Research paper

Genetic diversity in the C-terminus of merozoite surface protein 1 among Plasmodium knowlesi isolates from Selangor and Sabah Borneo, Malaysia

Nan Jiun Yap a,b, Xiang Ting Goh a, Anson V. Koehler b, Timothy William c, Tsin Wen Yeo d,f, Indra Vythilingam a, Robin B. Gasser b,e, Yvonne A.L. Lim a,b⁎

a Department of Parasitology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
b Melbourne Veterinary School, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Victoria, Australia
c Jesselton Medical Centre, Kota Kinabalu, 88300, Sabah, Malaysia
d Communicable Diseases Centre, Institute of Infectious Disease and Epidemiology, Tan Tock Seng Hospital, Moulmein Road, 308433 Singapore, Singapore
e State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan, China
f Lee Kong Chian School of Medicine, Nanyang Technological University, 11 Mandalay Road, Singapore, 308232, Singapore

ARTICLE INFO

Article history:
Received 25 March 2017
Received in revised form 12 June 2017
Accepted 16 June 2017
Available online 17 June 2017

Keywords:
Malaria
Malaysia
Plasmodium knowlesi
Merozoite surface protein-1
Genetic variation

ABSTRACT

Plasmodium knowlesi, a malaria parasite of macaques, has emerged as an important parasite of humans. Despite the significance of P. knowlesi malaria in parts of Southeast Asia, very little is known about the genetic variation in this parasite. Our aim here was to explore sequence variation in a molecule called the 42 kDa merozoite surface protein-1 (MSP-1), which is found on the surface of blood stages of Plasmodium spp. and plays a key role in erythrocyte invasion. Several studies of P. falciparum have reported that the C-terminus (a 42 kDa fragment) of merozoite surface protein-1 (MSP-142; consisting of MSP-119 and MSP-133) is a potential candidate for a malaria vaccine. However, to date, no study has yet investigated the sequence diversity of the gene encoding P. knowlesi MSP-142 (comprising Pk-msp-119 and Pk-msp-133) among isolates in Malaysia. The present study explored this aspect. Twelve P. knowlesi isolates were collected from patients from hospitals in Selangor and Sabah Borneo, Malaysia, between 2012 and 2014. The Pk-msp-142 gene was amplified by PCR and directly sequenced. Haplotype diversity (Hd) and nucleotide diversity (θ) were studied among the isolates. There was relatively high genetic variation among P. knowlesi isolates; overall θd and θ were 1 ± 0.034 and 0.01132 ± 0.00124, respectively. A total of nine different haplotypes related to amino acid alterations at 13 positions, and the Pk-MSP-119 sequence was found to be more conserved than Pk-msp-133. We have found evidence for negative selection in Pk-msp-42 as well as the 33 kDa and 19 kDa fragments by comparing the rate of non-synonymous versus synonymous substitutions. Future investigations should study large numbers of samples from disparate geographical locations to critically assess whether this molecule might be a potential vaccine target for P. knowlesi.

© 2017 Published by Elsevier B.V.

1. Introduction

Malaria, caused by protists of the genus Plasmodium, is transmitted by mosquitoes and remains one of the most important parasitic diseases in the world, with >3 billion people at risk of infection and about 200 million human cases of malaria occurring each year (WHO, 2015). The annual malaria-associated mortality is approaching a staggering 43, 800 people, with mortality primarily in children of less than five years of age (WHO, 2015). The recent recognition of Plasmodium knowlesi, a zoonotic malaria parasite, as one of the causative agents of human malaria has made the abatement of global incidence of this disease more challenging (Singh et al., 2004; Cox-Singh et al., 2008). Plasmodium knowlesi is transmitted by mosquitoes to humans from two major monkey reservoir hosts, the long-tailed (Macaca fascicularis) and pig-tailed (M. nemestrina) (Knowles and Das Gupta, 1932; Garnham, 1966). To date, human-to-human transmission has not been reported. The parasite is unique in that it has a short life cycle of 24 h, enabling a rapid progression of disease (Daneshwar et al., 2009; Cox-Singh et al., 2010). Importantly, it can be very virulent in human patients, associates with high parasitaemia, and can cause severe complications and death (Cox-Singh et al., 2008, 2010). Although globally the prevalence and incidence of human infection with P. knowlesi are much less than those caused by P. falciparum and P. vivax (80–95%) (Stéphien, 2014), recent reports of severe and fatal consequences of knowlesi malaria in humans (Cox-Singh et al., 2010; Lau et al., 2011) highlight the public importance.
of this simian parasite, particularly in Malaysia. In 2016, P. knowlesi comprised 69% of the malaria cases reported in Malaysia (Ministry of Health, Malaysia, unpublished reports).

