)	0.94	0.48-1.85	0.86

Data are presented as odd ratios (OR) and 95% confidence interval (CI) determined by bivariate logistic regression analyses controlling for age, sex, HIV-1 and bacteremia status, $\alpha^{3.7}$-thalassemia, G-6-PD deficiency, and sickle-cell status; Parasitemic children were categorized into uncomplicated malaria (UM, Hb≥5.0 g/dL, n=718) and severe malaria anemia (SMA, Hb<5.0 g/dL, n=165) and efficient erythropoiesis (Reticulocyte production index (RPI)≥2, n=142) and inefficient erythropoiesis (RPI<2.0, n=741); Ref-reference group; Individuals with the variant in the clinical group [N (%)]; and P-values ≤0.05 were considered significant (indicated in bold).

4.2. Specific objective 2: To Establish the Influence of Genetic Variants in CSF2 (-7032 G>A and 64544 T>C) on the Rate of Malaria infections and SMA Episodes

This section (4.2) presents the demographic, clinical and laboratory characteristics of the study participants for objectives 2, 3, and 4; distribution of *CSF2* genotypes, haplotypes and diplotypes in the clinical groups; and the outcome of the influence of genetic variants in *CSF2* on the rate of malaria infections and SMA episodes to complete objective 2.

4.2.1. Demographic, clinical, and laboratory characteristics of study participants

These data are summarized in Table 4.4. The distribution of sex was comparable (P=0.69) across the groups. Children with SMA were younger (P=0.02) relative to aparasitemic and non-SMA groups, but the difference was not significant after correction for multiple testing. Axillary temperature differed across the groups and was higher in children with non-SMA (P<0.01), compared to the aparasitemic or SMA. Children were stratified *a priori* based on Hb concentrations. As such, relative to both the aparasitemic and non-SMA groups, children with SMA presented with significantly decreased hematocrit levels (P<0.01). White blood cell (WBC) counts progressively increased across the groups (P<0.01), as did the mean corpuscular volume (P<0.01), with the SMA group having the highest levels. Parasite densities were lower in children with SMA relative to non-SMA (P<0.01). The presence of HIV-1 did not significantly differ across groups (P=0.07), nor did bacteremia status (P=0.15). However, the distribution of sickle-cell genotypes differed across groups (P<0.01), with a lower frequency of HbAS carriage in children with non-SMA and lowest frequencies among children with SMA. The distribution of $\alpha^{3.7}$-thalassemia deletion variants were also different across the groups (P=0.04), with the highest frequencies of single ($-\alpha^{3.7}/\alpha\alpha$) and double-deletions ($-\alpha^{3.7}/-\alpha^{3.7}$) in children with SMA

Table 4.4. Demographic, clinical and laboratory characteristics of study participants

Characteristics	N	Total (n=1654)	Aparasitemic (n=335)	Non-SMA (n=1029)	SMA (n=290)	*P*-value
Demographic parameters						
Sex [n, (%)]						
Female	1654	822 (49.70)	168 (50.15)	504 (48.98)	150 (51.72)	0.69[a]
Male		832 (50.30)	167 (49.85)	525 (51.02)	140 (48.28)	
Age, months	1651	13.85 (8.14)	13.82 (8.67)	14.20 (7.88)	12.64 (8.34)	**0.02[b]**
Clinical parameters						
Axillary temperature,°C	1636	37.67 (1.06)	37.27 (1.00)	37.80 (1.09)	37.69 (0.89)	**<0.01[b]**
Laboratory Parameters						
Hemoglobin, g/dL	1639	7.47 (2.51)	9.40 (2.63)	7.79 (1.82)	4.13 (0.73)	NA
Hematocrit, %	1609	24.34 (7.72)	29.52 (8.0)	25.40 (5.76)	14.22 (3.88)	**<0.01[b]**
White Blood Cells, $\times 10^3/\mu L$	1602	13.40 (6.98)	13.06 (8.11)	12.77 (5.52)	16.16 (9.39)	**<0.01[b]**
Mean corpuscular volume, fL	1603	70.10 (9.30)	69.87 (8.41)	68.89 (8.48)	74.90 (11.46)	**<0.01[b]**
Parasite density, MPS/µL	1644	55716.67 (111089.67)	0.00 (0.00)	72983.96 (124507.3)	59082.88 (103946.74)	**<0.01[e]**
HIV-1, [n (%)]						
Negative	1641	1573 (95.86)	315 (96.04)	990 (96.49)	268 (93.38)	0.07[c]
Positive		68 (4.14)	13 (3.96)	36 (3.51)	19 (6.62)	
Bacteremia						
Negative	1639	1518 (92.62)	295 (90.77)	960 (93.57)	263 (91.32)	0.15[a]
Positive		121 (7.38)	30 (9.23)	66 (6.43)	25 (8.68)	
Sickle cell trait, [n (%)]						
Hb AA	1619	1358 (83.88)	252 (78.26)	848 (83.38)	258 (92.14)	**<0.01[c]**
Hb AS		243 (15.01)	60 (18.63)	165 (16.22)	18 (6.43)	
Hb SS		18 (1.11)	10 (3.11)	4 (0.39)	4 (1.43)	
α^+-thalassemia deletion, [n (%)]						
αα/αα	1405	596 (42.42)	120 (43.96)	375 (42.86)	101 (39.3)	**0.04[a]**
-$\alpha^{3.7}$/αα		526 (37.44)	84 (30.77)	335 (38.29)	107 (41.63)	
-$\alpha^{3.7}$/-$\alpha^{3.7}$		283 (20.14)	69 (25.27)	165 (18.86)	49 (19.07)	

Data are presented as mean (standard error of mean) unless otherwise stated; MPS-malaria parasites; HIV-human immunodeficiency virus; Aparasitemic (without MPS); Parasitemic children were categorized into non-SMA (Hb≥5.0 g/dL) or SMA (Hb<5.0 g/dL); [a]Pearson's Chi-squared test; [b]One-way analysis of variance (ANOVA); [c]Fisher's Exact Test (used when cell count is <20); [e]Student's t test; and P-values ≤0.05 were considered significant (indicated in bold).

4.2.2. Distribution of CSF2 genotypes, haplotypes and diplotypes in the clinical groups

These results are summarized in Table 4.5. Following genotyping, SNP calling was successful for the two genetic variants in *CSF2* tested with both genotypes present in 1,203 out of 1,654 study participants. To increase the power of analysis, no imputations were performed to infer genotypes, haplotypes or diplotypes. As such, only data with both SNPs present for each participant were utilized for subsequent statistical analyses. The genotypic proportions for *CSF2*-7032G>A (P=0.99), and 64544T>C (P=0.22) were comparable across the groups. The MAF for the study population is comparable to those from the International HapMap Project and 1000 Genomes Project. The HWE analyses showed a departure for the 64544T>C SNP for the overall population, and in the aparasitemic, non-SMA, and SMA groups. There was no departure from HWE for the *CSF2*-7032G>A polymorphic variant. The distribution of haplotypes (P=0.55), and diplotypes (P=0.32) were not significantly different across the study groups.

Table 4.5. Distribution of *CSF2* genotypes, haplotypes and diplotypes in the clinical groups

Variants	Total n=1203	Aparasitemic n=250	Non-SMA n=772	SMA n=181	*P* value
Genotypes					
***CSF2*-7032 G>A**					
GG	636 (52.87)	134 (53.60)	407 (52.72)	95 (52.49)	0.99[a]
GA	476 (39.57)	98 (39.20)	305 (39.51)	73 (40.33)	
AA	91 (7.56)	18 (7.20)	60 (7.77)	13 (7.18)	
Allele frequency (p/q)	0.72/0.28	0.73/0.27	0.72/0.28	0.72/0.28	
HWE, *P* value	0.88	1.00	0.78	1.00	
***CSF2* 64544 T>C**					
TT	984 (81.80)	196 (78.40)	636 (82.38)	152 (83.98)	0.22[a]
TC	125 (10.39)	35 (14.00)	77 (9.97)	13 (7.18)	
CC	94 (7.81)	19 (7.60)	59 (7.64)	16 (8.84)	
Allele frequency (p/q)	0.87/0.13	0.87/0.13	0.85/0.15	0.87/0.13	
HWE, *P* value	<0.01	<0.01	<0.01	<0.01	
Haplotypes					
GT	1470 (61.10)	297 (59.40)	949 (61.50)	224 (61.90)	0.55[a]
GC	278 (11.60)	69 (13.80)	170 (11.00)	39 (10.80)	
AT	623 (25.90)	130 (26)	400 (25.90)	93 (25.70)	
AC	35 (1.50)	4 (0.80)	25 (1.60)	6 (1.70)	
Diplotypes					
GT/GT	500 (41.56)	93 (37.20)	331 (42.88)	76 (41.99)	0.32[a]
GT/GC	71 (5.90)	25 (10)	38 (4.92)	8 (4.42)	
GC/GC	65 (5.40)	16 (6.40)	38 (4.92)	11 (6.08)	
GT/AT	399 (33.17)	86 (34.4)	249 (32.25)	64 (35.36)	
GT/AC	27 (2.24)	3 (1.20)	19 (2.46)	5 (2.76)	
GC/AC	85 (7.07)	17 (6.80)	56 (7.25)	12 (6.63)	
AT/AT	50 (4.16)	9 (3.60)	37 (4.79)	4 (2.21)	
AT/AC	4 (0.33)	1 (0.40)	2 (0.26)	1 (0.55)	
AC/AC	2 (0.17)	0 (0.00)	2 (0.26)	0 (0.00)	

Data are presented as proportions [n, (%)]. Children were categorized into aparasitemic controls (n=250; no parasitemia) and parasitemic categorized into severe malaria anemia (SMA Hb≥5.0g/dL, n=772) or non SMA (Hb<5.0g/dL, n=181); [a]Fisher's Exact test; MPS–Malaria parasites; CSF2 -colony stimulating factor 2; P major allele and q minor allele; and HWE Hardy–Weinberg Equilibrium; and P-values ≤0.05 were considered significant (indicated in bold).

