Journal of Parasitology and Vector Biology

Volume 8 Number 3, March 2016

ISSN 2141-2510

ABOUT JPVB

The **Journal of Parasitology and Vector Biology (JPVB)** is published monthly (one volume per year) by Academic Journals.

Journal of Parasitology and Vector Biology (JPVB) provides rapid publication (monthly) of articles in all areas of the subject such as Parasitism, Helminthology, Cloning vector, retroviral integration, Genetic markers etc.

Contact Us

Editorial Office: jpvb@academicjournals.org

Help Desk: helpdesk@academicjournals.org

Website: http://www.academicjournals.org/journal/JPVB

Submit manuscript online http://ms.academicjournals.me/

Editors

Dr. Ratna Chakrabarti
Department of Molecular Biology and Microbiology,
University of Central Florida,
Biomolecular Research Annex,
12722 Research Parkway,
Orlando,
USA.

Dr. Rajni Kant
Scientist D (ADG),
(P&I Division)Indian Council of Medical Research
Post Box 4911, Ansari Nagar,
New Delhi-110029
India.

Dr. Ramasamy Harikrishnan
Faculty of Marine Science, College of Ocean Sciences
Jeju National University
Jeju city, Jeju 690 756
South Korea.

Dr. Rokkam Madhavi
Andhra University
Visakhapatnam - 530003
Andhra Pradesh
India.

Dr. Mukabana Wolfgang Richard
School of Biological Sciences
University of Nairobi
P.O. Box 30197 - 00100 GPO
Nairobi,
Kenya.

Dr. Lachhman Das Singla
College of Veterinary Science
Guru Angad Dev Veterinary and Animal Sciences University
Ludhiana-141004
Punjab
India.

Editorial Board

Dr. Imna Issa Malele
Tsetse & Trypanosomiasis Research Institute
Tanzania.

Dr. Mausumi Bharadwaj
Institute of Cytology & Preventive Oncology,
(Indian Council of Medical Research)
I-7, Sector - 39
Post Box No. 544
Noida - 201 301
India.

Dr. James Culvin Morris
Clemson University
214 Biosystems Research Complex
Clemson SC 29634
USA.

Journal of Parasitology and Vector Biology

Table of Content: Volume 8 Number 3 March 2016

ARTICLE

Characterization of the MSP-1 protein in field samples of *Plasmodium falciparum* and its homology to the *Plasmodium vivax* MSP-1 protein 27
G. N. Ndeta, L. A. Dickson and A. A. Winston

academicJournals

Vol. 8(3), pp. 27-38, March 2016
DOI: 10.5897/JPVB2015.0208
Article Number: D35B4DB57882
ISSN 2141-2510
Copyright © 2016
Author(s) retain the copyright of this article
http://www.academicjournals.org/JPVB

Journal of Parasitology and Vector Biology

Full Length Research Paper

Characterization of the MSP-1 protein in field samples of *Plasmodium falciparum* and its homology to the *Plasmodium vivax* MSP-1 protein

G. N. Ndeta*, L. A. Dickson and A. A. Winston

College of Health Science and Human Services, Department of Clinical Laboratory Sciences, The University of Texas-Rio Grande Valley, 1201 West University Drive Rm W 2.302, Edinburg Texas 78539, United States.

Received 25 June, 2015; Accepted 19 February, 2016

Merozoite surface protein-1 (MSP-1) is one of the several proteins on the surface of the asexual merozoite of malaria parasites. It undergoes a two stage proteolytic cleavage to form a C-terminal 19-kDa fragment or *Pf*MSP1$_{19}$ that is used for red blood cell invasion. A homologue of *Pf*MSP-1$_{19}$ of *Plasmodium falciparum* called *Pv*MSP-1$_{19}$ has been identified in *Plasmodium vivax*. Recently, attention has been focused on this C-terminal cysteine-rich region as an important determinant of haplotypes that characterize the malaria parasite in endemic areas. In this study, sequence specific primers were used to determine the *Pf*MSP-1$_{19}$ haplotypes from field samples and *PvMSP*-1$_{19}$ haplotypes from both field and *Aotus* adapted parasites. 163 *Pf*MSP-1$_{19}$ samples from Kisumu-Kenya were polymerase chain reaction (PCR) typed at position 1644 in the first epidermal growth factor (EGF)-like domain followed by sequencing of a random selection of positive samples so as to evaluate molecular changes at positions 1691, 1700 and 1701 in the second EGF-like domain. 121 selected samples gave bands, with 80% typing as GAA (E) and 20% as CAA (Q) at position 1644. From computer-alignment sequences, samples 96B216 confirmed as the Uganda-PA haplotype (E-KNG), 96B209 as the *Pf*MAD20 haplotype (E-TSR), 96B208 as the *Pf*K1/Wellcome haplotype (Q-KNG), 96B183 as the Uganda-PA haplotype (E-KNG) and 96B017 as the E-TSG haplotype. Next, genotyping and sequence analyses of the *Pv*MSP-1$_{19}$ fragment against reference sequences from GenBank demonstrated isolates as belonging to either one or the other of two parental haplotypes identified as Belem and Sal-1 strains. Comparison of *Pf*MSP-1$_{19}$ and *Pv*MSP-1$_{19}$ fragments demonstrated highly conserved cysteine amino acids in the two EGF-like domains.

Key words: Merozoite surface protein-1 (MSP-1), polymerase chain reaction (PCR), *Plasmodium falciparum*, epidermal growth factor (EGF), GenBank, *Plasmodium vivax*, malaria parasite.

INTRODUCTION

Malaria continues to be a global health problem with about 300 to 500 million cases diagnosed annually and approximately one million deaths occurring mostly in children and pregnant women (Mazumdar et al., 2010;

*Corresponding author. E-mail: george.ndeta@utrgv.edu. Tel: 956-665-2297. Fax: 956-665-5253.

Author(s) agree that this article remain permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Table 1. Primers used for the analysis of *P. falciparum* and *P. vivax* $MSP1_{19}$.

Type of amplification	Primer name	Primer sequences and orientation
Primary PCR	*Pf*MSP19S1	F- (5'-GCTTGGCAAATTACTTAGTACAGG-3').
	*Pf*MSP19A1	R-(5'-AATGATATTCCTAAGAAGTTAGAGG-3')
	*Pv*MSP19S1	F-(5'-TCTGGTCTTCTGGAAAAATTG-3')
	*Pv*MSP19A1	R-(5'-GAAGGACAAGCTTAGGAAGC-3')
Secondary/Nested PCR	*Pf*MSP19S	F- (5-'AACATTTCACAACACCAATGC-3')
	*Pf*MSP19A	R-(5'-ACTGCAGAAAATACCATCGAA-3')
	*Pv*MSP19S2	F-(5'-CCCAGCTGCTAAATGTGC-3')
	*Pv*MSP19A2	R-(5'-GCAACATGAGCAACAAGAAGG-3')
Competition PCR	*Pf*MSP19E	F-(5'-TGCGTAAAAAAACAATGTCCAGAA-3')
	*Pf*MSP19Q	F-(5'-TGCGTAAAAAAACAATGTCCACAA-3')
	*Pf*MSP19A	R- (5'-ACTGCAGAAAATACCATCGAA-3')

Nikodem and Davidson, 2000; Sorontou and Pakpahan, 2015). The disease results from red blood cells being invaded by merozoites (Barnwell et al., 1999; Barnwell and Galinski, 1995; Drew et al., 2004). Among the parasites that cause malaria, *Plasmodium falciparum* causes the most severe form of the disease as characterized by, a high level of morbidity and mortality, frequent antimalarial drug resistance and failed vaccine efficacy trials. On the surface of merozoites is a protein known as merozoite surface protein-1 (MSP-1). It is synthesized during schizogony as a 185 to 200 kDa precursor molecule that is held on the surface of merozoites by a glycosyl phosphatidyl inositol (GPI) anchor (Blackman and Holder, 1992; Blackman et al., 1991).

