GENETIC DIVERSITY OF *PLASMODIUM VIVAX* DUFFY BINDING PROTEIN (*PvDBP*) GENE IN SABAH, MALAYSIA

Farhan Saat[1], Myat Htut Nyunt[2], Noor Rain Abdullah[1], Umi Rubiah Sastu[1], Nor Azrina Norahmad[1], Prem Kumar Muniandy[1], Jenarun Jelip[3], Khairul Faiz Ruslan[4], Moizin Tikuson[5] and Zakiah Ismail[1]

[1]Herbal Medicine Research Centre, Institute for Medical Research, Kuala Lumpur, Malaysia; [2]Parasitology Research Division, Department of Medical Research, Yangon, Myanmar; [3]Sabah State Health Department, Rumah Persekutuan, Kota Kinabalu, Sabah; [4]Tawau Health Department, Tawau; [5]Vector Borne Disease Control Unit, Kota Marudu Hospital, Kota Marudu, Sabah, Malaysia

Abstract. *Plasmodium vivax* Duffy binding protein (PvDBP) is a member of the Duffy binding-like erythrocyte binding protein (DBL-EBP) family expressed on the surface of *P. vivax* merozoites. PvDBP consists of seven regions responsible for the invasion of the parasite into host reticulocyte. DBP type II region is highly polymorphic and is genetically diverse. The gene sequence and genetic polymorphism of this region were investigated among *P. vivax* samples collected in Sabah, Malaysia. DNA was extracted from 20 *P. vivax*-infected blood samples and subjected to nested PCR to amplify the *PvDBPII* region for subsequent sequencing of the 900-bp amplicons. Sequences were aligned and compared with that of Salvador-1 strain (Sal-1) as standard *PvDBPII,* and a phylogenetic tree was constructed employing sequences from neighboring countries. The samples from Sabah could be categorized into four haplotypes. The amplified PvDBPII fragment contained 288 amino acids, among which 36 are nonsynonymous and 11 synonymous silent mutations; no mutations involved the conserved cysteine residues. Phylogenetic analysis of *PvDBPII* indicated that the phylogenetic tree has nine clusters and samples from Sabah are categorized into four clusters: cluster 1 (6 samples), cluster 6 (1 sample from Sabah, and samples from Thailand and Myanmar and Sal-1), group 8 (8 samples) and group 9 (5 samples from Sabah and samples from Thailand and Myanmar). Phylogenetic analysis revealed that *PvDBPII* of Kalabakan samples are confined to that area.

Keywords: *Plasmodium vivax*, *PvDBPII*, Sabah, Malaysia

INTRODUCTION

Malaria is a major human parasitic disease caused by five *Plasmodium* species, *P. falciparum, P. vivax*, *P. malariae*, *P. ovale* and *P. knowlesi* (WHO, 2013). Vivax malaria has s wide distribution affecting millions of the world's population (Baird, 2009), mainly outside Africa. However it has been overshadowed by the burden of *P. falciparum* disease and currently there is an increase of *P. knowlesi* cases in some

Correspondence: Muhammad Nor Farhan Saat, Herbal Medicine Research Centre, Institute for Medical Research, Jalan Pahang, 50588 Kuala Lumpur, Malaysia.
Tel: +603 26162633; Fax: +603 26934114
E-mail: mnfarhan@imr.gov.my

areas of Malaysia (Ministry of Health Malaysia, 2015).

Even though *P. vivax* is considered a benign infection, recent studies reported clinical severity caused by this parasite species (Alexandre *et al,* 2010), drug resistance (Rijkin *et al,* 2011) and recurrent clinical episodes due to relapse (Krotoski *et al,* 1982), indicating that the parasites can cause more severe morbidity than previously thought.

A crucial step for parasite survival is invasion of red blood cells by merozoites (Chitnis, 2001). *P. vivax* has a preference for reticulocytes with Duffy blood group antigens (Miller *et al,* 1976). Merozoite recognizes and binds to these (and other) markers on the surface of the reticulocytes (Adam *et al,* 1992). *P. vivax* Duffy binding protein (PvDBP) is one of the reticulocyte binding proteins (Adam *et al,* 1992), which plays a key role in invasion of the parasite into the reticulocyte (Horuk *et al,* 1993). The corresponding receptor on the reticulocytes is Duffy antigen receptor for chemokines (DARC). Individuals who lack the DARC on their red blood cells are resistant to *P. vivax* infection (Miller *et al,* 1976). Antibodies produced against PvDBP can prevent binding to DARC and hence interfere with the invasion process (Souza-Silva *et al,* 2010). This protein has been considered as a potential vaccine candidate for asexual stage of *P. vivax* (Ceravolo *et al,* 2008).

PvDBP is divided into 7 regions and the DARC binding site is located at region II (PvDBPII) (Ranjan and Chitnis, 1999), which has high genetic polymorphism (Cole-Tobain and King, 2003) but the cysteine (C) residues in this region are conserved within and between *P. vivax* populations from different geographical areas (Gosi *et al,* 2008).

In Malaysia since 2011, the majority of malaria cases are caused by *P. vivax* (2,422 cases, 45.6%) and *P. falciparum* (973 cases, 18.3%), followed by *P. malariae* (903 cases, 17.0%), *P. knowlesi* (854 cases, 16.1%) and mixed infections (153 cases, 3%) (Ministry of Health Malaysia, 2015). As Malaysia is approaching its malaria elimination target of 2020 for Sabah and Sarawak, it will require innovative and effective ways to prevent plasmodium transmission and maintain monitoring malaria cases. Understanding the genetic profile of the parasite locally is a component of a pre-elimination program (WHO, 2012). This genetic approach will help to determine genetic profiles of malaria parasites from different geographical areas. To the best of the authors' knowledge there is no published data available on the genetic profile and sequence of *PvDBPII* in Malaysia. Therefore, an understanding the genetic profiles of *P. vivax* in the country is important, as Malaysia moves towards its malaria elimination goal.

The present study was conducted to determine gene sequences and genetic polymorphisms of *PvDBPII* region among *P. vivax* isolates collected in Sabah, Malaysia. The genetic relationship was then constructed based on gene sequences to determine the relationship among the samples and also with isolates from neighboring countries.

