Inter simple sequence repeats (ISSRs): Neglected DNA markers for molecular dissection of *Plasmodium* species in long-tailed Macaque (*Macaca fascicularis*)

Parastoo Khajeaian, Soon Guan Tan, Alireza Valdiani, Christina Seok Yien Yong, Mohammad Zareian Jahromi, Noorjahan Banu Alitheen, Jeffrine Japning Rovie-Ryan

Understanding the genetic variation of the *Plasmodium* parasites could play an important role in controlling and preventing this lethal infection. Inter simple sequence repeat (ISSR) markers have successfully been tested for investigating the genetic diversity of malaria vectors. It is hypothesized that ISSRs could lead to fruitful results in studying the genetic variation of *Plasmodium* species, as well. To illustrate the genetic diversity of two infectious *Plasmodium* species, including *Plasmodium knowlesi* and *Plasmodium cynomolgi*, infected and uninfected monkey blood samples were separately collected on filter papers (FTA cards), and used for DNA extraction. A total of 103 and 95 polymorphic ISSR loci were detected in infected and uninfected samples, respectively. Cluster analysis of the *Plasmodium* and *Macaca fascicularis* accessions both resulted in the generation of three clusters. However, the most significant result of the cluster analysis was revealing the high efficiency of ISSR markers in the discrimination of the two *Plasmodium* species from each other. The cluster analysis showed a wide range of genetic diversity among both *Plasmodium* and the long-tailed Macaque accessions. The principal component analysis (PCA) also confirmed the cluster analysis results.

Inter simple sequence repeats (ISSRs): Neglected DNA markers for molecular dissection of *Plasmodium* species in long-tailed Macaque (*Macaca fascicularis*)

Parastoo Khajeaian[1], Soon Guan Tan[1], Alireza Valdiani[2*], Christina Yong Seok Yien[3], Mohammad Zareian Jahromi[4], Noorjahan Banu Alitheen[1], Jeffrine Japning Rovie-Rian[5]

[1] Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia
[2] Department of Biochemistry, Faculty of Biotechnology and Molecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia
[3] Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia
[4]Institute of Bioscience, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia
[5] Wildlife Genetic Resource Bank (WGRB) Laboratory, Ex-Situ Conservation Division, Department of Wildlife and National Parks (DWNP), Kuala Lumpur, Malaysia

[*]Corresponding author:
Alireza Valdiani
E-mail: alireza.valdiani@gmail.com
Tel: +603 8946 6698
Fax: +603 8943 0913

Abstract

Understanding the genetic variation of the *Plasmodium* parasites could play an important role in controlling and preventing this lethal infection. Inter simple sequence repeat (ISSR) markers have successfully been tested for investigating the genetic diversity of malaria vectors. It is hypothesized that ISSRs could lead to fruitful results in studying the genetic variation of *Plasmodium* species, as well. To illustrate the genetic diversity of two infectious *Plasmodium* species, including *Plasmodium knowlesi* and *Plasmodium cynomolgi*, infected and uninfected monkey blood samples were separately collected on filter papers (FTA cards), and used for DNA extraction. A total of 103 and 95 polymorphic ISSR loci were detected in infected and uninfected samples, respectively. Cluster analysis of the *Plasmodium* and *Macaca fascicularis* accessions both resulted in the generation of three clusters. However, the most significant result of the cluster analysis was revealing the high efficiency of ISSR markers in the discrimination of the two *Plasmodium* species from each other. The cluster analysis showed a wide range of genetic diversity among both *Plasmodium* and the long-tailed Macaque accessions. The principal component analysis (PCA) also confirmed the cluster analysis results.

Keywords: ISSRs, Genetic pattern, *Plasmodium cynomolgi*, *Plasmodium knowlesi*, *Macaca fascicularis.*

1. Introduction

It has been more than a century since long- tailed Macaques (*Macaca fascicularis*) are being used in research and analysis of human diseases because of their similar and related immunological, neurological, and reproductive structures in humans (Fooden, 1995). They are the second most numerous primates and one of our closest living relatives (Barr et al., 2003). Monitoring studies of malaria parasites among the long-tailed Macaques in Malaysia have

revealed that this species has the maximum prevalence of malaria compared to the other Macaque species (Eyles, 1963; Lee et al., 2011). Recent studies have found that this species of Macaque is the natural host of *Plasmodium knowlesi* infection (Jeslyn et al., 2011).

Malaria still remains a major cause of human death globally in spite of a century of research (WHO, 2013). Numerous factors such as global travel and immigration of people from areas with high prevalence of malaria are the reasons for the presence of the disease in developed countries. Indeed, importation of the infection by travellers has been reported from non-endemic parts of the world such as Europe (Bronner et al., 2009; Tang et al., 2010), New Zealand (Hoosen and Shaw, 2011), Australia (Figtree et al., 2010), and the United States (Ennis et al., 2009). A simian malaria parasite, *P. knowlesi* has recently been reported to be infectious to humans (Singh et al., 2004). Most likely, on this occasion, the Southeast Asian region has become a major point of attention of *P. knowlesi* malaria, and an interesting work field for many scientists. As a consequence, significant findings regarding *P. Knowlesi* malaria has been revealed in Sarawak, Malaysian Borneo (Vythilingam et al., 2006; Vythilingam et al., 2008; Tan et al., 2008). For instance, these findings approve that *P. knowlesi* specifically belongs to the Southeast Asia region because *An. leucosphyrus* is the only group of mosquitoes shown to be the invertebrate vector for this parasite (Collins, 2012).

It is for sure that several mosquito species are capable of naturally transferring malaria in both monkeys and humans, whereas only three of them, namely, *An. balabacensis*, *An. Cracensand* and *An. Lateens* have been shown as simian vectors in Malaysia (Vythilingam et al., 2006; Tan et al., 2008; Jiram et al., 2012; Vythilingam, 2010). Mosquitoes belonging to the *Anopheles leucosphyrus* group have been incriminated for transmitting *P. knowlesi* in nature. Three species, including *An. hackeri*, *An. cracens* (Vythilingam et al., 2008) in Peninsular Malaysia

and *An. latens* in Sarawak, Malaysian Borneo (Vythilingam et al., 2006; Tan et al., 2008) have been found to transmit *P. knowlesi* and several other species of this group have also been found to transmit other simian malaria parasites under natural or experimental conditions (Coatney, 1971).

Some other species of simian malaria parasites such as *P. cynomolgi* and *P. inui*, have currently been proven to be infectious to humans as well (Galinski and Barnwell, 2009). According to the latest report of naturally acquired *P. cynomolgi* in human, the urgent need arises to find novel diagnostic methods for detecting and recognizing the species (Ta et al., 2014). Molecular markers are efficient tools for genetic analyses, taxonomic classification, studying the phylogenetic relationship, as well as prognostic studies in different taxa (Valdiani et al., 2014). Recent achievements in molecular methods, especially the polymerase chain reaction (PCR) enabled scientists to genotype malaria parasites directly from the patient's blood samples without any prior *in vitro* culture needed (Brito and Ferreira, 2011). During the 1980s RFLP (Restriction fragment-length polymorphisms) was the accepted method for mapping malaria infections (Wellems et al., 1990). Microsatellites or simple sequence repeats (SSR) were established in the 1990s to map the chloroquine (CQ) -resistant gene (Su and Wellems, 1996; Su et al., 1997). Advances in single-nucleotide polymorphism (SNP) genotyping method brought in a new aspect to marker technology in recent years (Maresso and Broeckel, 2008; Su et al., 2007). High mutation rates, multiple alleles and high polymorphisms are the advantages of SSRs compared to SNPs in population studies of closely related parasites (Anderson et al., 2000).

Considering the above-mentioned trends in the molecular studies of malaria, application of an intermediate form of markers is urgently required. The inter-simple sequence repeats (ISSRs)

106 can be useful alternatives with many benefits. The technique overcomes the limitations of the
107 PCR-based DNA marker systems like the low reproducibility of random amplified
108 polymorphic DNA (RAPD), high cost of amplified fragment length polymorphism (AFLP) and
109 the necessity of having enough information about the flanking sequences to develop primers for
110 simple sequence repeats (SSRs) (Gupta et al., 1994; Wu et al., 1994; Zietkiewicz et al., 1994).
111 Accordingly, the ISSRs are valuable tools for studies on genetic diversity, phylogeny, gene
112 tagging, genome mapping and evolutionary biology (Reddy et al., 2002; Becker and Heun,
113 1995; Joshi et al., 2000; Charters and Wilkinson, 2000).
114 On the other hand, the dominant nature of the ISSRs markers can lead to generating a diverse
115 and informative DNA profile in both *M. fascicularis* and *Plasmodium* malaria parasite. This
116 feature of the marker makes it suitable for population investigation of the two different taxa,
117 simultaneously. Since, the recent use of the microsatellites (SSRs), as a co-dominant marker,
118 has resulted in the detection of a low to moderate genetic diversity in Malaysian long-tailed
119 Macaque population (Nikzad et al., 2014); alternatively, the application of ISSRs as a dominant
120 marker can lead to unexpected results on the diversity of the Malaysia Macaque population.
121 The objective of this research was a concurrent study of the genetic variation of the
122 *Plasmodium*-infected and uninfected population of *M. fascicularis* in Malaysia. The present
123 investigation is the first attempt to study the two infectious *Plasmodium* species using ISSR
124 markers. This process could more specifically lead to getting a deeper insight into the
125 phylogenetic relationship of the *P. knowlesi* and *P. cynomolgi* malaria parasites.