The MSP-1 protein is composed of 17 blocks, 5 of which are conserved and flanked by semi-conserved or variable blocks. It undergoes an initial proteolytic cleavage by a protease known as subtilisin 1 (Sub1) to produce four major polypeptide fragments of approximately 83, 28 to 30, 38, and 42-kDa (denoted $MSP\text{-}1_{83}$, $MSP\text{-}1_{30}$, $MSP\text{-}1_{38}$, $MSP\text{-}1_{42}$) (Putaporntip et al., 1997). These fragments are held together on the surface of the merozoite by non-covalent interactions. Subsequently, the 42-kDa polypeptide fragment undergoes a secondary proteolytic cleavage by another protease known as subtilis in 2 (Sub 2) to produce an N-terminal 33-kDa fragment ($MSP\text{-}1_{33}$) and a C-terminal 19-kDa fragment ($MSP\text{-}1_{19}$). The N-terminal 33-kDa fragment is shed as a complex along with the other polypeptides derived from the first proteolytic cleavage. The C-terminal 19-kDa fragment which is used for red blood cell attachment and invasion is held onto the surface of the merozoite by a glycosyl phosphatidyl inositol anchor (Blackman and Holder, 1992; Holder, 2009; Jongwutiwes et al., 1993; Kadekoppala and Holder, 2010). The C-terminal 19-kDa fragment is highly conserved and contains two cysteine-rich epidermal growth factor (EGF)-like domain that play an important role in red blood cell invasion (Drew et al., 2004; Fenton et al., 1989; Wooden et al., 1993). The C-terminal 19-kDa fragment sequences also belong to one or the other of only two main allelic types (Sutton et al., 2010). These two main allelic types are: *Pf*MAD20 and *Pf*K1/Wellcome. The *Pf*K1/Wellcome is a designation for two parasitic isolates which are the K1 isolate from Thailand and Wellcome isolate from West Africa. The *Pf*MAD20 is a designation for a parasitic isolate from Papua New Guinea (Takala et al., 2007). In this study, polymerase chain reaction (PCR) product of the *Pf*$MSP\text{-}1_{19}$ fragment from field samples, when sequenced, displayed non-synonymous single nucleotide polymorphisms in codons that code for amino-acid at positions 11, 58, 67 and 68 that are comparable to amino acids at positions 1644, 1691, 1700, and 1701 in the *Pf*$MSP\text{-}1_{19}$ sequence from reference parasites in GeneBank (Jongwutiwes et al., 1993). The resulting amino-acids at these positions are E-TSR that characterizes the *Pf*MAD20 allelic form and Q-KNG that characterizes the *Pf*K1/Wellcome allelic form (Table 1). Amino acid E or Q at position 1644 in the first EGF-like domain is inherited as a result of a single mutational event in the mosquito, while amino-acids TSR or KNG at positions 1691, 1700 and 1701in the second EGF-like domain are inherited together to produce different haplotypes.

Meanwhile, the rate of meiotic recombination to generate different haplotypes of *Pf*$MSP\text{-}1_{19}$ has been shown to be influenced by the proportion of mixed infections as well as the number of clones that exist in each infected individual. This means that in an area like Kisumu where the entomological inoculation rate (EIR) is high, this could translate into a high number of allelic variants. If that is the case, then genetic diversity as a dynamic process will result in a continuous generation of novel haplotypes due to meiotic recombination with no dominant haplotype or it could be an exhaustive process that over time will result in the generation of a finite number of haplotypes. Hence,

genotyping was used in this study to characterize PCR amplified products at position 1644 in the first EGF-like domain and the role this activity assumes in haplotype formation during meiotic recombination, while random selection of some typed PCR products were sequenced and the sequences computer analyzed as a means to also understand the role that meiotic recombination at positions 1691, 1700 and 1701 in the second EGF-like domain plays in haplotypes formation.

Meanwhile, *P. vivax* also has a merozoite surface protein (*Pv*MSP-1) which is analogous to *Pf*MSP-1 (Bouyou-Akotet et al., 2015; Del Portillo et al., 1991; Conway and McBride, 1991). As compared to the *Pf*MSP-1_{19} fragment of *P. falciparum*, very little is known about the *Pv*MSP-1_{19} fragment of *Plasmodium vivax* (Egan et al., 2000) as a result of the inability to date to maintain the *vivax* parasite in a continuous *in-vitro* culture system (Galinski et al., 1992) except in non-human primates like the *Aotus* or *Saimiri* monkeys (Jongwutiwes et al., 1993; Yang et al., 1999). Like in *P. falciparum*, the *Pv*MSP-1 in *P. vivax* is attached to the surface of the merozoite by a glycosyl phosphatidyl inositol anchor. It also undergoes a two stage proteolytic cleavage to give *Pv*MSP-1_{19} which is involved in red blood cell attachment and invasion. In terms of meiotic recombination and haplotype formation of the *Pv*MSP-1_{19} fragment, very little is known about this process. Hence, genotyping and sequence analyses were used in this second study to not only evaluate the relatedness of *Pv*MSP-1_{19} and *Pf*MSP-$_{19}$ fragments, but to look for possible *Pv*MSP-1_{19} haplotypes that may exist in some geographical areas of the world.

MATERIALS AND METHODS

Chelex-100 extraction techniques *Pf*DNA and *Pv*DNA

Blood was collected on carbon fiber filter paper from infected patients in Kisumu, a holoendemic area that is located in the western part of Kenya in East Africa. Blood spots were cut out from filter papers and placed into 1.8 ml Eppendorf tubes. To lyse the red blood cells, 1.0 ml of a 1.0% Saponin (Sigma, St. Louis, MO) was added following the manufacturer's instruction. Next DNA was extracted with 200 µl of a 5.0% w/v Chelex-100 (Bio-Rad Laboratories, Hercules, CA) following the manufacturer's instruction. The supernatant was carefully removed and transferred to new clean Eppendorf tubes, and samples not used immediately were stored for later use at -20°C.

*Pf*MSP-1_{19} primers

The primary reaction sense primer was *Pf*MSP19S1 (5'-GCTTGGCAAATTACTTAGTACAGG-3'). The anti-sense primer for the primary reaction was *Pf*MSP19A1 (5'-AATGATATTCCTAAGAAGTTAGAGG-3'). The primers for the nested reaction started and ended reactions in the conserved block seventeen. The sense primer was *Pf*MSP19S (5-'AACATTTCACAACACCAATGC-3') and the anti-sense primer was *Pf*MSP19A (5'-ACTGCAGAAAATACCATCGAA-3'). The primers for the competition reaction were: PfMSP19E (5'-TGCGTAAAAAAACAATGTCCAGAA-3') for the *Pf*MAD20 allele and *Pf*MSP19Q (5'-TGCGTAAAAAAACAATGTCCACAA-3') for the *Pf*K1/Wellcome allele. The antisense primer for the competition reaction was *Pf*MSP19A (5'-ACTGCAGAAAATACCATCGAA-3').

PCR amplification

Two sets of primers were used in each of the DNA amplification phases. The primers that were used in the primary amplification phase were: *Pf* MSP19S1 and *Pf*MSP19A1. The set of primers that were used in the secondary (nested) amplification phase were: *Pf*MSP19S and *Pf*MSP19A.

Primary *Pf*MSP-1_{19} PCR amplification

For the primary PCR amplification, a sample concentration of 50 to 100 ng was used. The PCR reaction mix for each sample composed of 2.0 µl of 10X buffer, 0.8 mM of dNTP, 0.5 nmole of the forward primer, 0.5 nmole of the reverse primer, 0.5 units of Taq polymerase (Perkin Elmer Cetus, Norwalk, CT) and 13.4 µl of diethylpyrocarbonate (DEPC) treated H_2O. The MSP-1 gene product was amplified using the following protocol of denaturation at 95°C for 15 s, annealing at 50°C for 15 s and elongation at 72°C for 1 min for a total of 44 cycles. Each reaction was heated to 95°C for 5 min before the start of the amplification cycles and 72°C for 5 min after the end of the amplification cycles.