MATERIALS AND METHODS

Study sites

The study sites were located in Kalabakan and Kota Marudu, Sabah, Malaysia. The sites were selected based on the high number of malaria cases reported in these areas prior to the study (Gething *et al,* 2012). The study populations were from villagers, palm oil estates workers, rubber plantations and road construction workers.

The study protocol was reviewed and approved by the institutional review board of IMR, Kuala Lumpur and the Medical Review and Ethics Committee (MREC) of the Ministry of Health, Malaysia (JPP-IMR 07-047), which comply with the Declaration of Helsinki.

Study samples and DNA preparation

In brief, everyone in the villages, palm oil estates and road construction sites were briefed on the objectives of the study. Information sheets and informed consent forms were distributed and explained. Individuals who consented to participate in the study were screened for malaria infection using a rapid diagnostic test (RDT) (Paramax- 3™; Zephyr Biomedicals, Goa, India). In addition, blood films (BFMP) were prepared to determine parasite density. Study participants diagnosed positive for malaria infection by RDT were selected for blood collection. Approximately one ml of whole blood was collected by venepuncture and then were spotted on 3MM Whatman® filter paper (Whatman; GE Healthcare, Buckinghamshire, UK). The filter papers were allowed to dry completely, labelled and transferred into individual plastic bags before transportation to the Institute for Medical Research (IMR) in Kuala Lumpur.

Speciation for *P. vivax*, *P. falciparum*, *P. malariae* and *P. knowlesi* were undertaken using a modified version of a published method (Imwong *et al*, 2009; Fuehrer *et al*, 2011). In brief, PCR mixture contained was 50 ng of DNA template, 1X *Taq* DNA PCR buffer, 0.4 mM dNTPs, 2.0 mM $MgCl_2$, 0.2 μM each primer (rPLU1/rPLU5 for the first PCR and rPLU3/rPLU4 for the genus-specific second PCR) and 2.5 U *Taq* DNA polymerase (Bio-rad, Hercules, CA). When the second PCR showed positive results, the following species-specific primers were used: rVIV1/rVIV2 (*P. vivax*), rFAL1/rFAL2 (*P. falciparum*), rMAL1/rMAL2 (*P. malariae*), and Pmk8/Pmkr9 (*P. knowlesi*). Thermal cycling was performed in an Eppendorf Mastercycler Gradient instrument (Eppendorf, Hamburg, Germany) as follows: 98°C for 4 minutes; 25 cycles of 65°C for 2 minutes, 72°C for 2 minutes and 94°C for 1 minute; and a final step at 72°C for 4 minutes. Only samples confirmed positive for *P. vivax* were used. DNA was extracted from filter paper using QIAamp® DNA mini kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. DNA samples were kept at -20ºC until used.

Amplification and sequencing of *PvDBPII*

The *PvDBPII* region was amplified using a PCR employing specific primers Nest 1 forward primer 5'-GATAAAACTGGGGAGGAAAAAGAT-3' and Nest1 reverse primer 5'- CTTATCGGATTTGAATTGGTGGC-3' (Xainli *et al*, 2000). The Nest 1 PCR mixture was conducted using 90 ng of genomic DNA, 1.5 mM $MgCl_2$, 200 μM dNTPs, 5 pmol of each primer and 0.5 U *Taq* DNA polymerase (Bio-Rad Laboratories). Thermocycling conducted using Eppendorf Mastercycler Gradient instrument (Eppendor, Hamburge, Germany) was as follows: 94ºC for 1 minute; followed by 35 cycles of 94ºC for 2 minutes, 61ºC for 2 minutes, and 72ºC for 2 minutes; and a final step at 72ºC for 10 minutes. One μl of the Nest 1 reaction mixture was used in Nest 2 PCR employing Nest 2 forward primer 5'- CCTCGAATGGTGGCAATCCT-3' and reverse primer 5'-TACGATACCTGCCGTCGTAAC-3'. The reaction mixture and thermocycling conditions were the same as for Nest 1 except that the annealing temperature was 54ºC. Amplicons were analysed using 2% agarose gel-electrophoresis, stained with

GelRed™ stain (Biotium, Hayward, CA) and visualized under UV illumination.

Amplicons were gel purified using a QIAquick® gel extraction kit (QIAGEN, Hilden, Germany), inserted into pCR™2.1-TOPO® vector (Invitrogen, Life Technologies, Carlsbad, CA) and used to transform One Shot® TOP10 Chemically Competent *E. coli* (Invitrogen). Transformants were cultured in Luria broth (Gibco, Life Technologies, Grand Island, NY). Plasmid DNA was extracted using PureLink® Quick Plasmid Miniprep Kit (Invitrogen) followed by digestion with *Eco*RI (New England Biolabs®). Gel purified insert DNA then was sequenced using an ABI prism™ BigDye™ Terminator Cycle Sequencing Ready Reaction kit v3.1 in an automated Applied Biosystems 3730xl capillary DNA sequencer (Thermo Fisher Scientific, Rockford, IL). Sequences were deposited with GenBank under the accession numbers MF197474 – 197493.

***PvDBPII* sequence analysis**

PvDBPII sequence data were edited and were assembled using a MegAlign software of DNAstar Lasergene® program (DNASTAR, Madison, WI). The sequences then were aligned and compared to the Salvador-1 strain (Sal-1) PVX_110810 (Plasmo DB). The deduced amino acid sequences were obtained using Mega5.2 software. Nucleotide diversity (π), an estimate of the average number of substitutions between any two sequences, and S, a measurement of the number of polymorphic site in the samples, were determined using the Mega5.2 software. *P*-value is set at 95% confidence interval (CI). The phylogenetic relationship among the samples was analysed using neighbor-joining method (MegAlign software of DNAstar Lasergene® program; DNASTAR, Madison, WI). Bootstrap proportions were used to assess the robustness of the constructed phylogenetic tree with 1,000 bootstrap replications.

RESULTS

***P. vivax* samples**

Twenty confirmed *P. vivax* samples were used in the study. DNA samples were successfully extracted and nested PCR of the *PvDBPII* showed that all samples had the expected amplicon size of 900 bp (data not shown).