A study conducted by William et al. (2014) described the changing epidemiology of malaria in Sabah Borneo, and suggested a more than ten-fold increase in incidence of P. knowlesi infection in humans (from 59 in 2004 to 703 in 2011, 815 in 2012 and 996 in 2013). Besides Malaysian Borneo (Singh et al., 2004; Cox-Singh et al., 2008; Daneshvar et al., 2009; Lau et al., 2011; William et al., 2011) and Peninsular Malaysia (Cox-Singh et al., 2008; Kantele et al., 2008; Vythilingam et al., 2008; Lee et al., 2010), P. knowlesi infection in humans has also been reported in Southeast Asian regions such as Singapore (Ng et al., 2008; Ong et al., 2009; Jeslyn et al., 2011), Thailand (Putaporntip et al., 2009; Sermwittayawong et al., 2012), Myanmar (Zhu et al., 2006; Jiang et al., 2010), the Philippines (Luchavez et al., 2008), Indonesia (Figtree et al., 2010; Sulistyaningsih et al., 2010), Vietnam (Edede et al., 2009) and Cambodia (Khim et al., 2011).

Despite the significance of P. knowlesi malaria in these geographical regions, there is limited information about the genetic variation within the parasite (Putaporntip et al., 2013). Recent multilocus microsatellite genotyping of P. knowlesi from diverse regions of Malaysia indicated the presence of three major subpopulations of P. knowlesi, including two divergent clusters of human cases in Malaysian Borneo (associated with long-tailed macaques and pig-tailed macaques) and a third cluster in humans in Peninsular Malaysia, with most of the infections from wild long-tailed macaques sampled in Kelantan (Divis et al., 2017). Our focus here was to explore sequence variation in an immunogenic molecule called the merozoite surface protein-1 (MSP-1), which is found on the surface of blood stages of Plasmodium spp. and plays a key role in erythrocyte invasion (Holder and Freeman, 1984; Holder et al., 1992). Consequently, MSP-1 has been recognised as a vaccine candidate (Holder, 2009).

However, Plasmodium MSP-1 is known to exhibit sequence diversity among isolates, which might be the result of selective pressure by host immune responses (Tanabe et al., 1987; Putaporntip et al., 2002; Miahipour et al., 2012). In Plasmodium species, msp-1 gene encodes a 190 kDa precursor protein which undergoes a two-step proteolytic cleavage during merozoite maturation. It is cleaved into four major fragments of 83-, 42-, 38-, 30-kDa (MSP-183, MSP-142, MSP-138 and MSP-130), which remain on the merozoite surface as a glycosylphosphatidylinositol-anchored complex. Before erythrocyte invasion, the MSP-142 fragment undergoes a second cleavage, resulting in the 33- and 19-kDa (MSP-133

### Table 1

Sources of blood samples containing Plasmodium knowlesi from Malaysia.

