MOLECULAR CHARACTERIZATION OF THE APICAL MEMBRANE ANTIGEN 1 POLYMORPHISMS IN *PLASMODIUM FALCIPARUM* ISOLATES FROM KILIFI COUNTY, KENYA.

BY:

OUKO DAVID BOSIRE (B.Sc.)

I56/68747/2011

A THESIS SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF MASTER OF SCIENCE (APPLIED PARASITOLOGY)

UNIVERSITY OF NAIROBI.

2015.

DECLARATION

This thesis is my original work, and has never been presented for a degree in any other institution.

DAVID BOSIRE OUKO

Signature: ____________________Date_________________________

This thesis has been submitted by our approval as supervisors:

Dr. David Odongo

School of Biological Sciences

University of Nairobi

Signature: ____________________Date______________________

Dr. Isabella Oyier

Centre for Geographic Medicine – Kilifi Kenya Medical Research Institute

Centre for Biotechnology & Bioinformatics

University of Nairobi

Signature: ____________________Date_________________________

DEDICATION

To all the families living within the malaria endemic region in Kilifi County at the Kenyan coast.

AKNOWLEDGEMENT

I would like to acknowledge Dr. David Odongo and Dr. Isabella Oyier, my supervisors, for the great work they did of helping me to conceptualize this project from an idea and for their guidance throughout the process of developing protocols to data analysis. I would also like to acknowledge Prof. James Ochanda, who made it possible for me to conduct my lab work at Centre for Biotechnology and Bioinformatics. A big thank you goes to the CEBIB, for providing an enjoyable learning environment. A special thank you goes to Ann and Rono for their help in the Molecular Biology laboratory. Finally, I would like to thank my father Andrew, mum Priscilla and my siblings for their extensive support.

TABLE OF CONTENTS **Page**

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

aa	Amino acid
AMA	Apical Membrane Antigen
C1-L	Cluster 1 loop
dNTPs	Deoxynucleotide triphosphates
ddNTPs	Dideoxynucleotides
EBA	Erythrocyte Binding Antigen
EDTA	Ethylenediaminetetraacetic acid
ExoSAP	Exonuclease Shrimp Alkaline Phosphatase
PCR	Polymerase Chain Reaction
PfRh	*Plasmodium falciparum* Reticulocyte binding-like Protein Homologue
RBCs	Red Blood Cells
SNP	Single Nucleotide Polymorphism
TBE	Tris Borate EDTA
WHO	World Health Organization
RON2	Rhoptry neck protein 2
DNA	Deoxyribonucleic acid

MSP	Merozoite surface protein
RH5	Reticulocyte binding homologue
ELISA	Enzyme-linked immunosorbent assay
DALY	Daily adjusted life years
pH	Potential of hydrogen
FASTA	Fast all format
min	Minute
ml	Millimetre
sec	Second
μl	Microlitre
%	Per cent

ABSTRACT

Of the five *Plasmodium* species that cause human malaria, *Plasmodium falciparum* is the leading cause of morbidity and death with about 300 million medical cases every year. Erythrocyte invasion is an essential process in the life cycle of the malaria parasite hence parasite survival in the human host needs successful invasion of merozoites into uninfected red blood cells. Several parasite proteins such as apical membrane antigen (AMA1) can accomplish similar roles in the invasion process. *Plasmodium* parasites have developed a number of distinct evasion responses such as varying sequence and maintaining functions employed by merozoite surface proteins (MSPs) and AMA1, reduced antigenicity by reticulocyte binding homologue 5 (RH5) and redundancy in multi-gene families depicted by erythrocyte binding like antigens (EBAs) and RHs .The interaction between *P. falciparum* AMA1 and another parasite protein called rhoptry neck protein 2 (RON2) is essential for tight junction formation, which commits the merozoite for invasion. Currently there is no vaccine effective against the blood-stages of *P. falciparum* though RTS, S is the most advanced candidate malaria vaccine but it is only partially protective .The aim of this study was to understand the genetic diversity of AMA1 within *P. falciparum* isolates. A total of 241 human blood samples were obtained from Junju location, Kilifi County. The AMA1 gene was successfully amplified by polymerase chain reaction (PCR) for 37 samples. Following successful sequence analysis, 14 haplotypes were identified. Analysis of the cluster one loop (C1-L) codon regions encompassing position 187-207 revealed polymorphisms ranging between 2-4 different amino acids, with position 197 being the most polymorphic while comparing to the reference AMA1 sequence. Analysis of AMA1 deoxyribonucleic acid (DNA) sequence using the Tajima's D statistic test for neutrality showed that the identified single nucleotide polymorphisms were not under selection and mutations occurring in this gene are neutral. These observed polymorphisms are in agreement with previous studies on genetic diversity of AMA1 hence making it a possible vaccine candidate.

Key words: Cluster 1 loop, invasion, erythrocyte, polymorphisms, AMA1 and *P .falciparum.*

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

1.0. General Introduction

P. falciparum is the leading cause of morbidity and death of the five *Plasmodium* species that cause human malaria. About 300 million medical cases occur every year causing between 1.5 - 2.7 million deaths annually, mainly in Sub-Saharan Africa (World Health Organisation report, 2011). Children below the age of five and expectant mothers are the most susceptible. The percentage of the global population at risk has reduced from 77% at the turn of the 20th century to a low of 46% in 1994 as per the World Health Organisation, 2011.

Malaria remains a significant parasitic disease affecting humans and the efforts to come up with an effective vaccine has taken more than 60 years (Hill, 2011). People living in malaria prone areas tend to acquire immunity to the disease, which is evident from the fact that the burden of disease falls on young children. Though natural immunity to malaria develops in most inhabitants of endemic regions, this commonly takes some years of exposure and is deficient. Older children and adults are resistant to severe morbidity and death, though they remain susceptible to infection (Marsh and Kinyanjui, 2006).

The non-existence of an effective vaccine remains one of the most significant challenges in the portfolio of tools being developed to eliminate *P. falciparum* malaria (Bustamante *et al.*, 2013). Vaccines targeting erythrocyte invasion i.e. an essential step for both parasite development and malaria pathogenesis, have faced the specific task of genetic diversity (Bustamante *et al.*, 2013). The immune epidemiological studies have given inadequate information on the best antigen to include in vaccine development: natural immunity aims a widespread diversity of blood stage antigens but then no one antigen appears to be particularly significant in providing protection (Marsh & Kinyanjui, 2006).

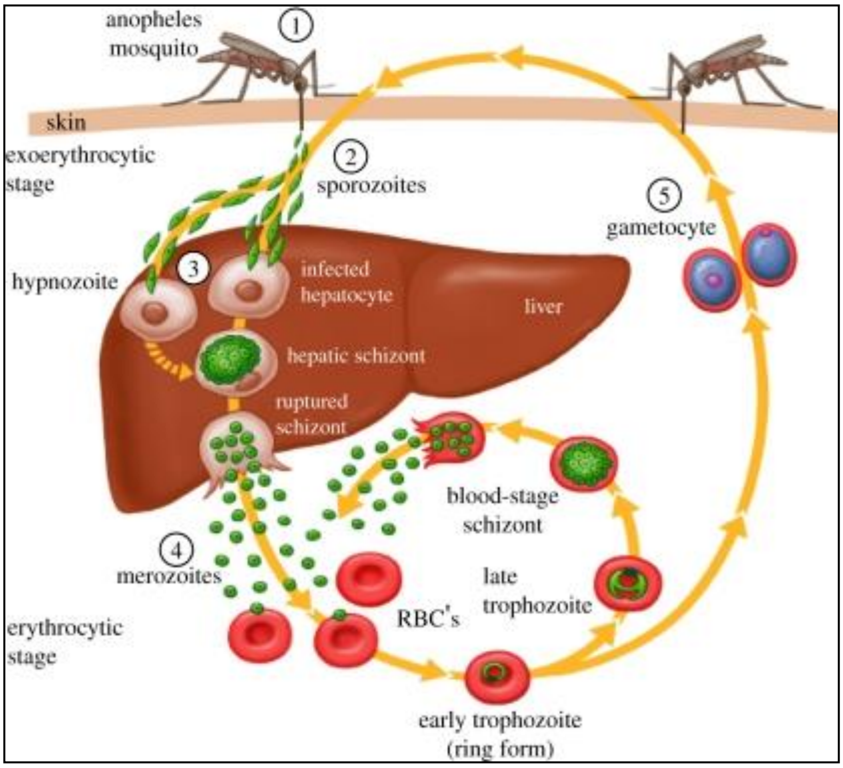

Figure 1: Summarized liver and blood stages in the parasite's life cycle (Hill, 2011).

Through their very complex life cycle (Figure 1), *Plasmodium* parasites have to traverse an extensive range of intracellular and extracellular surroundings in both the human and insect host. In order to be successful in this the parasite has to present itself in different physical appearances which are also known as zoites, each of which faces a specific biological challenge (Wright *et al.*,2014). Replication of merozoites always occurs intracellularly, thus they invade the red blood cells and this is possible after undergoing several challenges. This was first visualized by video microscopy around three decades ago (Aikawa *et al.*,1981).

Merozoites are in extracellular phase of the *Plasmodium* life cycle and are consequently open to a range of immune attack mechanisms (Figure 2). Merozoite antigens are recognized to be the target of antibody responses, which function both by opsonisation leading to phagocytosis and by simple steric hindrance of receptor–ligand interactions important for invasion, complement deposition on the merozoite surface may also play a role in parasite clearance. To evade these attack mechanisms, *Plasmodium* parasites have developed a number of distinct evasion responses such as varying sequence and maintaining functions employed by MSPs and AMA1, reduced antigenicity by RH5 and redundancy in multi-gene families depicted by EBAs and RHs (Wright & Rayner, 2014).

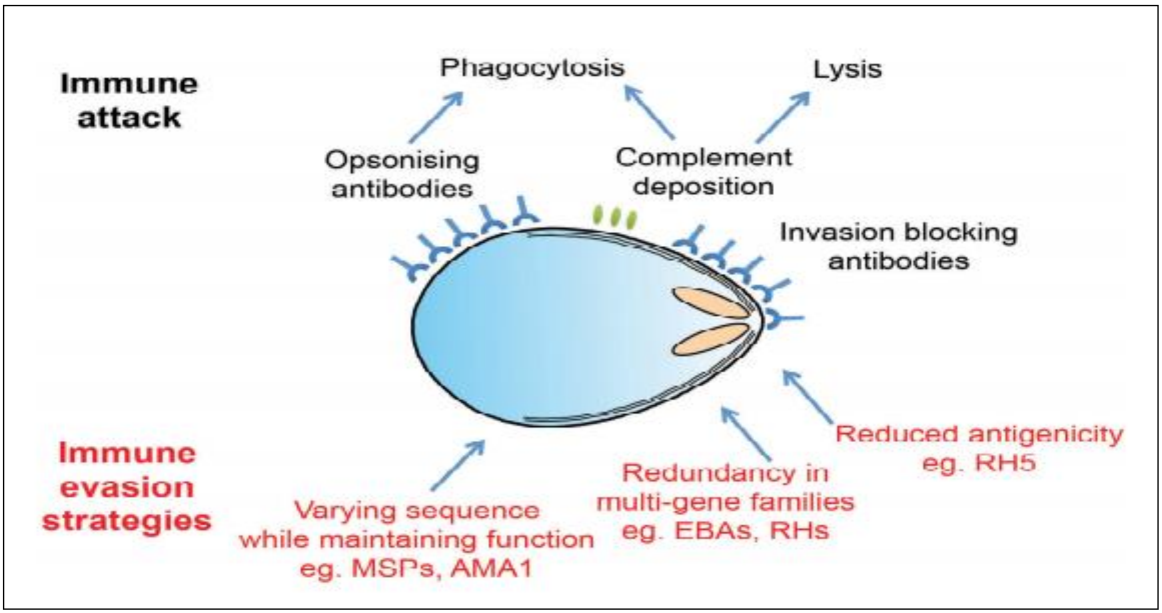

Figure 2: *Plasmodium* merozoite facing an array of immunological challenges (Wright & Rayner, 2014).

P. falciparum proteins that are hypothesised to somehow be involved in the invasion process, although in the vast majority of cases their precise function is unknown. The most well-studied of these have been organised into distinct functional classes: MSPs (merozoite surface proteins), which form a structurally complex coat around the merozoite surface, the

PfEBAs (*P. falciparum* erythrocyte binding antigens) and PfRHs (*P. falciparum* reticulocyte binding protein homologues), which are stored in specialised apical organelles, the rhoptries and micronemes (Wright & Rayner, 2014).

The merozoite has the conventional organelle repertoire of eukaryotic cells with the overall cytoskeletal architecture of an apicomplexan cell (Morrissette & Sibley, 2002), the phylum to which malaria parasites belong (Figure 3).

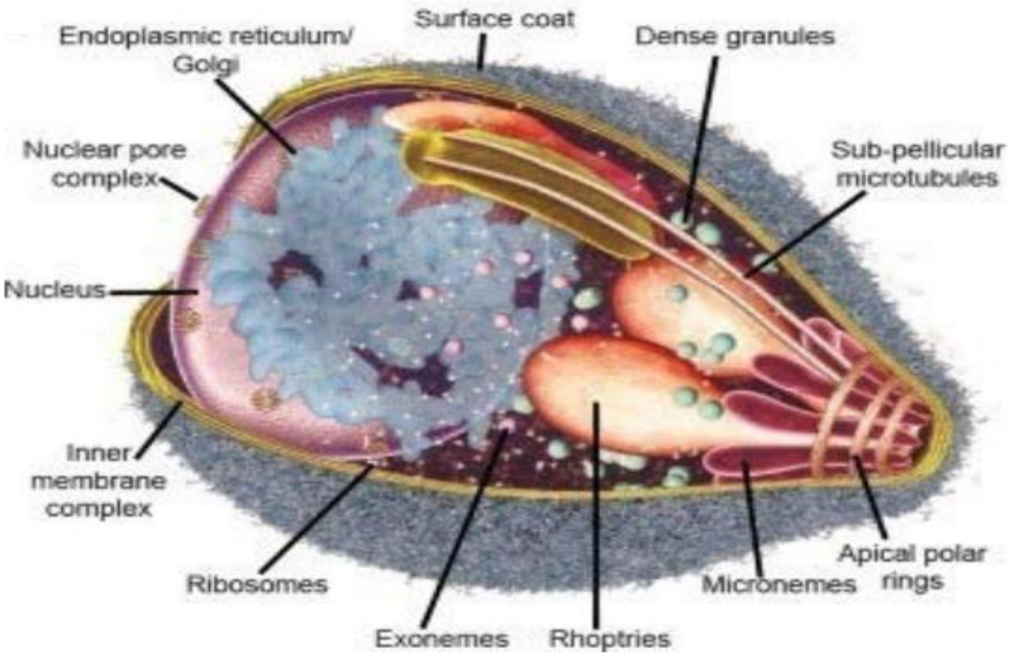

Figure 3: Three-dimensional diagram of *Plasmodium* merozoite showing its core secretory organelles (Cowman *et al*., 2012).