Sequences analysis

The nucleotide sequence of *PvDBPII* samples from Kalabakan and Kota Marudu were categorized into haplotypes A, B C and D (Fig 1). Haplotype A consisted of 6 samples (30%) (95% CI: 9.9 - 50.1) (5 Kalabakan samples and 1 Kota Marudu samples), haplotype B 1 sample (5%) (95% CI: -4.5 - 14.5) from Kota Marudu, haplotype C 8 samples (40%) (95% CI: 18.5 - 61.5) (4 Kalabakan and 4 Kota Marudu samples), and haplotype D 5 samples (25%) (95% CI: 6.0 - 44.0) from Kalabakan (Fig 1). Alignment of these four haplotypes yielded a majority sequence of 864 bp.

Sequences were analyzed to determine nucleotide and amino acid changes, revealing 47 mutation sites, which resulted in 36 non-synonymous and 11 synonymous amino acid changes. None of the non-synonymous mutations involved C. More than 30% of the samples had nucleotide changes resulting in amino acid substitutions R308S, N375D, S379C, D384G, E385K, K386N, R390H, N417K/S, L424I, W437R, I503K and T512Q (Table 1). Nucleotide diversity π was estimated at 0.0100 ± 0.0011 ($p < 0.05$), while the number of variable sites was 47 and the mean number of nucleotide differences 8.69 when compared to Sal-1 *PvDBPII* sequence.

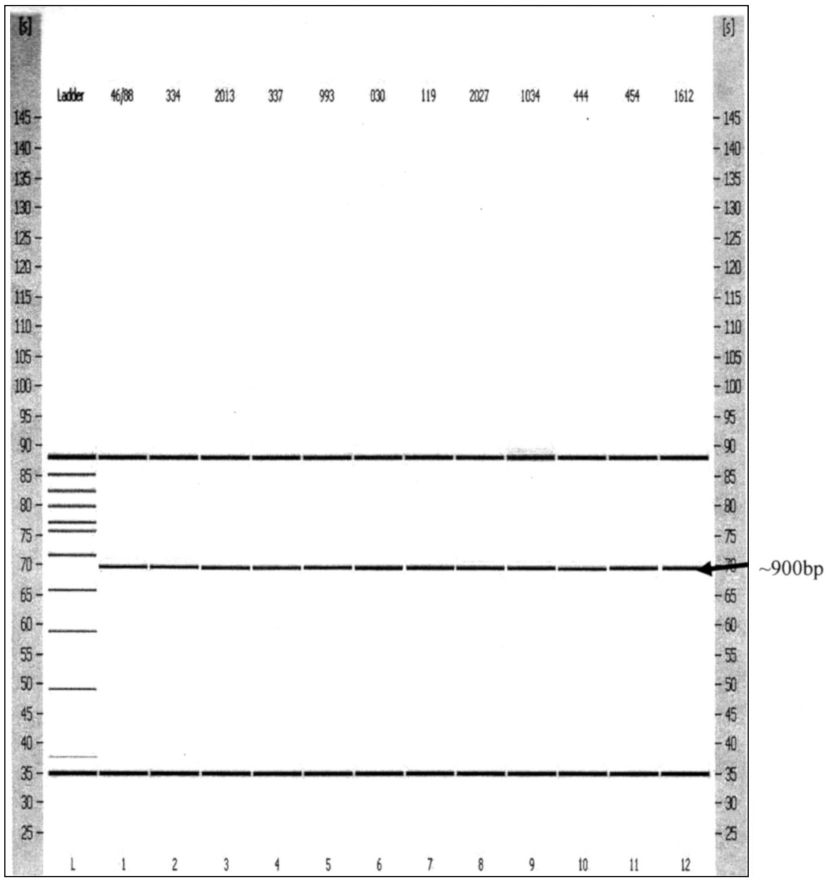

Fig 1–Sequence alignment of four *PvDBPII* haplotypes of samples from Sabah, Malaysia compared to that of Sal-1 (PVX_110810). The sequence alignment was generated by DNASTAR Program using MegAlign Software. All four haplotype sequence length is 864 bp. Alignment of the sample sequences compared to that of Sal -1 is between nucleotide position 687 and 1551. Differences in nucleotide compared to Sal -1 are marked in red.

High frequencies of variant amino acids were T512Q (100%) (95% CI: 100 - 100), D384G (95%) (95% CI: 85.4 - 104.5), R390H (95%) (95% CI: 85.4 - 104.5), L424I (70%) (95% CI: 9.9 - 50.1), I503K (70%) (95% CI: 49.9 - 90.1), W437R (60%) (95% CI: 38.5 - 81.5), N417K/S (35%) (95% CI: 14.1 - 55.9), R308S (30%) (95% CI: 9.9 - 50.1), N375D (30%) (95% CI: 9.9 - 50.1), S379C (30%) (95% CI: 9.9 - 50.1), E385K (30%) (95% CI: 9.9 - 50.1), and K386N (30%) (95% CI: 9.9 - 50.1) (Table 1). The same amino acid changes in PvDBPII samples from Sabah and neighboring countries (Myanmar, Papua New Guinea, Sri Lanka and Thailand) were R308S, K371E, D384G, E385K, K386N, H39N417K, L424I, W437R and I503K (Table 2). Although Q512 was common for all samples from Sabah, this variant is not present in samples from other neighboring countries.

Phylogenetic analysis

A phylogenetic tree was constructed based on the *PvDBPII* sequences alignment of the 20 samples from Sabah, Sal-1 (PVX_110810) and from published sequences from the neighboring countries, namely, India (Lim and Ayala, 2005), Myanmar (Ju *et al,* 2012), Thailand (Gosi *et al,* 2008), Sri Lanka (Premaratne *et al,* 2011), Indonesia (Ntumngia *et al,* 2009) and Vietnam (Lim *et al,* 2005) (Fig 2). The phylogenetic tree has 9 clusters with samples from Sabah categorized into 4 clusters, mainly cluster 1 (6 samples) and 6 (1 sample), while Thai, Myanmar and Sal-1 samples are in group 8 (8 samples) and group 9 (5 samples) (Fig 2). Samples from Kalabakan are distinct while those from