<table>
<thead>
<tr>
<th>Amino acid haplotype</th>
<th>Isolate</th>
<th>Accession number</th>
<th>Year of collection</th>
<th>Place of collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PeninKKB2</td>
<td>KX881363</td>
<td>2012</td>
<td>Hospital Kuala Kubu Bahru, Selangor</td>
<td></td>
</tr>
<tr>
<td>2 PeninKKB4</td>
<td>KX881364</td>
<td>2012</td>
<td>Hospital Kuala Kubu Bahru, Selangor</td>
<td></td>
</tr>
<tr>
<td>3 SBorneo3</td>
<td>KX894505</td>
<td>2014</td>
<td>Hospital Queen Elizabeth II, Kota Kinabalu, Sabah</td>
<td></td>
</tr>
<tr>
<td>4 PeninKKB1</td>
<td>KX881365</td>
<td>2012</td>
<td>Hospital Kuala Kubu Bahru, Selangor</td>
<td></td>
</tr>
<tr>
<td>5 PeninKKB5</td>
<td>KX881366</td>
<td>2012</td>
<td>Hospital Kuala Kubu Bahru, Selangor</td>
<td></td>
</tr>
<tr>
<td>6 PeninSgB1</td>
<td>KX881367</td>
<td>2013</td>
<td>Hospital Surgai Buloh, Selangor</td>
<td></td>
</tr>
<tr>
<td>7 SBorneo1</td>
<td>KX881368</td>
<td>2014</td>
<td>Hospital Queen Elizabeth II, Kota Kinabalu, Sabah</td>
<td></td>
</tr>
<tr>
<td>8 SBorneo2</td>
<td>KX881369</td>
<td>2014</td>
<td>Hospital Queen Elizabeth II, Kota Kinabalu, Sabah</td>
<td></td>
</tr>
<tr>
<td>9 SBorneo5</td>
<td>KX894507</td>
<td>2014</td>
<td>Hospital Queen Elizabeth II, Kota Kinabalu, Sabah</td>
<td></td>
</tr>
<tr>
<td>10 PeninSely1</td>
<td>KX881370</td>
<td>2013</td>
<td>Hospital Selayang, Selangor</td>
<td></td>
</tr>
<tr>
<td>11 PeninKKB3</td>
<td>KX881371</td>
<td>2012</td>
<td>Hospital Kuala Kubu Bahru, Selangor</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Nucleotide variants of Pk-msp-142 (940 bp) representing Plasmodium knowlesi isolates from Malaysia. In total, 37 nucleotide alterations were detected in 12 isolates. Identical nucleotides are marked as dots, while polymorphic with parsimony informative sites are shaded in light green.
and MSP-119) fragments, the latter of which remains on the surface as the merozoite (Blackman et al., 1991, 1996), and MSP-142, which has been reported to be highly antigenic (Cheong et al., 2013). MSP-142 and MSP-119 fragments have received attention as immunogens, given that antibodies directed against MSP-142 and MSP-119 were shown to interrupt merozoite invasion in vitro (Patino et al., 1997; Nvubu et al., 2002). Sequence variation in the central repeat region of MSP-142 of P. falciparum (see Mehrizi et al., 2008; Zamani et al., 2009; Pan et al., 2010; Viputtigul et al., 2013) and P. vivax (see Dias et al., 2011; Kang et al., 2012) has been relatively well studied, but nothing is known about the structure, function or genetic variation in MSP-1 in P. knowlesi. In the present study, we explored, for the first time, sequence variation in P. knowlesi MSP-142 (designated Pk-MSP-142) from Peninsular Malaysia and Sabah Borneo, Malaysia.

### 2. Materials and methods

#### 2.1. Collection of blood samples

Genomic DNAs (n = 12) from blood samples from confirmed knowlesi malaria patients were subjected to molecular analysis. Samples were collected in Selangor (n = 7), state of peninsular Malaysia, and in Sabah (n = 5), state of Malaysia Borneo. For samples from Selangor, filter blood spots were collected between 2012 and 2014 from Kuala Kubu Bharu Hospital (n = 5), Selayang Hospital (n = 1) and Hospital Sungai Buloh (n = 1). For samples from Sabah, whole blood samples were collected from humans in 2014 at Queen Elizabeth II Hospital, which is located in Kota Kinabalu. The details of samples collected in this study are shown in Table 1.

