Genetic diversity of the full length apical membrane antigen-1 of *Plasmodium knowlesi* clinical isolates from Peninsular Malaysia

Ng, Y.L.1, Fong, M.Y.1, Lau, Y.L1*

1Department of Parasitology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
*Corresponding author: lauyeeling@um.edu.my

ABSTRACT

The *Plasmodium knowlesi* apical membrane antigen-1 (PkAMA-1) plays an important role in the invasion of the parasite into its host erythrocyte, and it has been regarded as a potential vaccine candidate against human knowlesi malaria. This study investigates genetic diversity and natural selection of the full length PkAMA-1 of *P. knowlesi* clinical isolates from Peninsular Malaysia. Blood samples were collected from *P. knowlesi* malaria patients from Peninsular Malaysia. The PkAMA-1 gene was amplified from DNA samples using PCR, cloned into a plasmid vector and sequenced. Results showed that nucleotide diversity of the full length PkAMA-1 from Peninsular Malaysia isolates (π: 0.006) was almost similar to that of Sarawak (π: 0.005) and Sabah (π: 0.004) isolates reported in other studies. Deeper analysis revealed Domain I (π: 0.007) in the PkAMA-1 had the highest diversity as compared to Domain II (π: 0.004) and Domain III (π: 0.003). Z-test indicated negative (purifying) selection of the gene. Combined alignment analysis at the amino acid level for the Peninsular Malaysia and Sarawak PkAMA-1 sequences revealed 34 polymorphic sites. Thirty-one of these sites were dimorphic, and 3 were trimorphic. The amino acid sequences could be categorised into 31 haplotypes. In the haplotype network, PkAMA-1 from Peninsular Malaysia and Sarawak were separated into two groups.

Keywords: *Plasmodium knowlesi*; apical membrane antigen-1; genetic diversity.

Malaria is a disease caused by intra-erythrocytic protozoan parasites of the genus *Plasmodium*. In 2019, an estimated 229 million cases of malaria occurred globally, which caused approximately 409,000 deaths (WHO, 2020). Five species have been recognised to cause human malaria which are *Plasmodium falciparum*, *Plasmodium Malariae*, *Plasmodium vivax* and *Plasmodium ovale* and *Plasmodium knowlesi*.

*P. knowlesi* is a simian parasite which is prevalent in Southeast Asia (Jongwutiwes *et al.*, 2004; Ng *et al.*, 2008; Luchavez *et al.*, 2008; Eede *et al.*, 2009). It is currently the main cause of human malaria in Malaysia, especially Malaysian Borneo (Singh *et al.*, 2004; Yusof *et al.*, 2014). Previous study reported the total *P. knowlesi* cases notified in Malaysia were mainly attributed to human knowlesi malaria in Malaysian Borneo (77.1%) from 2013 to 2017. Besides, the average mortality rate of malaria in Sabah and Sarawak were higher than the states of Peninsular Malaysia (except for Perlis) (Hussin *et al.*, 2020). *P. knowlesi* malaria can be a life-threatening disease due to the parasite’s 24-hour asexual erythrocytic stage development (Knowles & Gupta, 1932; Cox-Singh *et al.*, 2008).

The *Plasmodium* parasite infection is dependent on the recognition and then invasion of erythrocytes. The invasion steps include attachment, apical reorientation, tight-junction formation and entry of the merozoite into a parasitophorous vacuole in the erythrocyte (Cowman & Crabb, 2006). These processes rely on the interactions between specific receptors on the erythrocyte membrane and the parasite’s surface molecules such as merozoite surface protein-1 (MSP-1) and apical membrane antigen-1 (AMA-1). These merozoite surface-associated proteins are potentially vaccine candidates against malaria (Anders & Saul, 2000).

The AMA-1 is found in all *Plasmodium* species. It is involved in merozoite invasion into erythrocytes and sporozoite invasion into hepatocytes (Bai *et al.*, 2005). Studies suggested that AMA-1 is as central molecule for the invasion of erythrocytes (Triglia *et al.*, 2000). AMA-1 is highly immunogenic and thus is suggested as a promising vaccine candidate for malaria (Yandar *et al.*, 2008; Gentil *et al.*, 2010; Mahdi Abdel Hamid *et al.*, 2011). Its vaccine potential was demonstrated in various animal models (Remarque *et al.*, 2008).