Immunity against *P. falciparum* malaria ultimately advances after frequent exposure to infection, and is characterised by control of blood-stage parasitemia and prevention of clinical illness and severe complications (Doolan *et al*., 2009). Antibodies play a major role in acquired immunity, though the main targets and mechanisms of action of defensive human antibodies are not understood well (Gardner *et al*., 2002). *P. falciparum* merozoites invasion of erythrocytes is during blood stage replication, and antibodies that hinder attack by targeting merozoite antigens are believed to be vital for acquired immunity (Drew *et al*., 2012). Identifying targets of defensive antibodies in humans and understanding the mechanisms by which antibodies to merozoite antigens shield against malaria is important for the development of blood-stage malaria vaccines, as well as for developing approaches to monitor immunity in populations, assess the impact of malaria control interventions on immunity, and classify populations at high risk of malaria (Drew *et al*., 2012).

Factors involved in malaria transmission are complex. As per Mackinnon *et al*., 2000, transimission of malaria was found not to be stable consistently but then in patches and depended on predisposing factors. Asymptomatic infections often go undetected and untreated, resulting in a major source of gametocytes for local mosquito vectors (Alves *et al*., 2002). Even in conditions where the possibility of re-infection is not expected, *P. falciparum* infection has been shown to persist asymptomatically in semi-immune individuals for more than 18 months. Asymptomatic *P. falciparum* infections can also persist inter-seasonally in regions with seasonal transmission (Osier *et al*., 2008). For example, even though the incidence of clinical malaria in Senegal is significantly lower during the dry season, a considerable proportion of the population remain parasitaemic throughout the year (Males *et al*., 2008).

Erythrocyte invasion is an essential process in the life cycle of the malaria parasite (Baum *et al*., 2009). Parasite survival in the human host needs successful invasion of merozoites into uninfected red blood cells. This is an active and sophisticated process, and needs numerous steps of interaction between receptors on the erythrocyte and parasite ligands (Baum *et al*., 2006). *P. falciparum* has established the capacity to invade erythrocytes by numerous parasite ligand-erythrocyte receptor interactions that are known as alternative invasion pathways (Hadley *et al*., 1987). Several parasite proteins such as erythrocyte binding ligands can accomplish similar roles in the invasion process and hence any successful malaria vaccine will have to target all alternative pathways of invasion.

1.1. Asymptomatic malaria

Notwithstanding a wealth of studies on the medical severity of disease, asymptomatic malaria infections are not well understood. Asymptomatic malaria remains a challenge for malaria control programs as it significantly influences transmission dynamics. A detailed understanding of the interaction between hosts and parasites in the development of different clinical outcomes is essential. Parasite prevalence, period prevalence of clinical attacks and period prevalence of severe life threatening attacks must all result from exposure to infection and be a measure of susceptibility, all show evidence of acquisition of resistance with increasing age, but it is striking that the indicators have quite different age relationships (Marsh & Kinyanjui, 2006).

Asymptomatic malaria is widespread in malaria prevalent areas and has become a serious cause for concern as efforts are increasing towards eliminating the parasite. Particularly, sub-patent malaria is still transmissible and will complicate elimination of malaria in high transmission regions. For example, a study in Senegal proposed that more than 90% of exposed individuals are likely infected with chronic asymptomatic malaria, a situation in which the majority of this population can then inadvertently act as a reservoir for malaria transmission (Laishram *et al*., 2012).

The main hindrance in the study of asymptomatic malaria is the absence of standard diagnostic criteria. For example, infected persons may be in a pre-symptomatic period with parasitaemia, and present with clinical manifestations at a subsequent date (Rodrigues *et al*,. 2006). In turn, studies that do not incorporate comprehensive medical history surveys may not capture individuals that may have experienced symptoms for a brief period and then taken medication that suppressed parasitaemia and symptoms. The most widely-used criteria for diagnosis of asymptomatic malaria are presence of parasites in peripheral thick blood smears, an axillary temperature <37.5°C, and an absence of malaria-related symptoms (De Mast *et al*., 2010). Some studies do contain other measures, such as longitudinal follow-up and parasite quantification. Longitudinal follow-up is particularly important for differentiating between infections that appear asymptomatic at time of detection, but may become symptomatic after the initial detection (Rottmann *et al*., 2006).

Antibodies to *P. falciparum* AMA1 may contribute to protective immunity against clinical malaria by inhibiting blood stage growth of *P. falciparum*, and AMA1 is a leading malaria vaccine candidate. Currently, there is limited knowledge of the acquisition of strain-specific and cross-reactive antibodies to AMA1 in humans, or the acquisition of invasion-inhibitory antibodies to AMA1.

1.2. Ligand-receptor interaction

The interaction between PfAMA1 and another parasite protein called rhoptry neck protein 2 (RON2) is essential for tight junction formation, which commits the merozoite for invasion

(Srinivasan *et al*., 2011). RON2 is part of a larger RON complex that also contains RON4 and RON5 (Alexander *et al*., 2005). The RON complex appears to be released from the merozoite's rhoptry organelles prior to penetration and embeds in the erythrocyte surface where it serves as an attachment point for AMA1 (Lamarque *et al*., 2011).

Several lines of evidence support this AMA1–RON2 model of the tight junction. The AMA1–RON2 interaction can be disrupted with small peptides such as RON2L, which competes with native RON2 protein and leads to inhibition of merozoite invasion (Srinivasan *et al*., 2011). AMA1-binding monoclonal antibodies inhibit merozoite invasion by blocking the AMA1–RON2 interaction.

AMA1 is an important vaccine candidate and seems to be a vital target of acquired immunity, hence it plays a key role in erythrocyte invasion (Lamarque *et al*., 2011) and antibodies raised against AMA1 or affinity-purified AMA1 antibodies from naturally exposed individuals inhibit merozoite invasion *in vitro* (Remarque *et al*., 2008). Immunization of animals with AMA1 can protect against blood stage challenge with the homologous strain, but less effectively against heterologous strains due to antigenic diversity (Doolan *et al*., 2009).

Antibodies to AMA1 are typically highly prevalent amongst people in malaria endemic populations. Some longitudinal studies have associated antibodies to recombinant AMA1 measured by ELISA with reduced risk of malaria (Osier *et al*., 2008). In a recent clinical trial of the vaccine FMP2.1/$AS02_A$ containing recombinant AMA1 of the 3D7 strain, there was no significant protection against clinical malaria overall, but there was a significant reduction in risk of clinical malaria caused by parasites expressing vaccine-like AMA1 alleles, suggesting strain-specific protective efficacy (Ouattara *et al*., 2013a). These results support the development of AMA1 as a malaria vaccine, but highlight the need to better understand antigenic diversity of AMA1 and the functional activity of antibodies against AMA1.

Antibodies to AMA1 are thought to contribute to protective immunity by inhibiting erythrocyte invasion and blood-stage replication of *P. falciparum*. However, to date, it has not been possible to directly measure AMA1-specifc inhibitory antibodies among individuals in relation to protection from clinical malaria, and better understand the acquisition of inhibitory antibodies. Furthermore, knowledge on the acquisition of antibodies to polymorphic and conserved epitopes in relation to immunity is limited (Osier *et al*., 2008).

AMA1 is undoubtedly an important parasite invasion ligand (Figure 4). Readily identifiable AMA1 orthologues exist across the genus Apicomplexa, and genetic deletion experiments have been unsuccessful and largely shown that AMA1 is essential (Sheehy *et al*., 2013). AMA1 is a micronemal type I transmembrane protein that translocates to the surface of invasive zoites, including the *P. falciparum* merozoite, and is localised at the moving junction during invasion (Narum & Thomas, 1994).

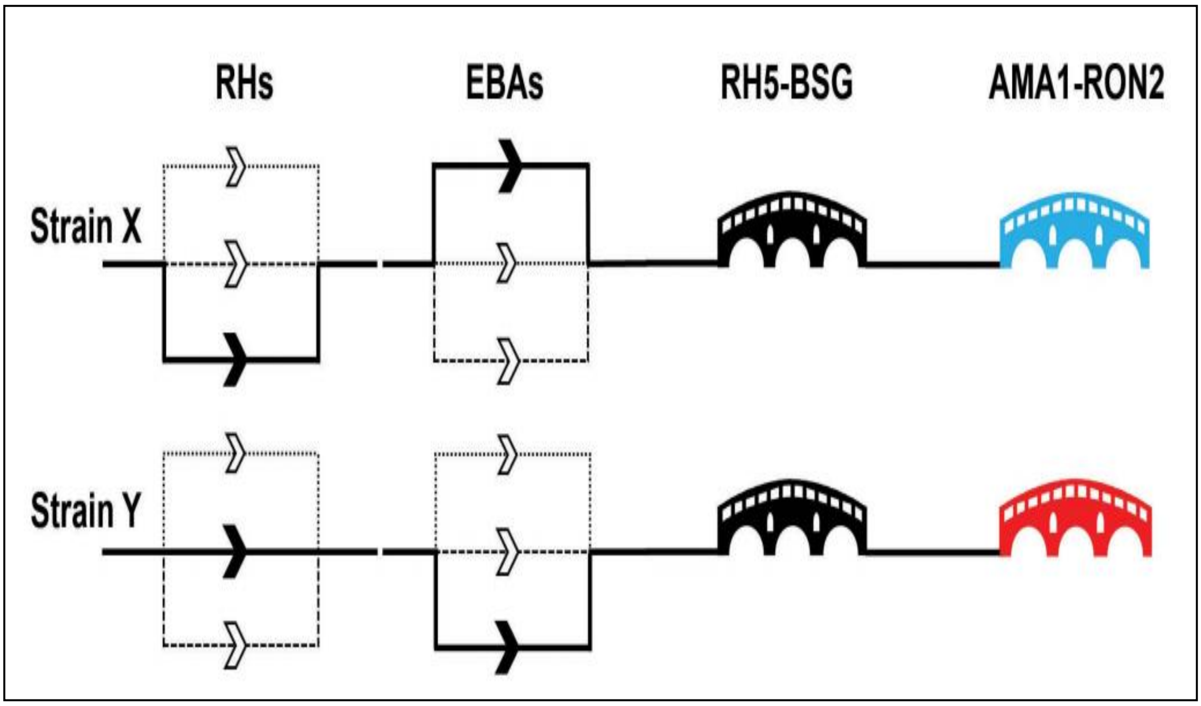

Figure 4: A molecular understanding of invasion leading to the identification of critical target points. RH5–basigin and AMA1–RON2 interactions are represented by critical ‘‘bridges.’’ The immunogenic AMA1 protein is highly variable between strains and is therefore represented by different colours: neutralising host antibodies elicited by one AMA1 variant would not protect against a strain containing a different AMA1 variant (Wright & Rayner, 2014).

Multiple lines of evidence indicate that polymorphisms in the *P. falciparum* AMA1 domain I result from selective pressures exerted by protective host immune responses (Corte *et al*., 2003) In a study in Papua New Guinea, a pattern of geographical diversity and the particular substitutions found were suggestive of strong constraints acting on the evolution of AMA1 at the population level. In addition, differences between the sequences of AMA1 domain I from symptomatic and asymptomatic infections implicate AMA1 as a possible determinant of the morbidity associated with a particular *P. falciparum* strain (Cortes *et al*., 2003).

1.3. Crystal structure of AMA1

The crystal structure of AMA1 reveals a long hydrophobic trough in domain I that appears to be a binding site for proteins creating an erythrocyte invasion complex comprised of AMA1 and RON proteins (Srinivasan *et al*., 2011). One end of this trough is flanked by several of the most polymorphic residues in the protein. These polymorphisms appear to have arisen due to diversifying selection and presumably allow the parasite to avoid invasion-inhibitory antibodies (Corte *et al*., 2003).

Knowledge of the distribution of the polymorphic sites on the surface of AMA1 is necessary to obtain a detailed understanding of their significance for vaccine development. The central two-thirds of AMA1 is relatively conserved among *Plasmodium* species as well as more distantly related apicomplexan parasites, and contains two clusters of disulfide-bonded cysteines termed domains I and II.

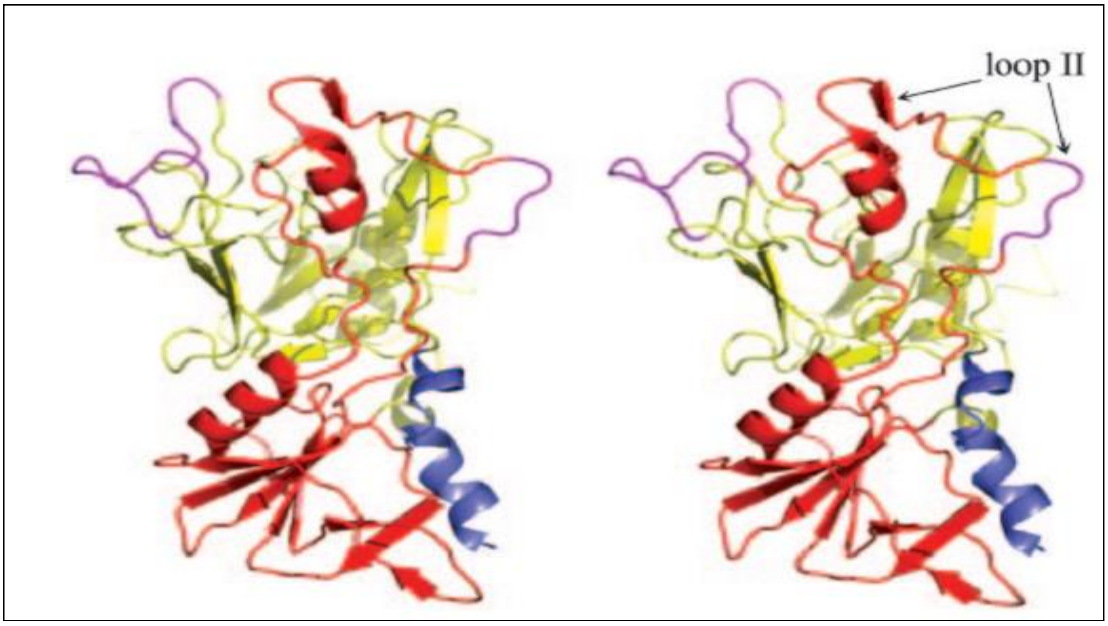

Figure 5: Stereo view of the AMA1 domain I+II structure showing the two interconnected domains. The 20 aa from the N-terminal extension are coloured blue, domain I is yellow, domain II is red, and loops that are disordered in the structure are violet (Bai *et al*., 2005).

The crystal structure of this fragment of AMA1 (Figure 5) reveals that domains I and II consists of two intimately associated PAN domains. PAN domain I contains many long loops that extend from the domain core and form a scaffold for numerous polymorphic residues. This extreme adaptation of a PAN domain reveals how malaria parasites have introduced significant flexibility and variation into AMA1 to evade protective human antibody responses. The polymorphisms on the AMA1 surface are exclusively located on one side of the molecule, presumably because this region of AMA1 is most accessible to antibodies reacting with the parasite surface. Moreover, the most highly polymorphic residues surround a conserved hydrophobic trough that is ringed by domain I and domain II loops. Examples set by viral receptor proteins would suggest that this is likely to be the AMA1 receptor binding pocket (Bai *et al*., 2005).