126 **2. Materials and Methods**

127 *2.1. Sample collection*

Whole blood samples and filter papers with dried blood spots were provided by the Department of Wildlife and National Parks (DWNP), Malaysia. Sampling was done on the conflict animals that invaded the human settlement areas (Fig. 1), as part of a Wildlife Disease Surveillance program (WDSP) conducted by DWNP. The collected samples with the common geographical origins were grouped under two categories; *Plasmodium*-infected and uninfected long-tailed Macaque (*Macaca fascicularis*) specimens (Tables 1 and 2).

2.2. Ethical clearance

The investigation detailed in this manuscript complied with the protocols approved by the Institutional Animal Care and Use Committee (IACUC), University of California, Davis, USA as adopted by the PREDICT program in Malaysia under which DWNP is working collaboratively with the EcoHealth Alliance, the Ministry of Health Malaysia, and the Veterinary Services Department, Malaysia. The related approval letter has been provided as a Supplemental Information file with the reference number of JPHL&TN(IP): 90-4/11.1 JID5 (9).

2.3. Extraction of genomic DNA

DNA was extracted from the blood samples by using the QIAamp DNA mini kit (Qiagen, Germany) according to the manufacturer's instructions. The DNA extraction was carried out using an approximate volume of 200 µL of whole blood or 4 punches from filter papers. The DNA samples were extracted in different final volumes due to the concentration of the used genetic materials. The final template volume obtained from the filter papers after extraction was 100 µL, while the amount of whole blood samples was 150 µL. One microliter of each genomic DNA sample was loaded on 1% (w/v) molecular biology grade agarose gel (Sigma-Aldrich, USA), and run for 55 min at 80 V. The gel was stained using ethidium bromide (130 µL L^{-1}), and visualized under UV light using an HD2 UVIdoc (Uvitec, UK). The intensity of

151 the genomic DNA bands was checked, and no additional band was observed on the gel. A
152 further quality evaluation of the extracted DNAs was done by calculating the OD260/OD280
153 ratio using a Nanodrop spectrophotometer, model ND1000 (NanoDrop Technologies, Inc.,
154 USA). The DNA extracts were then transferred to a –20°C freezer until further use.

155 *2.4. Detection of Plasmodium species using the nested PCR assay*

156 Prior to the ISSR-PCRs, a Nested Polymerase Chain Reaction (Nested PCR) was conducted
157 using specific primers for the detection of *P. knowlesi*, *P. inuii*, *P. cynomolgi*, *P. fieldi* and *P.*
158 *coatneyi* in the *Plasmodium*-infected samples. As a result, the presence of *P. cynomolgi* and *P.*
159 *knowlesi* was confirmed by nested PCR in the infected samples (Table 1). The same procedure
160 using the same primers was repeated on the *Plasmodium*-free (control) samples to confirm the
161 absence of the above-mentioned five *Plasmodium* species in the control blood samples (Table
162 2). The optimized protocol and the specific primers had been described by Lee et al. (2011),
163 and Sing et al. (1999), previously. In the present experiment, the PCR was aimed to target the
164 small subunit ribosomal RNA (ssrRNA) genes with some minor modifications. Negative
165 controls were included to ensure the lack of contamination.

166 *2.5. Amplification of Inter Simple Sequence Repeats (ISSRs)*

167 A total of 26 ISSR primers were chosen randomly for DNA fingerprinting of the 20
168 *Plasmodium* accessions of the *P. cynomolgi* and *P. knowlesi* species. The initial optimization of
169 all these primers was done using gradient PCR. The gradient PCR was performed at 94°C for 2
170 min, followed by 35 cycles consisting of 30 Sec at 94°C (denaturation), 45 Sec at ±6°C of the
171 melting temperature (T_m) of each primer, 1 min at 72°C (extension) and 7 min at 72°C (final
172 extension). The PCR was carried out in a total volume of 25 µL for each reaction, while the
173 final concentration of the PCR master mixes was adjusted at Green GoTaq® Flexi Buffer (1X),

PCR nucleotide mix or dNTP (0.2 mM each dNTP), $MgCl_2$ (1.5 mM), primer (0.4 µM), 0.7 units of Taq DNA polymerase (GoTaq® PCR Core Systems, Promega, USA) and 50 ng of the genomic DNA (Valdiani et al., 2012).

PCR thermal cycler condition was optimized to an initial denaturation at 94°C for 2 min, followed by 35 cycles consisting of 30 Sec at 94°C (denaturation), 45 Sec at desired annealing temperature for each primer (ranging from 50-62°C), 1 min at 72°C (extension) and 7 min at 72°C (final extension). The amplified fragments were separated using a 2% (w/v) molecular biology grade agarose gel (Sigma-Aldrich, USA). The gels were stained using ethidium bromide (130 µL L^{-1}), and documented under UV light using an HD2 UVIdoc (Uvitec, UK). To improve the reliability of the results, the whole procedure was repeated twice (Valdiani et al., 2013), using the same infected and uninfected samples.

The PCR revealed that only 19 of those 26 primers are capable of producing polymorphic patterns that are presented in Tables 3 and 4, while seven ISSR primers, including UBC 807, UBC 827, UBC 834, UBC 848, UBC 861, UBC 874 and UBC 878 were found to produce monomorphic banding patterns. Therefore, to maximize the reliability of the results and make them comparable, only the 19 polymorphic primers were employed for DNA profiling of the 20 uninfected long-tailed Macaque (*Macaca fascicularis*) samples.

2.6. Statistical analysis

The size of the generated ISSR bands was scored using the UVIdoc software version 99.02. The presence (1) and absence (0) of the bands was recorded in an Excel file.

The JMP version 8 software (SAS Institute Inc.) was utilized to generate the dendrograms and their color maps (SAS Institute Inc., 2009). The calculations of Jaccard's similarity indices, as well as the principal components analysis (PCA), were done by using the NTSYS-PC version

2.1 software (Rohlf, 2000). The data were analyzed using the SIMQUAL (Similarity for Qualitative Data) option to generate the similarity matrices. Jaccard's similarity coefficient was used to show the similarity of the individuals (Jaccard, 1901). Cluster analysis was performed based on the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) (Sneath and Sokal, 1973). The Shannon indices were calculated by the POPGENE 32 software (Yeh et al., 1997).

3. Results

3.1. ISSR-PCR analysis of the two Plasmodium species

Overall, 26 ISSR primers were employed for screening and profiling of the DNA patterns of the two *Plasmodium* species, namely *P. knowlesi* and *P. cynomolgi*. Out of the 26 primers screened, 19 successfully produced polymorphic banding patterns while seven primers, including UBC 807, UBC 827, UBC 834, UBC 848, UBC 861, UBC 874 and UBC 878 produced monomorphic ones. The 19 primers which amplified polymorphic bands generated a total of 153 ISSR loci by an average of 8.2 loci per primer. Of these, a sum of 103 loci was detected as polymorphic loci. The size of the amplified loci ranged from 200 to 1700 bp. (Table 3). The highest polymorphism (90%) was generated by primer UBC 842 and the lowest (37.5%) by UBC 840. The level of polymorphism detected in these two *Plasmodium* species and the Shannon indices for each primer is summarized in Table 3. UBC 815 generated the highest number of loci (14) followed by UBC 844 (12) and UBC 808 and UBC 811 (11). In addition, UBC 824 generated the least number of loci (3). The Shannon indices ranged from 0.1825 ± 0.1066 to 0.6771 ± 0.0086. The highest Shannon index (0.1825 ± 0.1066) was related

to primer UBC 842 as the most-polymorphic primer, and the lowest (0.1825 ± 0.1066) was related to UBC 840 as the least-polymorphic primer (Table 3).

3.2. Genetic diversity of the 20 Plasmodium accessions inferred from the ISSR-based cluster analysis

The ISSR-based cluster analysis of the 20 *Plasmodium* accessions using the UPGMA method resulted in the generation of three clusters shown with red, blue and green colors (Fig. 2). However, the red and blue clusters could be considered as one cluster due to their color maps' patterns, referring to the ISSR profiles of the accessions. In other words, the red and blue clusters were considered the sub-clusters of a single cluster. Accordingly, the first cluster consisted of the eight accessions of *P. cynomolgi* from Penang state with two red and blue sub-clusters. The second cluster (green cluster) was comprised of the 12 *P. knowlesi* accessions, all from Selangor state (Fig. 2). The genetic similarity of the *Plasmodium* accessions was shown by the Jaccard's similarity coefficient, which ranged from 0.25 to 0.81 showing the existence of high diversity. Jaccard's coefficient was perfectly matched to the present data due to excluding the negative co-occurrence (0-0) phenomenon.

Interestingly, the clustering was in a complete agreement with the geographical distribution of the samples as all the individuals belonging to the state of Penang were located in cluster one (sub-cluster red and blue), while the samples of Selangor were unexceptionally placed in the green cluster (Fig. 2). Remarkably, the color map of the dendrogram revealed that the separation of the two *Plasmodium* species happened mainly due to the absence of the amplified loci produced by the UBC 808, UBC 809, UBC 810, UBC 811, UBC 815, UBC 822, UBC 842, UBC 844, UBC 845, UBC 855 and UBC 856 primers in the species *P. knowlesi*. The absence of these loci was indicated as the unified plane green color surface in the color map of Figure 2.