AMA-1 consists of a signal sequence, transmembrane region, cytoplasmic tail and cysteine-rich ectodomain. The ectodomain consists of Domains I, II and III. The maintenance of the three-dimensional ectodomain structure depends on the eight intramolecular disulphide bonds formed by 16 cysteine residues (Peterson *et al.*, 1989).
Although AMA-1 is found in all Plasmodium species, but different species show different level of polymorphism of this protein. Studies showed high level of polymorphism in AMA-1 of P. falciparum and P. vivax (Escalante et al., 2001; Rajesh et al., 2007), whereas low level of polymorphism was found in P. knowlesi (Faber et al., 2015; Fong et al., 2015c; Chua et al., 2017).

The genetic polymorphism and natural selection of full length P. knowlesi AMA-1 (PkAMA-1) has been reported for P. knowlesi isolates from the Malaysian Borneo state of Sarawak (Faber et al., 2015). A minor study on the PkAMA-1 Domain 1 of Peninsular isolates has also been reported (Fong et al., 2015c). However, to date, no study has been conducted on the full length PkAMA-1 of Peninsular Malaysia. Therefore, in this study, full length PkAMA-1 sequences of isolates from Peninsular Malaysia is investigated, and compared with sequences from Malaysian Borneo.

This study was approved by the Medical Research Committee of the Ministry of Health (NMRR-15-67223975). Blood samples from patients with knowlesi malaria (n=20) were collected from hospitals in Peninsular Malaysia states of Pahang, Kelantan, Johor, Kedah, Terengganu, Perak, Negeri Sembilan, Selangor and Federal Territory of Kuala Lumpur. The majority of the blood samples for this study were collected in 2019-2020. Each blood sample was assigned a reference code for laboratory record. Further confirmation of P. knowlesi malaria infection in the samples was carried out by microscopic examination of Giemsa-stained thick and thin blood smears, and nested polymerase chain reaction (PCR) based on the Plasmodium small subunit ribosomal RNA gene (Snounou et al., 1993; Imwong et al., 2009).

Genomic DNA was extracted from the blood samples using the QIAGEN Blood and Tissue Kit (QIAGEN, Hilden, Germany). One hundred μl of blood were used for extraction and the DNA was eluted using 50 μl of EB Buffer. The DNA was stored at -20°C until use. PCR of PkAMA-1 gene was carried out using forward primer PkAMAfull-F: 5’-TACCTGAAGGAGACTTATGCTAC-3’ and reverse primer PkAMAfull-R: 5’-CTCGGTTGCTAC-3’. Approximately 0.5 μg of genomic DNA was used in a final volume of 20 μl which consisted of 2 mM MgCl2, 0.25 μM of forward and reverse primers, 0.2 mM of dNTPs, and 1 unit of Taq DNA polymerase in a buffer provided in the commercial kit (Promega, Madison, Wisconsin, USA). Cycling conditions for PCR were as follows: initial denaturation of one cycle at 95°C for 3 min, 35 cycles of 94°C for 1 min, 53°C for 1 min, and 72°C for 100 sec, followed by a 10 min extension at 72°C. The PCR products were resolved by gel electrophoresis on a 1% agarose gel stained with SYBR® Safe DNA gel stain (Invitrogen, Eugene, USA).

PCR products were purified using QIAquick PCR purification kit (QIAGEN, Hilden, Germany) following the manufacturer’s instructions. The purified PCR products were ligated into the pGEM-T® TA cloning vector (Promega, Madison, Wisconsin, USA), followed by transformation into Escherichia coli TOP10F’ competent cells. This transformation colonies were screened for the presence of PkAMA-1 DNA fragment. The plasmids of positive colonies were extracted and then sent to a commercial laboratory (Apical Scientific Sdn. Bhd., Malaysia) for DNA sequencing.

In order to perform polymorphism analysis on the sequences (1689 bp), the sequences were submitted to GenBank database (GenBank Accession Numbers MW484863-MW484882). The genetic diversity and natural selection pressure index of the full length, Domain I, Domain II and Domain III were calculated. The Z-test (P <0.05) was used to determine the haplotype diversity and natural selection of PkAMA-1.

Multiple nucleotide and deduced amino acid sequence alignments of PkAMA-1 sequences from the Peninsular Malaysia (n=20) (GenBank Accession Numbers MW484863-MW484882), Sarawak (n=20) (GenBank Accession Numbers KP067834-KP067843, KP067876-KP067885) and reference sequence (P. knowlesi strain H, GenBank Accession Number XM_002259303) were performed using BioEdit sequence alignment editor ver. 7.2.0. PkAMA-1 sequences from Sabah were not included in the analysis as the deposited sequences are not full length sequences (1664 bp). The NETWORK program v4.6.1.2 was used to construct the haplotype network of PkAMA-1 amino acid sequences.