#### 2.2. Genomic DNA isolation, amplification of Pk-msp-142 and sequencing

Genomic DNA was isolated from dried blood spots on filter paper or from 200 μl of blood using QIAamp DNA Blood Mini Kit (Qiagen, Germany), according to the manufacturer’s protocol, and then frozen at −20 °C until use. PCR-amplification of the msp-142 gene region from P. knowlesi was performed using primers msp-142_F (forward: 5’-CACTGCAGACTGCTCACTCA-3’) and msp-142_R (reverse: 5’-CTAGCTGGAGGAGGACCA-3’) as described previously (Cheong et al., 2013). For samples that did not yield a sufficient amount of ampiclon for DNA sequencing, a semi-nested PCR was performed using internal primers msp-142_F (forward: 5’-GAAATACCTGCGGCTGCTCACTC-3’) and msp-142_R (reverse: 5’-CTAGCTGGAGGAGGACCA-3’). The amplification conditions for both external and internal PCRs were as follows: an initial cycle of 94 °C for 5 min, followed by 35 cycles of 94 °C for 45 s, 61 °C for 45 s and 72 °C for 1 min (primary PCR); 94 °C for 30 s, 53 °C for 30 s and 72 °C for 30 s (secondary PCR), with a final extension of 72 °C for 10 min. PCR products were examined on 1.5% agarose gels; then, ampiclon aliquots were treated with ExoSAP-IT® (Fermentas, USA), according to the manufacturer’s instructions, and subjected to direct, automated sequencing (BigDye Terminator v.3.1 chemistry, Applied Biosystems, USA). Sequence electropherograms were carefully inspected for the presence of multiple sequence types.

### Table 2

Sequence variation in the Pk-foo 142 DNA sequence among Malaysian isolates (N = 12).

| Mass (Da) | Average number of pairwise nucleotide differences (N) | K | H | HD (S.D.) | Ds | Dn | Dn – Ds ± S.E. | Z-test Tajima’s D test Fu & Li’s D test Fu & Li’s F test |
|-----------|---------------------------------|---|---|------------|---|----|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 42 kDa    | Entire population (N = 12)       | 0.01320 ± 0.00214 | 10.6364 | 12 | 1.000 (0.034) | 0.039 | 0.005 | −0.034 ± 0.009 | p < 0.05 | −0.70334 (p < 0.10) | −0.62146 (p < 0.10) | −0.73276 (p < 0.10) |
|           | Peninsular Malaysia (N = 7)      | 0.01185 ± 0.00151 | 11.1429 | 12 | 1.000 (0.076) | 0.034 | 0.006 | −0.028 ± 0.009 | p < 0.05 | −0.14230 (p < 0.10) | −0.08241 (p < 0.10) | −0.10587 (p < 0.10) |
|           | Sabah Borneo (N = 5)             | 0.00936 ± 0.00237 | 8.8000  | 12 | 1.000 (0.126) | 0.039 | 0.002 | −0.037 ± 0.010 | p < 0.05 | −0.061605 (p < 0.10) | −0.61605 (p < 0.10) | −0.66002 (p < 0.10) |
| 33 kDa    | Entire population (N = 12)       | 0.01262 ± 0.00156 | 8.6667  | 12 | 1.000 (0.034) | 0.041 | 0.006 | −0.035 ± 0.010 | p < 0.05 | −0.57399 (p < 0.10) | −0.61349 (p < 0.10) | −0.68712 (p < 0.10) |
|           | Peninsular Malaysia (N = 7)      | 0.01442 ± 0.00192 | 9.9048  | 12 | 1.000 (0.076) | 0.043 | 0.008 | −0.035 ± 0.012 | p < 0.05 | −0.16628 (p < 0.10) | −0.15097 (p < 0.10) | −0.17043 (p < 0.10) |
|           | Sabah Borneo (N = 5)             | 0.00873 ± 0.00251 | 6.0000  | 12 | 1.000 (0.126) | 0.033 | 0.003 | −0.030 ± 0.011 | p < 0.05 | −0.708089 (p < 0.10) | −0.78089 (p < 0.10) | −0.83074 (p < 0.10) |
| 19 kDa    | Entire population (N = 12)       | 0.00779 ± 0.00143 | 1.9697  | 7  | 0.909 (0.056) | 0.032 | 0.002 | −0.030 ± 0.013 | p < 0.05 | −1.02416 (p < 0.10) | −0.53139 (p < 0.10) | −0.74693 (p < 0.10) |
|           | Peninsular Malaysia (N = 7)      | 0.00489 ± 0.00093 | 1.2381  | 4  | 0.857 (0.102) | 0.014 | 0.002 | −0.012 ± 0.011 | p < 0.05 | 0.05031 (p > 0.10)  | 0.38925 (p > 0.10)  | 0.33832 (p > 0.10)  |
|           | Sabah Borneo (N = 5)             | 0.01107 ± 0.00290 | 2.8000  | 4  | 0.900 (0.161) | 0.054 | 0.000 | −0.054 ± 0.023 | p < 0.05 | −0.190992 (p > 0.10)| −0.19092 (p > 0.10)| −0.19784 (p > 0.10) |