4.2.3. The Influence of genetic variants in *CSF2* (7032G>A and 64544T>C) on the rate of malaria Infections and SMA episodes

4.2.3.1. The influence of genetic variants in *CSF2* on the rate of malaria infections

Factors associated with the rate of malaria episodes over a 36 month follow-up period were determined using a generalized linear model (Poisson regression). Results for the covariates that emerged for the 6029 recorded malaria episodes over follow-up period are shown in Table 4.6. Protective effects against the risk of malaria were present in heterozygous carriers of the *CSF2*64544TC genotype [IRR=0.33 (CI=0.12-0.88), *P*=0.02] relative to the wild type genotype TT. Similarly, *CSF2*AC/GC diplotype was associated with decreased malaria episodes [IRR = 0.73 (CI=0.60-0.88), *P*<0.01] compared to wild type diplotype GT/GT. Conversely, co-inheritance of the *CSF2*AT/GC diplotype increased the rate of malaria episodes [IRR=3.13 (1.16-8.42), *P*=0.02]. Carriers of the *CSF2*GC/GT diplotype, however, had a borderline statistical significance [IRR=2.52 (CI=0.94-6.78), *P*=0.06].

Table 4.6. The influence of genetic variants in *CSF2* on the rate of malaria infections

Variable names	z value	IRR (95% CI)	*P* value
Age at enrollment (younger)	-18.12	0.96 (0.96 - 0.97)	**<0.01**
HIV-1 (+)	-3.04	0.72 (0.59 - 0.89)	**<0.01**
Sex (Female)	-3.22	0.91 (0.86 - 0.96)	**<0.01**
$-\alpha^{3.7}/-\alpha^{3.7}$ (double deletion)	-2.70	0.90 (0.84 - 0.97)	**<0.01**
Hb SS (sick cell anemia)	-4.15	0.47 (0.33 - 0.67)	**<0.01**
Hb AS (Sickle cell trait)	-5.45	0.79 (0.73 - 0.86)	**<0.01**
*CSF2*64544TC vs TT genotype	-2.20	0.33 (0.12 - 0.88)	**0.02**
*CSF2*GC/GT vs GT/GT diplotype	1.83	2.52 (0.94 – 6.78)	0.06
*CSF2*AT/GC vs GT/GT diplotype	2.26	3.13 (1.16 - 8.42)	**0.02**
*CSF2*AC/GC vs GT/GT diplotype	-3.17	0.73 (0.60 - 0.88)	**<0.01**

Poisson regression model fit was determined using Akaike's information criterion (AIC), with the heuristic approach being performed on an iterative manner to exclude potentially irrelevant variables; The relevant genetic variables that emerged were compared against the wild type; Data are ranked per variables as follows; metric variables (age), categorical variables (cohort, --1), genetic variables (sex, $-\alpha^{3.7}/-\alpha^{3.7}$,Hb SS, Hb AS) and CSF2 gene variants; A positive z value indicates a reduced risk, whereas a negative z value indicates an increased risk; IRR-Incidence rate ratio; CI-Confidence interval; HIV-1-Human immunodeficiency virus-1; and CSF2 -colony stimulating factor 2; and P-values ≤0.05 were considered significant (indicated in bold).

4.2.3.2. The influence of genetic variants in *CSF2* on the rate of SMA episodes

The effect of the genetic variants in *CSF2* on the rate of SMA episodes over 36 months follow-up period was also investigated using a Poisson regression. Factors associated with the rate of SMA for the 297 episodes that occurred over the follow-up period are shown in Table 4.7. The only *CSF2* variant found to influence the rate of SMA episodes was carriage of the *CSF2*AT haplotype, which increased the risk of SMA episodes [IRR=1.33 (CI=1.05-1.69), *P*=0.02] compared to the wild type GT haplotype.

Table 4.7. The influence of genetic variants in *CSF2* on rate of SMA episodes

Variable names	z value	IRR (95% CI)	*P* value
Age at enrollment (younger)	-6.11	0.94 (0.92- 0.96)	**<0.01**
HIV- 1 (+)	4.43	2.88 (1.80 - 4.61)	**<0.01**
Hb SS (sickle cell anemia)	3.03	2.81 (1.44 - 5.49)	**<0.01**
Hb AS (sickle cell trait)	-3.01	0.51 (0.32 - 0.79)	**<0.01**
CSF2 AT vs GT haplotype	2.39	1.33 (1.05 - 1.69)	**0.01**

Poisson regression model fit was determined using Akaike's information criterion (AIC), with the heuristic approach being performed on an iterative manner to exclude potentially irrelevant variables; The relevant genetic variables that emerged were compared against the wild type; Data are ranked per variables as follows; metric variables (age), categorical variables (cohort, HIV-1), genetic variables (sex, -$\alpha^{3.7}$/-$\alpha^{3.7}$,Hb SS, Hb AS) and CSF2 gene variants; A positive z value indicates a reduced risk, whereas a negative z value indicates an increased risk; IRR-Incidence rate ratio; CI-Confidence interval; HIV-1-Human immunodeficiency virus-1; and CSF2 -colony stimulating factor 2; and P-values ≤0.05 were considered significant (indicated in bold).

4.3. Specific objective 3: To determine the Relationship Between Genetic Variants in CSF2 and the Risk of Malaria Infections and SMA Episodes

4.3.1. The relationship between genetic variants in *CSF2* and the risk of malaria infections

After determining the impact of *CSF2* variants on the rate of malaria and SMA episodes, the effect of the variants on the time between events (risk) was then determined using an ordered-events, generalized Cox proportional hazard model. Table 4.7 shows the relevant covariates that emerged from the model. Although retained in the model, co-inheritance of the *CSF2* GC/GT [HR=0.85 (CI=0.75-0.97), *P*=0.11] compared to wild type (GT/GT) were not statistically significant predictors for the time between malaria infections. However, carriage of the *CSF2* AC haplotype was protective against malaria infections [HR=0.72 (CI=0.60-0.86), *P*<0.01] relative to wild type haplotype (GT).

Table 4.8. The relationship between genetic variants in CSF2 and the risk of malaria infections

Variable names	Coefficient	HR (95% CI)	*P* value
Age at first hospital visit (younger)	-0.04	0.96 (0.95 - 0.96)	**<0.01**
HIV- 1 (+)	-0.19	0.82 (0.67 - 1.00)	0.17
Sex (Female)	-0.08	0.92 (0.87 - 0.97)	0.05
$-\alpha^{3.7}/-\alpha^{3.7}$ (double deletion)	-0.09	0.91 (0.85 - 0.97)	0.08
Hb SS (sickle cell anemia)	-0.72	0.48 (0.34 - 0.68)	**<0.01**
Hb AS (sickle cell trait)	-0.24	0.78 (0.73 - 0.85)	**<0.01**
CSF2 AC vs GT haplotype	-0.32	0.72 (0.60 - 0.86)	**<0.01**
CSF2 GC/GT vs GT/GT diplotype	-0.15	0.85 (0.75 - 0.97)	0.11

Cox proportional hazard model fit to investigate on the time-to-event of covariates for malaria infections determined using Akaike's information criterion (AIC), with the heuristic approach being performed on an iterative manner to exclude potentially irrelevant variables; The relevant genetic variables that emerged were compared against the wild type; Data are ranked per variables as follows: the metric variables (age at first hospital visit), followed by categorical variables (cohort, HIV-1), genetic variables (sex, $-\alpha^{3.7}/-\alpha^{3.7}$, HbSS, HbAS) and CSF2 genetic variants; A positive coefficient indicates a worse prognosis, whereas a negative coefficient indicates a better prognosis; HR-Hazard ratio; CI-confidence interval; HIV-1-Human immunodeficiency virus-1; and CSF2 -colony stimulating factor 2; and P-values ≤0.05 were considered significant (indicated in bold).

4.3.2. The relationship between genetic variants in *CSF2* and the risk of SMA episodes

Cox proportional hazard model was also used to examine the influence of the two *CSF2* genetic variants on the risk of malaria infections that culminated in the development of SMA (Table 4.7). There were 248 SMA events during the 36 month follow-up period. Consistent with the Poisson models examining the rates of SMA, time-to-event modeling revealed that although not statistically significant, there was a trend towards increased risk of SMA in children with the *CSF2* AT haplotype [HR=1.26 (CI=0.98-1.62), *P*=0.08] relative to the wild type GT.

Table 4.9. The relationship between genetic variants in CSF2 and the risk of SMA episodes

Variable names	Coefficient	HR (95% CI)	*P* value
Age at first hospital visit (younger)	-0.04	0.95 (0.93 - 0.97)	**<0.01**
HIV -1 (+)	1.05	2.85 (1.76 - 4.64)	**<0.01**
Hb SS (sickle cell anemia)	0.99	2.68 (1.32 - 5.46)	**0.03**
Hb AS (sickle cell trait)	-0.70	0.49 (0.30 - 0.78)	**0.01**
CSF2 AT vs GT haplotype	0.23	1.26 (0.98 - 1.62)	0.08

Cox proportional hazard model fit to investigate on the time-to-event of covariates for malaria infections determined using Akaike's information criterion (AIC), with the heuristic approach being performed on an iterative manner to exclude potentially irrelevant variables; The relevant genetic variables that emerged were compared against the wild type; Data are ranked per variables as follows: the metric variables (age at first hospital visit), followed by categorical variables (cohort, HIV-1), genetic variables (sex, -$\alpha^{3.7}$/-$\alpha^{3.7}$, HbSS, HbAS) and CSF2 genetic variants; A positive coefficient indicates a worse prognosis, whereas a negative coefficient indicates a better prognosis; HR-Hazard ratio; CI-confidence interval; HIV-1-Human immunodeficiency virus-1; and CSF2 -colony stimulating factor 2; and P-values ≤0.05 were considered significant (indicated in bold).