Secondary *Pf*MSP-1_{19} PCR amplification

Primary PCR product (1.5 µl) was used as template in a 18.5 µl volume of the secondary PCR reaction. The secondary PCR reaction mix was setup, thus: 2.0 µl of the 10X buffer, 1.6 µl of dNTP, 0.2 µl of the forward primer (*Pf*MSP-19S), 0.2 µl of the reverse primer (*Pf*MSP19A), 0.5 units of Taq polymerase and 14.4 µl of DEPC treated H_2O. If the secondary PCR reaction was to be a competition reaction, then the volume of DEPC treated H_2O needed was adjusted to 14.3 µl. At the completion of the secondary PCR reaction, the PCR product was purified on a 2% agarose gel following the addition of a sample buffer to each PCR reaction product. Ethidium bromide was added to the gel to facilitate DNA viewing.

Preparing samples for competition studies

Positive samples were prepared for competition studies in a PCR mix with the following: 2.0 µl of 10X buffer, 1.6 µl of dNTP, 0.25 µl of *Pf*MSP19Q or *Pf*MSP19E, 0.05 µl of *Pf*MSP19S, 0.2 µl of *Pf*MSP19A and 14.3 µl of DEPC treated H_2O and finally 0.1 µl of Taq polymerase. The template volume was 1.5 µl, for a PCR mix volume of 18.5 µl. The denaturation temperature was 95°C, annealing temperature 50°C and elongation temperature 72°C for a total of 44 cycles. An agarose gel was prepared by dissolving 2 g of high resolution agar in 100 ml of TAE buffer. To this solution was added 2.0 µl of 0.5 µg/ml Ethidium bromide. Electrophoresis was done at 120 volts, 400 mA and for 38 min. Each sample was assayed in duplicate with one set of primers being used to test for the amino acid E and another set of primers being used to test for the amino acid Q.

Cloning

PCR product (1.5 µl) was added to 2.5 µl of ligation buffer, 0.5 µl of the pGEM-T and pGEM-T easy vector and 0.5 µl of enzyme. This mixture was incubated overnight at 4°C. The next day, 1.0 µl of this DNA/pGEM-T and pGEM-T easy vector mix was added to 4.0 µl of dH_2O. 1.0 µl of this 1:4 mix was added to 20.0 µl of DH10B strain of *Escherichia coli* and transformed by electroporation. Samples from the electroporation machine were placed into round bottom tubes containing 1.0 ml of super optimal broth with catabolite repression

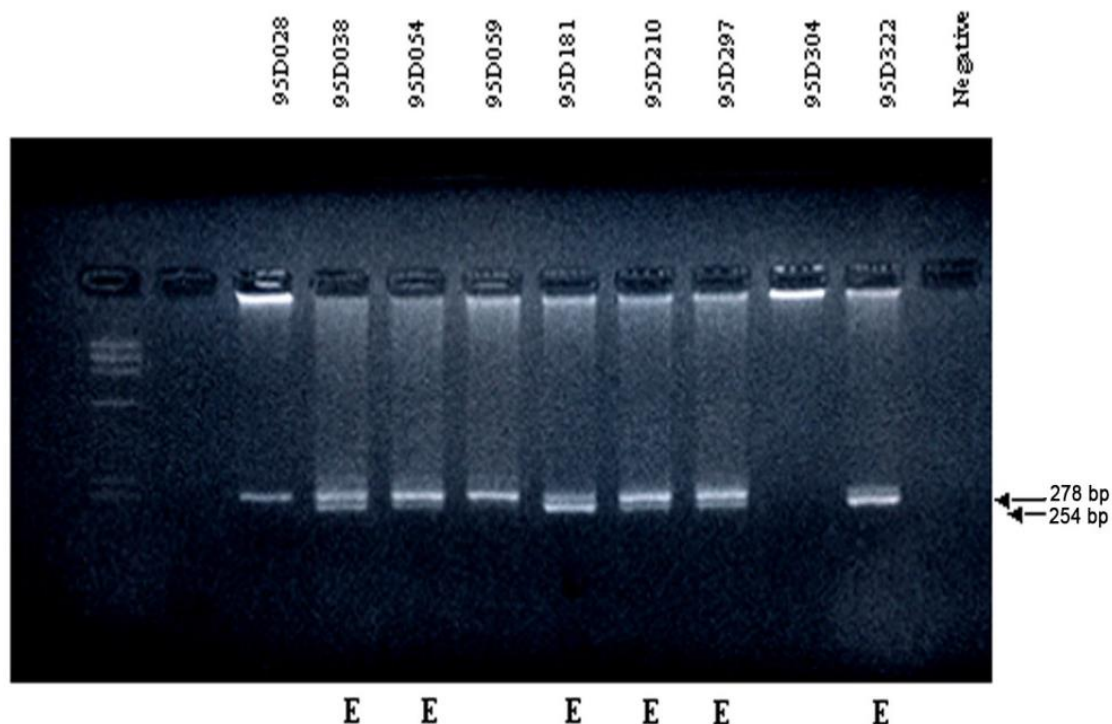

Figure 1. The size of PCR amplified products from the carboxyl terminal of the MSP-1 gene of *P .falciparum*. The competition reaction was only to test for the amino acid E at position 1644 of the MSP-1_{19} fragment. Bands of the molecular weight marker are shown in lane one. Double bands (one that is 278bp in addition to another that is 254bp) indicate an E at position 1644.

(SOC) medium. Tubes with SOC medium (Life Technologies, Grand Island, N.Y) and samples were placed in a 37°C incubator and shaking speed set at 200 rpm for 1.0 h. At the same time, 20.0 µl of X-gal (Promega, Madison, W.I) was applied to agar plates so as to facilitate using the blue-white screening system to distinguish transgenic cells from those that failed to take up foreign DNA. Bacterial colonies that had taken up the foreign DNA ligated to a vector (white colonies) were isolated and transferred to six well plates that contained 5.0 ml of LB broth in Ampicillin (10 mg/ml) and incubated at 37°C overnight with continuous shaking at 150 rpm. 1 to 3 ml of the DH10B culture was pelleted for 1 to 2 min at a speed of 10,000 ×g and the supernatant discarded the cloned product-DNA from field isolates ligated to pGEM-T and pGEM-T easy vector in DH10B were re-suspended in 200.0 µl of re-suspension solution (Promega, Madison WI) and the cloned product recovered according to the manufacturer's protocol. Cloned product was eluted with 50.0 µl of DEPC treated H_2O and DNA inserts from the pGEM-T and pGEM-T Easy vector was separated with EcoR I restriction enzyme.

*Pv*MSP-1_{19} primers

Oligonucleotides were designed to amplify the conserved block 17 of the *P. vivax* parasite. There were two sets of primers for the primary amplification reaction. The sense primer for the first primary amplification reaction was *Pv*MSP19S1 (5'-TCTGGTCTTCTGGAAAAATTG-3'), while the anti-sense primer was *Pv*MSP19A1 (5'-GAAGGACAAGCTTAGGAAGC-3'). The sense primer for the second primary amplification reaction was *Pv*MSP19S2 (5'-CCCAGCTGCTAAATGTGC-3'), while the anti-sense primer was *Pv*MSP19A2 (5'-GCAACATGAGCAACAAGAAGG-3'). The sense primer for the secondary amplification reaction was *Pv*MSP19S (5'-ATGAGCTCCGAGCACACA-3'), while the anti-sense primer for this reaction was *Pv*MSP19 A (5'-GGAGCTACAGAAAACTCCCT-3').

RESULTS

Primary PCR amplification was done on a total of 163 finger-prick samples. Of this number, 121 finger-prick samples gave positive results during the nested or secondary PCR amplification reaction, while 42 finger-prick samples were characterized as negative. Products from these PCR reactions were purified on a 2% agarose gel to which 2 µl of Ethidium bromide had been added. Analyses of the PCR products viewed on an Epifluorescence transilluminator demonstrated product sizes of 399 bp for the primary PCR assay, 278 bp for the nested or secondary PCR assay and 254 bp for the competition assay.