1.4. Overview of erythrocyte invasion by *Plasmodium falciparum* merozoites

Invasion of *P. falciparum* merozoites into erythrocytes starts with an initial weak attachment of the merozoite to the erythrocyte surface through parasite receptor–RBC ligand interactions, followed closely by a reorientation that ultimately brings up the apical end of the merozoite into close apposition with the red blood cell (RBC) surface (Srinivasan *et al.*, 2011). The merozoite then triggers the formation of a junction with the red blood cell that by electron microscopy appears as a dense area below the erythrocyte membrane at the site of the merozoite's apposed apical end. In addition, the merozoite secretes its rhoptry contents into the RBC that may enable the invasion of the merozoite (Aikawa *et al.*, 1978) (Figure 6).

The merozoite then travels through the junction as it pulls itself into the RBC through connections flanked by parasite surface proteins and its actin–myosin motor (Baum *et al.*, 2006). Hence, the formation of the junction and its connection with the molecular motor through the cytoplasmic tail of parasite receptors is critical for invasion (Buscaglia *et al.*, 2003). Formation of the parasitophorous vacuole, created by the inward flow of the RBC membrane occurs co-ordinately with the entry of the parasite into the RBC.

At the end of invasion, the electron-dense junction becomes part of the parasitophorous vacuole that surrounds the newly invaded parasite (Aikawa *et al.*, 1978).

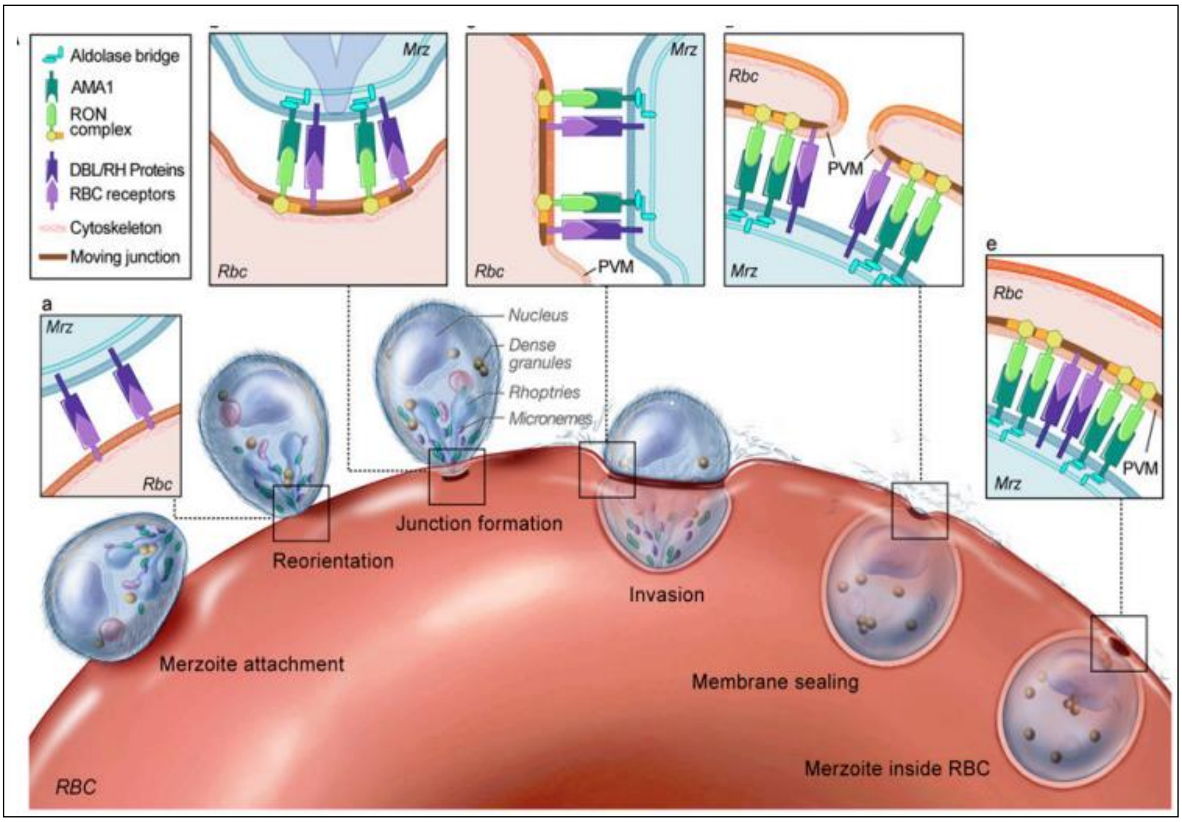

Figure 6: Schematic model of the steps involved in *P.falciparum* merozoite invasion (Srinivasan *et al*., 2011).

1.5 Justification

The burden of *P. falciparum* to individuals, households and national economies is enormous with the economic impact of malaria in Africa estimated at USD $12 billion every year. The economic impact includes the cost of health care, working days lost due to sickness, days lost in education, decreased productivity due to brain damage from cerebral malaria, and loss of investment and tourism (Greenwood *et al*., 2005). In 2001, malaria was ranked the 8th highest contributor to the global Disability Adjusted Life Year (DALY) and 2nd in Africa (WHO, 2002). The malaria DALY was largely estimated from the combined effects of *P. falciparum* infection as a direct cause of death and the much smaller contributions of short duration, self-limiting or treated surviving mild morbid events, malaria-specific anaemia and neurological disability following cerebral malaria (Murray & Lopez, 1997).

Malaria also has a devastating economic and social effect as it perpetuates poverty. It is both a root cause and consequence of poverty, burdening endemic countries and contributing to the cycle of poverty (WHO, 2011). Malaria affects the most isolated groups, such as poor women and children, in the most aggressive manner (WHO, 2011). There have been several historic attempts to provide an estimate of the numbers of deaths that occur each year due to malaria in Africa. Perhaps the most significant and influential was proposed by Leonard Bruce Chwatt of a million deaths each year (Bruce-Chwatt, 1952).

P. falciparum is responsible for 85% of the malaria cases while chronic infections may cause debilitating anaemia. The mortality levels are as high as 1 million per year with 90% of these deaths in sub Saharan Africa (WHO, 2011) and the cost of treatment borne by the victims especially when there are other accompanying complications is overwhelming.

The primary goal for malaria control would benefit from the development of a highly efficacious vaccine that protects against disease and interrupts transmission of *P. falciparum*. It is likely that such a vaccine will be multi-component, with antigens from different stages of the parasite life cycle.

The Apical membrane antigen 1 (AMA1) is considered as one of the potential candidates for inclusion in a vaccine against blood stages of *P. falciparum*. The polymorphisms in AMA1 have been attributed to the diversifying selection pressure due to immune responses (Drew *et al*., 2012). It was therefore important to investigate the genetic diversity in *P. falciparum* AMA1 in a malaria endemic population.

1.6. Problem Statement

Currently there is no vaccine effective against the blood-stages of *P. falciparum,* which causes the symptoms and severe manifestations of malaria. There are a myriad of challenges that are occasioned by a *P. falciparum* infection. The emergence of resistance to key antimalarial drugs such as chloroquine and pyrimethamine has been a setback for malaria control programs based primarily on prompt and effective treatment. Recent reports on the resistance towards the artemisinin drug have compounded the problem (WHO, 2011).

RTS, S is the most advanced candidate malaria vaccine but it is only partially protective and the causes of inter-individual variation in efficacy are poorly understood (Warimwe *et al*., 2013). RTS, S is currently in phase III trials in 6- to 12-week-old infants and 5- to 17-month-old children in Africa. In previous phase II trials conducted across 11 geographical sites in Africa, RTS, S, efficacy ranged between 34% and 65% (Warimwe *et al*., 2013). Pooled analysis of these phase II studies, as well as preliminary phase III data, found that RTS,S efficacy varied between individuals according to age at vaccination and the intensity of malaria transmission (Agnandji *et al*.,2012). Two *P. falciparum* blood-stage antigens, merozoites surface protein 1 (PfMSP1) and apical membrane antigen 1 (PfAMA1), have dominated blood-stage vaccine development, but appear to require high antibody concentrations to induce protection and suffer antigenic diversity rendering vaccine-induced antibodies strain-specific (Ouattara *et al*., 2013). There has never been a systematic head-to-head comparison of these and other candidate antigens delivered using the same human-compatible vaccine platform. Estimating the malaria public health burden continues to be driven by informed approximations, in part because of the paucity of reliable and accurate data but also due to the inherent difficulties of unique diagnosis (Snow *et al*., 2003). *Plasmodium falciparum* is known to show broad genetic diversity, mainly among surface antigens that have been under selective immune pressure and generally reflected as the main targets of subunit vaccines. Unfortunately, this great genetic diversity poses a key task for effective vaccine development since it could lead to vaccine-resistant malaria with non-vaccine type parasites growing in frequency within vaccinated populations (Takala *et al.*, 2009).

Although there are a large number of different AMA1 alleles circulating in human populations, recent studies have suggested that the extent of antigenic diversity may be limited, as evidenced by substantial cross-inhibitory activity of antibodies to isolates expressing different AMA1 alleles (Drew *et al.*, 2012), and sequence analyses suggesting that AMA1 alleles may be clustered into a small number of related groups.

The aim of this study was to understand the genetic diversity in Apical Membrane Antigen 1 in a malaria endemic population, define the polymorphisms and assess if there are any differences in allele frequencies between 2007 to 2010 in a community cross-sectional population. Since, AMA1 is a potential vaccine candidate; the study aims to describe the allelic types of AMA1 in a malaria endemic population and their temporal distribution.

1.7. Objectives

1.7.0. Overall objective:

To define the temporal distribution of *P. falciparum* Apical Membrane Antigen 1 (AMA1) haplotypes in a cross-sectional study of asymptomatic individuals living in a malaria endemic region in Kilifi, Kenya.

1.7.1. Specific objectives:

1) To define the genetic polymorphisms within the full length AMA1 gene.
2) To assess differences in allele frequencies between the years 2007 and 2010 in asymptomatic individuals from annual cross-sectional blood surveys.
3) To determine whether the polymorphisms are associated with the regions involved in evading the host immune response.

1.8. Hypothesis

The apical membrane antigen 1 (AMA1) gene in *Plasmodium falciparum* is not genetically diverse among different isolates.

CHAPTER TWO

MATERIALS AND METHODS

2.0. Study Area

The study was conducted in Kilifi County, Kenya. This region located at the Kenyan Coast adjacent to the Indian Ocean covers an area of 12,245.90 km^2 and has a population of 1,109,735 as per Kenyan census in 2009 (Figure 7).

Figure 7: Map of Kenya showing location of Kilifi County shaded in green (Wikipedia.org).

2.1. Collection of Blood Samples

Between 2007 and 2010, a number of households within Junju location in Kilifi were visited by a field-worker once a week, and axillary temperature obtained from every consenting participant using an electronic digital thermometer. The samples were collected every year before the rainy season (transmission season), and it was 1 sample per participant per year unless they fell ill with malaria then another sample is taken. Any participant with a temperature <36°C was tested twice more, to ensure that the apparent low temperature was not the result of poor placement of the thermometer. Any participant with a fever (axillary temperature ≥37.5 °C) or a history of fever was given local bus fare to travel to the study clinic at the Kilifi District Hospital, where a blood smear was performed. The study clinic was open daily, and participants were encouraged to seek treatment whenever they were ill. All investigations and treatments provided at the study clinic were free. Thick and thin blood smears were air-dried, and the thin blood smears were fixed in 100% methanol. Slides were then stained in 2% Giemsa (diluted in a buffer with a pH of 7.2) for 30 min and were analysed immediately. The parasitaemia range that was considered fell between 32-740,000 parasites/ml. Two hundred and forty one venous blood samples were collected and archived from annual cross-sectional sampling and were used in this study. Asymptomatic *Plasmodium* infections are highest among the school going children as depicted by Kimbi *et al.*, 2005. The participants of 12 years and below who were asymptomatic formed the study population. DNA extraction from the whole blood samples was carried out using the Qiagen DNA blood mini kit according to the manufacturer's instructions.

2.2. Oligonucleotide Primers

The following primer pairs were used to amplify the AMA1 gene; AMA1 F1 + AMA1 R1, AMA1 F143 + AMA1 R2, AMA1 F344 + AMA1 R2. The sequences of the AMA1 primers used were originally described Polley & Conway, 2001 and were designed to flank the predicted open reading frame covering position 29 to 1843 of the AMA1 gene. The positions of the primer sequence relative to the AMA1 gene sequence are indicated in figure 8.

ATGAGAAAATTATACTGCGTATTATTATT**GAGCGCCTTTGAGTTTAC**ATATATGATAAAC
TTTGGAAGAGGACAGAATTATTGGGAACATCCATATCAAAATAGTGATGTGTATCGTCCA
ATCAACGAACATAGGGAACATCCAAAAGAATACGAATATCCATTACACCAGGAACATACA
TACCAACAAGAAGATTCAGGAGAAGACGAAAATACATTACAACACGCATATCCAATAGAC
CACGAAGGTGCCGAACCCGCACCACAAGAACAAAATTTATTTTCAAGCATTGAAATAGTA
GAAAGAAGTAATTATATGGGTAATCCATGGACGGAATATATGGCAAAATATGATATTGAA
GAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTGAAGTAGCTGGAACT
CAATATA**GACTTCCATCAGGGAAATGTCC**AGTATTTGGTAAAGGTATAATTATTGAGAAT
TCAAATACTACTTTTTTAACACCGGTAGCTACGGGAAATCAATATTTAAAAGATGGAGGT
TTTGCTTTTCCTCCAACAGAACCTCTTATGTCACCAATGACATTAGATGAAATGAGACAT
TTTTATAAAGATAATAAATATGTAAAAAATTTAGATGAATTGACTTTATGTTCAAGACAT
GCAGGAAATATGATTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTAT
GATGACAAAGATAAAAAGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCT
AGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAA
GATATATCATTTCAAAACTATACATATTTAAGTAAGAATGTAGTTGATAACTGGGAAAAA
GTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAATTGT
GAAGATATACCACATGTAAATGAATTTCCAGCAATTGATCTTTTTGAATGTAATAAATTA
GTTTTTGAA**TTGAGTGCTTCGGATCAACCTAA**ACAATATGAACAACATTTAACAGATTAT
GAAAAAATTAAAGAAGGTTTCAAAAATA**AGAACGCTAGTATGATCAAAAG**TGCTTTTCTT
CCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGTCATGGTAAGGGTTATAATTGGGGA
AATTATAACACAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATT
AACAATTCATCATACATTGCTACTACTGCTTTGTCCCATCCCATCGAAGTTGAAAACAAT
TTTCCATGTTCATTATATAAAGATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGA
ATTAAATTAAATGATAATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTT
ATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGT
ACATGTCGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAAT
GAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAAACCA
ACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACT
ATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAATATGATAAAATGGAT
GAACCACAAGATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGCA
TCTTTTTGGGGGGAAGAAAAAAGAGCATCACATA**CAACACCAGTTCTGATGGA**AAAACCA
TACTATTAA

Figure 8: AMA1 gene indicating the relative positions of the primers.