4.4. Specific objective 4: To Establish the Relationship Between Genetic Variants in CSF2 and the Risk of All-cause Mortality

The influence of covariates on the risk of all-cause mortality was investigated using a Cox proportional hazard model. The results are presented in Table 4.7. Although the burden of malaria and co-infections is high in Siaya community, the current study reported an all-cause mortality rate of 5.98% which is lower than earlier reports and may be associated with improvement in healthcare services provision. Most deaths (64%) occurred at the children's homes and were recorded during the quarterly visits over the 36 months follow-up period. The data also revealed that inheritance of the *CSF2*-7032GA genotype significantly increased the risk of dying [HR=1.88 (CI=1.05-3.34), *P*=0.03], compared to the wild type AA. Carriage of the *CSF2* GC/GC diplotype, [HR=2.57 (CI=0.96-6.86), *P*=0.05] also significantly increased the risk of dying relative to the wild type GT/GT.

Table 4.10. The relationship between genetic variants in CSF2 and risk of all-cause mortality

Variable names	Coefficient	HR (95% CI)	*P* value
Age at enrollment (younger)	-0.07	0.93 (0.89 - 0.97)	**<0.01**
HIV -1 (+)	2.80	16.45 (8.79 - 30.77)	**<0.01**
Hb SS (sickle cell anemia)	1.88	6.56 (2.00 - 21.50)	**<0.01**
Hb AS (sickle cell trait)	-1.21	0.29 (0.09 - 0.96)	**0.04**
CSF2 -7032 GA genotype	0.63	1.88 (1.05 - 3.34)	**0.03**
CSF2 GC/GC diplotype	0.94	2.57 (0.96 - 6.86)	**0.05**

Cox proportional hazard model fit to predict all-cause mortality outcomes determined using Akaike's information criterion (AIC), with the heuristic approach being performed on an iterative manner to exclude potentially irrelevant variables; The relevant genetic variables that emerged were compared against the wild type; Data are ranked per variables as follows: the metric variables (age at first hospital visit), followed by categorical variables (cohort, HIV-1), genetic variables (sex, -$\alpha^{3.7}$/-$\alpha^{3.7}$, HbSS, HbAS) and CSF2 genetic variants; A positive coefficient indicates a worse prognosis, whereas a negative coefficient indicates a better prognosis; HR-Hazard ratio; CI-confidence interval; HIV-1-Human immunodeficiency virus-1; and CSF2 -colony stimulating factor 2; and P-values ≤0.05 were considered significant (indicated in bold).

CHAPTER FIVE

DISCUSSION

5.1. Introduction: Advances in human gene mapping, along with an increased understanding of the molecular mechanisms of protective immunity, illustrate that susceptibility to malaria and its clinical outcomes is conditioned by genotypic variation *(Perkins et al. 2013).* In holoendemic regions like Siaya in western Kenya, prior to the development of naturally-acquired anti-malarial immunity, children experience multiple episodes of malaria infections that often culminate in life-threatening complications that include bone marrow suppression and the resultant SMA leading to numerous deaths *(Anyona et al. 2020; Ouma et al. 2012; Ouma et al. 2010).* This study focused on determining how genetic variation in cytokine genes *(IL7* and *CSF2)* identified through a pilot GWAS study *(Achieng et al. 2019)* influence the risk and rate of malaria infections, ineffective erythropoiesis, SMA and all-cause mortality in young children. The findings from this study are discussed below as per the study objectives.

5.2. Association of Genetic Variants in IL7 with Inefficient Erythropoiesis and SMA

The importance of *IL7* gene in conditioning inflammatory disease outcomes has been previously described *(Ghavimi et al. 2014; Zhang et al. 2016; Zuvich et al. 2010).* The current study is the first pediatric study to examine the association between the *IL7* polymorphisms and susceptibility to severe disease outcomes in malaria. First, the variant frequencies of the two loci were compared separately, and the prevalence of individual genotypes at both loci were found to be comparable between the UM and SMA groups. In addition, the single-locus MAFs (0.33 for 72194C and 0.09 for -2440G) in this population were the same as those for the Yoruban Nigerian population in the HapMap (i.e., 0.33 and 0.09, respectively). These results suggest that the investigated alleles remain largely

unchanged over time in different African ethnic groups. However, since the overall genotypic frequencies for the 72194T>C and -2440A>G SNPs in the UM and SMA groups departed from HWE, it is feasible that population selection pressures from malaria may be responsible for this observation.

To date, only one study has reported an association between *IL7* genetic variation and malaria. This study, in pregnant women from Mozambique, revealed an association between *IL7* SNPs (rs2583764 and rs2583762, 33 kb apart) and increased susceptibility to placental malaria *(Sikora et al. 2011).* These two SNPs are in close proximity to SNP 72194T>C (rs2583759) genotyped in the present study. This particular SNP (72194T>C), along with additional SNPs in the same haploblock (rs2583760, rs2583764 and rs6993386), were associated with susceptibility to osteoarthritis in the Chinese Han population *(Zhang et al. 2016).* Although there were significant associations between genotypes and SMA in the GWAS pilot data, this was not replicated in the validation sample. Thus, the signal for SMA in the pilot study may have been a false-positive and could be attributed to the relatively small sample size in the pilot GWAS study (n=144). Alternatively, the pilot study may have examined children with extreme phenotypes. By including children in the validation study that were parsed into two groups above and below Hb of 5.0 g/dL, there are many children with low Hbs, yet slightly above 5.0 g/dL in the UM group. This may have weakened the ability to identify a significant signal, and is consistent with the reality that there is not a clear-cut clinical difference between Hb levels of 6.0 vs. 5.0 g/dL *(Novelli et al. 2010).*

Results presented here indicate that carriage of the C allele at *IL7*72194T>C loci significantly increased susceptibility to inefficient erythropoiesis (i.e., RPI<2.0) in both heterozygous (TC) and homozygous (CC) carriers. In addition, the trend of reduced susceptibility (TT< TC<CC) in cases with inefficient erythropoiesis was maintained. However, variation at *IL7*72194T>C

was not significantly associated with susceptibility to SMA, although carriage of the CC allele marginally increased the risk of SMA. This finding suggests that the association between *IL7*72194 C allele carriage and erythrocyte production are more closely linked than the phenotypic expression of SMA in which reduced hemoglobin concentrations are a central feature. This result may not be unexpected since inefficient erythropoiesis is an overlapping feature of both UM (89.50%) and SMA (94.40%), albeit slightly higher in SMA. Although a primary cause of SMA in the region is suppression of erythropoiesis, hemolysis of parasitized and non-parasitized RBCs also contributes to the reduced hemoglobin levels (White 2018). Consistent with the importance of inefficient erythropoiesis in SMA, there was a significant relationship between RPI<2.0 and SMA. Further, there was no significant association between *IL7*–2440A>G inefficient erythropoiesis or SMA.

Since haplotypic carriage can provide additional insight, beyond analysis of the individual loci, *IL7* haplotypes were constructed from the two SNPs (72194T>C and -2440A>G). Previous studies on SMA demonstrated that haplotypes are highly informative allelic markers for identifying associations with disease outcomes, not identifiable with single polymorphisms *(Ouma et al. 2012; Ouma et al. 2008).* Findings from this study show that carriage of the TA haplotype (72194T/-2440A) was significantly associated with a reduced risk of inefficient erythropoiesis. This finding reflects the fact that carriage of the wild-type alleles at the two loci each trended towards better erythropoietic responses. Conversely, the relationship between carriage of the C allele at *IL7*72194 and inefficient erythropoiesis was reflected in the haplotypes in which the CA haplotype (72194C/-2440A) was significantly associated with an RPI<2.0 and the CG haplotype (72194C/-2440G) was associated with a non-significant increase in the risk of inefficient erythropoiesis. Similar to the findings for the individual genotypes, none of the haplotypes were associated with susceptibility to SMA.

However, the directionality of the risk profiles for inefficient erythropoiesis and SMA were consistent.

Since there were not enough children in the current study for which circulating IL-7 levels were available, the study could not explore the functional association between the *IL7* variants and cytokine production. Previous investigations have demonstrated that circulating levels of IL-7 are reduced in children with SMA, but are not significant predictors of hemoglobin concentrations *(Ong'echa et al. 2011).* Moreover, IL-7 promotes T cell-dependent activation of monocyte/macrophages which, in turn, release a number of cytokines (i.e., IL-1, IL-3, IL-6, IL-8, GM-CSF, macrophage inflammatory protein -1β, TNF-α, and interferon gamma) some of which if deregulated are associated with SMA *(Perkins et al. 2011; Winer et al. 2022).* Additionally, since IL-7 induces erythropoiesis, reduced levels in children with SMA may contribute to delayed and/or inappropriate erythroid development *(Aiello et al. 2007).*

This study presents the first report on the association between genotypes and haplotypes of *IL7* (72194T>C and -2440A>G) and disease outcomes in children with severe malaria: profoundly low hemoglobin levels and inefficient erythropoiesis. Results presented here suggest that the selected variants are associated with erythropoietic responses.