In the first competition assay, the experimental setup was to type for the amino acid glutamate (E) or glutamine (Q) separately. The result shown in the agarose-gel of Figure 1 is of nine samples that were subjected to a first PCR competition study so as to amplify only for samples encoding the amino acid E. Double bands characterized by sizes of 278 and 254 bp can be seen in samples identified as: 95D038, 95D054, 95D181, 95D210,

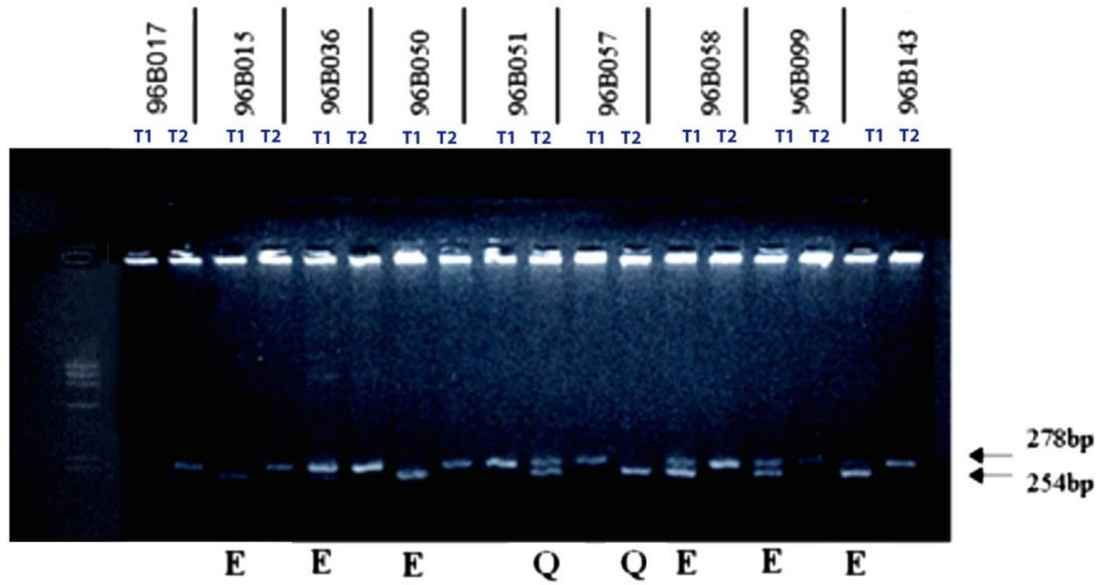

Figure 2. Agarose gel showing results of polymerase chain reaction of samples from *Plasmodium falciparum* that are tested in duplicate. Two sets of tubes were set up for each sample. The first tube (T1) was to type for the amino acid E at position 1644, while the second tube (T2) was to type for the amino acid Q also at position 1644.

95D297, and 95D322. Two samples identified as: 95D028 and 95D059 gave only single bands of size 278 bp, while sample 95D304 gave a negative result. The result shows an overall complex infection as 67% of patient samples have more than one allelic type.

Next, the competition reaction was setup in duplicate to simultaneously type for the amino acids E and Q at position 1644 in the first EGF-like domain. The gel of Figure 2 shows results of this second competition experiment. Of the assayed samples, eight of nine samples display double bands. The eight samples are 96B017, 96B015, 96B036, 96B050, 96B058, 96B099 and 96B143. Six of the eight double bands were for glutamic acid, while two of the double bands were for glutamine. Sample identified as 96B017 typed as indeterminate at this position.

Figure 3 is a computer alignment of sequencing result from the selected positive samples. The positive samples were randomly selected from the 121 positive samples that were characterized by band sizes of 278 and 254 bp. Non-synonymous single nucleotide substitution at position1644 results in amino acid E being replaced with amino acid Q or vice-versa and products that are 254 bp in size. The selected samples are: 96B017, 96B183, 96B208, 96B209, and 96B216. The sequences were aligned so as to compare permutations in the first and second EGF-like domains that can be used to characterize *P. falciparum* haplotypes from the blood of infected individuals in a holoendemic area like Kisumu. DNA sequence analysis was made after reconciling discrepancies in the sense and anti-sense strands. At position 41 which is located in the first EGF-like domain, non-synonymous single nucleotide polymorphism results in GAA being changed to CAA in sample 96B208. The GAA codon represents the amino acids glutamate (E), while the CAA codon is for the amino acid glutamine (Q). Non-synonymous single nucleotide polymorphism is not observed at position 41 in samples 96B017, 96B183, 96B209 and 96B216 as the codon in these samples remain as GAA

In the second EGF-like domain, non-synonymous single nucleotide substitutions are noted at positions 183, 210 and 213. In sample 96B017, single nucleotide substitution is from AGA to GGA at position 212. In sample 96B183, nucleotide substitution are from ACA to AAA at position 183, AGC to AAC at position 210 and AGA to GGA at position 212. In sample 96B208, nucleotide substitutions are from ACA to AAA at position 183, AGC to AAC at position 210 and AGA to GGA at position 212. In sample 96B209, no nucleotide substitution is observed while in sample 96B216 nucleotide substitutions are from ACA to AAA at position 183, AGC to AAC at position 210 and AGA to GGA at position 212 (Figure 3).

Furthermore, a computer alignment of deduced amino acid sequences was made based on the nucleotide sequences in Figure 3. Highlighted at position 11 in the first EGF-like domain of samples 96B017AA, 96B183AA, 96B209AA and 96B216AA is the amino acid E or glutamate. At this same position in sample 96208AA as a result of a non-synonymous single nucleotide substitution, glutamate (E) is substituted with the amino

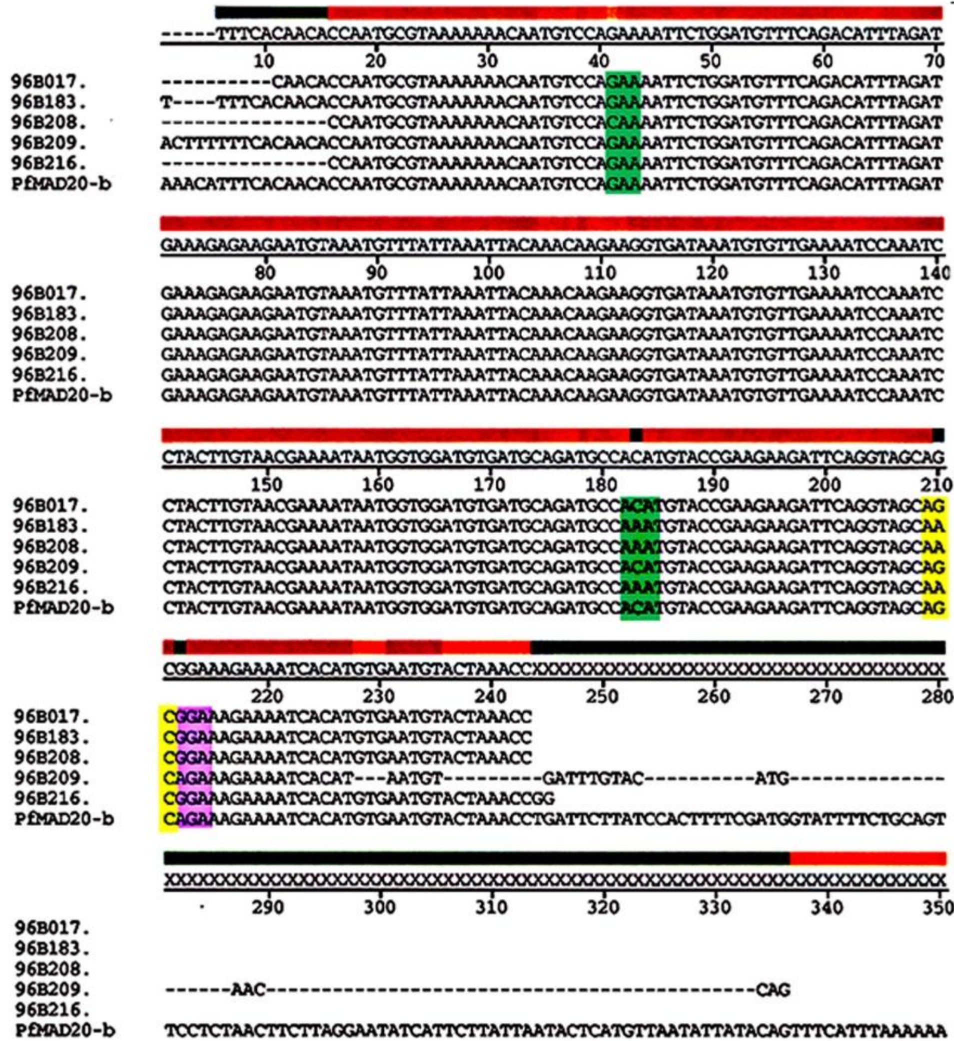

Figure 3. Sequenced results of samples that were typed at position 1644. The results confirm the PCR technique results of using sequence specific primers to type parasites from infected individuals. The samples were transported as blood spots on filter paper and DNA was isolated by the use of the Chelex-100 and the Gentra techniques. Sample 96B219 as a negative control, gave no sequencing result. Sequencing was done with the sense and anti-sense primers. Both strands were computer aligned, and the consensus sequence is what is shown here.