Key:

Primer ID	5′-3′ Primer sequence	Description
AMA1 F1	**GAGCGCCTTTGAGTTTAC**	Forward outer primer sequence
AMA1 R1	**CTTTTGATCATACTAGCGTTCT**	Reverse outer primer sequence
AMA1 F143	**GACTTCCATCAGGGAAATGTCC**	Forward inner primer sequence
AMA1 F344	**TTGAGTGCTTCGGATCAACCTAA**	Forward inner primer sequence
AMA1 R2	**TCCATCAGAACTGGTGTTG**	Reverse inner primer sequence

2.3. Amplification of Target DNA sequences by Polymerase Chain Reaction

A 10 µl reaction was set up in two parts with the Taq polymerase on a different tube since it is so efficient and could kickstart the reaction before the addition of the DNA sample thus causing primer dimers as a result of primers annealing to themselves. In one micro centrifuge tube, 1 µl of 25 mM $MgCl_2$ stock solution, 0.2 µl dNTPs (Bioline, London UK) 100Mm , 0.3 µl of forward and reverse primers (10 µM), 0.5 µl of DNA template, and 2.7 µl DNase free water to give a final volume of 5.0 µl. In a second tube, 0.14 µl of Expand High Fidelity Taq Polymerase (Roche diagnostics, Mannheim Germany), 1µl of Expand High Fidelity Buffer (10X) with 15 mM $MgCl_2$ and 3.86 µl DNase free water to make a final volume of 5 µl and primers. The contents of the two tubes were then mixed to give a final volume of 10 µl per PCR reaction. PCR was carried out in 96 well plates using PTC-1196 thermocycler (BioRAD,USA). Conditions consisted of an initial denaturation at 94^oC for 2 min; followed by 10 cycles of denaturation at 94^oC for 15 sec, annealing at 56^oC for 30 sec, extension at 72^oC for 4 minutes; and 25 cycles of denaturation at 94^oC for 15 sec, annealing at 56^oC for 30 sec, extension at 72^oC for 4 min with an increment of 5 sec/cycle to give more room for the enzyme extension time since it tends to be weak during the last cycles; a final extension at 72^oC for 7 min.

2.4. Analysis of PCR products by gel electrophoresis

Following PCR amplification, agarose gel electrophoresis was used to resolve and visualize amplified fragments. Briefly, a 1% agarose gel was prepared using 0.5X Tris/Borate/EDTA (TBE) buffer. A total volume of a 100 ml stock of 10X TBE buffer was prepared. A 1:20 dilution of the TBE stock solution was made to get a working solution of 0.5X. In order to make a 1% agarose gel, 1g of agarose powder was weighed, added to 100 ml of 0.5X TBE and heated to boiling. The solution was then left to cool before adding 2 µl of Ethidium Bromide from a stock of 0.5 g/ml and pouring in a gel tray with combs to set. One µl of each PCR product, as well as negative and positive control samples were individually mixed with 1 µl of 6X Blue Orange loading dye (Bioline,London UK) and loaded into wells on the gel. A total of 1.5 µl each of 1kb Hyperladder 1 (Bioline,London UK) was loaded into the first and final wells on the gel and the samples electrophoresed, in 0.5X TBE buffer, for 45 min at 100 volts and the gel was viewed under ultraviolet light and gel images captured on a Molecular Imager Gel Doc (BioRAD ,Universal hood II).

2.5. Purification of PCR products by ExoSAP-IT reagent

The amplified PCR products were purified by ExoSAP-IT reagent (Affymetrix, Inc.USA). Briefly, 3.6 µl of ExoSAP-IT reagent was mixed directly with 9.0 µl of the amplified PCR fragments. The mixture was loaded onto a 96 well thermal cycler and incubated at 37^{o}C for 15 min to degrade the remaining primers and nucleotides. The products were then incubated at 80^{o}C for 15 min to inactivate the ExoSAP-IT enzyme then cooled down to 15^{o}C for 5 min and stored at -20^o C till required.

2.6. Big Dye Sequencing PCR reaction

Each reaction mixture was set up by combining 0.5 µl of Big Dye terminator (Applied biosystems, USA) ready reaction mix, 1.75 µl of 5X sequencing buffer, 1 µl of 10 µM sequencing primer, 3.75 µl of DNase free water and 3 µl of ExoSAP-IT cleaned PCR product to give a final volume of 10 µl per reaction. Each sequencing primer used in the reaction was added into a different master mix tube. The plates were then loaded onto the thermo cycler and a sequencing program was set up as follows: 25 cycles of denaturation at 96^{o}C for 30 sec, annealing at 50^oC for 15 sec and extension at 60^oC for 4 min, with a ramp rate of 1^oC/ sec between the different temperatures.

2.7. Purification of PCR products

Purification was carried out using Ethanol/Sodium Acetate precipitation in 96 well plates. Briefly, 3 µl of Sodium acetate, pH 5.2, 62.5 µl of 95% ethanol and 24.5 µl of distilled water to make a final volume of 90 µl per well. The premix was added to each well containing the PCR products. The plates were sealed with micro-seals and incubated at -20^OC for 30 min. After incubation, the plates were centrifuged at 3000 xg for 30 min at 4^oC using a 5810R bench centrifuge (Eppendorf). The seals were then removed and the plates inverted on paper towels. The inverted plates were centrifuged at 50 x g for 1 min at 4^OC. A total of 150 µl of ice cold (-20^OC) 70% ethanol was then added into each well, the plate sealed and centrifuged at 3000 xg for 10 min at 4^oC. The plates were inverted over paper towels and excess fluid gently drained and again centrifuged at 50 xg for 1 min at 4^OC while still inverted on paper towel. The plates were then covered with fresh paper towels and left on the bench to air dry.

2.8. Capillary electrophoresis

Once the plates were completely dry in approximately 30 min, 10 µl of Hi-Di formamide was added into each well. Capillary electrophoresis was performed in an automated 3130xl sequencer from Applied 29 Biosystems, UK. The sequencer was able to separate DNA fragments that differ by just one base pair. Each of the four ddNTPs had a special fluorescent dye of a different colour attached to it. These dyes gave light at a different wavelength when excited using a laser beam. The resulting fluorescence was picked out by a charge coupled device camera and converted into a chromatogram. As the fluorescently labelled extension products from the sequencing reaction migrated through the polymer passed the laser detector, each base was detected as a colour signal**.**

2.9. Sequence editing

Sequence data obtained from the sequencer was analysed using the Seqman program from DNASTAR Lasergene software version 11. The program was used to align contigs and identify polymorphic sites. Mis-aligned sequences were corrected manually. The sequences were then saved as consensus files. DNA sequences were aligned using the Clustal W multiple alignment function to identify SNPs. The files were saved in a FASTA format.

2.10. Statistical analysis

Using DnaSP software Tajima’s D statistic was calculated to determine DNA sequences evolving in a random manner and those evolving in a process that is non-random. A negative Tajima's D signifies an excess of low frequency polymorphisms relative to expectation and this could be evidence of purifying selection or a recent population expansion. A positive Tajima's D signifies low levels of both low and high frequency polymorphisms and can be due to balancing selection acting on the population. The software allows the user to use the sliding window option which calculates some measures or parameters (for example the nucleotide diversity) across a DNA region. In this method a window (segment of DNA) is moved along the sequences in steps. The parameter is calculated in each window, and the value is assigned to the nucleotide at the midpoint of the window. The results can also be presented graphically (by a line chart). In the graph the parameter (Y axis) is plotted against the nucleotide position (X axis).

CHAPTER THREE

RESULTS

3.0. PCR Amplification of DNA samples

Of the 241 DNA samples (collected between 2007 and 2010) that were used in the study, successful PCR amplification was achieved for 37 samples. Amplification of the rest of the samples was unsuccessful even after the several attempts of increasing DNA sample volumes. These samples were those that had low DNA concentrations and some were assumed to have degraded during storage in the freezer (-20 ^{o}C). The positive and negative controls were included in the different PCR reactions to check whether there was any contamination after optimizing the PCR conditions (Figure 9).

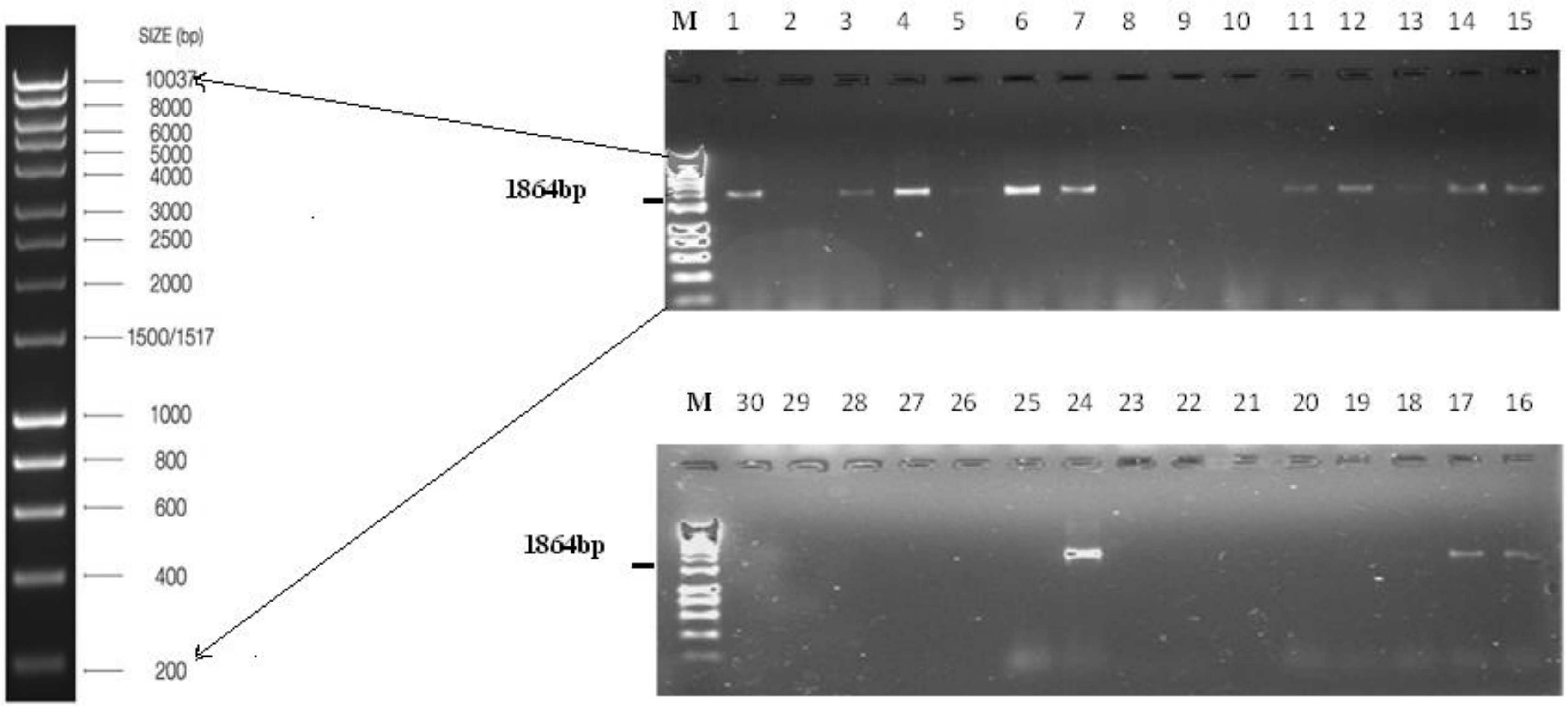

Figure 9: Gene amplification products of AMA1 at the first lane (M) in the gel picture represent Bioline's Hyperladder I used to determine amplicon size. Bands between lanes 1 to 30 represent amplified products of different samples with lane 24 indicating the positive control while lane 25 had the negative control. Positions that show no bands represent samples that were not able to be amplified. AMA1 F1 and AMA1 R1 primers were used in the amplification process to obtain the full length gene.

3.1. AMA1 sequence analysis

Seqman application (DNASTAR Lasergene Suite, Version 11) was used for analysis. For each sample, sequences generated from the different primer extensions were aligned into contigs and each primer trace file assessed for the quality of peaks and base calling. The reference sequence (AMA1)-PF3D7_1133400 (PlasmoDB) was used to scaffold the trace data generated from each primer. Corrections to base calling were done on the basis of the peaks of the electropherogram independent of the reference sequence.

Primer trace files were generated for 24 samples of AMA1, 13 samples generated poor data hence the electropherograms could not be read as some of them contained short sequences while others had multiple peaks at the same positions, making accurate conclusions and analysis impossible. The samples that failed to yield good sequences corresponded to be those that did not have distinct PCR amplification bands.

Clusters of polymorphisms that are likely to bring about antibody escape have been identified i.e CI-L, spanning amino acids 196 to 207. A MegAlign alignment of the 24 samples used in the study showed that amino acids 187,191-195,198-199 and 205 were conserved in all samples (Table 1).The haplotypes were also identified.

As shown in Table 1 some sequences had ‘X’ in certain positions within the sequences and this was as a result of mixed nucleotides hence they were excluded since this would make it impossible to determine which nucleotide belongs to a particular haplotype. A total of 14 haplotypes were determined. Sequences from sample 230_1_08 and 243_6_08 (Table 2) had similar amino acid alignment thus forming one haplotype.

Table 1: AMA 1 amino acids 187 to 207

	SAMPLES	AMINO ACIDS																					
		C187	C188	C189	C190	C191	C192	C193	C194	C195	C196	C197	C198	C199	C200	C201	C202	C203	C204	C205	C206	C207	
	3D7	**E**	**P**	**L**	**M**	**S**	**P**	**M**	**T**	**L**	**D**	**E**	**M**	**R**	**H**	**F**	**Y**	**K**	**D**	**N**	**K**	**Y**	**EPLMSPMTLDEMRHFYKDNKY**
1	111_6_09	K	P	L	I	S	P	M	T	L	N	G	M	R	D	L	Y	K	N	N	E	Y	KPLISPMTLNGMRDLYKNNEY
2	170_1_09	N	P	L	M	S	P	M	T	L	N	G	M	R	D	X	Y	K	N	N	E	Y	NPLMSPMTLNGMRDXYKNNEY
3	288_6_09	E	P	L	I	S	P	M	T	L	D	Q	M	R	H	F	Y	K	D	N	E	Y	EPLISPMTLDQMRHFYKDNEY
4	457_1_09	E	P	H	M	S	P	M	T	L	D	E	M	R	H	F	Y	K	D	N	K	Y	EPHMSPMTLDEMRHFYKDNKY
5	462_8_09	E	P	H	M	S	P	M	T	L	D	E	M	R	H	F	Y	K	D	N	K	Y	EPHMSPMTLDEMRHFYKDNKY
6	152_3_09	N	P	P	M	S	P	M	T	L	X	X	M	R	D	L	Y	K	N	N	E	Y	NPPMSPMTLXXMRDLYKNNEY
7	153_8_08	N	P	L	M	S	P	M	T	L	N	G	M	R	D	L	Y	K	N	N	E	Y	NPLMSPMTLNGMRDLYKNNEY
8	208_4_08	E	P	L	I	S	P	M	T	L	D	D	M	R	D	F	Y	K	N	N	E	Y	EPLISPMTLDDMRDFYKNNEY
9	220_4_08	N	P	P	M	S	P	M	T	L	X	D	M	R	D	L	Y	K	N	N	E	Y	NPPMSPMTLXDMRDLYKNNEY
10	223_1_08	E	P	L	M	S	P	M	T	L	D	D	M	R	D	F	Y	K	X	N	E	Y	EPLMSPMTLDDMRDFYKXNEY
11	230_1_08	N	P	P	M	S	P	M	T	L	N	G	M	R	D	L	Y	K	N	N	E	Y	EPLMSPMTLDDMRDFYKXNEY
12	243_6_08	N	P	P	M	S	P	M	T	L	N	G	M	R	D	L	Y	K	N	N	E	Y	NPPMSPMTLNGMRDLYKNNEY
13	259_2_08	X	P	L	M	S	P	M	T	L	N	G	M	R	H	L	Y	K	X	N	E	N	XPLMSPMTLNGMRHLYKXNEN
14	261_2_08	N	P	L	I	S	P	M	T	L	N	G	M	R	D	L	Y	K	N	N	E	D	NPLISPMTLNGMRDLYKNNED
15	264_1_08	X	P	L	M	S	P	M	T	L	X	X	M	R	D	F	Y	K	N	N	E	Y	XPLMSPMTLXXMRDFYKNNEY
16	271_3_08	E	P	L	M	S	P	M	T	L	D	E	M	R	H	F	Y	K	D	N	K	Y	EPLMSPMTLDEMRHFYKDNKY
17	310_6_08	K	P	L	M	S	P	M	T	L	D	Q	M	R	H	F	Y	K	D	N	E	D	KPLMSPMTLDQMRHFYKDNED
18	343_2_08	E	P	L	M	S	P	M	T	L	D	D	M	R	X	F	Y	K	D	N	E	Y	EPLMSPMTLDDMRXFYKDNEY
19	649_5_08	N	P	L	M	S	P	M	T	L	N	X	M	R	D	L	Y	K	X	N	E	Y	NPLMSPMTLNXMRDLYKXNEY
20	277_1_08	N	P	L	X	S	P	M	T	L	N	X	M	R	D	F	K	X	N	N	E	Y	NPLXSPMTLNXMRDFYKXNEY
21	287_0_10	E	P	L	M	S	P	M	T	L	D	D	M	R	D	F	Y	K	N	N	E	Y	EPLMSPMTLDDMRDFYKNNEY
22	337_5_10	N	P	P	M	S	P	M	T	L	D	Q	M	R	H	F	Y	K	D	N	K	Y	NPPMSPMTLDQMRHFYKDNKY
23	369_8_10	E	P	L	M	S	P	M	T	L	D	D	M	R	H	F	Y	K	D	N	E	Y	EPLMSPMTLDDMRHFYKDNEY
24	677_2_07	N	P	L	M	S	P	M	T	L	N	G	M	R	Y	F	Y	K	D	N	E	D	NPLMSPMTLNGMRYFYKDNED