5.3. Influence of Genetic Variants in CSF2 on the Rate of Malaria infections and SMA episodes

Previous studies have shown that the host releases both pro- and anti-inflammatory cytokines, chemokines, growth factors, and effector molecules as part of the innate immune response to malaria infections *(Perkins et al. 2011).* Among the growth factors is GM-CSF, which has been shown to be highest among children with uncomplicated malaria relative to those with non-SMA and SMA *(Ong'echa et al. 2011).* However, children with SMA have higher levels

of circulating GM-CSF than individuals with non-SMA, suggesting a complicated pattern of production during acute infection *(Ong'echa et al. 2011).* Multiple studies have shown divergent results on the influence of GM-CSF on clinical outcomes during malaria infection. For example, elevated GM-CSF levels has been associated with severe malaria complications (i.e., splenomegaly and leukocytosis) in some investigations, while others have found a protective role for GM-CSF *(Kumaratilake et al. 1996).* Recent studies suggest that the production of TLR 7/8-driven GM-CSF in cord blood is an independent predictor of enhanced malaria risk over the first year of life, suggesting that GM-CSF indeed plays an important role in malarial immunity *(Natama et al. 2018).*

The allele frequencies for the two SNPs in the study cohort were comparable to those in the Yoruba population (HapMap Project) and the African population (1000 Genome Project), suggesting that the two loci have been steadily maintained in ethnic groups of African descent *(Anyona et al. 2011).* The *CSF2*-7032G>A variant displayed HWE, whereas the *CSF2*64544T>C locus had a significant departure from HWE in the overall population, and in each of the clinical groups investigated. Although a departure from HWE at *CSF2*64544T>C locus could be attributed to historical pressure from malaria (Eybpoosh 2018), there were comparable frequencies of the genotypes, as well as haplotypes and diplotypes in combination with *CSF2*-7032G>A across the clinical groups (i.e., aparasitemic, non-SMA, and SMA), suggesting an influence of a large sample size. This finding suggests that the two variants selected for investigation are likely not to be under strong selection pressure. The LD measures for the two loci with a $D'=0.41$ and $r^2=0.01$ indicate that the two variants are not strongly linked (Hosking *et al.* 2004)

A Poisson regression model estimating the rate of malaria episodes over 36 months revealed that the presence of the *CSF2*64544 TC heterozygote genotype and *CSF2* AC/GC diplotype

decreased the incidence of malaria, while the *CSF2* AT/GC diplotype increased the incidence of malaria episodes. Additional Poisson analysis identified significantly more episodes of SMA among children who inherited the *CSF2* AT haplotype. From these findings, *CSF2* variants could be influencing the rate of malaria and SMA episodes through imparting changes in soluble immune mediator production, a phenomenon that has been observed for many other immune response genes in severe malaria studies *(Anyona et al. 2011; Ouma et al. 2012; Ouma et al. 2010; Phawong et al. 2010).* The current results also complement previous studies illustrating that children living in regions with high transmission rates for *P. falciparum* experience multiple infections prior to developing malarial immunity, with the greatest burden of severe disease manifesting in children less than 5 years of age *(Achieng et al. 2019; Anyona et al. 2020; Raballah et al. 2022).* As such, longitudinal studies have a greater ability to detect genetic factors, and covariates that can influence both the rate and timing of repeated malaria episodes and the development of severe disease overtime.

5.4. The relationship between Genetic Variants in CSF2 and the Risk of Malaria Infections and SMA episodes

To further elucidate the impact of the *CSF2* genetic variants on the longitudinal risk of malaria, a Cox proportional hazard model was used and it revealed that carriage of the *CSF2* AC haplotype was associated with a lower hazard risk for malaria infections. Additionally, inheritance of the *CSF2* AT haplotype that significantly increased the incidence rate for SMA was associated with an elevated longitudinal hazard for SMA, albeit non-significant. Results obtained in this study support previous findings that showed innate immune response genes influence longitudinal susceptibility to malaria and subsequent development of SMA in this *P. falciparum* holoendemic region *(Achieng et al. 2019; Anyona et al. 2020; Raballah et al. 2022).*

5.5. Relationship Between genetic Variants in CSF2 and the Risk of All-cause Mortality

This study also investigated predictors of all-cause mortality across 36 months using Cox proportional hazard modeling and found that several *CSF2* variants were associated with increased susceptibility to all-cause mortality, including carriage of the *CSF2*-7032GA genotype (1.8 times higher) and the *CSF2* GC/GC diplotype (2.6 times higher). Interestingly, neither of these variants significantly impacted susceptibility to either malaria or SMA, suggesting that their impact on mortality may be due to non-malaria-related causes.

Based on current knowledge, this is the first study examining the impact of two genetic variants flanking the 5q31.1 gene region on longitudinal malaria disease outcomes and all-cause mortality. In summary, results presented here illustrate that the two *CSF2* variants investigated are associated with both the rate and timing of malaria and SMA episodes; and all-cause mortality during the development of natural-acquired antimalarial immunity in children living in a high malaria transmission region. However, the underlying molecular mechanism(s) for the disease associations remain largely unknown.

5.6. Study Limitations

Given the relatively low frequency of the G allele at the *IL7*-2440A>G locus, the analyses of this locus could not be performed using the regression analyses.

CHAPTER SIX

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

6.1. Summary of Findings

The current study findings demonstrate that SNPs in *IL7* [72194C>T and -2240A>G] and *CSF2* [-7032G>A and 64544T>C] are associated with susceptibility to pediatric malaria and SMA. In summary, carriage of the TC and CC genotypes at *IL7*72194 was associated with susceptibility to inefficient erythropoiesis. Additionally, carriers of CA haplotype [72194C/-2240A] had an increased risk of inefficient erythropoiesis whereas those with TA haplotype [72194T/-2240A] had marginal protection. For *CSF2* investigations, carriage of TC genotype at *CSF2*64544 and AC/GC diplotype decreased the rate of malaria episodes while presence of AT/GC diplotype and AT haplotype [-7032A/64544T] increased the rate of SMA episodes. Furthermore, carriage of the AC haplotype [-7032A/64544C] was associated with a lower risk of malaria episodes while the presence of AG genotype and the *CSF2*GC/GC diplotype was associated with increased risk to all-cause mortality.

6.2. Conclusions

i. An increased susceptibility to inefficient erythropoiesis was observed among carriers of TC and CC genotypes at *IL7*72194 and *IL7*CA haplotype.

ii. Decreased rates of falciparum malaria infections were observed in carriers of *CSF2*64544TC genotype and *CSF2*AC/GC diplotype while increased rates were observed for carriers of *CSF2*AT/GC diplotype. Additionally, *CSF2*AT haplotype increased the rate SMA.

iii. A decreased risk of *falciparum* malaria infections was observed among *CSF2*AC haplotype carriers.

iv. An increased risk of all-cause mortality was observed among *CSF2*GA genotype carriers with *falciparum* malaria infection.

6.3. Recommendations from the Current Study

i. *IL7* polymorphisms may be considered potential genetic modulators of susceptibility to inefficient erythropoiesis in children with *P. falciparum* malaria.
ii. *CSF2* polymorphisms may be considered as potential genetic predictors of the rate of malaria infections, SMA and mortality in children with *P. falciparum* malaria.
iii. *CSF2* polymorphisms may be considered as potential genetic predictors of the risk of malaria infections, SMA and mortality in children with *P. falciparum* malaria.
iv. *CSF2* polymorphisms may be considered as potential genetic predictors of the risk of all-cause mortality in children with *P. falciparum* malaria.

6.4. Recommendations for Future Studies

Based on the findings from this study, the following are suggested for subsequent studies:

i. Determine if the assessed genetic variants in *IL7* and *CSF2* functionally alter IL-7 and GM-CSF production, respectively in these children exposed to *P. falciparum* malaria.

ii. Find out, using *in vtiro* assays, if products derived from malaria parasite including antigens, soluble parasitic molecules and hemozoin can alter IL-7 and GM-CSF production in children exposed to *P. falciparum* malaria..

iii. Evaluate and validate specific genetic variants in IL7 and GM-CSF as biomarkers for clinical diagnosis of severe malaria outcomes.

REFERENCES

Achieng, A. O., N. W. Hengartner, E. Raballah, Q. Cheng, S. B. Anyona, N. Lauve, B. Guyah, I. Foo-Hurwitz, J. M. Ong'echa, B. H. McMahon, C. Ouma, C. G. Lambert, and D. J. Perkins. 2019. 'Integrated OMICS platforms identify LAIR1 genetic variants as novel predictors of cross-sectional and longitudinal susceptibility to severe malaria and all-cause mortality in Kenyan children', *EBioMedicine*, 45: 290-302.

Aiello, F. B., J. R. Keller, K. D. Klarmann, G. Dranoff, R. Mazzucchelli, and S. K. Durum. 2007. 'IL-7 induces myelopoiesis and erythropoiesis', *J Immunol*, 178: 1553-63.

Ait Yahia S, Azzaoui I, Everaere L, Vorng H, Chenivesse C, Marquillies P, Duez C, Delacre M, Grandjean T, Balsamelli J, Fanton d'Andon M, Fan Y, Ple C, Werts C, Boneca IG, Wallaert B, Chamaillard M, Tsicopoulos A. CCL17 production by dendritic cells is required for NOD1-mediated exacerbation of allergic asthma. Am J Respir Crit Care Med. 2014 Apr 15;189(8):899-908. doi: 10.1164/rccm.201310-1827OC. PMID: 24661094.