Consensus	TCGAATGGTG	GCAATCCTTA	CGATATTGAT	CATAAGAAAA	CGATCTCTAG	TGCTATTATA	AATCATGCTT	TTCTTCAAAA	
Haplotype A	TCGAATGGTG	GCAATCCTTA	CGATATTGAT	CATAAGAAAA	CGATCTCTAG	TGCTATTATA	AATCATGCTT	TTCTTCAAAA	80
Haplotype B	TCGAATGGTG	GCAATCCTTA	CGATATTGAT	CATAAGAAAA	CGATCTCTAG	TGCTATTATA	AATCATGCTT	TTCTTCAAAA	80
Haplotype C	TCGAATGGTG	GCAATCCTTA	CGATATTGAT	CATAAGAAAA	CGATCTCTAG	TGCTATTATA	AATCATGCTT	TTCTTCAAAA	80
Haplotype D	TCGAATGGTG	GCAATCCTTA	CGATATTGAT	CATAAGAAAA	CGATCTCTAG	TGCTATTATA	AATCATGCTT	TTCTTCAAAA	80
Salvador-1	TCGAATGGTG	GCAATCCTTA	CGATATTGAT	CATAAGAAAA	CGATCTCTAG	TGCTATTATA	AATCATGCTT	TTCTTCAAAA	767
Consensus	TACTGTAATG	AAAAACTGTA	ATTATAAGAG	AAAACGTCGG	GAAAGAGATT	GGGACTGTAA	CACTAAGAAG	GATGTTTGTA	
Haplotype A	TACTGTAATG	AAAAACTGTA	ATTATAAGAG	AAAACGTCGG	GAAAGAGATT	GGGACTGTAA	CACTAAGAAG	GATGTTTGTA	160
Haplotype B	TACTGTAATG	AAAAACTGTA	ATTATAAGAG	AAAACGTCGG	GAAAGAGATT	GGGACTGTAA	CACTAAGAAG	GATGTTTGTA	160
Haplotype C	TACTGTAATG	AAAAACTGTA	ATTATAAGAG	AAAACGTCGG	GAAAGAGATT	GGGACTGTAA	CACTAAGAAG	GATGTTTGTA	160
Haplotype D	TACTGTAATG	AAAAACTGTA	ATTATAAGAG	AAAACGTCGG	GAAAGAGATT	GGGACTGTAA	CACTAAGAAG	GATGTTTGTA	160
Salvador-1	TACTGTAATG	AAAAACTGTA	ATTATAAGAG	AAAACGTCGG	GAAAGAGATT	GGGACTGTAA	CACTAAGAAG	GATGTTTGTA	847
Consensus	TACCAGATCG	AAGATATCAA	TTATGTATGA	AGGAACTTAC	GAATTTGGTA	AATAATACAG	ACACAAATTT	TCATAGGGAT	
Haplotype A	TACCAGATCG	AAGATATCAA	TTATGTATGA	AGGAACTTAC	GAATTTGGTA	AATAATACAG	ACACAAATTT	TCATAGTGAT	240
Haplotype B	TACCAGATCG	AAGATATCAA	TTATGTATGA	AGGAACTTAC	GAATTTGGTA	AATAATACAG	ACACAAATTT	TCATAGGGAT	240
Haplotype C	TACCAGATCG	AAGATATCAA	TTATGTATGA	AGGAACTTAC	GAATTTGGTA	AATAATACAG	ACACAAATTT	TCATAGGGAT	240
Haplotype D	TACCAGATCG	AAGATATCAA	TTATGTATGA	AGGAACTTAC	GAATTTGGTA	AATAATACAG	ACACAAATTT	TCATAGGGAT	240
Salvador-1	TACCAGATCG	AAGATATCAA	TTATGTATGA	AGGAACTTAC	GAATTTGGTA	AATAATACAG	ACACAAATTT	TCATAGGGAT	927

Consensus	ATAACATTTC	GAAAATTATA	TTTGAAAAGG	AAACTTATTT	ATGATGCTGC	AGTAGAGGGC	GATTTATTAC	TTAAGTTGAA	
Haplotype A	ATAACATTTC	GAAAATTATA	TTTGAAAAGG	AAACTTATTT	ATGATGCTGC	AGTAGAGGGC	GATTTATTAT	TTAAGTTGAA	320
Haplotype B	ATAACATTTC	GAAAATTATA	TTTGAAAAGG	AAACTTATTT	ATGATGCTGC	AGTAGAGGGC	GATTTATTAC	TTAAGTTGAA	320
Haplotype C	ATAACATTTC	GAAAATTATA	TTTGAAAAGG	AAACTTATTT	ATGATGCTGC	AGTAGAGGGC	GATTTATTAC	TTAAGTTGAA	320
Haplotype D	ATAACATTTC	GAAAATTATA	TTTGAAAAGG	AAACTTATTT	ATGATGCTGC	AGTAGAGGGC	GATTTATTAC	TTAAGTTGAA	320
Salvador-1	ATAACATTTC	GAAAATTATA	TTTGAAAAGG	AAACTTATTT	ATGATGCTGC	AGTAGAGGGC	GATTTATTAC	TTAAGTTGAA	1007
Consensus	TAACTACAGA	TATAACAAAG	ACTTTTGCAA	GGATATAAGA	TGGAGTTTGG	GAGATTTTGG	AGATATAATT	ATGGGAACGG	
Haplotype A	TAACTACAGA	TATAACAAAG	ACTTTTGCAA	GGATATAAGA	TGGAGTTTGG	GAGATTTTGG	AGATATAATT	ATGGGAACGG	400
Haplotype B	TAACTACAGA	TATAACAAAG	ACTTTTGCAA	GGATATAAGA	TGGAGTTTGG	GAGATTTTGG	AGATATAATT	ATGGGAACGG	400
Haplotype C	TAACTACAGA	TATAACAAAG	ACTTTTGCAA	GGATATAAGA	TGGAGTTTGG	GAGATTTTGG	AGATATAATT	ATGGGAACGG	400
Haplotype D	TAACTACAGA	TATAACAAAG	ACTTTTGCAA	GGATATAAGA	TGGAGTTTGG	GAGATTTTGG	AGATATAATT	ATGGGAACGG	400
Salvador-1	TAACTACAGA	TATAACAAAG	ACTTTTGCAA	GGATATAAGA	TGGAGTTTGG	GAGATTTTGG	AGATATAATT	ATGGGAACGG	1087
Consensus	ATATGGAAGG	CATCGGATAT	TCCAAAGTAG	TGGAAAATAA	TTTGCGCAGC	ATCTTTGGAA	CTGGTGAAAA	GGCCCAACAG	
Haplotype A	ATATGGAAGG	CATCGGATAT	TCCAAAGTAG	TGGAACATAA	TTTGCGTAGC	ATCTTTGGAA	CTGGTAAAAA	TGCCCAACAG	480
Haplotype B	ATATGGAAGG	CATCGGATAT	TCCAAAGTAG	TGGAAAATAA	TTTGCGCAGC	ATCTTTGGAA	CTGATGAAAA	GGCCCAACAG	480
Haplotype C	ATATGGAAGG	CATCGGATAT	TCCAAAGTAG	TGGAAAATAA	TTTGCGCAGC	ATCTTTGGAA	CTGGTGAAAA	GGCCCAACAG	480
Haplotype D	ATATGGAAGG	CATCGGATAT	TCCAAAGTAG	TGGAAAATAA	TTTGCGCAGC	ATCTTTGGAA	CTGGTGAAAA	GGCCCAACAG	480
Salvador-1	ATATGGAAGG	CATCGGATAT	TCCAAAGTAG	TGGAAAATAA	TTTGCGCAGC	ATCTTTGGAA	CTGATGAAAA	GGCCCAACAG	1167