N: Number of isolates; K: average number of pairwise nucleotide differences; H: number of haplotypes; HD: haplotype diversity; α: observed average pairwise nucleotide diversity; Ds: nucleotide diversity of synonymous mutation per synonymous site; Dn: nucleotide diversity of non-synonymous mutation per non-synonymous site; Dn – Ds: the difference of Dn and Ds with their standard deviation estimated by bootstrap with 1000 pseudo replicates. Z-test was applied to test significance.
2.3. Data analysis

Sequences determined (GenBank accession numbers: KX881363-KX881371 and KX894505-KX894507) were aligned, using the program MUSCLE (Edgar, 2004), with the reference sequence encoded by H strain of *P. knowlesi* accession no. (XM002258546), and the resultant alignments were adjusted manually using the program Mesquite v.3.04 (Maddison and Maddison, 2015). The numbers of segregating sites (S) and haplotypes (H) as well as haplotype diversity (Hd), nucleotide diversity (\( \pi \)) and average number of pairwise nucleotide differences within the population (K) were calculated using DnaSP ver. 5.10.00 (Librado and Rozas, 2009). The numbers of synonymous nucleotide substitutions per synonymous site (Ds), the number of nonsynonymous substitutions per non-synonymous site (Dn), and the difference between the non-synonymous and synonymous substitutions (Dn − Ds) were estimated using Nei and Gojobori’s method (Nei and Gojobori, 1986) with the Jukes and Cantor (JC) correction to get evidence of natural selection in MEGA 7. Tajima's D (Tajima, 1989) and Fu and Li's D and F test (Fu and Li, 1993) were further applied using DnaSP ver 5.10.00 to evaluate the neutral theory of evolution. Custom R scripts and the adegenet 2.0.0 package were used to identify single nucleotide polymorphisms (SNPs) (Darriba et al., 2012). We retrieved 26 additional *Pk-msp-142* sequences derived from *P. knowlesi*-infected samples (11 from humans, Thailand; one from a human, India; 12 from macaques, Thailand; *P. knowlesi* strain Hackeri; and *P. knowlesi* strain Nuri) from GenBank and included these sequences into the construction of phylogenetic tree and haplotype network tree. Phylogenetic tree was constructed using the neighbor-joining method as described in MEGA 7 (Kumar et al., 2016). Bootstrap replicates of 1000 were used to test the robustness of the tree. Haplotype networks for *Pk-msp-142* were constructed based on their polymorphic sites.
by using the median-joining method in NETWORK version 4.6.1.2 software (Fluxus Technology Ltd., Suffolk, UK).