Table 2: Polymorphic regions within C1-L of AMA1 samples

		C187	C189	C190	C196	C197	C200	C201	C203	C204	C206	C207	Haplotypes
	3D7	E	L	M	D	E	H	F	K	D	K	Y	ELMDEHFKDKY
1	111_6_09	K	L	I	N	G	D	L	K	N	E	Y	KLINGDLKNEY
2	288_6_09	E	L	I	D	Q	H	F	K	D	E	Y	ELIDQHFKDEY
3	457_1_09	E	H	M	D	E	H	F	K	D	K	Y	EHMDEHFKDEY
4	462_8_09	E	H	M	D	E	H	F	K	D	K	Y	EHMDEHFKDKY
5	153_8_08	N	L	M	N	G	D	L	K	N	E	Y	NLMNGDLKNEY
6	208_4_08	E	L	I	D	D	D	F	K	N	E	Y	ELIDDDFKNEY
7	230_1_08	N	P	M	N	G	D	L	K	N	E	Y	NPMNGDLKNEY
8	243_6_08	N	P	M	N	G	D	L	K	N	E	Y	NPMNGDLKNEY
9	261_2_08	N	L	I	N	G	D	L	K	N	E	D	NLINGDLKNED
10	271_3_08	E	L	M	D	E	H	F	K	D	K	Y	ELMDEHFKDKY
11	310_6_08	K	L	M	D	Q	H	F	K	D	E	D	KLMDQHFKDED
12	287_0_10	E	L	M	D	D	D	F	K	N	E	Y	ELMDDDFKNEY
13	337_5_10	N	P	M	D	Q	H	F	K	D	K	Y	NPMDQHFKDKY
14	369_8_10	E	L	M	D	D	H	F	K	D	E	Y	ELMDDHFKDEY
15	677_2_07	N	L	M	N	G	Y	F	K	D	E	D	NLMNGYFKDED

Samples from 2008 recorded the highest number of haplotypes due to their larger sample size, with 2007 recording the least (Table 3).

Table 3: Number of identified haplotypes

Year	No of samples	Haplotypes
2007	1	1
2008	14	6
2009	06	4
2010	3	3
Total	24	14

The frequency of a particular amino acid in different samples differed in specific regions within the amino acids alignment compared to the reference sequence, 3D7 (Table 2). These frequencies were calculated in excel (Microsoft office). Tyrosine(Y) had the highest frequency i.e. 80% at codon 207 with Tyrosine again recording the lowest frequency of 6% at codon 200. Six positions in C1-L were dimorphic i.e. 2 amino acids per locus; Codon190, Codon196, Codon201, Codon204, Codon206 and Codon207. Three of the positions within C1-L were trimorphic while position 197 had four different amino acids (Table 4).

Table 4: Amino acid frequency in each polymorphic codon

CODONS	**Frequency of amino acid per codon (%)**			
C187	E-47%	K-13%	N-40%	
C189	L-67%	H-13%	P-20%	
C190	M-73%	I-27%		
C196	D-60%	N-40%		
C197	E-20%	D-20%	G-40%	Q-20%
C200	D-47%	H-47%	Y-6%	
C201	L-33%	F-67%		
C204	D-53%	N-47%		
C206	E-73%	K-27%		
C207	D-20%	Y-80%		

The calculated frequencies from table 4 depict position 197 as the highly polymorphic region within the C1-L.

3.3. Statistical test for neutrality

To determine whether the single nucleotide polymorphisms were under any selection a statistical test for neutrality was conducted, Tajima's D, using DnaSP software on samples from 2008 and 2009 due to their fairly large sample size as compared to 2007 and 2010. Tajima's D values calculated were not statistically significant ($p>0.10$); 2008 n=14 (D: 0.58553), 2009 n=6 (D: 1.37177), this means AMA1 sequence was evolving randomly. Tajima's D was also calculated using a sliding window 100 sites long with a step size of 25 bases. The results were used to create DnaSP graphs (Figure 10).

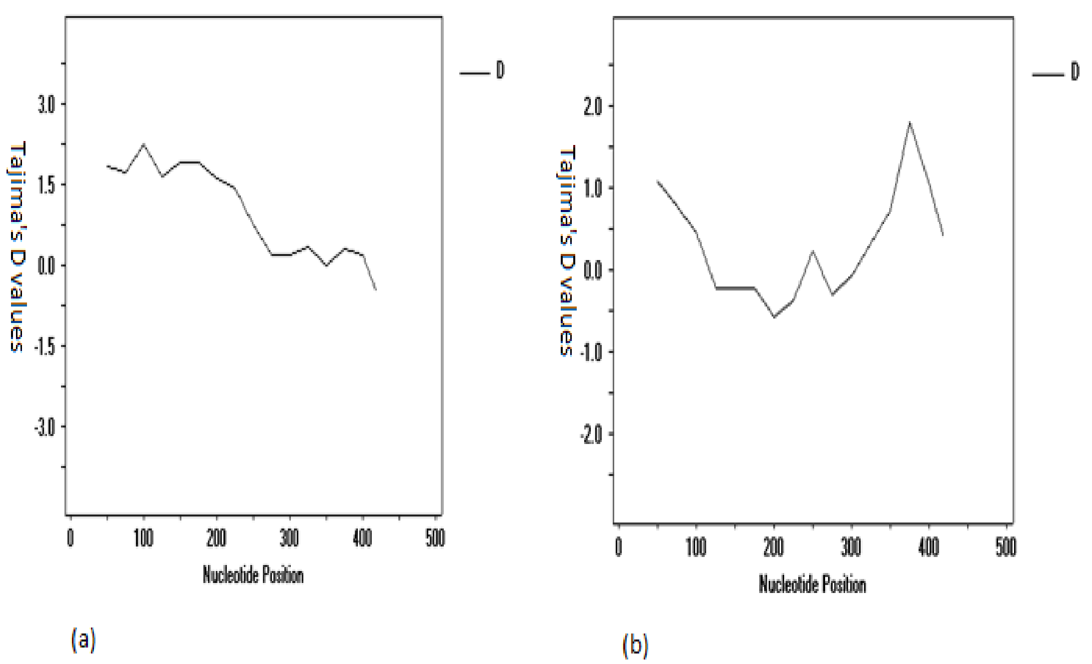

Figure 10: Tajima's D sliding window graphs for (a) 2008 and (b) 2009 samples. The graph line shows the relationship between the nucleotide position on the x-axis and the Tajima's D values on the y-axis.

CHAPTER FOUR

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

4.0: Discussion

Apical membrane antigen 1 is an important vaccine candidate that is expressed in mature stage parasites and is thought to be essential for invasion (Peterson *et al*., 1989). *P. falciparum* AMA1 is a type I integral membrane protein that is produced as an 83kDa 98/precursor and is localized initially to the micronemes, apical organelles of the merozoite (Bannister *et al*., 2003). Eight conserved intramolecular disulfide bonds constrain this protein into three distinct domains (Crewther, 1996). Shortly after synthesis, this precursor is cleaved to a 66-kDa product which is translocated onto the merozoite surface where much of the ectodomain is shed during invasion (Howell *et al*., 2003).

An effective malaria vaccine is expected to confer similar or better immunity to malaria susceptible individuals compared to that of adults who are resident in endemic areas, but over a shorter period of time (Kusi *et al*., 2011**).** PfAMA1 is an essential protein for merozoite invasion in *P.falciparum* and either directly or indirectly plays a role in resealing of the red blood cell at the posterior end of the invasion event (Yap *et al*., 2014). While AMA1 is a major candidate for a blood-stage vaccine, the issue of polymorphism of AMA1 in various isolates of *P. falciparum* is potentially a formidable problem for a vaccine based on this protein (Conway, 1997). This problem is compounded by the partial information on the biological roles of the various areas of the AMA1 molecule and on the epitopes in those sections which present effective antibody targets (Miura *et al*., 2007).

In this study, it was shown that the polymorphisms within the *P. falciparum* AMA1 gene occur at particular regions. Though the whole gene was sequenced codons 187-207 were the focus of the study while comparing to the reference AMA1 sequence-PF3D7_1133400 that was obtained from PlasmoDB (Appendix3). Within the population of *P. falciparum* AMA1 sequences there were five ‘‘highly polymorphic’’ residues (positions 187, 197, 200, 230, and 243). Interestingly a study conducted by Ouattara *et al.*, 2013 identified position 197 as the most important polymorphic site harbouring four different amino acids thus glycine (G), glutamine (Q), glutamic acid (E) and aspartic acid (D) which agrees to the findings of this study too as depicted in Table 4.

Most AMA1 polymorphic sites in this region were dimorphic, with only two amino acid residues found in the population (Bai *et al*., 2005). Likewise in this study, six positions in C1-L were dimorphic i.e. 2 amino acids per locus; Codon190, Codon196, Codon201, Codon204, Codon206 and Codon207. Three of the positions within C-1L were trimorphic while position 197 had four different amino acids (Table 4). The hydrophobic trough in AMA1 surface contains several polymorphic residues which are exclusively within this particular region of the AMA1 protein molecule and apparently it is because this region is widely exposed to the antibodies reacting with the surface of the parasite as explained by Bai *et al*,. 2005. Thus, the highly polymorphic residues are located around a highly hydrophobic trough.

Clusters of polymorphisms that might contribute to antibody escape have been identified on all three domains of AMA1 (Dutta *et al*., 2007) although domain 1 appears to be the major target of inhibitory antibodies (Healer *et al*., 2005). One cluster known as C1-L, spans amino acids 196 to 207 of domain 1. As explained by a study conducted by Ouattara *et al*., 2013, the strongest barrier induced by the malaria vaccine was elicited at amino acid position 197, which is located in C1-L of the domain 1. Alleles defined on the basis of this position depicted vaccine efficacy which was identical to that for the whole C1-L, suggesting that this position may be the most critical amino acid in antibody binding.

A study was conducted in Bandiagara, Mali based on allelic-specific efficacy on blood stage malaria vaccine. It showed that position 197 was the most important polymorphic site for characterizing AMA1 allelic identity by all 3 methods of analysis used to assess the role of specific amino acid positions. Amino acid at this position was used to define alleles to assess the time to the first clinical malaria episode with a 3D7 allele. The allele-specific efficacy data calculated using only position 197 to define alleles revealed vaccine efficacy identical to that for the whole C1-L haplotype against 5 known C1-L haplotypes alleles. Instead of considering all polymorphic locations of AMA1 to define haplotypes, position 197 alone might be used to describe which alleles to include in a vaccine (Ouattara *et al*., 2013). Position 197 was shown to be highly polymorphic as compared to other positions within the C1-L The amino acids in C1-L are located within a hydrophobic pocket in AMA1.

Bai *et al*., 2005 conducted a study to reveal a clustering of polymorphisms that surrounds the hydrophobic trough of *P falciparum* AMA1. It was revealed that the selective acquisition of several loops on AMA1 domains I and II during evolution of the PAN domains suggested that the loops served a purpose, possibly that of ‘‘protecting’’ a functionally critical portion of the molecule. Examination of the region between the loops revealed the presence of an extended pocket with a base that contained a series of hydrophobic side chains. This hydrophobic trough consisted of nine hydrophobic amino acid side chains that were solvent exposed and hydrophobic in all *Plasmodium* AMA1 sequences. Tyrosine 251, at the centre of the trough, rised above the floor of the trough, and is identical in all AMA1 sequences, even those of the more distantly related *Toxoplasma gondii* and *B. bovis* parasites). The overall features of the hydrophobic trough are therefore conserved in all AMA1 molecules. Consistent with the functional importance of the hydrophobic trough is the presence of polymorphic residues on the loops that surround the trough.

Since AMA1 is a potential vaccine candidate, it's in order to describe the allelic types of AMA1 in a malaria endemic population and their temporal distribution. A Tajima's D test was done to evaluate whether the sequenced data showed randomly evolving mutations thus 'neutral' or mutations under selection i.e 'non neutral'. Tajima's values for 2008 and 2009 were 0.58553 and 1.37177 respectively. None of the values was significant ($p > 0.10$). The AMA1 sequences obtained from 2008 and 2009 seem to be evolving in a random mode conferring to the values obtained from Tajima's D statistic thus not to interfere with their natural roles. This test for neutrality showed that the SNPs were not under selection implying that they were not affected by natural selection and mutations occurring in this gene are neutral.

The large number of polymorphic sites in a study conducted by Conway *et al*., 1997 allowed an analysis of individual AMA1 domains with Domain I clearly showing a significant excess of intraspecific nonsynonymous polymorphisms. This positive trend is also seen, although not significantly, in the other two domains, which have fewer polymorphic nucleotides.

P. falciparum AMA1gene is under selection according to (Polley & Conway, 2001), the strongest selection appears to act on sequences encoding domains I and III, and they hypothesized that this is produced by the host immune system mounting an effective response to epitopes within these domains of the protein. They further predicted that immunological studies will demonstrate that human antibodies to polymorphic sequences in one or both of these domains inhibit parasites and protect against malaria.