All newly-generated PkAMA-1 sequences from Peninsular Malaysia were deposited in the GenBank database (GenBank Accession Numbers MW484863-MW484882). The genetic diversity and natural selection pressure index of the full length, Domain I, Domain II and Domain III were presented in Table 1. Domain I showed the highest nucleotide diversity (π: 0.004) and Domain II (π: 0.003) was found to be close. The haplotype diversity of Domain II and Domain III (π: 0.003) was found to be close. The haplotype diversity of Domain II and Domain III was 0.795 and 0.637 respectively.

A detailed nucleotide diversity analysis of the entire PkAMA-1 sequence was determined in a sliding window plot with a window length of 100 bp and a step size of 25 bp. The nucleotide diversity ranged from 0.001 to 0.015 (Figure 1). The highest peak diversity was within nucleotide positions 26-150, whereas conserved regions were within nucleotide positions 901-1000, 1076-1200 and 1226-1325.

The Z-test revealed negative (purifying) selection on the Peninsular Malaysia PkAMA-1 (dn < ds, P = 0.000), with Domain I (dn < ds, P = 0.005) most likely the main reason for this negative (purifying) selection since Domain II (dn > ds, P = 0.440; dn < ds, P = 1.000) and Domain III (dn > ds, P = 1.000; dn < ds, P = 0.081) showed neutral selection.

The PkAMA-1 sequences of Peninsular Malaysia (including of strain H) and Sarawak were translated into amino acid sequences for analysis of polymorphism at the protein level (Figure 2). The analysis, the PkAMA-1 amino acid sequence of strain H was used as reference. Close

<table>
<thead>
<tr>
<th>Gene/Domain</th>
<th>Nucleotide positions</th>
<th>Nucleotide diversity (π ± SD)</th>
<th>Haplotype diversity (Hd ± SD)</th>
<th>Z-test P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full length PkAMA-1</td>
<td>1-1689</td>
<td>0.006 ± 0.001</td>
<td>1.000 ± 0.016</td>
<td>1.000</td>
</tr>
<tr>
<td>Domain I</td>
<td>127-744</td>
<td>0.007 ± 0.001</td>
<td>0.974 ± 0.025</td>
<td>1.000</td>
</tr>
<tr>
<td>Domain II</td>
<td>745-1155</td>
<td>0.004 ± 0.001</td>
<td>0.795 ± 0.088</td>
<td>0.440</td>
</tr>
<tr>
<td>Domain III</td>
<td>1156-1461</td>
<td>0.003 ± 0.001</td>
<td>0.637 ± 0.116</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: π – nucleotide diversity, Hd – haplotype diversity, SD – standard deviation, dn – non-synonymous mutation rate, ds – synonymous mutation rate; * significant at P < 0.05.
Figure 1. Nucleotide polymorphism in the PkAMA-1 of Peninsular Malaysia. Sliding window plot of the nucleotide diversity (\( \pi \)) along the PkAMA-1, generated with a window length of 100 bp and step size of 25 bp.

Figure 2. Amino acid sequence polymorphism in PkAMA-1 from Peninsular Malaysia and Sarawak. Polymorphic amino acid residues are listed for each haplotype. Amino acid residues identical to those of the reference sequence [strain H (haplotype 1)] are marked by dots. Dimorphic and trimorphic amino acid positions are marked in yellow and green shadings respectively. Haplotype frequency for each haplotype is listed in the right panel.
examination revealed 34 polymorphic sites, in which 31 were dimorphic, and 3 were trimorphic. Overall, the amino acid sequences could be categorised into 31 haplotypes (H1-H31) with haplotype H19 having the highest frequency (n=4).

Haplotype network (Figure 3) indicates geographical separation of the PkAMA-1 sequences from Peninsular Malaysia and Sarawak. It is noteworthy to highlight that haplotype H17 was from a Peninsular Malaysia patient, but this haplotype was placed in the Sarawak group. This can be explained by the fact that the patient acquired the infection in Sarawak’s neighbour state of Sabah, but was diagnosed positive only upon returning to Peninsular Malaysia.