3. Results and discussion

High quality C-terminal sequences of Pk-msp-142 (940 bp) were obtained for 12 isolates (seven from Selangor, Peninsular Malaysia, and five from Sabah Borneo). This fragment contained a region coding for a protein sequence of 313 amino acids. Analysis and comparison at the nucleotide level against the reference sequence for a protein sequence of 313 amino acids. Analysis and comparison at the nucleotide level against the reference sequence for P. knowlesi strain H (GenBank accession no. XM_002258546) displayed nucleotide alterations at 37 positions among the isolates in the 940 bp region. P. knowlesi msp-133 (Pk-msp-133) contained more segregating sites as compared to P. knowlesi msp-19 (Pk-msp-19) (30 versus 7). Of the 940 nucleotides, 903 were monomorphic and 37 were polymorphic.

We performed genetic analyses for the entire Pk-MSP-142 region as well as for its 33 and 19 kDa fragments separately for the two study locations (i.e. Selangor and Sabah Borneo), and in the entire population (Table 2). All 12 Pk-msp-142 were found to be unique, resulting in 12 different haplotypes (H). The average number of pairwise nucleotide differences (K) in Pk-msp-142 for all 12 sequences was 10.6364, while the haplotype diversity (Hd) and nucleotide diversity (θs) were 1 ± 0.034 and 0.01132 ± 0.00124, respectively. The high haplotype and nucleotide diversities in Pk-msp-142 indicated that the sequences exhibited extensive genetic polymorphism. This was similar to the results for Pk-1-msp-42 reported by Pacheco et al. (2012) whereby the haplotype diversity of their P. knowlesi-infected orangutan samples was 0.0103 ± 0.0031. In addition, our result was also found to be more divergent than Pk-msp-19 (Table 2). A similar pattern was seen in both P. falciparum and P. vivax isolates (Mehrizi et al., 2008; Zamani et al., 2009; Pan et al., 2010; Viputtigul et al., 2013). A sliding window analysis (window: 200 bp) across Pk-msp-142 revealed a SNP density ranging from 0 to 0.006 (Fig. 2). The highest peak diversity was between nucleotide positions 1–200 bp in Pk-msp-133, whilst the most conserved region in Pk-msp-19 was between nucleotide positions 51–150.

In order to explore the role of natural selection, we further analyzed the genetic polymorphism in the Pk-msp-142 by estimating the number of synonymous (Ds) and non-synonymous (Dn) substitutions per site using the Nei and Gojobori method with the Jukes and Cantor correction. The difference between the rates of non-synonymous (Dn) and synonymous (Ds) mutations for the entire 940 bp region was −0.034 ± 0.009 S.E. (Z-test, p < 0.05), suggesting that purifying selection acts in this part of the molecule (Table 2), and a possibility of population expansion. The values for Tajima’s D and Fu and Li’s D and F test further supported the proposal for a negative selection pressure on the Pk-msp-142, fragment (Table 2), even though support was not significant. The same finding was also seen in the P. knowlesi-infected orangutan samples reported by Pacheco et al. (2012) whereby negative selection.