Fig 2

Consensus	CATCGTAAAC AGTGGTGGAA TGAATCTAAA GCACAAATTT GGACAGCAAT GATGTACTCA GTTAAAAAAA GATTAAAGGG	
Haplotype A	CATCGTAAAC AGTGGTGGAA TGAATCTAAA GCACAAATTT GGACAGCAAT GATGTACTCA GTTAAAAAAA GATTAAAGGG	560
Haplotype B	CGTCGTAAAC AGTGGTGGAA TGAATCTAAA GCACAAATTT GGACAGCAAT GATGTACTCA GTTAAAAAAA GATTAAAGGG	560
Haplotype C	CATCGTAAAC AGTGGTGGAA TGAATCTAAA GCACAAATTT GGACAGCAAT GATGTACTCA GTTAAAAAAA GATTAAAGGG	560
Haplotype D	CATCGTAAAC AGTGGTGGAA TGAATCTAAA GCACAAATTT GGACAGCAAT GATGTACTCA GTTAAAAAAA GATTAAAGGG	560
Salvador-1	CGTCGTAAAC AGTGGTGGAA TGAATCTAAA GCACAAATTT GGACAGCAAT GATGTACTCA GTTAAAAAAA GATTAAAGGG	1247
Consensus	GAATTTTATA TGGATTTGTA AATTAAATGT TGCGGTAAAT ATAGAACCGC AGATATATAG ATGGATTCGA GAATGGGGAA	
Haplotype A	GAAATTTTATA TGGATTTGTA AATAAAATGT TGCGGTAAAT ATAGAACCGC AGATATATAG ACGGATTCGA GAATGGGGAA	640
Haplotype B	GAATTTTATA TGGATTTGTA AATTAAATGT TGCGGTAAAT ATAGAACCGC AGATATATAG ATGGATTCGA GAATGGGGAA	640
Haplotype C	GAATTTTATA TGGATTTGTA AAATAAATGT TGCGGTAAAT ATAGAACCGC AGATATATAG ACGGATTCGA GAATGGGGAA	640
Haplotype D	GAATTTTATA TGGATTTGTA AATTAAATGT TGCGGTAAAT ATAGAACCGC AGATATATAG ATGGATTCGA GAATGGGGAA	640
Salvador-1	GAATTTTATA TGGATTTGTA AATTAAATGT TGCGGTAAAT ATAGAACCGC AGATATATAG ATGGATTCGA GAATGGGGAA	1327
Consensus	GGGATTACGT GTCAGAATTG CCCACAGAAG TGCAAAAACT GAAAGAAAAA TGTGATGGAA AAATCAATTA TACTGATAAA	
Haplotype A	GGGATTACGT GTCAGAATTG CCCACAGAAG TGCAAAAACT GAAAGAAAAA TGTGATGGAA AAATCAATTA TACTGATAAA	720
Haplotype B	GGGATTACGT GTCAGAATTG CCTACAGAAG TGCAAAAACT GAAAGAAAAA TGTGATGGAA AAATCAATTA TACTGATAAA	720
Haplotype C	GGGATTACGT GTCAGAATTG CCCACAGAAG TGCAAAAACT GAAAGAAAAA TGTGATGGAA AAATCAATTA TACTGATAAA	720
Haplotype D	GGGATTACGT GTCAGAATTG CCCACAGAAG TGCAAAAACT GAAAGAAAAA TGTGATGGAA AAATCAATTA TACTGATAAA	720
Salvador-1	GGGATTACGT GTCAGAATTG CCCACAGAAG TGCAAAAACT GAAAGAAAAA TGTGATGGAA AAATCAATTA TACTGATAAA	1407

Consensus	AAAGTATGTA AGGTACCACC ATGTCAAAAT GCGTGTAAAT CATATGATCA ATGGATAACC AGAAAAAAAA ATCAATGGGA	
Haplotype A	AAAGTATGTA AGGTACCACC ATGTCAAAAT GCGTGTAAAT CATATGATCA ATGGATAACC AGAAAAAAAA ATCAATGGGA	800
Haplotype B	AAAGTATGTA AGGTACCACC ATGTCAAAAT GCGTGTAAAT CATATGATCA ATGGATATCC AGAAAAAAAA ATCAATGGGA	800
Haplotype C	AAAGTATGTA AGGTACCACC ATGTCAAAAT GCGTGTAAAT CATATGATCA ATGGATAACC AGAAAAAAAA ATCAATGGGA	800
Haplotype D	AAAGTATGTA AGGTACCACC ATGTCAAAAT GCGTGTAAAT CATATGATGA ATGGATAACC AGAAAAAAAA ATCAATGGGA	800
Salvador-1	AAAGTATGTA AGGTACCACC ATGTCAAAAT GCGTGTAAAT CATATGATCA ATGGATAACC AGAAAAAAAA ATCAATGGGA	1487
Consensus	TGTTCTGTCA AATAAATTCA TAAGTGTAAA AAACGCAGAA AAGGTTACGA CGGCAGGTAT CGTA	
Haplotype A	TGTTCTGTCA AATAAATTCA AAAGTGTAAA AAACGCAGAA AAGGTTACGA CGGCAGGTAT CGTA	864
Haplotype B	TGTTCTGTCA AATAAATTCA TAAGTGTAAA AAACGCAGAA AAGGTTACGA CGGCAGGTAT CGTA	864
Haplotype C	TGTTCTGTCA AATAAATTCA AAAGTGTAAA AAACGCAGAA AAGGTTACGA CGGCAGGTAT CGTA	864
Haplotype D	TGTTCTGTCA AATAAATTCA TAAGTGTAAA AAACGCAGAA AAGGTTACGA CGGCAGGTAT CGTA	864
Salvador-1	TGTTCTGTCA AATAAATTCA TAAGTGTAAA AAACGCAGAA AAGGTTCAGA CGGCAGGTAT CGTA	1551