The most significant result of this finding was to reveal the high efficiency of ISSR markers in the discrimination of the two *Plasmodium* species from each other. Moreover, the ISSR-based cluster analysis revealed that despite the high similarity among the *P. cynomolgi* accessions, *P. cynomolgi* 63, 64, 65 and 70 slightly diverged from *P. cynomolgi* 66, 67, 68 and 69.

Another noticeable point was the problematic feature of the *P. knowlesi* species in the amplification of AGA- and CTC-based ISSR markers. The latter case could be taken into further consideration due to its importance in revealing the genomics, as well as the composition of nucleotides of this species.

3.3. ISSR-based PCA analysis of the two Plasmodium species

The principal components analysis of the two *Plasmodium* species using the ISSR data confirmed the cluster analysis results. The PCA graph demonstrated a clear separation between the two species of *Plasmodium* parasites by grouping those eight vectors of *P. cynomolgi* accessions on the right side of the graph, and placing the 12 vectors of *P. knowlesi* accessions on the left side of the graph (Fig. 3). Two accessions *P. cynomolgi* 63 and *P. cynomolgi* 69 were positioned in the upper sites compared to the other members of this species. This situation was similar to the ISSR-based cluster analysis of the *Plasmodium* species, in which these two accessions were clustered as the two most-distinctive accessions of *P. cynomolgi* (Fig. 2).

3.4. ISSR-PCR analysis of the M. fascicularis samples

As mentioned previously, 19 ISSR primers out of 26 were found polymorphic in the infected samples; thus to avoid any differences in the research trend, only the polymorphic primers were used for screening the uninfected monkeys. This strategy was successful, in that all the primers showed polymorphic banding patterns after optimization for the uninfected Macaque samples too. Consequently, a total of 109 ISSR loci by an average of 5.7 loci per primer were generated.

Of these, a sum of 95 loci was detected as polymorphic loci. The size of the amplified ISSR loci ranged between 380 to 1977 bp (Table 4).

The polymorphic banding patterns of the primers in the uninfected long-tailed Macaque samples was totally different from the *Plasmodium* accessions, in that the highest polymorphism (100%) was produced by the UBC 820, UBC 822, UBC 824, UBC 841, UBC 842, UBC 844, UBC 845, UBC 881 primers while the lowest polymorphism was produced by UBC 811 (57.1%) (Table 4). It is worthy to mention that some of these primers such as UBC 822, UBC 842, UBC 844 and UBC 845 did not produce even a single locus in *P. knowlesi* and some of them like UBC 809, UBC 810 and UBC 811 were amongst the least-polymorphic primers in the *Plasmodium* accessions (Table 3). Primer UBC 840 generated the highest number of loci (10) followed by UBC 815 with a total number of nine loci. The least number of loci in the uninfected Macaque samples was generated by UBC 824 (3), UBC 845 (2) and UBC 868 (3) (Table 4). The Shannon indices ranged from 0.3939 ± 0.3596 to 0.6763 ± 0.0000. The highest Shannon index (0.6763 ± 0.0000) was related to primer UBC 844 as the most-polymorphic marker, and the lowest (0.3939 ± 0.3596) was related to UBC 811 as the least-polymorphic marker (Table 4).

3.5. Genetic diversity and ISSR-based cluster analysis of the uninfected M. fascicularis samples

The ISSR-based cluster analysis of the 20 long-tailed uninfected Macaque accessions using the UPGMA method resulted in the generation of three entirely distinct clusters shown with red, blue and green colors (Fig. 4). The Jaccard's similarity coefficient ranged between 0.20 and 1.00. Observing such a wide range in Jaccard's similarity coefficient complied with the existence of a very high genetic variation among the long-tailed Macaque samples.

The red cluster was exclusively grouped the samples from Selangor while the green cluster only consisted of the samples from Penang state. Surprisingly, the blue cluster was a combination of the Selangor and Penang samples. In line with the mixed composition of the blue cluster, the color map also revealed a transitional ISSR pattern for the accessions of this cluster which was an intermediate form of the red and green clusters. Besides, four samples of the red cluster, including *M. fascicularis* P1, P2, P7 and P7 were separated at the sub-cluster level from M. *fascicularis* P3, P4 and P8. This intra-cluster separation was mainly due to the absence of ISSR loci related to the UBC 822 and UBC 842 primers, and the concurrent presence of ISSR loci related to the UBC 824, UBC 825, UBC 845, UBC 855 and UBC 868 primers in the *M. fascicularis* P1, P2, P7 and P7 samples.

Unlike the cluster analysis of the *Plasmodium* accessions, the clustering of the long-tailed Macaques was not in complete agreement with the geographical origins of the samples. The best example of this contradiction was the construction of the blue cluster, where three individuals belonging to the state of Selangor (*M. fascicularis* S3, S8, and S4) were clustered together with individuals from Penang (Fig. 4).

High level of genetic variation and the possibility of intra-specific hybridization among *M. fascicularis* individuals from the northern and southern parts of Peninsular Malaysia were the most important outcomes of the ISSR-based cluster analysis of the uninfected long-tailed Macaques.

3.6. ISSR-based PCA analysis of the uninfected M. fascicularis samples

The principal components analysis of the *M. fascicularis* samples using the ISSR data were in accordance with the cluster analysis results (Fig. 5). However, in the PCA graph the number of the vectors decreased to 12 (instead of 20). This happened due to the position of the monkey

samples in the sub-clusters as shown in Figure 4. Therefore, because of the nature of PCA graph, the two accessions of each sub-cluster were represented as a single accession due to their equal similarity matrices. In another word, the vectors of the similar accessions were outlined parallel.

4. Discussion

4.1. ISSRs filling the current gaps in Plasmodium pathogenesis

As mentioned, simian malaria parasites are readily infectious for long-tailed (*M. fascicularis*) and pig-tailed (*M. nemestrina*) Macaques (Eyles, 1963). On the other hand, the chance of switching the pathogenic species to humans (as their desired host) is not negligible. The above-mentioned situation could be due to the increasing trend of human populations, as well as ecological alterations caused by pollution or deforestation in recent years. Meanwhile, there are gaps and challenges in the current knowledge, especially on epidemiology, pathogenesis of *P. knowlesi* (Singh and Daneshvar, 2013). A very recent study shows that information on the diversity and transmission dynamics of the *P. vivax* population can potentially provide insights into the changing dynamics of decreasing malaria occurrence and the progression of elimination, as well as the effect of imported cases on the likelihood of successful elimination in Central China (Liu et al., 2014). Consequently, investigating the genetic diversification of this zoonotic infection appears to be supportive to prevent and control the prevalence of *P. knowlesi* malaria in Malaysia. However, the species-specific primers have already been developed to detect and differentiate the five types of infectious *Plasmodium* to humans, but studying the genetic variation of these *Plasmodium* species together using any of those specific primers is seriously hampered by their high specificity. Exactly, under such circumstances, the

necessity of exploiting ISSRs as the universal primer pairs to target specific sequences in a wide range of species is a helpful solution (Kumar et al., 2009). Contemporarily, ISSR markers have successfully been used to investigate the genetic diversity of malaria vectors such as *Anopheles annularis* (Das et al., 2014), and *Culex quinquefasciatus* (Mendki et al., 2011), in recent years.

At one glance, the high polymorphic feature of ISSR markers, and their potential in acting as randomly amplified microsatellites (RAMP) make them an intermediate candidate between dominant and co-dominant markers. Although, this in turn will provide concrete evidence to answer a part of ambiguities that lie ahead in studying the genetic variation of *Plasmodium* species. As a further step towards understanding the genetic structure of *Plasmodium* species, ISSR markers can be taken into consideration as a potent candidate. In this regard, the competence of ISSRs has been proved for investigating the genetic structure of both plant (Zimisuhara et al., 2015), and animal species (Shafiei-Astani et al., 2015), in very recent studies.

4.2. ISSR markers, geographic extent and genetic diversity of P. knowlesi and P. cynomolgi infections

Unlike the uninfected individuals, the ISSRs revealed a more specific pattern on the infected samples in which these markers could differentiate the samples based on geographical distribution and the taxonomy of the parasite. It should be noted that the observed introgressions among the uninfected samples should not be attributed to the low performance of the ISSR markers in analysing the uninfected samples. As an explanation, technically, the FTA card-based DNA extraction protocol was mainly optimized to extract the DNA of the parasites; and since the uninfected samples were free of parasites. Accordingly, two different taxa were

subjected to the ISSR analysis in the present study, one of them was a mammalian animal (long-tailed Macaque) capable of relocating up to 10 kilometres in three days (Hasan et al., 2013), and another one was a parasite transmitted by an insect (*Anopheles* ssp.) with a very low range of relocation (Spitzen et al., 2013). Differences in the relocation ranges of these two species can be introduced as a key factor in constructing a different pattern in each group of the monkeys.

Finally, taking the current results into consideration, it can be concluded that the most prevalent *Plasmodium* infection in Selangor is related to *P. knowlesi* species, while the species *P. cynomolgi* is the most prevalent malaria infection in the state of Penang.