acid glutamine (Q). At positions 58, 67 and 68 which are all in the second EGF-like domain, non-synonymous single nucleotide substitutions result in the amino acids threonine (T), serine (S), arginine (R), lysine (K), asparagine (N) and glycine (G) (Figure 4). In sample 96B017, in addition to glutamate at position 11 in the first EGF-like domain, there is substitution in the second EGF-like domain as arginine (R) in the E-TSR haplotype is replaced by glycine (G) in the Q-KNG haplotype to produce the haplotype identified as E-TSG at position 68.

Samples 96B183 and 96B216 as a result of a single mutational event, that is, a C/G transversion, display the E-KNG haplotype that is used to characterize the Uganda-PA allelic type. The single mutational event involves the amino acid glutamate at position 11 in the first EGF-like domain which in combination with KNG in the second EGF-like domain results in the E-KNG haplotype. In sample 96B208, the mutational event involves glutamine at position 11 in the first EGF-like domain which when combined with KNG in the second EGF-like domain results in the Q-KNG haplotype that is used to characterize the *Pf*K1/Wellcome type. In sample 96B209, no change is noted as the E-TSR haplotype that is used to characterize the *Pf*MAD20 parental strain remains the same.

Next to determine the extent of homology of *Pv*MSP-1_{19} to *Pf*MSP-1_{19}, designed primers were used to initially amplify DNA from samples that were isolated from *Aotus*

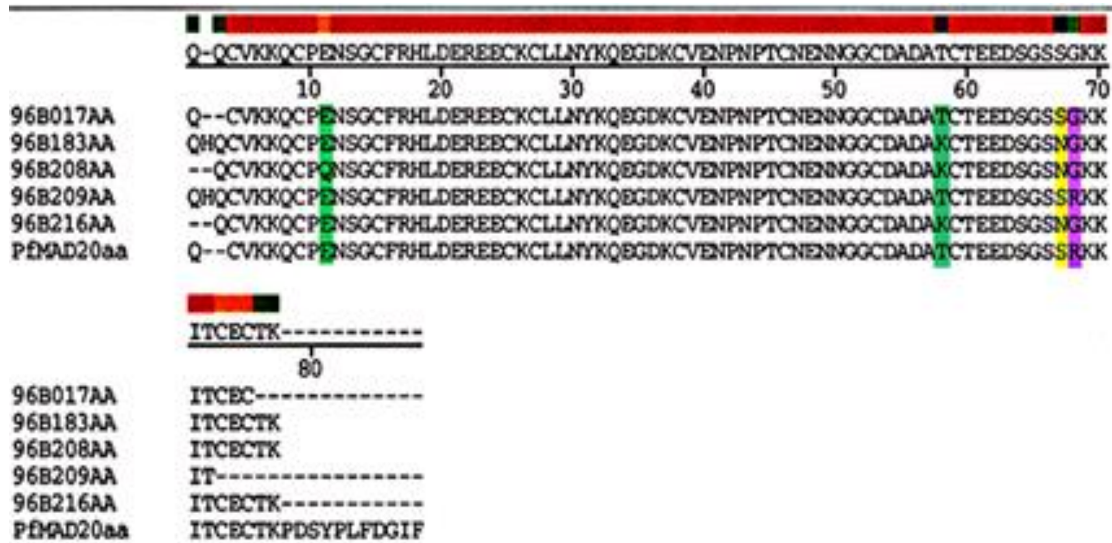

Figure 4. Computer aligned amino acid sequences of five field isolates of *Plasmodium falciparum*. Highlighted in the six sequences are the dimorphic amino acids of the first and second epidermal growth factor-like domains (EGF-like domain).

monkey adapted parasites and later from field isolates. The results in Figure 5 display homology between the sequence that belong to our field isolates and the published sequence from GenBank that belong to the Sal-1 parasite and polymorphism between the sequences of our field isolate and the published sequence from GenBank that belong to the Belem parasite. At position 31 in the sequences of the field isolates from Thailand and Western Pakistan is the amino acid threonine, which is different from the amino acid methionine (U/C transition) that is found in the sequence of the Belem strain. Also at position 50 in the sequences of the field parasites is the amino acid arginine which is different from the amino acid glycine that is found in the sequence of the Belem strain isolate.

Figure 6 is a computer alignment of deduced amino acid sequences that belong to *P. falciparum* and *P. vivax* parasites. Highlighted are nine conserved cysteine amino acids that are vital for demonstrating the structural integrity of the 19-kDa fragment of the MSP-1 protein in both parasites. In addition to the nine conserved cysteine amino acids, there is 86% amino acid homology between *P. falciparum* and *P. vivax*. In the *Pf*MSP-1_{19} amino acid sequences from the field isolates, 31 polar amino acids and 9 hydrophobic amino acids are observed. 67% of the nucleotides in the codons of these amino acids are adenine and thymine, while 33% are cytosine and guanine.

In the *Pv*MSP-1_{19} amino acid sequences from the field isolates, 20 hydrophobic to 31 polar amino acids are observed. Of these amino acids, 55% of the nucleotides in the codons of the amino acids are adenine and thymine while 45% of the nucleotides in the codons of the amino acids are guanine and cytosine.

DISCUSSION

Malaria as a heterogeneous disease is caused by several independently transmitted and antigenically distinct parasite subpopulation or strains that result from non-synonymous single nucleotide polymorphism due to meiotic combinational activities that occur in the mosquito (Thomas et al., 2012). These activities result in the formation of diverse haplotypes from alleles in both the first and second EGF-like domains. Elucidation of the molecular mechanisms responsible for the creation of these antigenically diverse alleles is of utmost importance towards the development of an effective control and prevention method for malaria. Consequently, the goal of the PCR technique used in this study was to identify how events at the molecular level resulted in the formation of alleles and how these alleles rearranged to form various haplotypes. To accomplish this task, DNA from field samples were extracted with Chelex 100 (Long et al., 1996) and subjected to PCR amplification using genus specific primers for the primary PCR reaction and species specific primers for the nested and competition PCR reactions.

In designing the genus specific and species specific primers for use in this study, the focus was on the 19 kDa carboxyl terminals. In nature, this carboxyl terminal exists as two basic versions; is located in block 17 and is highly conserve within species of the malaria parasite. Hence, the genus specific and the species specific primers were designed by using published sequences from GenBank that belong to *Pf*MAD20 and *Pf*K1/Wellcome. These two sequences from GenBank were also used as reference sequences in the alignment of nucleotide and deduced amino acid sequences from positive field isolates so as to