Akech, S., M. Chepkirui, M. Ogero, A. Agweyu, G. Irimu, M. English, and R. W. Snow. 2020. 'The Clinical Profile of Severe Pediatric Malaria in an Area Targeted for Routine RTS,S/AS01 Malaria Vaccination in Western Kenya', *Clin Infect Dis*, 71: 372-80.

Amek, N., N. Bayoh, M. Hamel, K. A. Lindblade, J. E. Gimnig, F. Odhiambo, K. F. Laserson, L. Slutsker, T. Smith, and P. Vounatsou. 2012. 'Spatial and temporal dynamics of malaria transmission in rural Western Kenya', *Parasit Vectors*, 5: 86.

Amek, N. O., F. O. Odhiambo, S. Khagayi, H. Moige, G. Orwa, M. J. Hamel, A. Van Eijk, J. Vulule, L. Slutsker, and K. F. Laserson. 2014. 'Childhood cause-specific mortality in rural Western Kenya: application of the InterVA-4 model', *Glob Health Action*, 7: 25581.

Ananth, N., B. V. Shetty, and D. M. Vasudevan. 2003. 'Possible role of Granulocyte Macrophage Colony Stimulating Factor receptor (GM-CSF R) in malaria', *Indian J Exp Biol*, 41: 357-9.

Angulo, I., and M. Fresno. 2002. 'Cytokines in the pathogenesis of and protection against malaria', *Clin Diagn Lab Immunol*, 9: 1145-52.

Anyona, S. B., N. W. Hengartner, E. Raballah, J. M. Ong'echa, N. Lauve, Q. Cheng, P. W. Fenimore, C. Ouma, C. G. Lambert, B. H. McMahon, and D. J. Perkins. 2020. 'Cyclooxygenase-2 haplotypes influence the longitudinal risk of malaria and severe malarial anemia in Kenyan children from a holoendemic transmission region', *J Hum Genet*, 65: 99-113.

Anyona, S. B., P. Kempaiah, E. Raballah, G. C. Davenport, T. Were, S. N. Konah, J. M. Vulule, J. B. Hittner, C. W. Gichuki, J. M. Ong'echa, and D. J. Perkins. 2012. 'Reduced systemic bicyclo-prostaglandin-E2 and cyclooxygenase-2 gene expression are associated with inefficient erythropoiesis and enhanced uptake of monocytic hemozoin in children with severe malarial anemia', *Am J Hematol*, 87: 782-9.

Anyona, S. B., P. Kempaiah, E. Raballah, C. Ouma, T. Were, G. C. Davenport, S. N. Konah, J. M. Vulule, J. B. Hittner, C. W. Gichuki, J. M. Ong'echa, and D. J. Perkins. 2011. 'Functional promoter haplotypes of interleukin-18 condition susceptibility to severe malarial anemia and childhood mortality', *Infect Immun*, 79: 4923-32.

Ashley, E. A., A. Pyae Phyo, and C. J. Woodrow. 2018. 'Malaria', *Lancet*, 391: 1608-21.

Awandare, G. A., J. J. Martinson, T. Were, C. Ouma, G. C. Davenport, J. M. Ong'echa, W. Wang, L. Leng, R. E. Ferrell, R. Bucala, and D. J. Perkins. 2009. 'MIF (macrophage migration inhibitory factor) promoter polymorphisms and susceptibility to severe malarial anemia', *J Infect Dis*, 200: 629-37.

Baro, B., K. Deroost, T. Raiol, M. Brito, A. C. Almeida, A. de Menezes-Neto, E. F. Figueiredo, A. Alencar, R. Leitao, F. Val, W. Monteiro, A. Oliveira, M. D. Armengol, C. Fernandez-Becerra, M. V. Lacerda, and H. A. Del Portillo. 2017. 'Plasmodium vivax gametocytes in the bone marrow of an acute malaria patient and changes in the erythroid miRNA profile', *PLoS Negl Trop Dis*, 11: e0005365.

Barreda, D. R., P. C. Hanington, and M. Belosevic. 2004. 'Regulation of myeloid development and function by colony stimulating factors', *Dev Comp Immunol*, 28: 509-54.

Becher, B., S. Tugues, and M. Greter. 2016. 'GM-CSF: From Growth Factor to Central Mediator of Tissue Inflammation', *Immunity*, 45: 963-73.

Brown, S. M., C. K. Grissom, M. T. Rondina, J. R. Hoidal, M. B. Scholand, R. K. Wolff, A. H. Morris, R. Paine, 3rd, and Nih Nhlbi Ards Network. 2015. 'Polymorphisms in key pulmonary inflammatory pathways and the development of acute respiratory distress syndrome', *Exp Lung Res*, 41: 155-62.

Burkhardt, J., H. Kirsten, G. Wolfram, E. Quente, and P. Ahnert. 2012. 'Differential allelic expression of IL13 and CSF2 genes associated with asthma', *Genet Mol Biol*, 35: 567-74.

Chen, D., T. X. Tang, H. Deng, X. P. Yang, and Z. H. Tang. 2021. 'Interleukin-7 Biology and Its Effects on Immune Cells: Mediator of Generation, Differentiation, Survival, and Homeostasis', *Front Immunol*, 12: 747324.

Chen, W. C., J. C. Wei, H. F. Lu, H. S. Wong, P. Y. Woon, Y. W. Hsu, J. D. Huang, and W. C. Chang. 2017. 'rs657075 (CSF2) Is Associated with the Disease Phenotype (BAS-G) of Ankylosing Spondylitis', *Int J Mol Sci*, 18.

Churchman, S. M., J. J. El-Jawhari, A. N. Burska, R. Parmar, V. Goeb, P. G. Conaghan, P. Emery, and F. Ponchel. 2014. 'Modulation of peripheral T-cell function by interleukin-7 in rheumatoid arthritis', *Arthritis Res Ther*, 16: 511.

Cokic, V. P., and A. N. Schechter. 2008. 'Effects of nitric oxide on red blood cell development and phenotype', *Curr Top Dev Biol*, 82: 169-215.

Cotterchio, M., E. Lowcock, Z. Bider-Canfield, M. Lemire, C. Greenwood, S. Gallinger, and T. Hudson. 2015. 'Association between Variants in Atopy-Related Immunologic Candidate Genes and Pancreatic Cancer Risk', *PLoS One*, 10: e0125273.

Croft, B., T. Ohnesorg, J. Hewitt, J. Bowles, A. Quinn, J. Tan, V. Corbin, E. Pelosi, J. van den Bergen, R. Sreenivasan, I. Knarston, G. Robevska, D. C. Vu, J. Hutson, V. Harley, K. Ayers, P. Koopman, and A. Sinclair. 2018. 'Human sex reversal is caused by duplication or deletion of core enhancers upstream of SOX9', *Nat Commun*, 9: 5319.

Dahan, S., A. Sharma, K. Cohen, M. Baker, N. Taqatqa, M. Bentata, E. Engal, A. Siam, G. Kay, Y. Drier, S. Elias, and M. Salton. 2021. 'VEGFA's distal enhancer regulates its alternative splicing in CML', *NAR Cancer*, 3: zcab029.

Davenport, G. C., J. B. Hittner, T. Were, J. M. Ong'echa, and D. J. Perkins. 2012. 'Relationship between inflammatory mediator patterns and anemia in HIV-1 positive and exposed children with Plasmodium falciparum malaria', *Am J Hematol*, 87: 652-8.

Desai, M., A. M. Buff, S. Khagayi, P. Byass, N. Amek, A. van Eijk, L. Slutsker, J. Vulule, F. O. Odhiambo, P. A. Phillips-Howard, K. A. Lindblade, K. F. Laserson, and M. J. Hamel. 2014. 'Age-specific malaria mortality rates in the KEMRI/CDC health and demographic surveillance system in western Kenya, 2003-2010', *PLoS One*, 9: e106197.

Dumarchey, A., C. Lavazec, and F. Verdier. 2022. 'Erythropoiesis and Malaria, a Multifaceted Interplay', *Int J Mol Sci*, 23.

Dunst, J., F. Kamena, and K. Matuschewski. 2017. 'Cytokines and Chemokines in Cerebral Malaria Pathogenesis', *Front Cell Infect Microbiol*, 7: 324.

Dybedal, I., S. Larsen, and S. E. Jacobsen. 1995. 'IL-12 directly enhances in vitro murine erythropoiesis in combination with IL-4 and stem cell factor', *J Immunol*, 154: 4950-5.

Eybpoosh, S. 2018. 'Hardy Weinberg Equilibrium Testing and Interpretation: Focus on infection', *J Med Microbiol Infect Dis*, 6: 35-36.

Felli, N., F. Pedini, A. Zeuner, E. Petrucci, U. Testa, C. Conticello, M. Biffoni, A. Di Cataldo, J. A. Winkles, C. Peschle, and R. De Maria. 2005. 'Multiple members of the TNF superfamily contribute to IFN-gamma-mediated inhibition of erythropoiesis', *J Immunol*, 175: 1464-72.

Fry, T. J., and C. L. Mackall. 2002. 'Interleukin-7: from bench to clinic', *Blood*, 99: 3892-904.

Ghavimi, R., M. Pourhossein, K. Ghaedi, F. Alesahebfosoul, M. A. Honardoost, and M. R. Maracy. 2014. 'Genetic association of rs1520333 G/A polymorphism in the IL7 gene with multiple sclerosis susceptibility in Isfahan population', *Adv Biomed Res*, 3: 238.

Gutierrez, L., N. Caballero, L. Fernandez-Calleja, E. Karkoulia, and J. Strouboulis. 2020. 'Regulation of GATA1 levels in erythropoiesis', *IUBMB Life*, 72: 89-105.