4.3. ISSR markers and genetic diversity of M. fascicularis population

The analysis of ISSR data generated two entirely different patterns among the infected and uninfected samples in this study. The uninfected samples were not separated according to their geographical distribution whilst three samples of Selangor state, namely *M. fascicularis* S3 and S8 from Jeram and *M. fascicularis* S4 from Puncak Alam were located in the same cluster as the individuals of Penang. Considering the location of the samples *M. fascicularis* S3 and S8 that both were from the northern part of Selangor state; and, the point that the accession *M. fascicularis* S4 was only 150 Kilometres (direct distance) away, heading south, the positioning of these three samples in a cluster is easily justifiable. On the other hand, the geographical proximity of the *M. fascicularis* S3 and S8 accessions to the southern borders of the Penang state makes it understandable to observe them among the Penang-related clusters; however, only *M. fascicularis* P9 belonged to the mainland of the Penang state (Ladang Byram area),

379 while the P5 and P6 belonged to Tanjung Tokong. This area is a suburb of George Town-
380 Penang Island. Therefore, relying on the geographical proximity and natural relocation of the
381 three P5, P6 and P9 Macaque accessions is not a rational explanation for their genetic similarity
382 together, as well as with the accessions *M. fascicularis* S3, S4 and S8. As a matter of fact, the
383 Penang Island was connected to the mainland of Penang through the man-made structures
384 (bridge) by 1985. Therefore, natural relocation of the Macaques is logically ruled out, and
385 instead, human factor should be considered as an inevitable justification for witnessing such a
386 complex pattern in long-tailed Macaque population using the ISSR-based clustering.
387 Interestingly, a similar pattern has been outlined in a recent SSR-based assessment while the
388 individual assignment test using STRUCTURE analysis revealed an affinity between the
389 Penang and Selangor populations (Nikzad et al., 2014). The source of such a genetic similarity
390 between these geographically distant sites probably is a result of activities back to 40–50 years
391 ago, when local traders transported long-tailed Macaques from Penang into Selangor for
392 exportation (Muda, 1982). Nevertheless, the ISSR markers were successful by grouping both
393 the infected and uninfected individuals.

394 *4.4. Justification of the clustering patterns in Macaque and Plasmodium*

395 Probably, Macaque's high capability in relocation along with their mating habit and sexual
396 desires facilitated the condition for exchanging their genetic materials. A part of these genetic
397 interactions were detected by the ISSR markers and featured as the blue cluster in Figure 4. The
398 mentioned function of the ISSR markers and cluster analysis has been repeatedly utilized to
399 detect and confirm the intra-specific hybridization in various plant and animal species such as
400 *Coffea* spp., sea star and oak (Ruas et al., 2003; Harper and Hart, 2007; Neophytou, 2014). The
401 present study also showed that the ISSR markers are capable of eliciting a higher genetic

variation of the Malaysian long-tailed Macaque more than the SSR markers that were used by Nikzad *et al*. (2014) for studying the genetic diversity of this species in Malaysia.

As mentioned previously, the presence of three anopheline mosquito species; namely, *Anopheles cracens* (Vythilingam et al., 2008), and *A. hackeri* (Wharton and Eyles, 1961), in Peninsular Malaysia, and *A. latens* in Sarawak, Malaysian Borneo, has already been confirmed (Vythilingam et al., 2006; Tan et al., 2008). These three species are capable of transmitting *P. knowlesi* parasite and the results of the present research complied with the high prevalence of two of the species in the peninsula. But it is rational to carry out an in-depth investigation about the prevalence of the other species of mosquito vectors adapted for each geographical part of Peninsular Malaysia.

4.5. Precision of the used PCR methods

Another reason which explains the reliability of our results to a high extent is the PCR method that was used to detect the samples with single infections in the present study. The PCR has provided a great opportunity to design highly sensitive methods of parasite detection (Bottius et al., 1996). The application of the nested PCR methods has improved the efficiency of the assays markedly (Rubio et al., 1999). Nested PCR includes two steps of amplifications in which, the product of the first PCR reaction serves as the template for the second amplification. It can enable a researcher to detect a single parasite genome practically and reproducibly in the DNA template purified from 10 µl of blood sample (0.000002% parasitemia) or a 2.0 mm diameter disc, punched from FTA filter paper (Snounou et al., 2002). It is also suggested to remove white blood cells from blood samples before DNA purification, or prior to apply and preserve them on FTA cards. It makes them even more amenable to nested PCR in the samples with lower parasitemia.

4.6. Nucleotides frequencies in Plasmodium and Macaque genomes

At one glance, the greater number of the total amplified ISSR loci using each ISSR primer in any species would imply that the frequency of the reverse-complementary sequence of that particular primer is predominantly distributed throughout the genomic DNA of that species. In other words, the high-occurrence of a particular ISSR sequence complies with the abundance of their ingredient (polynucleotides) in the genome of the species. The mentioned point could also lead to prediction of the microsatellite motifs in any genome (Ruas et al., 2003; Hu et al., 2003).

According to Blair et al. (1999), if microsatellites were randomly distributed along the genome, the length of the intervening regions between simple sequence repeats of the same motif should be a function of their frequency and of the number of bands produced by an SSR primer with a given microsatellite repeat would reflect the relative frequency of that motif. Since ISSRs act like the random amplified microsatellite polymorphism (RAMP), similar conditions may also prevail for ISSR markers.

Based on the above-mentioned concept, there are some amplified ISSR loci, with a high frequency among the uninfected monkey samples as follows; GAG AGA GAG AGA GAG AC, CTC TCT CTC TCT CTC TYG and TGT GTG TGT GTG TGT GYA. Likewise, the high abundance of GAG AGA GAG AGA GAG AC, GAG AGA GAG AGA GAG ARG, TCT CTC TCT CTC TCT CG and CTC TCT CTC TCT CTC TG can be predicted in the *Plasmodium* parasite's genome.

Undoubtedly, to specify the two *Plasmodium* species to which this prediction may possibly apply (taking note that some species have AT-rich genome, while others are GC-rich), a higher size of sequencing data is required (Chanda et al., 2005; Panneerselvam et al., 2011). Despite,

448 relatively high content of GC has been detected in the genome of *P. cynomolgi* (45%) and *P.*
449 *knowlesi* (40%) (Tachibana et al., 2012).
450
451 **5. Conclusions**
452 Developing potential biomarkers for rapid and reliable detection of malaria has become an
453 undeniable priority (Jain et al., 2014). This study provides information for further
454 investigations to design and develop diagnostic microsatellite markers for malaria infection in
455 the Macaque populations in upcoming researches. Furthermore, the present research will open
456 up a new horizon in this field, in which the dominant DNA markers such as ISSRs can be
457 regarded as an efficient tool with a noticeable potential in differentiating the Plasmodium
458 species, while it has been neglected in all this time. Further investigation using more ISSR
459 markers and bigger numbers of infected and uninfected samples may be beneficial to gain
460 better insights into the relationships between the geographical distributions of the samples in
461 the light of the patterns associated with their genetic diversity.
462 **Competing interests**
463 The authors declare that they have no competing interests.
464 **Authors' contributions**
465 Parastoo Khajeaian, Soon Guan Tan and Alireza Valdiani conceived and designed the study.
466 Soon Guan Tan supervised the project. Parastoo Khajeaian conducted the study and collected
467 data. Alireza Valdiani and Parastoo Khajeaian wrote the manuscript. Alireza Valdiani
468 performed data analyses and interpretation. Christina Yong Seok Yien and Noorjahan Banu
469 Alitheen co-supervised the project. Jeffrine J. Rovie-Rian managed the sampling process.

Mohammad Zareian Jahromi collaborated in the lab works. Alireza Valdiani and Soon Guan Tan performed a final proofreading.

References

Anderson, T.J., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol. Biol. Evol. 17: 1467–1482.

Barr, C.S., Newman, T.K., Becker, M.L., Parker, C.C., Champoux, M., Lesch, K., Goldman, D., Suomi, S., Higley, J., 2003. The utility of the non-human primate model for studying gene by environment interactions in behavioral research. Genes. Brain Behav. 2, 336–340.

Becker, J., Heun, M., 1995. Mapping of digested and undigested random amplified microsatellite polymorphisms in barley. Genome 38: 991–998.

Blair, M., Panaud, O., McCouch, S., 1999. Inter-simple sequence repeats (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). Theor. Appl. Genet. 98, 780–792.

Bottius, E., Guanzirolli, A., Trape, J.F., Rogier, C., Konate, L., Druilhe, P., 1996. Malaria: even more chronic in nature than previously thought; evidence for subpatent parasitaemia detectable by the polymerase chain reaction. Trans. Roy. Soc. Trop. Med. Hyg. 90, 15–19.

Brito, C.F.A.D., Ferreira, M.U., 2011. Molecular markers and genetic diversity of *Plasmodium vivax*. Memórias Instit. Oswaldo Cruz 106, 12–26.

Bronner, U., Divis, P.C., Färnert, A., Singh, B., 2009. Swedish traveller with *Plasmodium knowlesi* malaria after visiting Malaysian Borneo. Malar. J. 8: 15.