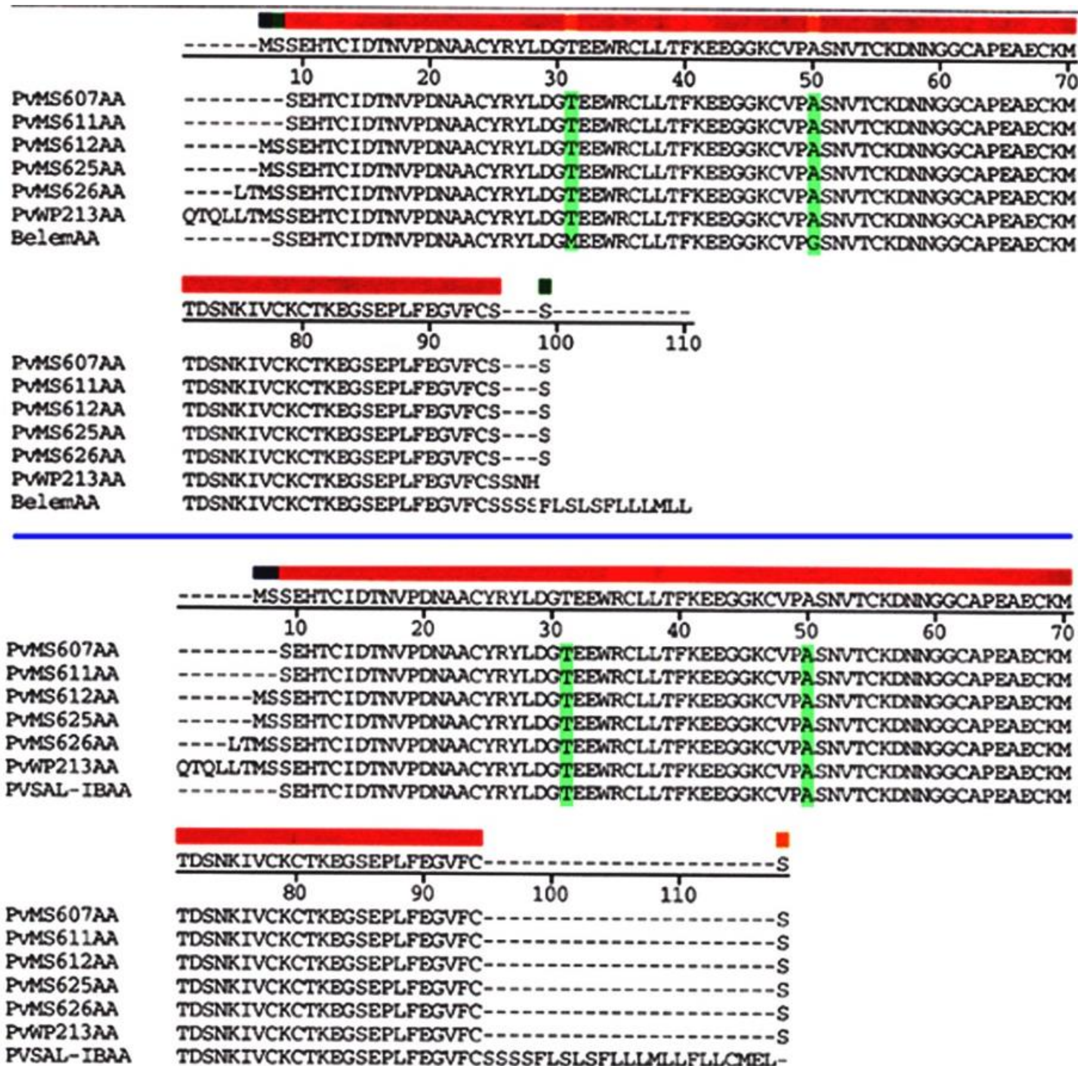

Figure 5. Computer aligned amino acid sequences of the field parasites and those of the Belem parasite of Del Portillo et al. (1991) and the Sal-1 parasite of Gibson et al. (1992). The top part shows polymorphism in the sequences when comparisons are made with the Belem parasite as oppose to the Sal-1 parasite of Gibson and others.

understand the molecular mechanisms that were involve in the creation of antigenically diverse alleles.

Identification of the antigenically diverse alleles was initiated by way of PCR competition typing at position 1644. In the first competition experiment, the design of primers was such that one experiment was set up on day one to only type for glutamate and then the PCR product was purified on a 2% agarose gel (Figure 1). This was followed on a later date by another experiment that was set up to only type for glutamine and the PCR product purified on a 2% agarose gel. The nature of this testing whereby the PCR typing was not done on the same day and at the same time, meant making a side by side comparison of the PCR product on a 2% agarose gel from each patient was going to be not only difficult but impractical. Compounding this situation was the fact that in the first competition typing, there was also the problem of the setup being rather time consuming, wasteful on reagents and supplies and most importantly it was impossible to determine the total number of allelic types in the blood sample of a patient as being E, Q or both E and Q at position 41, especially as the goal of our PCR technique was to identify alleles in both the first and second EGF-like domains and how their arrangements leads to various haplotype formations. As a solution to this problem, two tubes were used in the next experiment set-up to simultaneously type each patient blood sample on the same day and at the same time for glutamate and glutamine at the dimorphic position of 1644. The decision to use two separate tubes to type for the E and Q alleles

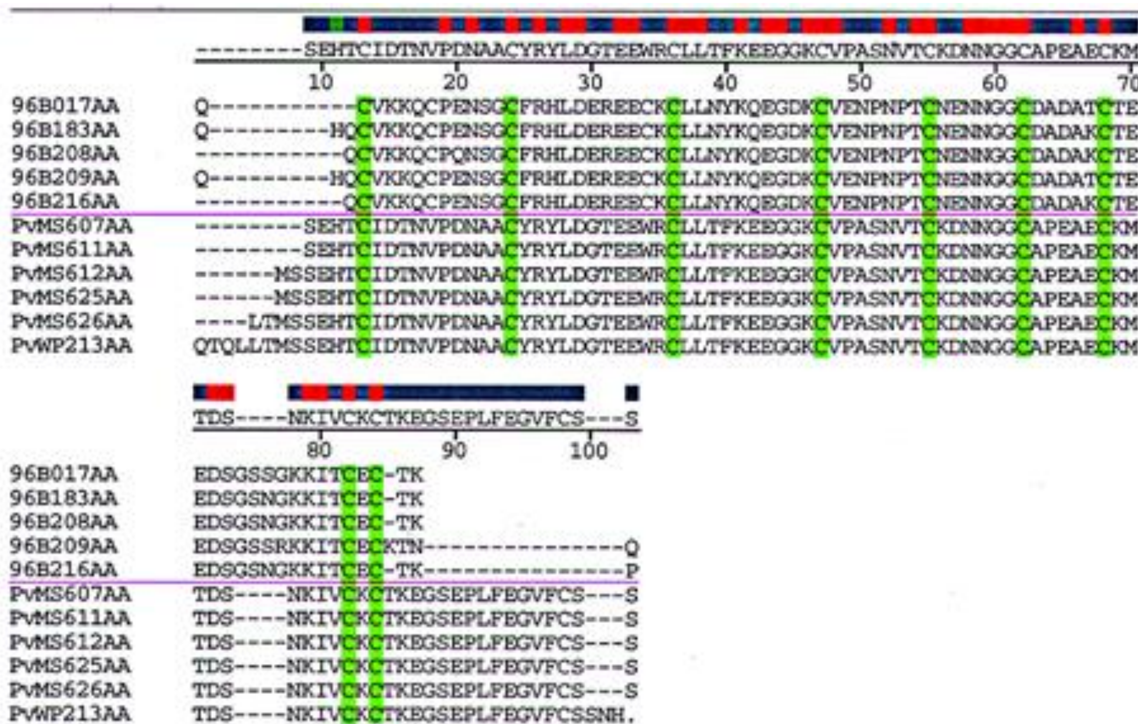

Figure 6. Computer aligned amino acid sequences for field samples of *Plasmodium falciparum* isolates shown above the purple line and amino acid sequences for field samples of *Plasmodium vivax* isolates shown below the purple line. Alignment is done to determine the extent of homology in the MSP-1_{19} fragment from these two malaria parasites.

stem from findings in previous studies, which demonstrated that when the typed PCR products that belong to glutamate and glutamine are purified on a 2% agarose gel, the typed PCR products gave similar band size of 254 bp.

Hence, in tube number one of this second competition assay, primers identified as *Pf*MSP19S, *Pf*MSP19E and *Pf*MSP19A were added so as to type for the amino acid E while in the second tube, primers identified as *Pf*MSP19Q, *Pf*MSP19S and *Pf*MSP19A were used so as to type for the amino acid Q. With this second competition experiment, it was now possible to decipher the total number of alleles in the blood of infected individuals as being E, Q or both E and Q at position 41 in the first EGF-like domain of our field samples.