Hardikar, S., L. G. Johnson, M. Malkki, E. W. Petersdorf, D. A. Galloway, S. M. Schwartz, and M. M. Madeleine. 2015. 'A population-based case-control study of genetic variation in cytokine genes associated with risk of cervical and vulvar cancers', *Gynecol Oncol*, 139: 90-6.

Hartgring, S. A., J. W. Bijlsma, F. P. Lafeber, and J. A. van Roon. 2006. 'Interleukin-7 induced immunopathology in arthritis', *Ann Rheum Dis*, 65 Suppl 3: iii69-74.

Hartgring, S. A., J. A. van Roon, M. Wenting-van Wijk, K. M. Jacobs, Z. N. Jahangier, C. R. Willis, J. W. Bijlsma, and F. P. Lafeber. 2009. 'Elevated expression of interleukin-7 receptor in inflamed joints mediates interleukin-7-induced immune activation in rheumatoid arthritis', *Arthritis Rheum*, 60: 2595-605.

He, J. Q., K. Shumansky, J. E. Connett, N. R. Anthonisen, P. D. Pare, and A. J. Sandford. 2008. 'Association of genetic variations in the CSF2 and CSF3 genes with lung function in smoking-induced COPD', *Eur Respir J*, 32: 25-34.

Helleberg, M., B. Q. Goka, B. D. Akanmori, G. Obeng-Adjei, O. Rodriques, and J. A. Kurtzhals. 2005. 'Bone marrow suppression and severe anaemia associated with persistent Plasmodium falciparum infection in African children with microscopically undetectable parasitaemia', *Malar J*, 4: 56.

Hogan, P. G. 2017. 'Calcium-NFAT transcriptional signalling in T cell activation and T cell exhaustion', *Cell Calcium*, 63: 66-69.

Hollowell, T., M. O. Sewe, J. Rocklov, D. Obor, F. Odhiambo, and C. Ahlm. 2023. 'Public health determinants of child malaria mortality: a surveillance study within Siaya County, Western Kenya', *Malar J*, 22: 65.

Hosking, L., S. Lumsden, K. Lewis, A. Yeo, L. McCarthy, A. Bansal, J. Riley, I. Purvis, and C. F. Xu. 2004. 'Detection of genotyping errors by Hardy-Weinberg equilibrium testing', *Eur J Hum Genet*, 12: 395-9.

Hotamisligil, G. S. 2017. 'Inflammation, metaflammation and immunometabolic disorders', *Nature*, 542: 177-85.

Jakeman, G. N., A. Saul, W. L. Hogarth, and W. E. Collins. 1999. 'Anaemia of acute malaria infections in non-immune patients primarily results from destruction of uninfected erythrocytes', *Parasitology*, 119 (Pt 2): 127-33.

Johnson, L. G., S. M. Schwartz, M. Malkki, Q. Du, E. W. Petersdorf, D. A. Galloway, and M. M. Madeleine. 2011. 'Risk of cervical cancer associated with allergies and polymorphisms in genes in the chromosome 5 cytokine cluster', *Cancer Epidemiol Biomarkers Prev*, 20: 199-207.

Khagayi, S., N. Amek, G. Bigogo, F. Odhiambo, and P. Vounatsou. 2017. 'Bayesian spatio-temporal modeling of mortality in relation to malaria incidence in Western Kenya', *PLoS One*, 12: e0180516.

Kim, K., S. O. Zakharkin, and D. B. Allison. 2010. 'Expectations, validity, and reality in gene expression profiling', *J Clin Epidemiol*, 63: 950-9.

Koay, T. W., C. Osterhof, I. M. C. Orlando, A. Keppner, D. Andre, S. Yousefian, M. Suarez Alonso, M. Correia, R. Markworth, J. Schodel, T. Hankeln, and D. Hoogewijs. 2021. 'Androglobin gene expression patterns and FOXJ1-dependent regulation indicate its functional association with ciliogenesis', *J Biol Chem*, 296: 100291.

Kreft, K. L., E. Verbraak, A. F. Wierenga-Wolf, M. van Meurs, B. A. Oostra, J. D. Laman, and R. Q. Hintzen. 2012. 'Decreased systemic IL-7 and soluble IL-7Ralpha in multiple sclerosis patients', *Genes Immun*, 13: 587-92.

Kumaratilake, L. M., A. Ferrante, T. Jaeger, and C. Rzepczyk. 1996. 'GM-CSF-induced priming of human neutrophils for enhanced phagocytosis and killing of asexual blood stages of Plasmodium falciparum: synergistic effects of GM-CSF and TNF', *Parasite Immunol*, 18: 115-23.

Lamikanra, A. A., D. Brown, A. Potocnik, C. Casals-Pascual, J. Langhorne, and D. J. Roberts. 2007. 'Malarial anemia: of mice and men', *Blood*, 110: 18-28.

Lee KMC, Achuthan AA, Hamilton JA. GM-CSF: A Promising Target in Inflammation and Autoimmunity. Immunotargets Ther. 2020 Oct 29;9:225-240. doi: 10.2147/ITT.S262566. PMID: 33150139; PMCID: PMC7605919.

Lupton, S. D., S. Gimpel, R. Jerzy, L. L. Brunton, K. A. Hjerrild, D. Cosman, and R. G. Goodwin. 1990. 'Characterization of the human and murine IL-7 genes', *J Immunol*, 144: 3592-601.

Luty, A. J., D. J. Perkins, B. Lell, R. Schmidt-Ott, L. G. Lehman, D. Luckner, B. Greve, P. Matousek, K. Herbich, D. Schmid, J. B. Weinberg, and P. G. Kremsner. 2000. 'Low interleukin-12 activity in severe Plasmodium falciparum malaria', *Infect Immun*, 68: 3909-15.

Macharia, P. M., E. Giorgi, A. M. Noor, E. Waqo, R. Kiptui, E. A. Okiro, and R. W. Snow. 2018. 'Spatio-temporal analysis of Plasmodium falciparum prevalence to understand the past and chart the future of malaria control in Kenya', *Malar J*, 17: 340.

Matsushita, N., Y. Kim, C. F. S. Ng, M. Moriyama, T. Igarashi, K. Yamamoto, W. Otieno, N. Minakawa, and M. Hashizume. 2019. 'Differences of Rainfall-Malaria Associations in Lowland and Highland in Western Kenya', *Int J Environ Res Public Health*, 16.

Miyashita, H., N. Katayama, A. Fujieda, T. Shibasaki, K. Yamamura, Y. Sugimoto, E. Miyata, K. Ohishi, K. Nishii, M. Masuya, and H. Shiku. 2005. 'IL-4 and IL-10 synergistically inhibit survival of human blood monocytes supported by GM-CSF', *Int J Oncol*, 26: 731-5.

Miyatake, S., T. Otsuka, T. Yokota, F. Lee, and K. Arai. 1985. 'Structure of the chromosomal gene for granulocyte-macrophage colony stimulating factor: comparison of the mouse and human genes', *EMBO J*, 4: 2561-8.

Mohandas, N., and X. An. 2012. 'Malaria and human red blood cells', *Med Microbiol Immunol*, 201: 593-8.

Munde, E. O., W. A. Okeyo, E. Raballah, S. B. Anyona, T. Were, J. M. Ong'echa, D. J. Perkins, and C. Ouma. 2017. 'Association between Fcgamma receptor IIA, IIIA and IIIB genetic polymorphisms and susceptibility to severe malaria anemia in children in western Kenya', *BMC Infect Dis*, 17: 289.

Natama, H. M., G. Moncunill, E. Rovira-Vallbona, H. Sanz, H. Sorgho, R. Aguilar, M. Coulibaly-Traore, M. A. Some, S. Scott, I. Valea, P. F. Mens, Hdfh Schallig, L.

Kestens, H. Tinto, C. Dobano, and A. Rosanas-Urgell. 2018. 'Modulation of innate immune responses at birth by prenatal malaria exposure and association with malaria risk during the first year of life', *BMC Med*, 16: 198.

Natividad, A., J. Hull, G. Luoni, M. Holland, K. Rockett, H. Joof, M. Burton, D. Mabey, D. Kwiatkowski, and R. Bailey. 2009. 'Innate immunity in ocular Chlamydia trachomatis infection: contribution of IL8 and CSF2 gene variants to risk of trachomatous scarring in Gambians', *BMC Med Genet*, 10: 138.

Novelli, E. M., J. B. Hittner, G. C. Davenport, C. Ouma, T. Were, S. Obaro, S. Kaplan, J. M. Ong'echa, and D. J. Perkins. 2010. 'Clinical predictors of severe malarial anaemia in a holoendemic Plasmodium falciparum transmission area', *Br J Haematol*, 149: 711-21.

Obonyo, C. O., J. Vulule, W. S. Akhwale, and D. E. Grobbee. 2007. 'In-hospital morbidity and mortality due to severe malarial anemia in western Kenya', *Am J Trop Med Hyg*, 77: 23-8.

Ong'echa, J. M., G. C. Davenport, J. M. Vulule, J. B. Hittner, and D. J. Perkins. 2011. 'Identification of inflammatory biomarkers for pediatric malarial anemia severity using novel statistical methods', *Infect Immun*, 79: 4674-80.

Ong'echa, J. M., C. C. Keller, T. Were, C. Ouma, R. O. Otieno, Z. Landis-Lewis, D. Ochiel, J. L. Slingluff, S. Mogere, G. A. Ogonji, A. S. Orago, J. M. Vulule, S. S. Kaplan, R. D. Day, and D. J. Perkins. 2006. 'Parasitemia, anemia, and malarial anemia in infants and young children in a rural holoendemic Plasmodium falciparum transmission area', *Am J Trop Med Hyg*, 74: 376-85.