Fig 2

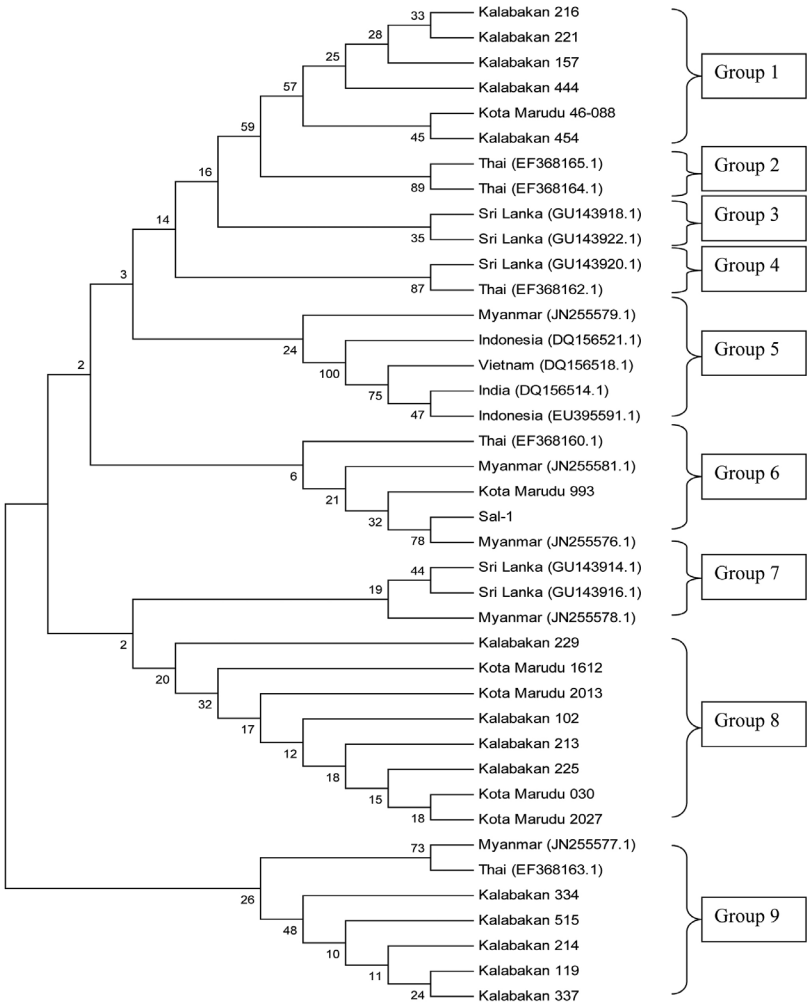

Fig 2–Phylogenetic tree of twenty *PvDBPII* sequences from Sabah, Malaysia compared to Sal-1 (PVX_110810) and those from neighboring countries. The phylogenetic tree was constructed with the neighbor-joining method using the MEGA 5.2 software. Percent replicate tree in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branch site.

Kota Marudu are present in a number of clusters that contain members from other neighboring countries.

DISCUSSION

Malaysia is pursuing a malaria elimination and has a national goal to eliminate malaria by 2015 in the Peninsular Malaysia and by 2020 in Sabah and Sarawak, Malaysia (APMEN, 2013). Malaria is still one of the most important vector-borne diseases in Malaysia after dengue (Alias *et al*, 2014). Although the number of malaria cases decreased in some areas, there are still cases in inland hilly rural areas especially in Sabah and Sarawak, where transportation and diagnosis are still a problem.

Although Duffy antigens are required for *P. vivax* merozoite invasion of reticulocytes (Ju *et al*, 2013), the PvDBPII binding region is highly polymorphic and varies from region to region (Patchanee *et al*, 2014). Polymorphism of *PvDBPII* may help the parasite to evade host immune response and facilitate invasion of parasite. In Malaysia, genetic polymorphism of *P. vivax* has not been examined closely and comprehensively unlike that of *P. falciparum* (Abdullah *et al*, 2013).

The present study is the first report on *PvDBPII* genetic polymorphism of samples collected in Sabah. There were four haplotypes, with haplotypes A and D dominant in Kalabakan and haplotype

Table 1
Sequence polymorphism of *PvDBPII* of twenty samples from S, Malaysia compared to that of Sal-1 (PVX_110810).