Moreover, because the use of the designed primers only provided information about non-synonymous single nucleotide polymorphism in the first EGF-like domain, there was a need for a method that will enable us to determine non-synonymous single nucleotide polymorphisms in the second EGF-like domain. The need for such a method is important in that its use will enable us to evaluate molecular changes in the second EGF-like domain that are key in the creation of antigenically diverse alleles at positions 183, 210 and 212 in the second EGF-like domain of our field samples. It will also enable us to compare molecular changes in the first and second EGF-like domain and how those changes result in haplotype formations especially as malaria prevalence in a geographical area is influence by frequency of meiotic recombination in the mosquito, entomological inoculation rate as well as the number of alleles circulating in the area. That method was found in first sequencing of the PCR typed products followed by aligning of the sequences from the positive PCR samples that produced double bands of 278 and 254bp. Prior to sequencing, the typed samples were randomly selected and cloned onto pGEM-T and pGEM-T easy vector. This was to enable sequencing of the PCR product to be performed while the DNA from our field samples was ligated to the vector. The objective of this type of sequencing technique is to avoid the problem of loss of between fifty to seventy-five nucleotide bases that usually occurs during the direct sequencing of PCR products.

Alignment of the sequences was followed by analysis using DNAstar software. Based on the analysis of the sequences, we noted as was the case in the reference sequences that antigenically diverse alleles were formed as a result of single nucleotide changes at only the dimorphic position of 41 in the first EGF-like domain as well as the dimorphic positions of 183, 210 and 213 that are all located in the second EGF-like domain of our field

Table 2. A summary of PCR typing results at one position and the confirmed sequencing results at four positions.

Anti-sense	PCR typing	Sequencing results at three positions that corresponds to positions 1644, 1691, 1700 and 1701 of the complete MSP-1 sequence of Tanabe							
sequence # and alleles	Position 1644	Position 1644	Codon 41-43	Position 1691	Codon 182-84	Position 1700	Codon 209-11	Position 1701	Codon 212-14
96B017 E-TSG	E	E	GAA	T	ACA	S	AGC	G	GGA
96B183 E-KNG	E	E	GAA	K	AAA	N	AAC	G	GGA
96B208 Q-KNG	Q	Q	CAA	K	AAA	N	AAC	G	GGA
96B209 E-TSR	E	E	GAA	T	ACA	S	AGC	R	AGA
96B216 E-KNG	E	E	GAA	K	AAA	N	AAC	G	GGA

The PCR typing at position 1644 was for the amino acid E (glutamate) or Q (glutamine). The codon for amino acid E is GAA, while that for Q is CAA. The confirmed sequencing results show codon at positions 41-43(GAA or CAA), 182-184 (ACA or AAA), 209-211 (AGC or AAC) and 212-214 (AGA or GGA), which corresponds to amino acids E, T, S, R or Q, K, N, G at positions 1644, 1691, 1700, 1701 in the sequence of Tanabe and others. The confirmed sequence for 96B183 and 96B216 identify the Uganda-PA type(E-KNG), 96B209 identify the *Pf*MAD20 type (E-TSR) and 96B208 identify the Wellcome type (Q-KNG). 96B017 identifies the E-TSG allelic type.

samples. These changes produce four alleles identified as: E and Q in the first EGF-like domain and KNG, TSR and TSG in the second EGF-like domain. Also, as a result of allelic change from TSR to TSG in the second EGF-like domain, the E-TSG haplotype is formed as a result of a double mutational event which involve E in the first EGF-like domain and TSG in the second EGF-like that is form from intra-genetic recombination between TSR and KNG. This then result in the following haplotypes of E-TSG, E-KNG, Q-KNG and E-TSR. Incidentally, the E-KNG and Q-KNG are pre-dominant haplotypes that have also been observed in other holoendemic areas like: Congo, Gabon Uganda and Tanzania that are in Central and East Africa, respectively (Mayengue et al., 2011; Viputtigul et al., 2013). Details of the nucleic and amino acid sequences for each sample are summarized in Table 2. This result confirms the suitability of the PCR technique followed by sequencing used in this study as methods that could be used to identify MSP-1_{19} permutations that are formed from non-synonymous single-nucleotide polymorphism in the blood of an infected individual that can result in varied haplotypes. It also shows that our PCR technique can be used as an important tool to monitor treatment failure due to recrudescence from treatment failure due to infection with new haplotypes, especially as it has been observed that antibodies to MSP-1_{19} haplotypes to which an individual has already been exposed to have no effect to subsequent haplotypes due to entomological inoculation.

These changes could possibly be caused by several different mechanisms like point mutation that will cause single nucleotide polymorphism, insertion or deletion and meiotic recombination. But because the nucleotide changes were non-synonymous and involved rearrangement of the E-TSR haplotype that is used to characterize the *Pf*MAD20 parental strain and Q-KNG haplotype that is used to characterize the *Pf*K1/Wellcome strain; this suggest that the variations in the MSP-1_{19} fragment is due partly to meiotic recombination in the mosquito. This finding clearly shows that in holoendemic regions, the propensity for genetic diversity in the *Plasmodium falciparum* parasite is very high. Based on the data from Figure 1 and 2 it can be seen that the majority of our samples when typed using allele specific primers shows that the inhabitants of Kisumu Kenya are infected by a highly diverse parasite population resulting in infected individuals harboring more than one haplotype. This is consistent with findings from other holoendemic areas of Africa which correlates high MSP-1_{19} haplotypes with high transmission as well as a high entomological inoculation rate (EIR) (Mayengue et al., 2011; Thomas et al., 2012). In the first EGF-like domain, 80% of the parasites typed as *PF*MAD20 while 20% of our isolates typed as *Pf*K1/welcome.

Furthermore, primers were formulated not only to determine the extent of polymorphism of the *Pf*MSP-1_{19} fragment in *P. falciparum* field samples, but they were also designed to determine how the *Pv*MSP-1_{19} fragment in *Plasmodium vivax* relates to the *Pf*MSP-1_{19} fragment. To design primers for the *Pv*MSP-1_{19} study, published sequences of the Belem and the Sal-1 strains of *P. vivax* from GenBank were used (Del Portillo et al., 1991; Gibson et al., 1992). The primers were to amplify the carboxyl terminal end of the gene that is located in block 17. These primers were used initially to amplify DNA from laboratory parasites (Chesson-*Pv*331 and Sal-1-*Pv*2339) and field isolates identified as *Pv*214 and *Pv*526. The PCR products when purified by electrophoresis on a 2% agarose gel gave bands of 267 bp in size. The PCR product was purified sequenced and computer aligned using the DNAstar program with the sequences from GenBank that belong to the Belem strain and Sal-1 strain serving as reference sequences. In the Belem reference sequence, codon polymorphism is observed at four positions, while in the Sal-1 reference sequence, there is 100% homology with the field isolates.

In a second *P. vivax,* study which was designed to still determine the extent of polymorphism in the *Pv*MSP-1_{19}

fragment, new field samples were PCR amplified and the isolated DNA sequenced. When the amino acid sequences of our field isolates were compared against the reference sequences from GenBank, polymorphism was observed in the amino acid sequence that belong to the Belem parasite while homology was observed with the amino acid sequence that belong to the Sal-I parasites (Figure 5).

Meanwhile, it should be noted that in the first study, the laboratory-maintained isolates (*Chesson-Pv*331 and Sal-1-*Pv*2339) as well as the field isolates (*Pv*214 and *Pv*526) were from different geographical areas of the world. The *Chesson* (*Pv*331) strain of the *P. vivax* parasite is an isolate from Papua New Guinea, which is an island in the Southwest Pacific Ocean that lies between the Equator and Australia. This strain of the *P. vivax* parasite was first isolated from an American soldier shortly after the World War II. The Sal-1 (*Pv*2339) strain is an isolate from El Salvador in South America. Both of these parasites have been maintained in the Owl (*Aotus*) monkey at the Walter Reed Army Institute of Research in Silver Spring, Maryland.

With difference in geographical origin, one would have expected that there would be sequence heterogeneity between the isolates than sequence homology (Gutierrez et al., 2000). On the contrary, when the sequences from our isolates were computer analyzed against two reference sequences from GeneBank, they identified more with the sequence of the Sal-1 strain from El Salvador than with the sequence of the Belem strain from Brazil.