Chanda, I., Pan, A., Dutta, C., 2005. Proteome composition in *Plasmodium falciparum*: higher usage of GC-rich nonsynonymous codons in highly expressed genes. J. Mol. Evol. 61, 513–23.

Charters, Y., Wilkinson, M., 2000. The use of self-pollinated progenies as' in-groups' for the genetic characterization of cocoa germplasm. Theor. Appl. Genet. 100, 160–166.

Coatney, G.R., 1971. The simian malarias: zoonoses, anthroponoses, or both? Am. J. Trop. Med. Hyg. 20, 795–803.

Collins, W.E., 2012. *Plasmodium knowlesi*: A malaria parasite of monkeys and humans. Annu. Rev. Entomol. 57, 107–21.

Das, B., Patra, A.P., Das, M., Mahapatra, N., Tripathy, H., Kar, S.K., Hazra, R.K., 2014. Vectorial capacity and genetic diversity of *Anopheles annularis* (Diptera: Culicidae) mosquitoes in Odisha, India from 2009 to 2011. Acta Trop. 137, 130–139.

Ennis, J., Teal, A., Habura, A., Madison-Antenucci, S., Keithly, J., Arguin, P., Barnwell, J., Collins, W., Mali, S., Slutsker, L., 2009. Simian malaria in a US traveler-New York, 2008. Morb. Mort. Wkly. Rep. 58, 229–232.

Eyles, D.E., 1963. The species of simian malaria: Taxonomy, morphology, life cycle, and geographical distribution of the monkey species. J. Parasitol. 49, 866–887.

Figtree, M., Lee, R., Bain, L., Kennedy, T., Mackertich, S., Urban, M., Cheng, Q., Hudson, B.J., 2010. *Plasmodium knowlesi* in human, Indonesian Borneo. Emerg. Infect. Dis. 16, 672–674.

Fooden, J., 1995. Systematic review of Southeast Asian longtail Macaques, *Macaca fascicularis* (Raffles, [1821]). Fieldiana Zool. 81, 1–206.

Galinski, M.R., Barnwell, J.W., 2009. Monkey malaria kills four humans. Trends Parasitol. 25, 200–204.

Gupta, M., Chyi, Y.S., Romero-Severson, J., Owen, J., 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. Theor. Appl. Genet. 89, 998–1006.

Harper, F.M., Hart, M.W., 2007. Morphological and phylogenetic evidence for hybridization and introgression in a sea star secondary contact zone. Invertebr. Biol. 126, 373–384.

Hasan, M.K., Aziz, M.A., Alam, S.R., Kawamoto, Y., Engel, L.J., Kyes, R.C., Akhtar, S., Begum, S., Feeroz, M.M., 2013. Distribution of rhesus Macaques (*Macaca mulatta*) in Bangladesh: inter-population variation in group size and composition. Primate Conserv. 26, 125–132.

Hoosen, A., Shaw, M., 2011. *Plasmodium knowlesi* in a traveller returning to New Zealand. Travel Med. Infect. Dis. 9, 144–148.

Hu, J., Nakatani, M., Lalusin, A.G., Kuranouchi, T., Fujimura, T., 2003. Genetic analysis of sweetpotato and wild relatives using inter-simple sequence repeats (ISSRs). Breeding Sci. 53, 297–304.

Jaccard, P., 1901. A comparative study of the floral distribution in Alps and Jura. Bullet. Walden Soc. Nat. Sci. 37, 547–579.

Jain, P., Chakma, B., Patra, S., Goswami, P., 2014. Potential biomarkers and their applications for rapid and reliable detection of malaria. BioMed Res. Int. doi.org/10.1155/2014/852645

534 Jeslyn, W.P.S., Huat, T.C., Vernon, L., Irene, L.M.Z., Sung, L.K., Jarrod, L.P., Singh, B.,
535 Ching, N.L., 2011. Molecular epidemiological investigation of *Plasmodium knowlesi* in
536 humans and Macaques in Singapore. Vector-Borne. Zoonot. Dis. 11, 131–135.

537 Jiram, A.I., Vythilingam, I., Noorazian, Y.M., Yusof, Y.M., Azahari, A.H., Fong, M.Y., 2012.
538 Entomologic investigation of *Plasmodium knowlesi* vectors in Kuala lipis, Pahang, Malaysia.
539 Malar. J. 11: 213.

540 Joshi, S., Gupta, V., Aggarwal, R., Ranjekar, P., Brar, D., 2000. Genetic diversity and
541 phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in
542 the genus *Oryza*. Theor. Appl. Genet. 100, 1311–1320.

543 Kumar, P., Gupta, V.K., Misra, A.K., Modi, D.R., Pandey, B.K., 2009. Potential of molecular
544 markers in plant biotechnology. Plant Omics 2, 141-162.

545 Lee, K.S., Divis, P.C., Zakaria, S.K., Matusop, A., Julin, R.A., Conway, D.J., Cox-Singh, J.,
546 Singh, B., 2011. *Plasmodium knowlesi*: reservoir hosts and tracking the emergence in humans
547 and Macaques. PLoS Pathog. 7: e1002015. doi: 10.1371/journal.ppat.1002015

548 Liu, Y., Auburn, S., Cao, J., Trimarsanto, H., Zhou, H., Gray, K.A., Clark, T.J., Price, R.N.,
549 Cheng, Q., Huang, R., Gao, Q., 2014. Genetic diversity and population structure of
550 *Plasmodium vivax* in central China. Malar. J. 13:262.

551 Maresso, K., Broeckel, U., 2008. Genotyping platforms for mass-throughput genotyping with
552 SNPs, including human genome-wide scans. Adv. Genet. 60, 107–139.

553 Mendki, M.J., Sharma, A.K., Vijay, V., Agrawal, O.P., Shri, P., Parashar, B.D., 2011. Population
554 genetic structure of *Culex quinquefasciatus* in India by ISSR marker. Asian Pacific J. Trop. Med.
555 2011, 357–362.

556 Muda, H., 1962. Perdagangan primat di Malaysia barat. Jabatan perlindungan hidupan liar
557 semenanjung Malaysia, Kuala Lumpur.

558 Neophytou, C., 2014. Bayesian clustering analyses for genetic assignment and study of
559 hybridization in oaks: effects of asymmetric phylogenies and asymmetric sampling scheme.
560 Tree. Genet. Genomes 10, 273–285.

561 Nikzad, S., Tan, S.G., Yong, S.Y.C., Ng, J., Alitheen, N.B., Khan, R., Rovie-Ryan, J.J.,
562 Valdiani, A., Khajeaian, P., Kanthaswamy, S., 2014. Genetic diversity and population structure
563 of long-tailed Macaque (*Macaca fascicularis*) populations in Peninsular Malaysia. J. Med.
564 Primatol. doi: 10.1111/jmp.12130

Panneerselvam, P., Bawankar, P., Kulkarni, S., Patankar, S., 2011. *In silico* prediction of evolutionarily conserved GC-rich elements associated with antigenic proteins of *Plasmodium falciparum*. Evol. Bioinform. Online 7, 235–255.

Reddy, M.P., Sarla, N., Siddiq, E., 2002. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica 128, 9–17.

Rohlf, F., 2000. NTSYSpc: Numerical Taxonomy System, ver. 2.1. New York: Exeter Publishing.

Ruas, P.M., Ruas, C.F., Rampim, L., Carvalho, V.P., Ruas, E.A., Sera, T., 2003. Genetic relationship in *Coffea* species and parentage determination of interspecific hybrids using ISSR (Inter-Simple Sequence Repeat) markers. Genet. Mol. Biol. 26, 319–327.

Rubio, J., Benito, A., Roche, J., Berzosa, P., Garcia, M., Mico, M., Edu, M., Alvar, J., 1999. Semi-nested, multiplex polymerase chain reaction for detection of human malaria parasites and evidence of Plasmodium vivax infection in Equatorial Guinea. Am. J. Trop. Med. Hyg. 60, 183–187.

SAS Institute Inc., 2009. JMP® 8 User Guide, Second edition. SAS Institute Inc., Cary, NC.

Shafiei-Astani, B., Ong, A.H.K., Valdiani, A., Tan, S.G., Yong, C.Y.S., Ahmady, F., Alitheen, N.B., Ng, W.L., Kuar, T., 2015. Molecular genetic variation and structure of Southeast Asian crocodile (*Tomistoma schlegelii*): comparative potentials of SSRs versus ISSRs. Gene doi: 10.1016/j.gene.2015.06.053

Singh, B., Bobogare, A., Cox-Singh, J., Snounou, G., Abdullah, M.S., Rahman, H.A., 1999. A genus-and species-specific nested polymerase chain reaction malaria detection assay for epidemiologic studies. Am. J. Trop. Med. Hyg. 60, 687–692.

Singh, B., Daneshvar, C., 2013. Human infections and detection of *Plasmodium knowlesi*. Clin. Microbiol. Rev. 28, 165–184.

Singh, B., Sung, L.K., Matusop, A., Radhakrishnan, A., Shamsul, S.S., Cox-Singh, J., Thomas, A., Conway, D.J., 2004. A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. Lancet 363, 1017–1024.

Sneath, P.H., Sokal, R.R., 1973. Numerical taxonomy. The principles and practice of numerical classification. H. Freeman and Co., San Francisco.