	Position of amino acid														
	248	250	251[a]	255	257	308	312	313	316[a]	320[a]	326[a]	333	342	358	359
sal-1	ATT	AAT	CAT	CAA	ACT	AGG	TTT	CGA	TAT	AAA	GCA	CTT	AAA	ATA	ATT
Sabah	ACT	GAT	CAC	CGA	ATT	AGT	TCT	CAA	TAC	AAG	GCG	TTT	GAA	AAA	GTT
	I	N	H	Q	T	R	F	R	Y	K	A	L	K	I	I
	T	D	H	R	I	S	S	Q	Y	K	A	F	E	K	V
Frequency	(1/20)	(1/20)	(1/20)	(1/20)	(1/20)	(6/20)	(1/20)	(1/20)	(1/20)	(1/20)	(1/20)	(4/20)	(1/20)	(1/20)	(1/20)
Percentage	5	5	5	5	5	30	5	5	5	5	5	20	5	5	5
	Position of amino acid														
	367	371	375	378[a]	379	384	385	386	390	391	396	405[a]	407	417[b]	424
sal-1	ATC	AAA	AAT	CGC	AGC	GAT	GAA	AAG	CGT	CGT	AAT	GCA	ATG	AAT	TTA
Sabah	ACC	GAA	GAT	CGT	TGC	GGT	AAA	AAT	CAT	TGT	GAT	GCG	ACG	AAA/AGT	ATA
	I	K	N	R	S	D	E	K	R	R	N	A	M	N	L
	T	E	D	R	C	G	K	N	H	C	D	A	T	K/S	I
Frequency	(1/20)	(1/20)	(6/20)	(6/20)	(6/20)	(19/20)	(6/20)	(6/20)	(19/20)	(1/20)	(1/20)	(1/20)	(1/20)	(7/20)	(14/20)
Percentage	5	5	30	30	30	95	30	30	95	5	5	5	5	35	70
	Position of amino acid														
	429	437	443	450[a]	452	462[a]	466[a]	486	488	489	490	503	505	506	512
sal-1	AAT	TGG	AGG	CCC	GAA	GGA	TAT	CAA	ATA	ACC	AGA	ATA	GTA	AAA	CAG
Sabah	TAT	CGG	GGG	CCT	AAA	GGG	TAC	GAA	GTA	TCC	GGA	AAA	GCA	AGA	ACG
	N	W	G	P	E	G	Y	Q	I	TCC	R	I	V	K	T
	Y	R	R	P	K	G	Y	E	V	S	G	K	A	R	Q
Frequency	(1/20)	(12/20)	(1/20)	(1/20)	(1/20)	(2/20)	(1/20)	(5/20)	(1/20)	(1/20)	(1/20)	(14/20)	(1/20)	(1/20)	(20/20)
Percentage	5	60	5	5	5	10	5	25	5	5	5	70	5	5	100

[a]Single nucleotide change leading to synonymous mutation.

Table 2
Percent frequency of common amino acid changes of PvDBPII sequences samples from Sabah, Malaysia and other neighboring countries.

AA position	% frequency									
	R308S	K371E	D384G	E385K	K386N	H390R	N417K	L424I	W437R	I503K
Sabah, Malaysia	30.0	5.0	95.0	30.0	30.0	95.0	30.0	70.0	60.0	70.0
Myanmar[a]	22.2	22.2	85.2	33.3	33.3	63.0	38.9	83.3	61.1	77.8
Papua New Guinea[b]	69.0	11.5	34.5	9.7	9.7	50.4	33.6	68.1	32.7	42.5
Sri Lanka[c]	13.0	34.0	94.0	20.0	20.0	66.0	36.0	49.0	37.0	55.0
Thailand[d]	26.7	20.0	76.7	46.6	40.0	56.6	36.6	86.7	63.3	56.7

[a]Ju *et al* (2012); [b]Hans *et al* (2005); [c]Premaratne *et al* (2011); [d]Gosi *et al* (2008).

C more common among samples from both Kalabakan and Kota Marudu. This finding is the first such observation in this region.

The 36 nonsynonymous mutations detected in the 20 PvDBPII sequences did not involve the conserved C residues. It is well known that the C residues within *PvDBPII* sequence are conserved within and between *P. vivax* populations from different geographic regions (Gosi *et al*, 2008; Premaratne *et al*, 2011; Babaeekho *et al*, 2009). The majority of the C residues are involved in reticulocyte binding (Hans *et al*, 2005). The non-synonymous amino acid changes seen in the Sabah samples also were found in neighboring countries including Thailand (Gosi *et al*, 2008), Sri Lanka (Premaratne *et al*, 2011), Myanmar (Ju *et al*, 2012) and Papua New Guinea (Nóbrega *et al*, 2011) indicating a commonality shared among *P. vivax* isolates in the region.

Interestingly, amino acid change T512Q was only observed in the Sabah samples and not in samples of neighboring countries, such as Myanmar (Ju *et al*, 2012), Papua New Guinea (Nóbrega *et al*, 2011), Sri Lanka (Premaratne *et al*, 2011) and Thailand (Gosi *et al*, 2008). Although the function of the residue at this position is unknown, it can be used as a molecular marker for *P. vivax* population in Sabah.

Mutations at amino acid positions 417, 437 and 503 has been reported to have the ability to influence the mechanism of inhibitory antibodies against reticulocyte binding (McHenry *et al*, 2011). The polymorphism causes the parasite to generate new variants of PvDBPII, which would allow the parasites to maintain the ability to bind to reticulocyte and evade from host inhibitory antibody system (VanBuskirk *et al*, 2004).

The high polymorphism of *PvDBPII* sequence is often associated with emergence of a new haplotype (Ju *et al*, 2013). Phylogenetic analysis of *PvDBPII* suggested that Sabah isolates belong to 4 clusters, with samples from Kalabakan being clustered together and distinct from samples from Kota Marudu, indicating that these samples are confined to this area.

In summary, the *PvDBPII* sequenced of Sabah samples constitute four haplotypes. Nucleotide and deduced amino acid changes resulted in 36 non-synonymous amino acid mutations that did not involve the conserved cysteine residues. Phylogenetic analysis shows that *PvDBPII* of Kalabakan samples were confined to that area.

ACKNOWLEDGEMENTS

The authors thank the Director General of Health Malaysia for permission to publish this paper. We also thank the Director of the Institute for Medical Research, Kuala Lumpur, for critical review and support in publishing this paper. The study was supported by the National Institute of Health, Ministry of Health Malaysia. We are grateful to all staff of Tawau Health Department and Kota Marudu Health Clinic for their logistic support and assistance during the study in the field and in sample collection. We are indebted to the individuals and their families for their willingness to participate in this study.

REFERENCES

Abdullah NR, Barber BE, William T, Norahmad NA, Satsu UR, Muniandy PK. *Plasmodium vivax* population structure and transmission dynamics in Sabah Malaysia. *PLOS One* 2013; 8: e82553.

Adams JH, Sim BK, Dolan SA, Fang X, Kaslow DC, Miller LH. A family of erythrocyte binding proteins of malaria parasites. *Proc Natl Acad Sci USA* 1992; 89: 7085-9.

Alexandre MA, Ferreira CO, Siqueira AM, *et al.* Severe *Plasmodium vivax* malaria, Brazilian Amazon. *Emerg Infect Dis* 2010, 16: 1611-4.

Alias H, Johari S, Rohela M, *et al*. Spatial distribution of malaria in Peninsular Malaysia from 2000 to 2009. *Parasit Vectors* 2014; 7: 186.