4.1. Conclusion

P. falciparum apical membrane antigen 1 (PfAMA1), a candidate malaria vaccine, is polymorphic. Several single nucleotide polymorphisms were detected in the AMA1 gene. The findings of this study are in agreement with other studies that had been conducted earlier on the genetic diversity within the AMA1 polymorphisms within the C1-L. The single nucleotide polymorphisms obtained from this study population in Kilifi County SNPs were not under selection and mutations occurring in this gene are neutral thus they had no effect in an individual's well-being. This implies that the mutations are not affected by natural selection.

These results suggest the polymorphisms detected are associated with the regions involved in evading the host immune response. In addition the polymorphisms on the AMA1 surface are strategically located on a particular side of the molecule, presumably because this region of AMA1 is more exposed to antibodies interacting with the parasite surface. Moreover, the most highly polymorphic residues surround a conserved hydrophobic trough. This hydrophobic pocket is the potential site for formation of the ligand-receptor junction which enhances the merozoite invasion. This part of the molecule is exposed to the antibodies hence its polymorphic nature potentially for evading immune attack.

4.2. Recommendations

The following recommendations were reached after the study:

1. The sample size was small in this research project and for future studies require a much bigger sample size should be employed for more conclusive and elaborate findings.
2. The function of AMA1 is unknown, as is the mechanism by which antibodies prevent merozoite invasion, but there is a general consensus that AMA1 plays an important role in the invasion process. More research work focusing on AMA1 structure and function should be carried out.

REFERENCES

Aikawa, M., Miller, L. H., Rabbege, J. R., & Epstein, N. (n.d.). Freeze-Fracture Study on the Erythrocyte Membrane during Malarial Parasite Invasion, 55–62.

Agnandji S, *et al*., 2012 Safety and efficacy of the RTS,S/AS01E candidate malaria vaccine given with expanded-programme-on-immunisation vaccines: 19 month follow-up of a randomised, open-label, phase 2 trial. Lancet Infectious Diseases 11:741-749.

Alexander, D. L., Mital, J., Ward, G. E., Bradley, P., & Boothroyd, J. C. (2005). Identification of the moving junction complex of *Toxoplasma gondii*: a collaboration between distinct secretory organelles. *PLoS Pathogens*, *1*(2), 17.

Alves, F. P., Durlacher, R. U. I. R., Menezes, M. J., Krieger, H., Pereira, L. H., & Camargo, E. P. (2002). High prevalence of asymptomatic *Plasmodium vivax* and *Plasmodium falciparum* infections in native amazonian populations, *66*(6), 641–648.

Bai, T., Becker, M., Gupta, A., Strike, P., Murphy, V. J., Anders, R. F., & Batchelor, A. H. (2005). Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(36), 12736–41.

Bannister, L. H., Hopkins, J. M., Dluzewski, A. R., Margos, G., Williams, I. T., Blackman, M. J. Mitchell, G. H. (2003). *Plasmodium falciparum* apical membrane antigen 1 (PfAMA-1) is translocated within micronemes along subpellicular microtubules during merozoite development. *Journal of Cell Science*, *116*(Pt 18), 3825–34.

Bannister, L. H., Mitchell, G. H., Butcher, G. A., Dennis, E. D., & Cohen, S. (1986). Structure and development of the surface coat of erythrocytic merozoites of *Plasmodium knowlesi*. *Cell and Tissue Research*, *245*(2).

Baum, J., Chen, L., Healer, J., Lopaticki, S., Boyle, M., Triglia, T, Cowman, A. F. (2009). Reticulocyte-binding protein homologue 5 - an essential adhesin involved in invasion of human erythrocytes by *Plasmodium falciparum*. *International Journal for Parasitology*, *39*(3), 371–80.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.-W.,Cowman, A. F. (2006). A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. *The Journal of Biological Chemistry*, *281*(8), 5197–208.

Bromide, U. S. E. E. (2012). Agarose gel electrophoresis (basic method) Brooke Jude Bard College 2012 To image the gel , use the BioRad imaging system . Get instructions from a faculty member or another student prior to use . Brooke Jude Bard College 2012.

Bruce-Chwatt Lj (1952). Malaria in African infants and children in Southern Nigeria *Tropical Medical Parasitology*,46 (2):173–200.

Buscaglia CA Coppens I,Hol WG, Nussenzweig V.(2003). Sites of interaction between aldolase and thrombospondin-related anonymous protein in *plasmodium .Molecular Biology of the Cell* Dec 14(12):4947-57.

Bustamante, L. Y., Bartholdson, S. J., Crosnier, C., Campos, M. G., Wanaguru, M., Nguon, C., Rayner, J. C. (2013). A full-length recombinant *Plasmodium falciparum* PfRH5 protein induces inhibitory antibodies that are effective across common PfRH5 genetic variants. *Vaccine*, *31*(2), 373–9.

Conway, D. J. (1997). Natural selection on polymorphic malaria antigens and the search for a vaccine. *Parasitology Today*, *13*(1), 26–29.

Corte, A., Mellombo, M., Mueller, I., Benet, A., Reeder, J. C., & Anders, R. F. (2003). Geographical Structure of Diversity and Differences between Symptomatic and Asymptomatic Infections for *Plasmodium falciparum* Vaccine Candidate AMA1, *71*(3), 1416–1426.

Cowman, A. F., Berry, D., & Baum, J. (2012). The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *The Journal of Cell Biology*, *198*(6), 961–71. Crewther, P. E. (1996). The Disulfide Bond Structure of Plasmodium Apical Membrane Antigen-1. *Journal of Biological Chemistry*, *271*(46), 29446–29452.

De Mast, Q., Syafruddin, D., Keijmel, S., Riekerink, T. O., Deky, O., Asih, P. B., van der Ven, A. J. (2010). Increased serum hepcidin and alterations in blood iron parameters associated with asymptomatic *P. falciparum* and *P. vivax* malaria. *Haematologica*, *95*(7), 1068–74.

Doolan, D. L., Dobaño, C., & Baird, J. K. (2009). Acquired immunity to malaria. *Clinical Microbiology Reviews*, *22*(1), 13–36, Table of Contents.

Drew, D. R., Hodder, A. N., Wilson, D. W., Foley, M., Mueller, I., Siba, P. M., Beeson, J. G. (2012). Defining the antigenic diversity of *Plasmodium falciparum* apical membrane antigen 1 and the requirements for a multi-allele vaccine against malaria. *PloS One*, *7*(12), e51023.

Dutta, S., Lee, S. Y., Batchelor, A. H., & Lanar, D. E. (2007). Structural basis of antigenic escape of a malaria vaccine candidate. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12488–93.

Fowkes, F. J. I., Richards, J. S., Simpson, J. a, & Beeson, J. G. (2010). The relationship between anti-merozoite antibodies and incidence of *Plasmodium falciparum* malaria: A systematic review and meta-analysis. *PLoS Medicine*, *7*(1), e1000218.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Barrell, B. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, *419*(6906), 498–511.

Gilson, P. R., & Crabb, B. S. (2009). Morphology and kinetics of the three distinct phases of red blood cell invasion by *Plasmodium falciparum* merozoites. *International Journal for Parasitology*, *39*(1), 91–6.

Healer, J., Triglia, T., Hodder, A. N., Gemmill, A. W., & Cowman, A. F. (2005). Functional Analysis of *Plasmodium falciparum* Apical Membrane Antigen 1 Utilizing Interspecies Domains, *73*(4), 2444–2451.

Hill, A. V. S. (2011). Vaccines against malaria. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1579), 2806–14.

Howell, S. a, Well, I., Fleck, S. L., Kettleborough, C., Collins, C. R., & Blackman, M. J. (2003). A single malaria merozoite serine protease mediates shedding of multiple surface proteins by juxtamembrane cleavage. *The Journal of Biological Chemistry*, *278*(26), 23890–8.

Kusi, K. a, Faber, B. W., van der Eijk, M., Thomas, A. W., Kocken, C. H. M., & Remarque, E. J. (2011). Immunization with different PfAMA1 alleles in sequence induces clonal imprint humoral responses that are similar to responses induced by the same alleles as a vaccine cocktail in rabbits. *Malaria Journal*, *10*(1), 40.

Laishram, D. D., Sutton, P. L., Nanda, N., Sharma, V. L., Sobti, R. C., Carlton, J. M., & Joshi, H. (2012). The complexities of malaria disease manifestations with a focus on asymptomatic malaria. *Malaria Journal*, *11*(1), 29.

Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., Lebrun, M. (2011). The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathogens*, *7*(2), e1001276.

Mackinnon, M. J., Gunawardena, D. M., Rajakaruna, J., Weerasingha, S., Mendis, K. N., & Carter, R. (2000). Quantifying genetic and nongenetic contributions to malarial infection in a Sri Lankan population, University of Oxford *97*(23), 3–8.

Males, S., Gaye, O., & Garcia, A. (2008). Long-term asymptomatic carriage of Plasmodium falciparum protects from malaria attacks: a prospective study among Senegalese children. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *46*(4), 516–22.

Marsh, K., & Kinyanjui, S. (2006). Immune effector mechanisms in malaria, 51–60. *Parasite Immunology*, September,2006, 28, 51–60

Miura, K., Zhou, H., Muratova, O. V, Orcutt, A. C., Giersing, B., Miller, L. H., & Long, C. a. (2007). In immunization with Plasmodium falciparum apical membrane antigen 1, the specificity of antibodies depends on the species immunized. *Infection and Immunity*, *75*(12), 5827–36.

Morrissette, N. S., & Sibley, L. D. (2002). Cytoskeleton of Apicomplexan Parasites, , 21–38. *Microbiology and Molecular biology Reviews*, Mar. 2002, *66*(1) p. 21–38

Murray, C. J., & Lopez, A. D. (1997). Mortality by cause for eight regions of the world: Global Burden of Disease Study. *Lancet*, *349*(9061), 1269–76.

Narum, D. L., & Thomas, A. W. (1994). Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of *Plasmodium falciparum* merozoites. *Molecular and Biochemical Parasitology*, *67*(1), 59–68.

Osier, F. H. a, Fegan, G., Polley, S. D., Murungi, L., Verra, F., Tetteh, K. K. a, Marsh, K. (2008). Breadth and magnitude of antibody responses to multiple *Plasmodium falciparum* merozoite antigens are associated with protection from clinical malaria. *Infection and Immunity*, *76*(5), 2240–8.

Ouattara, A., Mu, J., Takala-Harrison, S., Saye, R., Sagara, I., Dicko, A., Doumbo, O. K. (2010). Lack of allele-specific efficacy of a bivalent AMA1 malaria vaccine. *Malaria Journal*, *9*, 175.

Ouattara, A., Takala-Harrison, S., Thera, M. a, Coulibaly, D., Niangaly, A., Saye, R., Plowe, C. V. (2013). Molecular basis of allele-specific efficacy of a blood-stage malaria vaccine: vaccine development implications. *The Journal of Infectious Diseases*, *207*(3), 511–9.

Peterson, M. G., Marshall, V. M., Smythe, J. A., Crewther, P. E., Lew, A., Silva, A., Kemp, D. J. (1989). Integral Membrane Protein Located in the Apical Complex of *Plasmodium falciparum*, *Molecular and Cellular Biology*, July, *9*(7), 3151–3154.

Polley, S. D., & Conway, D. J. (2001). Strong Diversifying Selection on Domains of the *Plasmodium falciparum* Apical Membrane Antigen 1 Gene. *Genetics Society of America* Augus, p1505–1512 .

Remarque, E. J., Faber, B. W., Kocken, C. H. M., & Thomas, A. W. (2008). Apical membrane antigen 1: a malaria vaccine candidate in review. *Trends in Parasitology*, *24*(2), 74–84.

Richards, J. S., & Beeson, J. G. (2009). The future for blood-stage vaccines against malaria. *Immunology and Cell Biology*, *87*(5), 377–90.

Rodrigues, J., Martha, C., & Ladeia-andrade, S. (2006). A new challenge for malaria control in Brazil : asymptomatic *Plasmodium* infection - A Review, *101*(April), 229–237.

Rottmann, M., Lavstsen, T., Mugasa, J. P., Kaestli, M., Jensen, A. T. R., Müller, D., Beck, H.-P. (2006). Differential expression of var gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. *Infection and Immunity*, *74*(7), 3904–11.

Sheehy, S. H., Douglas, A. D., & Draper, S. J. (2013). Challenges of assessing the clinical efficacy of asexual blood-stage *Plasmodium falciparum* malaria vaccines. *Human vaccines & immunotherapeutics* 9:9, 1831–1840; September.

Snow, R. W., Hospital, J. R., & Newton, C. R. J. C. (2003). The Public Health Burden of *Plasmodium falciparum* Malaria in Africa : Deriving the Numbers, (11). *Disease Control Priorities Project*, working Paper No. 11.

Srinivasan, P., Beatty, W. L., Diouf, A., Herrera, R., Ambroggio, X., Moch, J. K., Miller, L. H. (2011). Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(32), 13275–80.

Warimwe, G. M., Fletcher, H. a, Olotu, A., Agnandji, S. T., Hill, A. V. S., Marsh, K., & Bejon, P. (2013). Peripheral blood monocyte-to-lymphocyte ratio at study enrollment predicts efficacy of the RTS,S malaria vaccine: analysis of pooled phase II clinical trial data. *BMC Medicine*, *11*(1), 184.

World Health Organisation Report 2011.

World Health Organisation Report 2002.

Wright, G. J., & Rayner, J. C. (2014). *Plasmodium falciparum* erythrocyte invasion: combining function with immune evasion. *PLoS Pathogens*, *10*(3), e1003943.

Yap, A., Azevedo, M. F., Gilson, P. R., Weiss, G. E., O'Neill, M. T., Wilson, D. W., Cowman, A. F. (2014). Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cellular Microbiology*, *16*(5), 642–56.