Othoro, C., A. A. Lal, B. Nahlen, D. Koech, A. S. Orago, and V. Udhayakumar. 1999. 'A low interleukin-10 tumor necrosis factor-alpha ratio is associated with malaria anemia in children residing in a holoendemic malaria region in western Kenya', *J Infect Dis*, 179: 279-82.

Otieno, R. O., C. Ouma, J. M. Ong'echa, C. C. Keller, T. Were, E. N. Waindi, M. G. Michaels, R. D. Day, J. M. Vulule, and D. J. Perkins. 2006. 'Increased severe anemia in HIV-1-exposed and HIV-1-positive infants and children during acute malaria', *AIDS*, 20: 275-80.

Ouma, C., G. C. Davenport, S. Garcia, P. Kempaiah, A. Chaudhary, T. Were, S. B. Anyona, E. Raballah, S. N. Konah, J. B. Hittner, J. M. Vulule, J. M. Ong'echa, and D. J. Perkins. 2012. 'Functional haplotypes of Fc gamma (Fcgamma) receptor (FcgammaRIIA and FcgammaRIIIB) predict risk to repeated episodes of severe malarial anemia and mortality in Kenyan children', *Hum Genet*, 131: 289-99.

Ouma, C., G. C. Davenport, T. Were, M. F. Otieno, J. B. Hittner, J. M. Vulule, J. Martinson, J. M. Ong'echa, R. E. Ferrell, and D. J. Perkins. 2008. 'Haplotypes of IL-10 promoter variants are associated with susceptibility to severe malarial anemia and functional changes in IL-10 production', *Hum Genet*, 124: 515-24.

Ouma, C., C. C. Keller, G. C. Davenport, T. Were, S. Konah, M. F. Otieno, J. B. Hittner, J. M. Vulule, J. Martinson, J. M. Ong'echa, R. E. Ferrell, and D. J. Perkins. 2010. 'A

novel functional variant in the stem cell growth factor promoter protects against severe malarial anemia', *Infect Immun*, 78: 453-60.

Owhashi, M., N. Kirai, and M. Asami. 1996. 'Temporary appearance of a circulating granulocyte-macrophage colony-stimulating factor in lethal murine malaria', *Southeast Asian J Trop Med Public Health*, 27: 530-4.

Parajuli, B., Y. Sonobe, J. Kawanokuchi, Y. Doi, M. Noda, H. Takeuchi, T. Mizuno, and A. Suzumura. 2012. 'GM-CSF increases LPS-induced production of proinflammatory mediators via upregulation of TLR4 and CD14 in murine microglia', *J Neuroinflammation*, 9: 268.

Penha-Goncalves, C. 2019. 'Genetics of Malaria Inflammatory Responses: A Pathogenesis Perspective', *Front Immunol*, 10: 1771.

Perkins, D. J., T. Were, S. B. Anyona, J. B. Hittner, P. Kempaiah, G. C. Davenport, and J. M. Ong'echa. 2013. 'The Global Burden of Severe Falciparum Malaria: An Immunological and Genetic Perspective on Pathogenesis.' in, *Dynamic Models of Infectious Diseases* (Springer).

Perkins, D. J., T. Were, G. C. Davenport, P. Kempaiah, J. B. Hittner, and J. M. Ong'echa. 2011. 'Severe malarial anemia: innate immunity and pathogenesis', *Int J Biol Sci*, 7: 1427-42.

Phawong, C., C. Ouma, P. Tangteerawatana, J. Thongshoob, T. Were, Y. Mahakunkijcharoen, D. Wattanasirichaigoon, D. J. Perkins, and S. Khusmith. 2010. 'Haplotypes of IL12B promoter polymorphisms condition susceptibility to severe malaria and functional changes in cytokine levels in Thai adults', *Immunogenetics*, 62: 345-56.

Raballah, E., S. B. Anyona, Q. Cheng, E. O. Munde, I. F. Hurwitz, C. Onyango, C. Ndege, N. W. Hengartner, M. A. Pacheco, A. A. Escalante, C. G. Lambert, C. Ouma, Hcjt Obama, K. A. Schneider, P. D. Seidenberg, B. H. McMahon, and D. J. Perkins. 2022. 'Complement component 3 mutations alter the longitudinal risk of pediatric malaria and severe malarial anemia', *Exp Biol Med (Maywood)*, 247: 672-82.

Ringwald, P., F. Peyron, J. P. Vuillez, J. E. Touze, J. Le Bras, and P. Deloron. 1991. 'Levels of cytokines in plasma during Plasmodium falciparum malaria attacks', *J Clin Microbiol*, 29: 2076-8.

Rossati, A., O. Bargiacchi, V. Kroumova, M. Zaramella, A. Caputo, and P. L. Garavelli. 2016. 'Climate, environment and transmission of malaria', *Infez Med*, 24: 93-104.

Ryan, S. J., C. A. Lippi, and F. Zermoglio. 2020. 'Shifting transmission risk for malaria in Africa with climate change: a framework for planning and intervention', *Malar J*, 19: 170.

Sharma, L., and G. Shukla. 2017. 'Placental Malaria: A New Insight into the Pathophysiology', *Front Med (Lausanne)*, 4: 117.

Shio, M. T., F. A. Kassa, M. J. Bellemare, and M. Olivier. 2010. 'Innate inflammatory response to the malarial pigment hemozoin', *Microbes Infect*, 12: 889-99.

Sieling, P. A., L. Sakimura, K. Uyemura, M. Yamamura, J. Oliveros, B. J. Nickoloff, T. H. Rea, and R. L. Modlin. 1995. 'IL-7 in the cell-mediated immune response to a human pathogen', *J Immunol*, 154: 2775-83.

Sikora, M., H. Laayouni, C. Menendez, A. Mayor, A. Bardaji, B. Sigauque, M. G. Netea, F. Casals, and J. Bertranpetit. 2011. 'A targeted association study of immunity genes and networks suggests novel associations with placental malaria infection', *PLoS One*, 6: e24996.

Singh, S. K., R. Bhardwaj, K. M. Wilczynska, C. I. Dumur, and T. Kordula. 2011. 'A complex of nuclear factor I-X3 and STAT3 regulates astrocyte and glioma migration through the secreted glycoprotein YKL-40', *J Biol Chem*, 286: 39893-903.

Szeliga, J., D. S. Daniel, C. H. Yang, Z. Sever-Chroneos, C. Jagannath, and Z. C. Chroneos. 2008. 'Granulocyte-macrophage colony stimulating factor-mediated innate responses in tuberculosis', *Tuberculosis (Edinb)*, 88: 7-20.

Team, R Core. 2018. 'R: A language and environment for statistical computing. R Foundation for Statistical Computing V, Austria. 2016', *URL http://www. R-project. org*.

van Leeuwen, B. H., M. E. Martinson, G. C. Webb, and I. G. Young. 1989. 'Molecular organization of the cytokine gene cluster, involving the human IL-3, IL-4, IL-5, and GM-CSF genes, on human chromosome 5', *Blood*, 73: 1142-8.

Waitumbi, J. N., M. O. Opollo, R. O. Muga, A. O. Misore, and J. A. Stoute. 2000. 'Red cell surface changes and erythrophagocytosis in children with severe plasmodium falciparum anemia', *Blood*, 95: 1481-6.

Were, T., G. C. Davenport, J. B. Hittner, C. Ouma, J. M. Vulule, J. M. Ong'echa, and D. J. Perkins. 2011. 'Bacteremia in Kenyan children presenting with malaria', *J Clin Microbiol*, 49: 671-6.

Were, T., J. B. Hittner, C. Ouma, R. O. Otieno, A. S. Orago, J. M. Ong'echa, J. M. Vulule, C. C. Keller, and D. J. Perkins. 2006. 'Suppression of RANTES in children with Plasmodium falciparum malaria', *Haematologica*, 91: 1396-9.

White, N. J. 2018. 'Anaemia and malaria', *Malar J*, 17: 371.

WHO. 2012. Management of Severe Malaria Anemia (2012).

WHO. 2022. World malaria report . Geneva: World Health Organization (2022)

Wickramasinghe, S. N., S. Abdalla, and D. J. Weatherall. 1982. 'Cell cycle distribution of erythroblasts in P. falciparum malaria', *Scand J Haematol*, 29: 83-8.

Wickramasinghe, S. N., S. Looareesuwan, B. Nagachinta, and N. J. White. 1989. 'Dyserythropoiesis and ineffective erythropoiesis in Plasmodium vivax malaria', *Br J Haematol*, 72: 91-9.

Winer, H., G. O. L. Rodrigues, J. A. Hixon, F. B. Aiello, T. C. Hsu, B. T. Wachter, W. Li, and S. K. Durum. 2022. 'IL-7: Comprehensive review', *Cytokine*, 160: 156049.