Conflict of Interests

The authors have not declared any conflict of interests.

REFERENCES

Barnwell JW, Galinski MR, Desimone SG, Perler F. Ingravallo, P (1999). *Plasmodium vivax*, *P. cynomolgi*, and *P. knowlesi*: Identification of homologue proteins associated with the surface of merozoites. Exp. Parasitol. 3:238-249.

Barnwell JW, Galinski RM (1995). *Plasmodium vivax*: A Glimpse into the Unique and Shared Biology of the Merozoites. Ann. Trop. Med. Parasitol. 89:113-130

Blackman MJ, Holder AA (1992). Secondary processing of the *Plasmodium falciparum* merozoite surface protein-1 (MSP-1) by a calcium-dependent membrane-bound serine protease: shedding of $MSP1_{33}$ as a noncovalently associated complex with other fragments of the MSP1. Mol. Biochem. Parasitol. 50:307-316.

Blackman MJ, Ling IT, Nicholls SC, Holder AA (1991). Proteolytic processing of the *Plasmodium falciparum* merozoite surface protein-1 produces a membrane-bound fragment containing two epidermal growth factor-like domains. Mol. Biochem. Parasitol. 49:29-34.

Bouyou-Akotet MK. M'Bondoukwe NP, Mawili-Mboumba DP (2015) Genetic polymorphism of merozoite surface protein-1 in *Plasmodium falciparum* isolates from patients with mild to severe malaria in Libreville, Gabon. Parasite 22:1-9.

Conway DJ, McBride JS (1991). Population genetics of *Plasmodium falciparum* within a malaria hyperendemic area. Parasitology 103:7-16.

Del Portillo HA, Longacre S, Khouri E, David PH (1991). Primary stucture of the merozoite surface antigen of *Plasmodium vivax* reveals sequences conserved between different *Plasmodium* species. Proc. Natl. Acad. Sci. 88:4030-4034.

Drew DR, O'Donnell RA, Smith BJ, Brendan CS (2004). A common cross-species function for the double Epidermal Growth Factor-like modules of the highly divergent Plasmodium surface proteins NSP-1 and MSP-8. J. Biol. Chem. 279:20147-20153.

Egan FA, Blackman MJ, Kaslow DC (2000). Vaccine efficiency of recombinant *Plasmodium falciparum* merozoite surface protein-1 in malaria-naïve, exposed and/or rechallenged *Aotus vociferans* monkeys. Infect. immun. 68:1418-1427.

Fenton B, Clark JT, Wilson CF, McBride JS, Walliker D (1989). Polymorphism of a 35-48-kDa *Plasmodium falciparum* merozoite surface antigen. Mol. Biochem. Parasitol. 34:79-86

Galinski RM, Medina CC, Ingravallo P, Barnwell JW (1992). A reticulocyte-Binding Protein Complex of *Plasmodium vivax* merozoites. Cell 69:1213-1226.

Gibson HL, Tucker JE, Kaslow DC, Krettli AU, Collins WE, Kiefer MC, Bathurst IC, Barr PJ (1992). Structure and expression of the gene for Pv200, a major blood-stage surface antigen of *Plasmodium vivax*. Mol. Biochem. Parasitol. 50:325-33.

Gutierrez A, Vicini J, Patarroyo ME, Murillo LA, Patrroyo MA (2000). *Plasmodium vivax*: Polymorphism in the Merozoite Surface Protein 1 Gene from wild Colombian isolates. Exp. Parasitol. 95:215-219.

Holder AA (2009). The carboxy-terminus of merozoite surface protein 1: structure, specific antibodies and immunity to malaria. Parasitology 136:1445-1456

Jongwutiwes S, Tanabe K, Kanbara H (1993). Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of *Plasmodium falciparum* from field isolates. Mol. Biochem. Parasitol. 59:95-100

Kadekoppala M, Holder AA (2010). Merozoite surface proteins of the malaria parasite: The MSP1 complex and the MSP7 family. Int. J. Parasitol. 40:1155-1161

Lim CS, Kim SH, Kim, SI, Song WJ, Song KJ, Lee KN (2000). Analysis of *Plasmodium vivax* merozoite surface protein –1 gene sequences from resurgent Korean isolates. Am. J. Trop. Med. Hyg. 62: 261-265.

Long GW, Fries L, Watt GH, Hoffman SL (1996). Polymerase chain reaction amplification from *Plasmodium falciparum* on dried blood spots. Am. J. Trop. Med. Hyg., 52:344-346.

Mayengue PS, Ndouga M, Malonga FV, Bitemo M, Noumi F (2011). Genetic polymorphism of merozoite surface protein-1 and merozoite surface protein-2 in *Plasmodium falciparum* isolates from Brazzaville, Republic of Congo. Malaria J. 10:276.

Mazumdar S, Mukherjee P, Yazdani SS, Jain SK, Mohmmed A, Chauhan VS (2010). *Plasmodium falciparum* merozoite surface protein 1 (MSP-1)-MSP-3 Chimeric protein: immunogenicity determined with human-compatible adjuvants and induction of protective immune response, Infect. Immun. 78(2):872-883.

Pachebat JA, Kadekoppala M, Grainger M, Dluzewski AR, Gunaratne RS, Scott-Finnigan TJ, Ogun SA, Ling IT, Bannister LH, Taylor HM Mitchell GH (2007). Extensive proteolytic processing of the malaria parasite merozoite surface protein 7 during biosynthesis and parasite release from erythrocytes. Mol. Biochem. Parasitol .151:59-69.

Putaporntip C, Jungwutiwes S, Tanabe K, Thaithong S (1997). Interallelic recombination in the merozoite surface protein 1 (MSP-1) gene of *Plasmodium vivax* from Thai isolates. Mol. Biochem. Parasitol. 84:49-56.

Sorontou Y, Pakpahan A (2015). Genetic diversity in MSP-1 gene of *Plasmodium falciparum* and its association with malaria severity, parasite density and host factors of asymptomatic and symptomatic patients in Papua, Indonesia.

Sutton PL, Clark EH, Silva C, Branch OH (2010). The *Plasmodium falciparum* merozoite surface protein-1 19 kD antibody response in the Peruvian Amazon predominantly targets the non-allele specific sites of the antigen. Malaria J. 9:3-13.

Takala SL, Smith DL, Thera MA, Coulibaly D, Doumbo OK, Plowe CV (2007). Short Report: Rare *Plasmodium falciparum* merozoite surface protein 1 19-KDA ($MSP\text{-}1_{19}$) haplotypes identified in Mali using High-throughput genotyping methods. Am. J. Trop. Med. Hyg. 76:855-859.

Thomas BN, Petrella CR, Crespo SR, Thakur TJ, Moulds JM, Diallo DA (2012). Geneticpolymorphism of *Plasmodium falciparum* merozoite

surface protein-1 and -2 and diversity of drug resistance genes in blood donors from Bamako, Mali. Infect. Dis. Res. Treat. 6:49-57.

Viputtigul K, Tungpukdee N, Ruangareerate T, Luplertlop N, Wilairatana P, Gaywee J, Krudsood S (2013). C-terminal polymorphism of *Plasmodium falciparum* merozoite surface protein-1 (MSP-1) from Tak province, Thailand. Southeast Asian. J. Trop. Med. Public Health 44:1-12.

Wooden J, Kyes S, Sibley CH (1993). PCR and strain identification in *Plasmodium falciparum.* Parasitol. Today 9:303-305.

Yang C, Collins WC, Sullivan JS, Kaslow DC, Xiao L, Lal AA (1999). Partial protection against *Plasmodium vivax* blood-stage infection in *Saimiri* monkeys by immunization with a recombinant C-terminal fragment of merozoite surface protein 1 in block copolymer adjuvant. Infect. Immun. 67:342-349.

Related Journals Published by Academic Journals

- *Journal of Diabetes and Endocrinology*
- *Journal of Veterinary Medicine and Animal Health*
- *Research in Pharmaceutical Biotechnology*
- *Journal of Physiology and Pathophysiology*
- *Journal of Infectious Diseases and Immunity*
- *Journal of Public Health and Epidemiology*

academicJournals