APPENDICES

Appendix 1 AMA1 sequences with 'X' depicting mixed bases

	SAMPLES	AMINO ACIDS											HAPLOTYPES
		C187	**C189**	**C190**	**C196**	**C197**	**C200**	**C201**	**C203**	**C204**	**C206**	**C207**	
	3D7	**E**	**L**	**M**	**D**	**E**	**H**	**F**	**K**	**D**	**K**	**Y**	**ELMDEHFKDKY**
1	111_6_09	K	L	I	N	G	D	L	K	N	E	Y	KLINGDLKNEY
2	152_3_09	N	P	M	X	X	D	L	K	N	E	Y	NLMXXDLKNEY
3	153_8_08	N	L	M	N	G	D	L	K	N	E	Y	NLMNGDLKNEY
4	170_1_09	N	L	M	N	G	D	X	K	N	E	Y	NLMNGDXKNEY
5	208_4_08	E	L	I	D	D	D	F	K	N	E	Y	ELIDDDFKNEY
6	220_4_08	N	P	M	X	D	D	L	K	N	E	Y	NPMXDDLKNEY
7	223_1_08	E	L	M	D	D	D	F	K	X	E	Y	ELMDDDFKXEY
8	230_1_08	N	P	M	N	G	D	L	K	N	E	Y	NPMNGDLKNEY
9	243_6_08	N	P	M	N	G	D	L	K	N	E	Y	NPMNGDLKNEY
10	259_2_08	X	L	M	N	G	H	L	K	X	E	N	XLMNGHLKXEN
11	261_2_08	N	L	I	N	G	D	L	K	N	E	D	NLINGDLKNED
12	264_1_08	X	L	M	X	X	D	F	K	N	E	Y	XLMXXDFKNEY
13	271_3_08	E	L	M	D	E	H	F	K	D	K	Y	ELMDEHFKDKY
14	277_1_08	N	L	X	N	X	D	F	X	N	E	Y	NLXNXDFXNEY
15	287_0_10	E	L	M	D	D	D	F	K	N	E	Y	ELMDDDFKNEY
16	288_6_09	E	L	I	D	Q	H	F	K	D	E	Y	ELIDQHFKDEY
17	310_6_08	K	L	M	D	Q	H	F	K	D	E	D	KLMDQHFKDED
18	337_5_10	N	P	M	D	Q	H	F	K	D	K	Y	NPMDQHFKDKY
19	343_2_08	E	L	M	D	D	X	F	K	D	E	Y	ELMDDXFKDEY
20	369_8_10	E	L	M	D	D	H	F	K	D	E	Y	ELMDDHFKDEY
21	457_1_09	E	H	M	D	E	H	F	K	D	K	Y	EHMDEHFKDEY
22	462_8_09	E	H	M	D	E	H	F	K	D	K	Y	EHMDEHFKDKY
23	649_5_08	N	L	M	N	X	D	L	K	X	E	Y	NLMNXDLKXEY
24	677_2_07	N	L	M	N	G	Y	F	K	D	E	D	NLMNGYFKDED

Appendix 2 alignment report of AMA1 sequences

```
Majority      --------------------------------------------------------------------------------------------------------------
                       10        20        30        40        50        60        70        80        90       100       110
111_6_09.seq  ---------------------------------------------------------------------------------------------TTTTCAAGCATGGAAAT 17
152_3_09.seq  -------------------------------------------------------------------------------------------------------------- 0
153_8_08.seq  -------------------------------------------------------------------------------------------------------------- 0
170_1_09.seq  -------------------------------------------------------------------------------------------------------------- 0
208_4_08.seq  --------------AGAAGACGAAAATACATTACAAGACGCATATCCAATAGACCACGAAGGTGCTGAACCCGCACCACAAGAACAAAATTTATTTTCAAGCATTGAAAT 96
220_4_08.seq  -------------------------------------------------------------------------------------------------------------- 0
223_1_08.seq  -------------------------------------------------------------------------------------------------------------- 0
230_1_08.seq  -------------------------------------------------------------------------------------------------------------- 0
243_6_08.seq  -------------------------------------------------------------------------------------------------------------- 0
259_2_08.seq  -------------------------------------------------------------------------------------------------------------- 0
261_2_08.seq  -----------------------------------------ATATCCAATAGACCACGAAGGTGCCGAACCCGCACCACAAGAACAAAATTTATTTTCAAGCATTGAAAT 69
264_1_08.seq  ------------------------------------------------------------------------------------------------------ATTGAAAT 8
271_3_08.seq  -------------------------------------------------------------------------------------------------------------- 0
277_1_08.seq  -------------------------------------------------------------------------------------------------------------- 0
287_0_10.seq  CAAGAAGATTCAGGAGAAGACGAAAATACATTACAACACGCATATCCAATAGACCACGAAGGTGCTGAACCCGCACCACAAGAACAAAATTTATTTTCAAGCATTGAAAT 110
288_6_09.seq  -------------------------------------------------------------------------------------------------------------- 0
310_6_08.seq  -------------------------------------------------------------------------------------------------------------- 0
337_5_10.seq  -------------------------------------------------------------------------------------------------------------- 0
343_2_08.seq  -------------------------------------------------------------------------------------------------------------- 0
```

(*Continued*) Appendix 2 alignment report of AMA1 sequences

```
Majority      ----------------------------------------------------------------------------------------XXXXXXXXXXXXXXXX--GCTG
                       120       130       140       150       160       170       180       190       200       210       220
111_6_09.seq  AGTAGAAAGAAGTAATTATATGGGTAATCCATGGACGGAATATATGGCAAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 127
152_3_09.seq  -------------------------------------------------------------------------------------------------------------- 0
153_8_08.seq  ----------------------------------------------------------------------------------------------------------GCTG 4
170_1_09.seq  -------------------------------------------------------------------------------------------------------------- 0
208_4_08.seq  ACCCCAAAGAAGTAATTATATGGGTAATCCATGGACGGAATATATGGCAAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 206
220_4_08.seq  --------------------------------------------TGGCAAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 66
223_1_08.seq  -------------------------------------------------------------------------------------------------------------- 0
230_1_08.seq  -------------------------------------------------------------------------------------------------------------- 0
243_6_08.seq  -------------------------------------------------------------------------------------------------------------- 0
259_2_08.seq  ----------------------------------------------------------ATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 52
261_2_08.seq  AGTAGAAAGAAGTAATTATATGGGTAATCCATGGACGGAATACCCGGCCAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGW-GCTG 178
264_1_08.seq  AGTAGAAAGAAGTAATTATATGGGTAATCCATGGACGGAATATATGGCAAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 118
271_3_08.seq  -------------------------------------------------------------------------------------------------------------- 0
277_1_08.seq  --------------------------------------------------------------------------------GTATAAGAGTAGATTTAGGAGAAGATGCTG 30
287_0_10.seq  AGTAGAAAGAAGTAATTATATGGGTAATCCATGGACGGAATATATGGCAAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 220
288_6_09.seq  -------------------------------------------------------------------------------------------------------------- 0
310_6_08.seq  ----------------------------------------------------------------------------------------GTAGATTTAGGAGAAGATGCTG 22
337_5_10.seq  -------------------------------------------------------------------------------------------------------------- 0
343_2_08.seq  -------------------------------------------------------------------------------------------------------------- 0
369_8_10.seq  -------------------------------------------------------------GAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 49
457_1_09.seq  ----------------------------------------------GCAAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 64
462_8_09.seq  AGTAGAAAGAAGTWATTATATGG-TAATCCATGGACGGAATATATGGCAAAATATGATATTGAAGAAGTTCATGGTTCAGGTATAAGAGTAGATTTAGGAGAAGATGCTG 142
649_5_08.seq  -------------------------------------------------------------------------------------------------------------- 0
677_2_07.seq  -------------------------------------------------------------------------------------------------------------- 0
```

(*Continued*) Appendix 2 alignment report of AMA1 sequences

```
Majority      AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA
                       230       240       250       260       270       280       290       300       310       320       330
111_6_09.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAAAACTACTTTTTTAACACCGGTAGCTACGGGA 237
152_3_09.seq  --------------------------------------------------------------------------------ACTACTTTTTTAACACCGGTAGCTACGGGA 30
153_8_08.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAAAACTACTTTTTTAACACCGGTAGCTACGGAA 114
170_1_09.seq  -----------------------------------------------------------------------------------------TTAACACCGGTAGCTACGGAA 21
208_4_08.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 316
220_4_08.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAAAACTACTTTTTTAACACCGGTAGCTACGGAA 176
223_1_08.seq  --------------------------------------------------------------ATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 48
230_1_08.seq  -------------------------------------------------------------------------------------------------------------- 0
243_6_08.seq  --------------------------------------------------------------------------------------------------------ACGGAA 6
259_2_08.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 162
261_2_08.seq  AAGTAGCTGGAACTCAATATAGACTTCCATMAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 288
264_1_08.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGAA 228
271_3_08.seq  -------------------------------------------------------------------------------------------------------------- 0
277_1_08.seq  AAGTAGCTGGAACTCAATATAGAATTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 140
287_0_10.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 330
288_6_09.seq  -----------------------------------------------------------------------------------------------CCGGTAGCTACGGGA 15
310_6_08.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGAA 132
337_5_10.seq  ---------------------------CATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGAA 83
343_2_08.seq  ------------------ATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATWATTATTGAGAATTCAAAAACTACTTTTTTAACACCGGTAGCTACGGAA 92
369_8_10.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAAAACTACTTTTTTAACACCGGTAGCTACGGAA 159
457_1_09.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 174
462_8_09.seq  AAGTAGCTGGAACTCAATATAGACTTCCATCAGGGAAATGTCCAGTATTTGGTAAAGGTATAATTATTGAGAATTCAAATACTACTTTTTTAACACCGGTAGCTACGGGA 252
649_5_08.seq  ---------------------------------------------------------------------------------------TTTTAACACCGGTAGCTACGGRA 23
677_2_07.seq  -------------------------------------------------------------------------------------------------------------- 0
```

(*Continued)* Appendix 2 alignment report of AMA1 sequences

```
Majority       AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCTTATGTCACCAATGACATTAGATGATATGAGAGATTTTTATAAAAATAATGAATATGTAAA
                        340       350       360       370       380       390       400       410       420       430       440
111_6_09.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAAACCTCTTATATCACCAATGACATTAAATGGTATGAGAGATTTATATAAAAATAATGAATATGTAAA 347
152_3_09.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCCTATGTCACCAATGACATTARATGRTATGAGAGATTTATATAAAAATAATGAATATGTAAA 140
153_8_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCTTATGTCACCAATGACATTAAATGGTATGAGAGATTTATATAAAAATAATGAATATGTAAA 224
170_1_09.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCTTATGTCACCAATGACATTAAATGGTATGAGAGATTTWTATAAAAATAATGAATATGTAAA 131
208_4_08.seq   AAACAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCTTATATCACCAATGACATTAGATGATATGAGAGATTTTTATAAAAATAATGAATATGTAAA 426
220_4_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCCTATGTCACCAATGACATTARATGATATGAGAGATTTATATAAAAATAATGAATATGTAAA 286
223_1_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCTTATGTCACCAATGACATTAGATGATATGAGAGATTTTTATAAARATAATGAATATGTAAA 158
230_1_08.seq   ------------------------TTTGCTTTTCCTCCAACAAATCCTCCTATGTCACCAATGACATTAAATGGTATGAGAGATTTATATAAAAATAATGAATATGTAAA 86
243_6_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCCTATGTCACCAATGACATTAAATGGTATGAGAGATTTATATAAAAATAATGAATATGTAAA 116
259_2_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACARATCCTCTTATGTCACCAATGACATTAAATGGTATGAGACATTTATATAAARATAATGAAAATGTAAA 272
261_2_08.seq   AAACAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCTTATATCACCAATGACATTAAATGGTATGAGAGATTTATATAAAAATAATGAAGATGTAAA 398
264_1_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAAWCCTCTTATGTCACCAATGACATTARATGRWATGAGAGATTTTTATAAAAATAATGAATATGTAAA 338
271_3_08.seq   ---------------------GGTTTTGCTTTTCCTCCAACAGAACCTCTTATGTCACCAATGACATTAGATGAAATGAGACATTTTTATAAAGATAATAAATATGTAAA 89
277_1_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCTTATRTCACCAATGACATTAAATGRTATGAGAGATTTTTATAAARATAATGAATATGTAAA 250
287_0_10.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCTTATGTCACCAATGACATTAGATGATATGAGAGATTTTTATAAAAATAATGAATATGTAAA 440
288_6_09.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCTTATATCACCAATGACATTAGATCAAATGAGACATTTTTATAAAGATAATGAATATGTAAA 125
310_6_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAAACCTCTTATGTCACCAATGACATTAGATCAAATGAGACATTTTTATAAAGATAATGAAGATGTAAA 242
337_5_10.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCCTATGTCACCAATGACATTAGATCAAATGAGACATTTTTATAAAGATAATAAATATGTAAA 193
343_2_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCTTATGTCACCAATGACATTAGATGATATGAGACRTTTTTATAAAGATAATGAATATGTAAA 202
369_8_10.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCTTATGTCACCAATGACATTAGATGATATGAGACATTTTTATAAAGATAATGAATATGTAAA 269
457_1_09.seq   AATCAATATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCATATGTCACCAATGACATTAGATGAAATGAGACATTTTTATAAAGATAATAAATATGTAAA 284
462_8_09.seq   AATCAATATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAGAACCTCATATGTCACCAATGACATTAGATGAAATGAGACATTTTTATAAAGATAATAAATATGTAAA 362
649_5_08.seq   AATCAAGATTTAAAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCTTATGTCACCAATGACATTAAATGRTATGAGAGATTTATATAAARATAATGAATATGTAAA 133
677_2_07.seq   ------------AAAGATGGAGGTTTTGCTTTTCCTCCAACAAATCCTCTTATGTCACCAATGACATTAAATGGTATGAGATATTTTTATAAAGATAATGAAGATGTAAA 98
```

(*Continued*) Appendix 2 alignment report of AMA1 sequence

```
Majority      AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA
                       450       460       470       480       490       500       510       520       530       540       550
111_6_09.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATRAAAATTCAAATTATAAATATCCAGCTGTTTATGATKACAAAGATAAAA  457
152_3_09.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  250
153_8_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  334
170_1_09.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATKACAAAGATAAAA  241
208_4_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATCAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACGAAGATAAAA  536
220_4_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  396
223_1_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATMAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACGAAKATAAAA  268
230_1_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  196
243_6_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  226
259_2_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAWTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAAATAAAA  382
261_2_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACAATGATAAAA  508
264_1_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATKACAAAGATAAAA  448
271_3_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGATTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  199
277_1_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATGAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  360
287_0_10.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATCAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACGAAGATAAAA  550
288_6_09.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACGAAGATAAAA  235
310_6_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGATTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  352
337_5_10.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGATTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  303
343_2_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACAATGATAATA  312
369_8_10.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATCATTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACAAAGATAAAA  379
457_1_09.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGATTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACGAAGATAAAA  394
462_8_09.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGATTCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATTACGAAGATAAAA  472
649_5_08.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGAATCCAGATAATGATAAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  243
677_2_07.seq  AAATTTAGATGAATTGACTTTATGTTCAAGACATGCAGGAAATATGATTCCAGATAATGATRAAAATTCAAATTATAAATATCCAGCTGTTTATGATGACAAAGATAAAA  208
```

(*Continued)* Appendix 2 alignment report of AMA1 sequences

```
Majority      AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA
                       560       570       580       590       600       610       620       630       640       650       660
111_6_09.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACSAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  567
152_3_09.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  360
153_8_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  444
170_1_09.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  351
208_4_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACAAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  646
220_4_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  506
223_1_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATATA  378
230_1_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  306
243_6_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  336
259_2_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  492
261_2_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  618
264_1_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  558
271_3_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  309
277_1_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACSAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  470
287_0_10.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  660
288_6_09.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  345
310_6_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  462
337_5_10.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  413
343_2_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  422
369_8_10.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATATA  489
457_1_09.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTATAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATATA  504
462_8_09.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  582
649_5_08.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACCAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  353
677_2_07.seq  AGTGTCATATATTATATATTGCAGCTCAAGAAAATAATGGTCCTAGATATTGTAATAAAGACGAAAGTAAAAGAAACAGCATGTTTTGTTTTAGACCAGCAAAAGATAAA  318
```