Snounou, G., Singh, B., 2002. Nested PCR analysis of *Plasmodium* parasites. Methods Mol. Med. 72, 189–203.

Spitzen, J., Spoor, C.W., Grieco, F., ter Braak, C., Beeuwkes, J., van Brugge, S.P., Kranenbarg, S., Noldus, L.P., van Leeuwen, J.L., Takken, W., 2013. A 3D analysis of flight behavior of

598 *Anopheles gambiae* sensu stricto malaria mosquitoes in response to human odor and heat. PLoS
599 ONE 8: e62995. doi: 10.1371/journal.pone.0062995

600 Su, X., Hayton, K., Wellems, T.E., 2007. Genetic linkage and association analyses for trait
601 mapping in *Plasmodium falciparum*. Nat. Rev. Genet. 8, 497–506.

602 Su, X., Kirkman, L.A., Fujioka, H., Wellems, T.E., 1997. Complex polymorphisms in an
603 approximately 330 kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast
604 Asia and Africa. Cell 91, 593–603.

605 Su, X., Wellems, TE., 1996. Toward a high-resolution *Plasmodium falciparum* linkage map:
606 Polymorphic markers from hundreds of simple sequence repeats. Genomics 33, 430–444.

607 Ta, T.H., Hisam, S., Lanza, M., Jiram, A.I., Ismail, N., Rubio, J.M., 2014. First case of a
608 naturally acquired human infection with *Plasmodium cynomolgi*. Malar. J. 13, 68.

609 Tachibana, S.I., Sullivan, S.A., Kawai, S., Nakamura, S., Kim, H.R., et al. 2012. *Plasmodium*
610 *cynomolgi* genome sequences provide insight into *Plasmodium vivax* and the monkey malaria
611 clade. Nat. Genet. 44, 1051–1055.

612 Tan, C.H., Vythilingam, I., Matusop, A., Chan, S.T., Singh, B., 2008. Bionomics of *Anopheles*
613 *latens* in Kapit, Sarawak, Malaysian Borneo in relation to the transmission of zoonotic simian
614 malaria parasite *Plasmodium knowlesi*. Malar. J. 7, 52.

615 Tang, T.H.T., Salas, A., Ali-Tammam, M., Martínez, M.C., Lanza, M., Arroyo, E., Rubio, J.M.,
616 2010. First case of detection of *Plasmodium knowlesi* in Spain by Real Time PCR in a traveller
617 from Southeast Asia. Malar. J. 9, 219.

618 Valdiani, A., Javanmard, A., Talei, D., Tan, S.G., Nikzad, S., Kadir, M.A., Abdullah, S.N.A.,
619 2013. Microsatellite-based evidences of genetic bottlenecks in the cryptic species
620 "*Andrographis paniculata* Nees": a potential anticancer agent. Mol. Biol. Rep. 40, 1775–1784.

621 Valdiani, A., Kadir, M.A., Saad, M.S., Talei, D., Tan, S.G., 2012. Intra-specific hybridization:
622 Generator of genetic diversification and heterosis in *Andrographis paniculata* Nees. A bridge
623 from extinction to survival. Gene 505, 23–36.

624 Valdiani, A., Talei, D., Javanmard, A., Tan, S.G., Kadir, M.A., Maziah, M., 2014. Morpho-
625 molecular analysis as a prognostic model for repulsive feedback of the medicinal plant
626 "*Andrographis paniculata*" to allogamy. Gene 542, 156–167.

627 Vythilingam, I., Noorazian, Y.M., Huat, T.C., Jiram, A.I., Yusri, Y.M., Azahari, A.H.,
628 NorParina, I., NoorRain, A., Lokman, H.S., 2008. *Plasmodium knowlesi* in humans, Macaques
629 and mosquitoes in peninsular Malaysia. Parasite Vectors 1, 26.

630 Vythilingam, I., Tan, C., Asmad, M., Chan, S., Lee, K., Singh, B., 2006. Natural transmission
631 of *Plasmodium knowlesi* to humans by *Anopheles latens* in Sarawak, Malaysia. Trans. Roy.
632 Soc. Trop. Med. Hyg. 100, 1087–1088.

633 Vythilingam, I., 2010. Review paper *Plasmodium knowlesi* in humans: A review on the role of
634 its vectors in Malaysia. Trop. Biomed. 27, 1–12.

635 Wellems, T.E., Panton, L.J., Gluzman, I.Y., Do Rosario, V.E., Gwadz, R.W., Walker-Jonah,
636 A., Krogstad, D.J., 1990. Chloroquine resistance not linked to mdr-like genes in a *Plasmodium*
637 *falciparum* cross. Nature 345, 253–355.

638 Wharton, R., Eyles, D.E., 1961. *Anopheles hackeri*, a vector of *Plasmodium knowlesi* in
639 Malaya. Science 134, 279–280.

640 W.H.O., 2013. Malaria entomology and vector control - Guide for tutors. *World Health*
641 *Organization*. Geneva, Switzerland.

642 Wu, K.S., Jones, R., Danneberger, L., Scolnik, P.A., 1994. Detection of microsatellite
643 polymorphisms without cloning. Nucleic Acids Res. 22, 3257–3258.

644 Yeh, F.C., Yang, R.C., Boyle, T.B., Ye, Z., Mao, J.X., 1997. POPGENE, the user-friendly
645 shareware for population genetic analysis. Molecular Biology and Biotechnology centre,
646 University of Alberta, Canada.

647 Zietkiewicz, E., Rafalski, A., Labuda, D., 1994. Genome fingerprinting by simple sequence
648 repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20, 176–183.

649 Zimisuhara, B., Valdiani, A., Shaharuddin, N.A., Qamaruz-Zaman, F., Maziah, M., 2015.
650 Structure and principal components analyses revealed an intervarietal fusion in Malaysian
651 mistletoe fig (*Ficus deltoidea* Jack) populations. Int. J. Mol. Sci. 16, 14369-14394.
652 doi:10.3390/ijms160714369

653

654

655

656

657

658

Figure legends:

Fig. 1. Sampling sites of the long-tailed Macaque (*M. fascicularis*) in Peninsular Malaysia.

Fig. 2. ISSR-based dendrogram of the *Plasmodium* accessions generated by the UPGMA clustering method. The green and yellow squares of the colour map indicate the absence (0) and presence (1) of the ISSR loci in each sample, respectively. The genetic similarity matrices are shown by Jaccard's similarity coefficient.

Fig. 3. Three-dimensional ISSR-based PCA graph of the *Plasmodium* accessions. The PCA shows a clear discrimination between the two species of *Plasmodium* parasite by grouping the eight vectors of *P. cynomolgi* on the right side of the graph, and replacing the 12 *P. knowlesi* on the left side of the graph.

Fig. 4. ISSR-based dendrogram of the uninfected *Macaca fascicularis* accessions generated by the UPGMA clustering method. The blue and yellow squares of the colour map indicate the absence (0) and presence (1) of the ISSR loci in each sample, respectively. The genetic similarity matrices are shown by Jaccard's similarity coefficient.

Fig. 5. Three-dimensional ISSR-based PCA graph of the uninfected *Macaca fascicularis* accessions. The present PCA has decreased the number of the vectors to 12. This was due to the position of the monkey accessions in the sub-clusters shown in Figure 3. Therefore, because of the nature of this PCA graph, the two accessions of each sub-cluster were represented as one accession due to their equal similarity matrices.

703

704

705 **Fig. 1.**

706

707

708

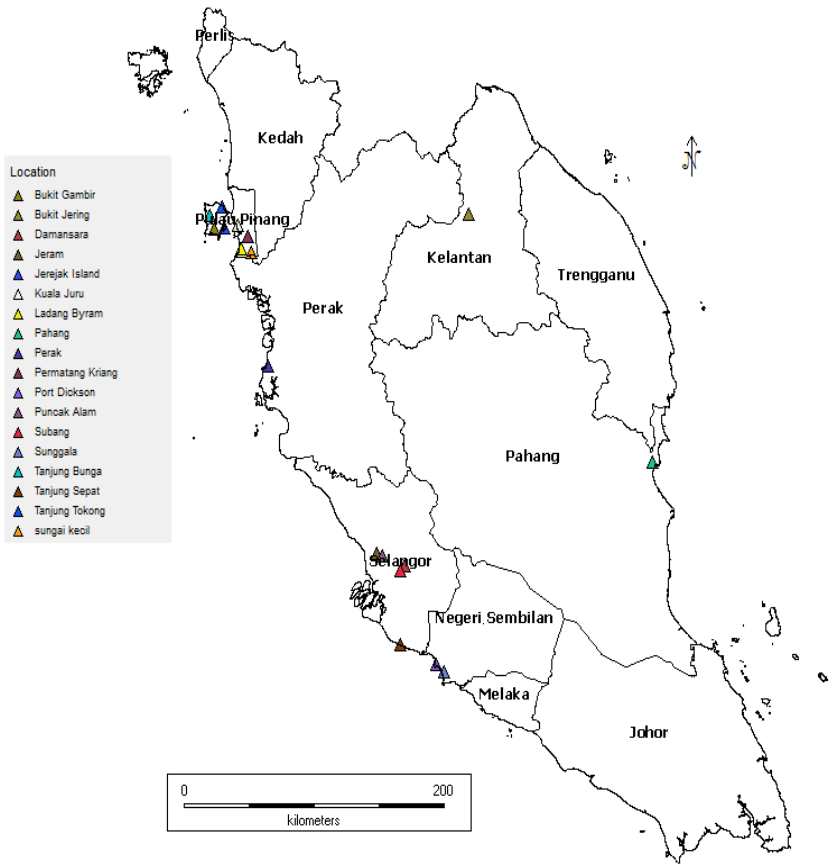


Fig. 2.