In the present study, the nucleotide diversity and haplotype diversity of Domain I (π: 0.007, Hd: 0.974) are close to the diversity of Domain I (π: 0.008, Hd: 0.941) reported previously for Peninsular Malaysia P. knowlesi isolates (Fong et al., 2015c). The nucleotide diversity of PkAMA-1 from Peninsular Malaysia (π: 0.006) was just slightly higher than that of Sarawak (π: 0.005) (Faber et al., 2015) and Sabah (π: 0.004) (Chua et al., 2017). Thus, it can be generally concluded that genetic polymorphism of PkAMA-1 is low.

The nucleotide diversity of PkAMA-1 was found to be much lower than that of other P. knowlesi genetic markers such as PkDBPαII (π: 0.012) (Fong et al., 2015a), PkMSP3 (π: 0.046) (De Silva et al., 2017) and PkRAP-1 (π: 0.013) (Rawo et al., 2016). However, the haplotype diversity of PkAMA-1 (Hd:1.000) was almost similar to that of PkDBPαII (Hd: 0.999) (Fong et al., 2015a), PkMSP3 (Hd: 0.997) (De Silva et al., 2017) and PkRAP-1 (Hd: 0.995) (Rawo et al., 2016).

In this study, the haplotype network revealed geographical separation of PkAMA-1. Similar separation has also been observed in some of the P. knowlesi genes such as Duffy binding protein cull (PkDBPαII) (Fong et al., 2015a), cytochrome oxidase subunit I protein (PkCOXI) and small subunit ribosomal 18S RNA type A (PkA-type 18S rRNA) (Yusof et al., 2016). However, other genes such as the CSP (PkCSP), normocyte binding protein (PkNBPXα), and merozoite surface protein 7D (PkMSP7D) did not show any geographical-based separation (Fong et al., 2015b; Ahmed et al., 2016; Ahmed & Quan, 2019). Therefore, geographical separation of Borneo Island from Peninsular Malaysia and subsequent genetic drift of the P. knowlesi populations may not be the reason for this unique PkAMA-1 separation.

The Z-test for all sequences was conducted to obtain a clearer picture of natural selection. Positive selection has been observed in AMA-1 of P. falciparum and P. vivax isolates (Escalante et al., 2001; Rajesh et al., 2007). On the contrary, the PkAMA-1 from both Peninsular Malaysia and Malaysian Borneo was under negative (purifying) selection. The reason for this finding is not clear, but it is probably related to functional constraints or absence of immune pressure in the parasite’s natural host. Further studies characterizing the immunological and functional validation of the PkAMA-1 would be necessary to validate its vaccine potential. It has been reported that the rhoptry-associated protein 1 (RAP1) of non-human primate malarial parasites such as P. knowlesi, P. cynomolgi, P. inui and P. fieldi were under negative (purifying) selection, whereas the same protein of human primate malarial parasites (P. falciparum and P. vivax) were under positive selection (Pacheco et al., 2010). Evidence of negative (purifying) selection was also observed in PkDBPαII, PkCSP and PkNBPXα of non-human primates (Fong et al., 2015a, 2015b; Ahmed et al., 2016).

In summary, although there was geographical separation of PkAMA-1 revealed by haplotype analysis, the diversity of PkAMA-1 was much lower when compared with other P.

Figure 3. Network analysis of 31 haplotypes of PkAMA-1. Red nodes indicate Sarawak haplotype members and yellow nodes indicate Peninsular Malaysia haplotypes. The size of the each node reflects the number of isolates in each haplotype.
knowlesi genes such as PkDBPα2l, PkCoxI and 185 RNA. This low level of polymorphism, which was observed in both Peninsular Malaysia and Malaysian Borneo, thus suggested PkAMA-1’s potential as a knowlesi malaria vaccine in the region.

ACKNOWLEDGEMENTS

This study was supported by the Long Term Research Grant Scheme (LR002A-2018) of the Ministry of Higher Education, Malaysia. We would like to express our gratitude to Dr. Mohd Hafizi Abdul Hamid, Dr. Jenarun Jelip, and Dr. Rose Nani Mudin from Vector Borne Disease Sector, Disease Control Division, Ministry of Health, Malaysia for their contribution in coordinating nationwide sample collection. Also, we would like to thank the public health officials from the State Health Departments and District Health Offices in Pahang, Kelantan, Johor, Kedah, Terengganu, Perak, Negeri Sembilan, Selangor, and Federal Territory of Kuala Lumpur for their assistance in sample collection.

Conflict of Interests

The authors declare that they have no conflict of interests.

REFERENCES