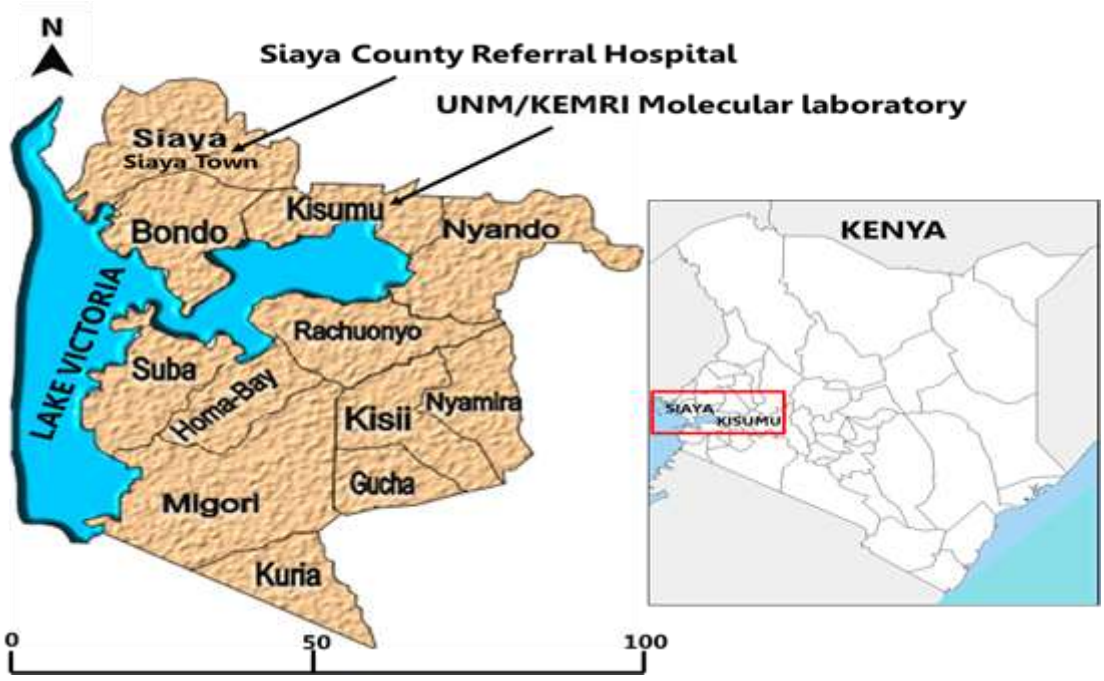
Xu, L., X. Zheng, K. Berzins, and A. Chaudhuri. 2013. 'Cytokine dysregulation associated with malarial anemia in Plasmodium yoelii infected mice', *Am J Transl Res*, 5: 235-45.

Zhang, H. X., Y. G. Wang, S. Y. Lu, X. X. Lu, and J. Liu. 2016. 'Identification of IL-7 as a candidate disease mediator in osteoarthritis in Chinese Han population: a case-control study', *Rheumatology (Oxford)*, 55: 1681-5.

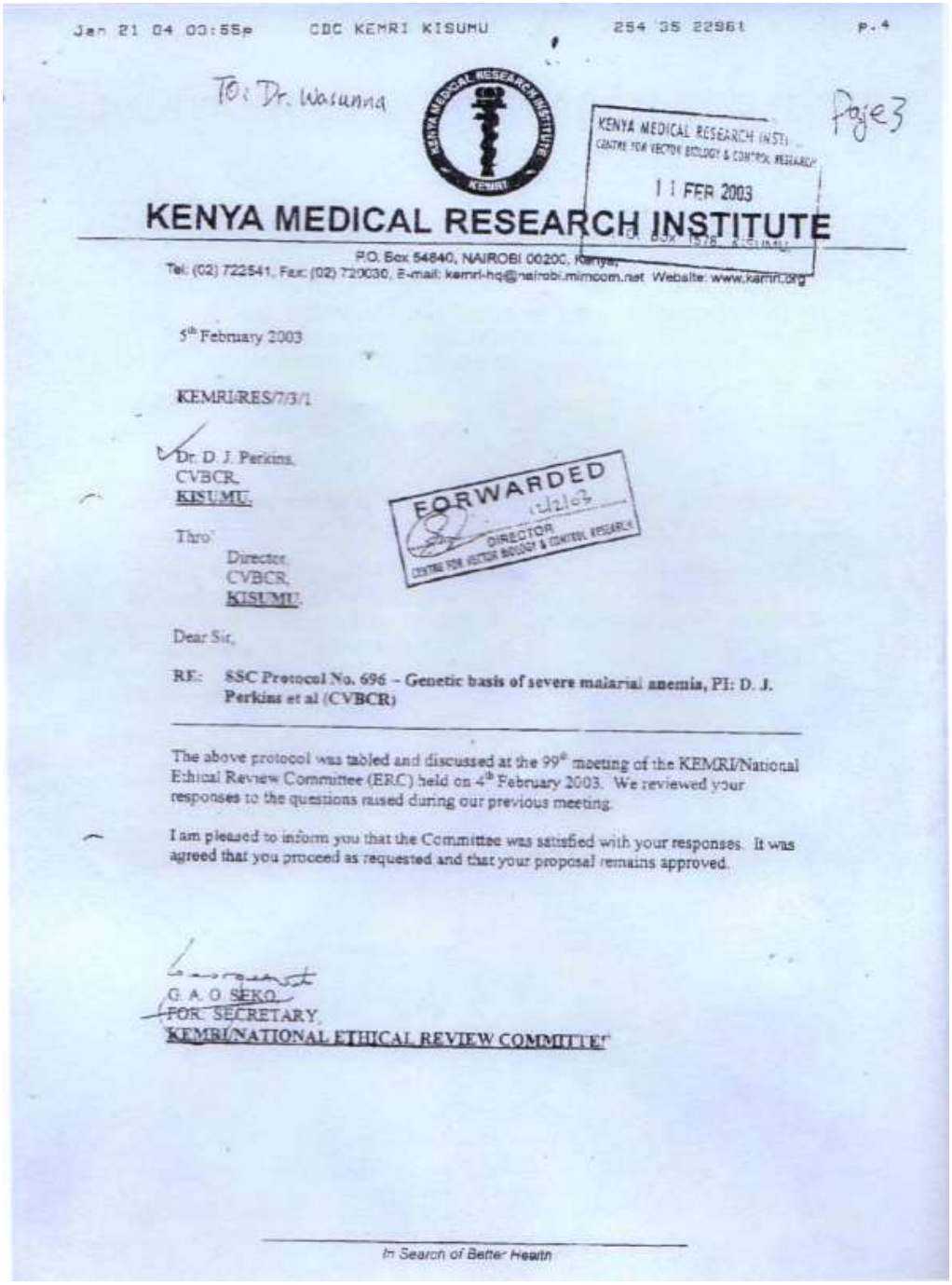
Zuvich, R. L., J. L. McCauley, J. R. Oksenberg, S. J. Sawcer, P. L. De Jager, Consortium International Multiple Sclerosis Genetics, C. Aubin, A. H. Cross, L. Piccio, N. T. Aggarwal, D. Evans, D. A. Hafler, A. Compston, S. L. Hauser, M. A. Pericak-Vance, and J. L. Haines. 2010. 'Genetic variation in the IL7RA/IL7 pathway increases multiple sclerosis susceptibility', *Hum Genet*, 127: 525-35.

APPENDICES

Appendix 1: Location of Siaya County Referral Hospital in western Kenya



Appendix 2: Ethical approval from the Scientific and Ethics Review Committee of the Kenya Medical Research Institute (KEMRI)



Jan 21 04 03:55p CDC KEMRI KISUMU 254 35 22961 P.4

To: Dr. Waiumna

Page 3

KENYA MEDICAL RESEARCH INSTITUTE

KENYA MEDICAL RESEARCH (INST.) CENTRE FOR VECTOR BIOLOGY & CONTROL RESEARCH 11 FEB 2003

P.O. Box 54840, NAIROBI 00200, Kenya.
Tel: (02) 722541, Fax: (02) 729030, E-mail: kemri-hq@nairobi.mimcom.net Website: www.kemri.org

5th February 2003

KEMRI/RES/7/3/1

Dr. D. J. Perkins,
CVBCR,
KISUMU.

FORWARDED 12/2/03 DIRECTOR CENTRE FOR VECTOR BIOLOGY & CONTROL RESEARCH

Thro'
Director,
CVBCR,
KISUMU.

Dear Sir,

RE: **SSC Protocol No. 696 – Genetic basis of severe malarial anemia, PI: D. J. Perkins et al (CVBCR)**

The above protocol was tabled and discussed at the 99th meeting of the KEMRI/National Ethical Review Committee (ERC) held on 4th February 2003. We reviewed your responses to the questions raised during our previous meeting.

I am pleased to inform you that the Committee was satisfied with your responses. It was agreed that you proceed as requested and that your proposal remains approved.

G. A. O. SEKO
FOR: SECRETARY,
KEMRI/NATIONAL ETHICAL REVIEW COMMITTEE

In Search of Better Health

KENYA MEDICAL RESEARCH INSTITUTE
CENTRE FOR GLOBAL HEALTH RESEARCH

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541 , 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: kemri-hq@nairobi.mimcom.net; director @ kemri. org; website: www.kemri.org

KEMRI/RES/7/3/1 **July 17, 2009**

TO: **DR. DOUGLAS J PERKINS (PRINCIPAL INVESTIGATOR)**
GLOBAL GEOGRAPHIC MEDICINE,
UNIVERSITY OF NEW MEXICO

'THRO': **DR. J. VULULE,**
THE DIRECTOR, CGHR,
KISUMU

FORWARDED
DIRECTOR
CENTRE FOR GLOBAL HEALTH RESEARCH
KISUMU

RE: **SSC PROTOCOL No. 696 (*REQUEST FOR 2nd AMENDMENT*):**
GENETIC BASIS OF SEVERE MALARIAL ANEMIA

This is to inform you that during the 168th meeting of KEMRI/National Ethics Review Committee held on Tuesday 14th July 2009, the provisional approval granted for the amendment by the Chair of the KEMRI/NERC on 24th June 2009 was **ratified** by the full committee.

You may proceed with your study,

Yours sincerely,

R. C. KITHINJI,
FOR: SECRETARY,
KEMRI/NATIONAL ETHICS REVIEW COMMITTEE

In Search of Better Health