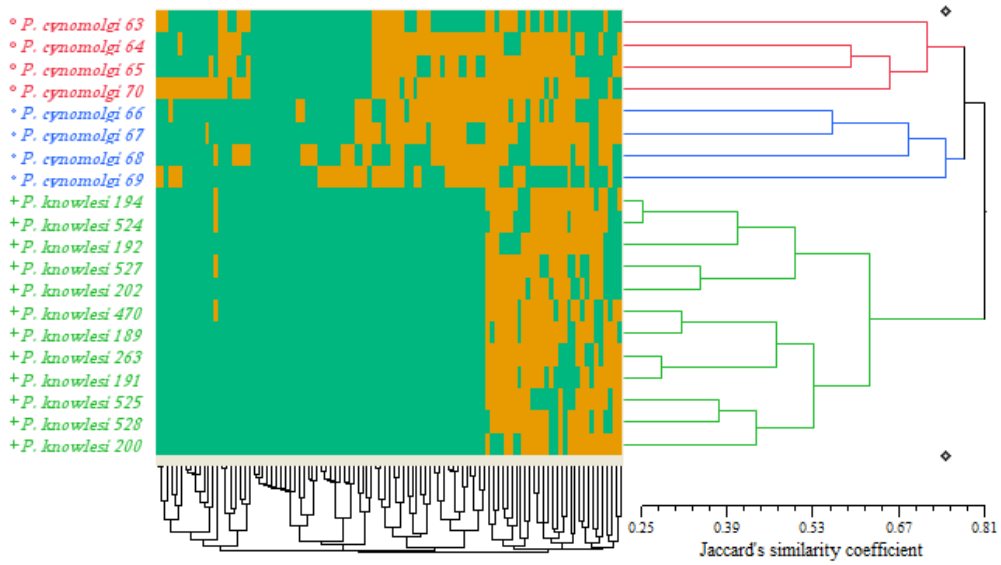
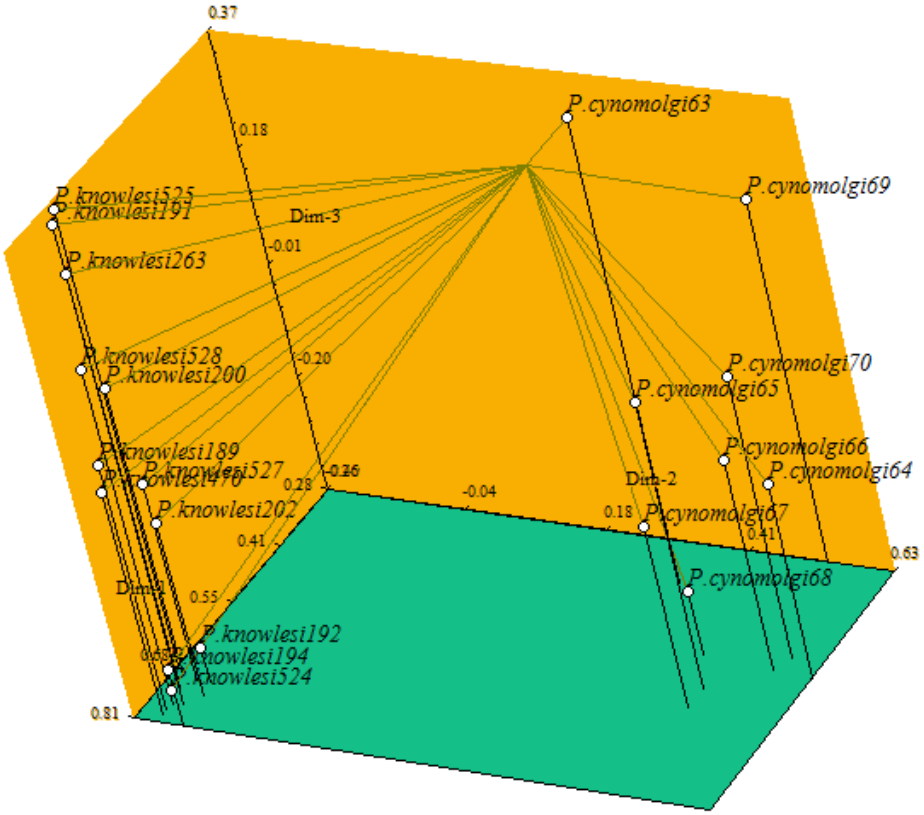


Fig. 3.



748

749

750 **Fig. 4.**

751

752 753 754 755 756 757 758

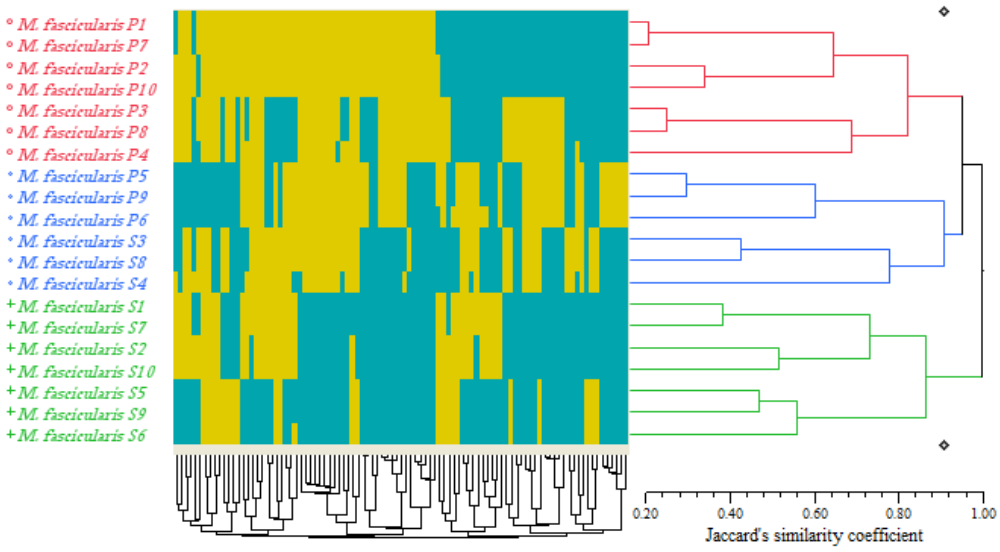


Fig. 5.

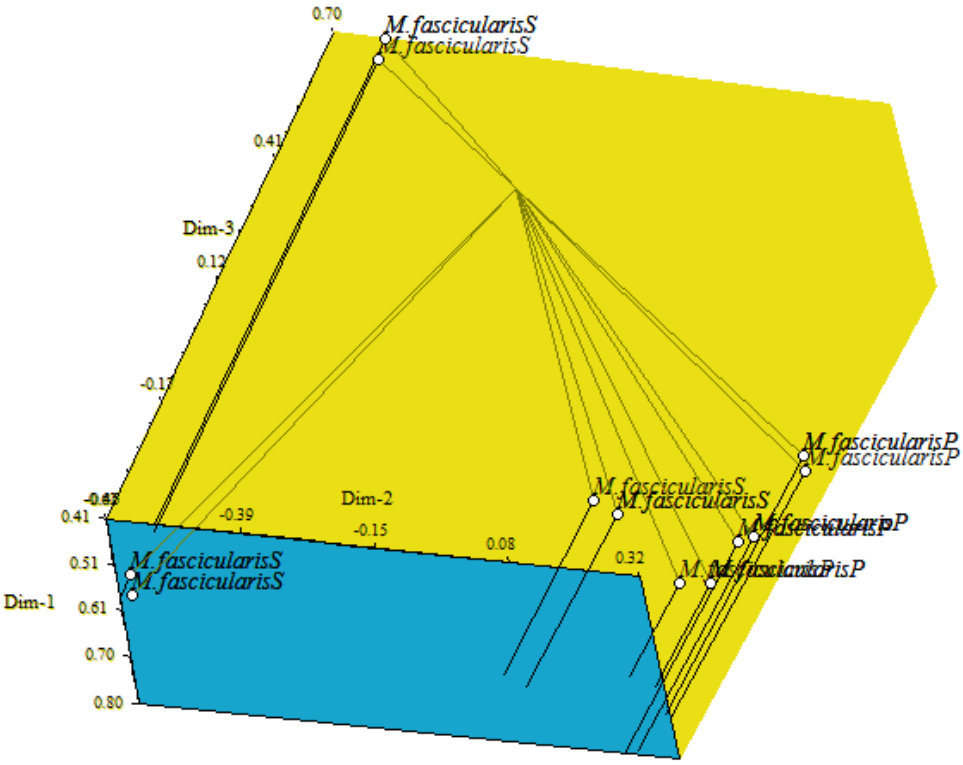


Table 1

Geographical distribution, gender and age of the *Plasmodium*-Infected long-tailed Macaque samples (*M. fascicularis*)