Babaeekho L, Zakeri S, Djadid N. Genetic mapping of the Duffy binding protein (DBP) ligand domain of *Plasmodium vivax* from unstable malaria region in the Middle East. *Am J Trop Med Hyg* 2009; 80: 112-8.

Baird J. Resistance to therapies for infection by *Plasmodium vivax*. *Clin Microbiol Rev* 2009; 22: 508-34.

Ceravolo IP, Souza-Silva FA, Fontes CJF, *et al.* Inhibitory properties of the antibody response to *Plasmodium vivax* Duffy binding protein in an area with unstable malaria transmission. *Scand J Immunol* 2008; 67: 270-8.

Chitnis CE. Molecular insights into receptors used by malaria parasites for erythrocyte invasion. *Curr Opin Hematol* 2001; 8: 85-91.

Cole-Tobian J, King CL. Diversity and natural selection in *Plasmodium vivax* Duffy binding protein gene. *Mol Biochem Parasitol* 2003; 127: 121-32.

Fuehrer HP, Fally MA, Habler VE, Starzengruber P, Swoboda P, Noedl H. Novel nested direct PCR technique for malaria diagnosis using filter paper samples. *J Clin Microbiol* 2011; 49: 1628-30.

Gething PW, Elyazar IR, Moyes CL, Smith DL, Battle KE. A long neglected world malaria map: *Plasmodium vivax* endemicity in 2010. *PLOS Negl Trop Dis* 2012; 6: e1814.

Gosi P, Khusmith S, Khalambaheti T, *et al*. Polymorphism patterns in Duffy-binding protein among Thai *Plasmodium vivax* isolates. *Malar J* 2008; 7: 112.

Hans D, Pattnaik P, Bhattacharyya A, *et al.* Mapping binding residues in the *Plasmodium vivax* domain that binds Duffy antigen during red cell invasion. *Mol Microbiol* 2005; 55: 1423-34.

Horuk R, Chitnis C, Darbonne W, *et al*. A receptor for the malarial parasite *Plasmodium vivax*: the erythrocyte chemokine receptor. *Science* 1993; 261: 1182-4.

Imwong M, Tanomsing N, Pukrittayakamee S, Day NP, White NJ, Snounou G. Spurious amplification of a *Plasmodium vivax* small-subunit RNA gene by use of primers currently used to detect *P. knowlesi*. *J Clin Microbiol* 2009; 47: 4173-5.

Ju HL, Kang JM, Moon SU, *et al.* Genetic polymorphism and natural selection of Duffy binding protein of *Plasmodium vivax* Myanmar isolates. *Malar J* 2012; 11: 60.

Ju HL, Kang JM, Moon SU, *et al.* Genetic diversity and natural selection of Duffy binding protein of *Plasmodium vivax* Korean isolates. *Acta Trop* 2013; 125: 67-74.

Krotoski WA, Collins WE, Bray RS, *et al.* Demonstration of hypnozoites in sporozoite-transmitted *Plasmodium vivax* infection. *Am J Trop Med Hyg* 1982; 31: 1291-3.

Lim C, Ayala FJ. Population structure and genetic diversity of *Plasmodium vivax* Duffy binding protein gene. Department of Ecology and Evolutionary Biology, University of California, 2005.

McHenry AM, Barnes SJ, Ntumngia FB, King CL, Adams JH. Determination of the molecular basis for a limited dimorphism, N417K, in the *Plasmodium vivax* Duffy-binding protein. *PLOS One* 2011; 6: e20192.

Miller LH, Mason SJ, Clyde DF, McGinniss MH. The resistance factor to *Plasmodium vivax* in blacks. *N Engl J Med* 1976; 295: 302-4.b

Ministry of Health Malaysia, the World Health Organization, the University of California, San Francisco. Eliminating malaria: case-study 8 | Progress towards elimination in Malaysia. Geneva: The World Health Organization, 2015.

Nóbrega de Sousa T, Carvalho LH, Alves de Brito CF. Worldwide genetic variability of the Duffy binding protein: insights into *Plasmodium vivax* vaccine development. *PLOS One* 2011; 6: e22944.

Ntumngia FB, McHenry AM, Barnwell JW, Cole-Tobian J, King CL, Adams JH. Genetic variation among *Plasmodium vivax* isolates adapted to non-human primates and the implication for vaccine development. *Am J Trop Med Hyg* 2009; 80: 218-27.

Patchanee C, Amy MM, Francis BN, Jetsumon S, John HA. The association of Duffy binding protein region II polymorphisms and its antigenicity in *Plasmodium vivax* isolates from Thailand. *Parasitol Int* 2014; 63: 858-64.

Premaratne P, Aravinda B, Escalante A, Udagama P. Genetic diversity of *Plasmodium vivax* Duffy Binding Protein II (*PvDBPII*) under unstable transmission and low intensity malaria in Sri Lanka. *Infect Genet Evol* 2011; 11: 1327-39.

Ranjan A, Chitnis CE. Mapping regions containing binding residues within functional domains of *Plasmodium vivax* and *Plasmodium knowlesi* erythrocyte-binding proteins. *Proc Natl Acad Sci USA* 1999; 96: 14067-72.

Rijken M, Boel M, Russell B, *et al.* Chloroquine resistant *vivax* malaria in a pregnant woman on the western border of Thailand. *Malar J* 2011; 10: 113.

Souza-Silva FA, da Silva-Nunes M, Sanchez BAM, *et al.* Naturally acquired antibodies to *Plasmodium vivax* Duffy Binding Protein (DBP) in rural Brazilian Amazon. *Am J Trop Med Hyg* 2010; 82: 185-93.

The Asia Pacific Malaria Elimination Network (APMEN). [Website]. [Cited 2013 Jul 30]. Available from: http://apmen.org

VanBuskirk KM, Cole-Tobian J, Baisor M, *et al.* Antigenic drift in the ligand domain of *Plasmodium vivax* duffy binding protein confers resistance to inhibitory antibodies. *J Infect Dis* 2004; 190: 1556-62.

World Health Organization (WHO). World malaria report 2012. Geneva: WHO, 2012.

World Health Organization (WHO). World malaria report 2013.Geneva: WHO, 2013.

Xainli J, Adams J, King C. The erythrocyte binding motif of *Plasmodium vivax* Duffy binding protein is highly polymorphic and functionally conserved in isolates from Papua New Guinea. *Mol Biochem Parasitol* 2000; 111: 253-60.