Fig. 4. Median-joining networks of Pk-1-msp-142 from Malaysia. The genealogical haplotype network shows the relationships among the 28 haplotypes present in the 12 sequences obtained from human from Peninsular Malaysia and Malaysian Borneo as well as 25 additional sequences derived from P. knowlesi-infected samples (11 from humans, Thailand, 1 from human, India, 19 from macaques, Thailand and P. knowlesi strain Nuri) from GenBank. Each distinct haplotype has been designated a number (H, n). Circle sizes represent the frequencies of the corresponding haplotype (the number is indicated for those that were observed >1 ×). Distances between nodes are arbitrary.
Fig. 5. Amino acid polymorphism in the Pk-msp-142 among geographically diverse malaria endemic areas. Sequence analysis revealed 9 distinct variants of Pk-msp-142 among 12 isolates, when compared with the reference sequence encoded by H strain of P. knowlesi (XM002258546). Identical amino acids are marked as dots, while the monomorphic and dimorphic changes at a particular amino acid position are shaded red and yellow, respectively. (A) Amino acid haplotypes from current study: H1 - Accession nos KX881363, JQ213367, H2 - Accession nos KX881364, KX894505, JF837348, H3 - Accession nos KX881365, KX894506, H4 - Accession nos KX881366, KX894506, H5 - Accession nos KX881367, JF837352, JF046795, H6 - Accession nos KX881368, H7 - Accession nos KX881369, KX894507, H8 - Accession nos KX881370, H9 - Accession nos KX881371. (B) Amino acid haplotypes of previously published global isolates: H11 - Accession no JF837343, H12 - Accession no JF837345, H13 - Accession nos JF837346, JF837347, H14 - Accession nos JF837349, JF837350, H15 - Accession nos JF837351, H16 - Accession no JC837353, H17 - Accession nos JX046791, JF837339, JF046792-046794, H18 - Accession no JF837342, H19 - Accession no JF837344, H20 - Accession no JF046796, H21 - Accession no JF046797, H22 - Accession no JF046798, H23 - Accession no X51855.
was suggested. When we analyzed \(Pk\text{-msp-1}_{133}\) and \(Pk\text{-msp-1}_{142}\) separately, most of the nonsynonymous mutations were found in the 33 kDa fragment and negative selection was suggested to be acting on these two fragments as well (Table 2).

Phylogenetic analysis of \(Pk\text{-msp-1}_{142}\) was constructed using \(P.\) knowlesi strain H as well as other previously published sequences as references (Fig. 3). Isolates from Sabah Borneo were found to form a cluster with one of the samples from Hospital Kuala Kubu Bharu, while the rest formed multiple clusters with human samples from Thailand. Unlike the study by Divis et al. (2017) which demonstrated the presence of three subpopulations of \(P.\) knowlesi using a multilocus genotyping approach, no obvious clustering into groups was seen. In order to clearly observe the DNA sequence variation, the same set of sequences was also subjected to a median-joining network analysis of haplotypes (Fig. 4). The network tree revealed “star-like” networks, suggesting a population expansion. In contrast to network analyses of two \(P.\) knowlesi genes, type A small subunit ribosomal 18S RNA gene and cytochrome c oxidase subunit 1 gene from a previous study (Yusof et al., 2016), no genetrically divergent clusters were observed.

At the protein level, the analysis of 12 \(Pk\text{-msp-1}_{142}\) sequences resulted in the classification of nine different variants, with amino acid alterations at 13 positions when compared with the reference sequence for \(P.\) knowlesi strain H (Fig. 5). Of the 13 polymorphic amino acid positions, only one had a dimorphic mutation, with changes resulting in two amino acid types (F1789Y/S); the others were monomorphic mutations (Fig. 5A). Most of the amino acid substitutions were found in the \(Pk\text{-msp-1}_{133}\) region, and only one dimorphic alteration (F1789Y/S) in the \(Pk\text{-msp-1}_{119}\) region. All nine amino acid haplotypes were also used to compare with the additional 26 published sequences retrieved from GenBank (Fig. 5). The human isolate from India was found represent haplotype 1. One isolate each from human and macaque from Thailand were found to be identical to haplotypes 2 and 3, respectively. On the other hand, one of the human isolates and two of the macaque isolates from Thailand corresponded to haplotype 4.

In conclusion, the present study reveals a high level of genetic variability in \(Pk\text{-msp-1}_{142}\); however, no significant association between the genetic diversity and geographical location was found. Future study should be conducted using large numbers of samples from disparate geographical locations to critically assess whether this molecule might be considered as a potential vaccine candidate for \(P.\) knowlesi.

Acknowledgements

Funding from the Australian Research Council (ARC), the National Health and Medical Research Council (NHMRC) of Australia, Melbourne Water Corporation (R.B.G. et al.) and University of Malaya/Ministry of Higher Education High Impact Research (UMC625/1/HIR/MOHE/ MED-23), University of Malaya [PG056-2013A] (Y.A.L.L. et al.) gratefully acknowledged. Support from the Endeavour program (ID 4731-2015) is also gratefully acknowledged (NYJ).

References