(*Continued)* Appendix 2 alignment report of AMA1 sequences

```
Majority      TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA
                          670       680       690       700       710       720       730       740       750       760       770
111_6_09.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  677
152_3_09.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  470
153_8_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  554
170_1_09.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  461
208_4_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTCATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  756
220_4_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  616
223_1_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAARAAGTTTGCCCTAGAAAGAATTTASAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  488
230_1_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  416
243_6_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  446
259_2_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAARAAGTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  602
261_2_08.seq  TTATTTGAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  728
264_1_08.seq  TYATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTCATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  668
271_3_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  419
277_1_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  580
287_0_10.seq  TTATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTCATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  770
288_6_09.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  455
310_6_08.seq  TTATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTCATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  572
337_5_10.seq  TTATTTGAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  523
343_2_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  532
369_8_10.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAAAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  599
457_1_09.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTAAAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  614
462_8_09.seq  TTATTTGAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAGAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  692
649_5_08.seq  TCATTTCAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAAAAAGTTTGCCCTAGAAAGAATTTACAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  463
677_2_07.seq  TCATTTSAAAACTATACATATTTAAGTAAAAATGTAGTTGATAACTGGGAARAAGTTTGCCCTAGAAAGAATTTAGAGAATGCAAAATTCGGATTATGGGTCGATGGAAA  428
```

(*Continued*) Appendix 2 alignment report of AMA1 sequences

```
Majority      TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC
                       780       790       800       810       820       830       840       850       860       870       880
111_6_09.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 787
152_3_09.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 580
153_8_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 664
170_1_09.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 571
208_4_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 866
220_4_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 726
223_1_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 598
230_1_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 526
243_6_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 556
259_2_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTYCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 712
261_2_08.seq  TTGTGAAGATATACCACATGTAAATAAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 838
264_1_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 778
271_3_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 529
277_1_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 690
287_0_10.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 880
288_6_09.seq  TTGTGAAGATATACCACATGTAAATGAATTTCCAGCAATTGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 565
310_6_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 682
337_5_10.seq  TTGTGAAGATATACCACATGTAAATGAATTTCCAGCAATTGATCTTTTTGAATGTAATAAATTAGTTTT----------------------------------------- 592
343_2_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTCCAGCAATTGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 642
369_8_10.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 709
457_1_09.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 724
462_8_09.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCT                                                                  737
649_5_08.seq  TTGTGAAGATATACCACATGTAAATGAATTTCCAGCAATTGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 573
677_2_07.seq  TTGTGAAGATATACCACATGTAAATGAATTTTCAGCAAATGATCTTTTTGAATGTAATAAATTAGTTTTTGAATTGAGTGCTTCGGATCAACCTAAACAATATGAACAAC 538
```

(*Continued)* Appendix 2 alignment report of AMA1 sequences

```
Majority      ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT

                       890       900       910       920       930       940       950       960       970       980       990

111_6_09.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 897
152_3_09.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 690
153_8_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 774
170_1_09.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 681
208_4_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 976
220_4_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 836
223_1_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 708
230_1_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 636
243_6_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 666
259_2_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 822
261_2_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 948
264_1_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 888
271_3_08.seq  ATTTAACAGATTAKGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 639
277_1_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 800
287_0_10.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 990
288_6_09.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 675
310_6_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 792
337_5_10.seq  ------------------------------------------------------------------------------------------AAAGCAGATAGATATAAAAGT 613
343_2_08.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 752
369_8_10.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 819
457_1_09.seq  ATTTAACAGATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 834
462_8_09.seq                                                                                                                    737
649_5_08.seq  ATTTAACAGATTAKGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 683
677_2_07.seq  ATTTAACAAATTATGAAAAAAATTAAAGAAGGTTTCAAAAATAAGAACGCTAGTATGATCAAAAGTGCTTTTCTTCCCACTGGTGCTTTTAAAGCAGATAGATATAAAAGT 648
```

(*Continued)* Appendix 2 alignment report of AMA1 sequences

```
Majority       CATGGTAAGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC

                        1000      1010      1020      1030      1040      1050      1060      1070      1080      1090      1100

111_6_09.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 1007
152_3_09.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 800
153_8_08.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 884
170_1_09.seq   CATGGTARGGGTTATAATTGGGGAAATTATAACACAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 791
208_4_08.seq   CATGGTCAGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 1086
220_4_08.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 946
223_1_08.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACACAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 818
230_1_08.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 746
243_6_08.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 776
259_2_08.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACACAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 932
261_2_08.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 1058
264_1_08.seq   TATGGTAAGGGTTATAATTGGGGAAATTATAACACAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 998
271_3_08.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACACAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 749
277_1_08.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACASAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 910
287_0_10.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 1100
288_6_09.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACACAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 785
310_6_08.seq   CGTGGTAAGGGTTATAATTGGGGAAATTATAACACAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 902
337_5_10.seq   GATGGTAAGGGTTATAATTGGGGAAATTATAACAGAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 723
343_2_08.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACACAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 862
369_8_10.seq   CGTGGTAAGGGTTATAATTGGGGAAATTATAACASAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 929
457_1_09.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACAGAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 944
462_8_09.seq                                                                                                                  737
649_5_08.seq   CATGGTAAGGGTTATAATTGGGGAAATTATAACACAGAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 793
677_2_07.seq   CATGGTAGGGGTTATAATTGGGGAAATTATAACAGAAAAACACAAAAATGTGAAATTTTTAATGTCAAACCAACATGTTTAATTAACAATTCATCATACATTGCTACTAC 758
```

(*Continued*) Appendix 2 alignment report of AMA1 sequences

Majority	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	
	1110 1120 1130 1140 1150 1160 1170 1180 1190 1200 1210	
111_6_09.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1117
152_3_09.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	910
153_8_08.seq	TGCTTTGTCCCATCCCATCGAAGT	908
170_1_09.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	901
208_4_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1196
220_4_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1056
223_1_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	928
230_1_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	856
243_6_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	886
259_2_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAAAACAATTTTCCATGTTCATTATATAAAGATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1042
261_2_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1168
264_1_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1108
271_3_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAAAACAATTTTCCATGTTCATTATATAAAGATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	859
277_1_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1020
287_0_10.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1210
288_6_09.seq	TGCTTTGTCCCATCCCATCGAAGTTGAAAACAATTTTCCATGTTCATTATATAAAGATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	895
310_6_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAAAACAATTTTCCATGTTCATTATATAAAGATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1012
337_5_10.seq	TGCTTTGTCCCATCCCATCGAAGTTGAAAACAATTTTCCATGTTCATTATATAAAAATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	833
343_2_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAAAACAATTTTCCATGTTCATTATATAAAGATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	972
369_8_10.seq	TGCTTTGTCCCATCCCAACGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1039
457_1_09.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	1054
462_8_09.seq		737
649_5_08.seq	TGCTTTGTCCCATCCCATCGAAGTTGAAAACAATTTTCCATGTTCATTATATAAAGATGAAATAATGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	903
677_2_07.seq	TGCTTTGTCCCATCCCATCGAAGTTGAACACAATTTTCCATGTTCATTATATAAAGATGAAATAAAGAAAGAAATCGAAAGAGAATCAAAACGAATTAAATTAAATGATA	868

(*Continued)* Appendix 2 alignment report of AMA1 sequences

```
Majority      ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT

                       1220      1230      1240      1250      1260      1270      1280      1290      1300      1310      1320

111_6_09.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1227
152_3_09.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1020
153_8_08.seq                                                                                                                  908
170_1_09.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1011
208_4_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1306
220_4_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1166
223_1_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGCCCCTGAAATTGTAAGTAATAGTACATGT  1038
230_1_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  966
243_6_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGCCCCTGAAATGGTAAGTAATAGTACATGT  996
259_2_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1152
261_2_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGCCCCTGAAATTGTAAGTAATAGTACATGT  1278
264_1_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1218
271_3_08.seq  ATGATGATGAAGGGA                                                                                                 874
277_1_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGCCCCTGAAATGGTAAGTAATAGTACATGT  1130
287_0_10.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGCCCCTGAAATTGTAAGTAATAGTACATGT  1320
288_6_09.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1005
310_6_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGCCCCTGAAATGGTAAGTAATAGTACATGT  1122
337_5_10.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGACCCTGAAATTGTAAGTAATAGTACATGT  943
343_2_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1082
369_8_10.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGCCCCTGAAATGGTAAGTAATAGTACATGT  1149
457_1_09.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATAGACAGTTTAAAATGCCCATGTGCCCCTGAAATTGTAAGTAATAGTACATGT  1164
462_8_09.seq                                                                                                                  737
649_5_08.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATAAAGACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  1013
677_2_07.seq  ATGATGATGAAGGGAATAAAAAAATTATAGCTCCAAGAATTTTTATTTCAGATGATATATACAGTTTAAAATGCCCATGTGACCCTGAAATGGTAAGTAATAGTACATGT  978
```

(*Continued)* Appendix 2 alignment report of AMA1 sequences

Majority	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	
	1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430	
111_6_09.seq	CATTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTA	1269
152_3_09.seq	CRTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1130
153_8_08.seq		908
170_1_09.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAAC	1117
208_4_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1416
220_4_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1276
223_1_08.seq	CATTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1148
230_1_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1076
243_6_08.seq	MATTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1106
259_2_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1262
261_2_08.seq	AATTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1388
264_1_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1328
271_3_08.seq		874
277_1_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1240
287_0_10.seq	CATTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1430
288_6_09.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1115
310_6_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1232
337_5_10.seq	AGTTTCTTTGTATGTAAATGTGTAG	968
343_2_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1192
369_8_10.seq	CATTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1259
457_1_09.seq	AATTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1274
462_8_09.seq		737
649_5_08.seq	CGTTTCTTTGTATGTAAATGTGTAGAAAAAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1123
677_2_07.seq	CDTTTCTTTGTATGTAAATGTGTAGAAAGAAGGGCAGAAGTAACATCAAATAATGAAGTTGTAGTTAAAGAAGAATATAAAGATGAATATGCAGATATTCCTGAACATAA	1088

(*Continued)* Appendix 2 alignment report of AMA1 sequences

Majority	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	
	1440 1450 1460 1470 1480 1490 1500 1510 1520 1530 1540	
111_6_09.seq		1269
152_3_09.seq	ACCAACTTATGATAAMATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1240
153_8_08.seq		908
170_1_09.seq		1117
208_4_08.seq	ACCAACTTATGATAACATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1526
220_4_08.seq	ACCAACTTATGATAACATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1386
223_1_08.seq	ACCAACTTATGATAAMATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1258
230_1_08.seq	ACCAACTTATGATAACATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1186
243_6_08.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTT	1188
259_2_08.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1372
261_2_08.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1498
264_1_08.seq	ACCAACTTATGATAACATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1438
271_3_08.seq		874
277_1_08.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1350
287_0_10.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1540
288_6_09.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1225
310_6_08.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1342
337_5_10.seq		968
343_2_08.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1302
369_8_10.seq	ACCAACTTATGATAACATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1369
457_1_09.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1384
462_8_09.seq		737
649_5_08.seq	ACCAACTTATGATAAAATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1233
677_2_07.seq	ACCAACTTATGATAACATGAAAATTATAATTGCATCATCAGCTGCTGTCGCTGTATTAGCAACTATTTTAATGGTTTATCTTTATAAAAGAAAAGGAAATGCTGAAAAAT	1198

(*Continued)* Appendix 2 alignment report of AMA1 sequences

Sequence	Alignment (positions 1550–1650)	End
Majority	ATGATAAAATGGATGAACCACAAXATTATGGGAAATCXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
111_6_09.seq		1269
152_3_09.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGCATCTTTTTGGGG	1325
153_8_08.seq		908
170_1_09.seq		1117
208_4_08.seq	ATGATAAAATGGATCAACCACAAGATTATGGGAAATCAAAATCAA	1571
220_4_08.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAATC	1423
223_1_08.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGC	1331
230_1_08.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAATCAAAATCAAGAAATGATGAAATGTTAGATCCTGAGGCATCTTTTTGGGGGGAAGAAAAAAGAGCATCACATA	1294
243_6_08.seq		1188
259_2_08.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAAT	1408
261_2_08.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGCATCTTTTTGGG	1582
264_1_08.seq	ATGATAAAATGGATGAACCAC	1459
271_3_08.seq		874
277_1_08.seq	ATGATAAAATGGATGAACCACAACATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGCATCTTT	1429
287_0_10.seq	ATGATAAAATGGATGAACCACAACATTATGGGAAATCAAATTCAAGAAATGATGAAAT	1598
288_6_09.seq	ATGATAAAATGGATGAACCACAACATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGCATCTTTTTGGGGGGAAGAAAAAAGAGCATCACATACA	1335
310_6_08.seq	ATGATAAAATGG	1354
337_5_10.seq		968
343_2_08.seq	ATGATAAAATGGATGAACCACAACATTATGGGAAATCAAATTCAAGAAATG	1353
369_8_10.seq		1369
457_1_09.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGCATCTTTTTGGGGGAA--GAAAAAGAGCATCAC	1487
462_8_09.seq		737
649_5_08.seq	ATGATAAAATGGATGAACCACAACATTATGGGAAATCAAATTCAAGAAATGATGAAATGTTAGATCCTGAGGCATCTTTTTGGGG	1318
677_2_07.seq	ATGATAAAATGGATGAACCACAAGATTATGGGAAATCAACATCAAGAAATGATGAAATGTTA	1260

Appendix 3 Apical membrane antigen 1 (AMA1)- PF3D7_1133400 (PlasmoDB)

MRKLYCVLLLSAFEFTYMINFGRGQNYWEHPYQNSDVYRPINEHREHPKEYEYPLHQEHTYQQEDSGEDENTLQH

AYPIDHEGAEPAPQEQNLFSSIEIVERSNYMGNPWTEYMAKYDIEEVHGSGIRVDLGEDAEVAGTQYRLPSGKCPV

FGKGIIIENSNTTFLTPVATGNQYLKDGGFAFPPTEPLMSPMTLDEMRHFYKDNKYVKNLDELTLCSRHAGNMIPD

NDKNSNYKYPAVYDDKDKKCHILYIAAQENNGPRYCNKDESKRNSMFCFRPAKDISFQNYTYLSKNVVDNWEKV

CPRKNLQNAKFGLWVDGNCEDIPHVNEFPAIDLFECNKLVFELSASDQPKQYEQHLTDYEKIKEGFKNKNASMIKS

AFLPTGAFKADRYKSHGKGYNWGNYNTETQKCEIFNVKPTCLINNSSYIATTALSHPIEVENNFPCSLYKDEIMKEIE

RESKRIKLNDNDDEGNKKIIAPRIFISDDKDSLKCPCDPEMVSNSTCRFFVCKCVERRAEVTSNNEVVVKEEYKDEY

ADIPEHKPTYDKMKIIIASSAAVAVLATILMVYLYKRKGNAEKYDKMDEPQDYGKSNSRNDEMLDPEASFWGEEK

RASHTTPVLMEKPYY

Appendix 4: One and three letter abbreviation of amino acids

SYMBOL		
1-Letter	3-Letter	AMINO ACID
Y	Tyr	Tyrosine
G	Gly	Glycine
F	Phe	Phenylalanine
M	Met	Methionine
A	Ala	Alanine
S	Ser	Serine
I	Ile	Isoleucine
L	Leu	Leucine
T	Thr	Threonine
V	Val	Valine
P	Pro	Proline
K	Lys	Lysine
H	His	Histidine
Q	Gln	Glutamine
E	Glu	Glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	Tryptophan
R	Arg	Arginine
D	Asp	Aspartic acid
N	Asn	Asparagine