Code	Capture location	Age	Gender	*P. knowlesi*	*P. inuii*	*P. cynomolgi*	*P.fieldi*	*P.coatneyi*
P. knowlesi 524	Kampung Seri Aman, Puchong, Selangor	J	M	1	0	0	0	0
P. knowlesi 525	Kampung Seri Aman, Puchong, Selangor	J	F	1	0	0	0	0
P. knowlesi 527	Kampung Seri Aman, Puchong, Selangor	A	F	1	0	0	0	0
P. knowlesi 528	Kampung Seri Aman, Puchong, Selangor	S	F	1	0	0	0	0
P. knowlesi 189	Taman Bukit Saga Subang, Selangor	A	M	1	0	0	0	0
P. knowlesi 191	Taman Bukit Saga Subang, Selangor	S	M	1	0	0	0	0
P. knowlesi 192	Taman Bukit Saga Subang, Selangor	S	M	1	0	0	0	0
P. knowlesi 200	Jeram, Selangor	J	F	1	0	0	0	0
P. knowlesi 202	Jeram, Selangor	A	F	1	0	0	0	0
P. knowlesi 194	Taman Bukit Saga Subang, Selangor	J	M	1	0	0	0	0
P. knowlesi 263	Damansara, Selangor	A	M	1	0	0	0	0
P. knowlesi 470	Kem Sg. Buloh Selangor	A	M	1	0	0	0	0
P. cynomolgi 63	Surau MK 2, Tg. Bungah, Pulau Pinang, Penang	S	M	0	0	1	0	0
P. cynomolgi 64	Surau MK 2, Tg. Bungah, Pulau Pinang, Penang	J	M	0	0	1	0	0
P. cynomolgi 65	Surau MK 2, Tg. Bungah, Pulau Pinang, Penang	J	M	0	0	1	0	0
P. cynomolgi 66	Surau MK 2, Tg. Bungah, Pulau Pinang, Penang	J	F	0	0	1	0	0
P. cynomolgi 67	Surau MK 2, Tg. Bungah, Pulau Pinang, Penang	J	F	0	0	1	0	0
P. cynomolgi 68	Bkt. Gambir, Gelugor, Pulau Pinang, Penang	A	M	0	0	1	0	0
P. cynomolgi 69	Bkt. Gambir, Gelugor, Pulau Pinang, Penang	J	M	0	0	1	0	0
P. cynomolgi 70	Bkt. Gambir, Gelugor, Pulau Pinang, Penang	J	M	0	0	1	0	0

J: Juvenile, A: Adult, and S: Subadult, 0 and 1: negative and positive infection, respectively.

Table 2

Geographical distribution, gender and age of the uninfected long-tailed Macaque samples (*M. fascicularis*)

Code	Capture location	Age	Gender	*P. knowlesi*	*P. inuii*	*P. cynomolgi*	*P.field i*	*P.coatneyi*
M. fascicularis P1	Jerejak Rainforest Resort, Pulau Jerejak, Pulau Pinang, Penang	A	M	0	0	0	0	0
M. fascicularis P2	Jerejak Rainforest Resort, Pulau Jerejak, Pulau Pinang, Penang	A	M	0	0	0	0	0
M. fascicularis P3	Jerejak Rainforest Resort, Pulau Jerejak, Pulau Pinang, Penang	A	M	0	0	0	0	0
M. fascicularis P4	Tanjung Tokong, Penang	S	M	0	0	0	0	0
M. fascicularis P5	Tanjung Tokong, Penang	A	F	0	0	0	0	0
M. fascicularis P6	Tanjung Tokong, Penang	J	M	0	0	0	0	0
M. fascicularis P7	Tanjung Tokong, Penang	J	M	0	0	0	0	0
M. fascicularis P8	Sg. Kecil, Seberang Perai, Pulau Pinang, Penang	A	F	0	0	0	0	0
M. fascicularis P9	Ladang Byram, Seberang Perai, Pulau Pinang, Penang	A	F	0	0	0	0	0
M. fascicularis P10	Ladang Byram, Seberang Perai, Pulau Pinang, Penang	A	F	0	0	0	0	0
M. fascicularis S1	Bukit Kerayong Jeram, Selangor	A	M	0	0	0	0	0
M. fascicularis S2	Sg. Kandis, Selangor	J	F	0	0	0	0	0
M. fascicularis S3	Jeram, Selangor	A	M	0	0	0	0	0
M. fascicularis S4	Puncak Alam, Selangor	A	F	0	0	0	0	0
M. fascicularis S5	Tanjung Sepat, Selangor	A	M	0	0	0	0	0
M. fascicularis S6	Tanjung Sepat, Selangor	J	M	0	0	0	0	0
M. fascicularis S7	Tanjung Sepat, Selangor	A	M	0	0	0	0	0
M. fascicularis S8	Jeram, Selangor	A	F	0	0	0	0	0
M. fascicularis S9	Sg. Kandis, Selangor	A	F	0	0	0	0	0
M. fascicularis S10	Bukit Kerayong Jeram, Selangor	S	F	0	0	0	0	0

J: Juvenile, A: Adult, and S: Subadult 0: negative *Plasmodium* infection

785 **Table 3**

786 Polymorphic content of the ISSR markers in *Plasmodium cynomolgi* and *Plasmodium knowlesi*

Primer	Sequence 5′→3′	T_a (˚C)	Size (bp)	TL	TPL	Shannon indices	Polymorphism (%)
UBC 808	AGA GAG AGA GAG AGA GC	56	470-1170	11	9	0.6496 ± 0.0334	81.8
UBC 809	AGA GAG AGA GAG AGA GG	55	200-540	7	4	0.3041 ± 0.0853	57.1
UBC 810	GAG AGA GAG AGA GAG AT	56	780-1300	6	4	0.3276 ± 0.1682	66.6
UBC 811	GAG AGA GAG AGA GAG AC	55	300-1200	11	5	0.2466 ± 0.1400	45.4
UBC 815	CTC TCT CTC TCT CTC TG	55	400-1700	14	9	0.3238 ± 0.1716	64.2
UBC 818	CAC ACA CAC ACA CAC AG	54	200-550	9	5	0.2963 ± 0.0724	55.5
UBC 820	GTG TGT GTG TGT GTG TC	59	240-620	4	3	0.5607 ± 0.1634	75.0
UBC 822	TCT CTC TCT CTC TCT CA	54	250-680	9	6	0.3668 ± 0.1372	66.6
UBC 824	TCT CTC TCT CTC TCT CG	55	350-700	3	2	0.3358 ± 0.1410	66.6
UBC 825	ACA CAC ACA CAC ACA CT	58	250-630	9	5	0.3019 ± 0.1592	55.5
UBC 840	GAG AGA GAG AGA GAG AYT	54	280-900	8	3	0.1825 ± 0.1066	37.5
UBC 841	GAG AGA GAG AGA GAG AYC	54	420-1200	9	6	0.3828 ± 0.1256	66.6
UBC 842	GAG AGA GAG AGA GAG AYG	56	250-1480	10	9	0.6771 ± 0.0086	90.0
UBC 844	CTC TCT CTC TCT CTC TRC	56	300-820	12	10	0.6625 ± 0.0511	83.3
UBC 845	CTC TCT CTC TCT CTC TRG	56	300-750	9	8	0.6440 ± 0.0191	88.8
UBC 855	ACA CAC ACA CAC ACA CYT	60	500-1150	5	4	0.6362 ± 0.0490	80.0
UBC 856	ACA CAC ACA CAC ACA CY	60	500-1200	5	4	0.6377 ± 0.0482	80.0
UBC 868	GAA GAA GAA GAA GAA GAA	52	230-600	6	3	0.2830 ± 0.1456	50.0
UBC 881	GGG TGG GGT GGG GTG	60	400-600	6	4	0.3304 ± 0.0050	66.6
Total / Mean	–	–	348.4-936.3	153 / 8.1	103 / 5.4	0.4289 ± 0.0963	67.2

787 T_a: Annealing temperature, TL: total number of loci, TPL: total number of polymorphic loci, Y: C/T, and R: A/G.

Table 4

Polymorphic content of the ISSR markers in long-tailed Macaque (*Macaca fascicularis*)

Primer	Sequence 5′→3′	T_a (˚C)	Size (bp)	TL	TPL	Shannon indices	Polymorphism (%)
UBC 808	AGA GAG AGA GAG AGA GC	56	650-900	4	3	0.4710 ± 0.2675	75.0
UBC 809	AGA GAG AGA GAG AGA GG	55	517-1100	5	4	0.5008 ± 0.0464	80.0
UBC 810	GAG AGA GAG AGA GAG AT	56	517-1185	6	5	0.5366 ± 0.0000	83.3
UBC 811	GAG AGA GAG AGA GAG AC	55	930-1200	7	4	0.3939 ± 0.3596	57.1
UBC 815	CTC TCT CTC TCT CTC TG	55	520-1977	10	9	0.5686 ± 0.0310	90.0
UBC 818	CAC ACA CAC ACA CAC AG	54	900-1100	4	3	0.4131 ± 0.1775	75.0
UBC 820	GTG TGT GTG TGT GTG TC	59	680-1210	4	4	0.5795 ± 0.1595	100
UBC 822	TCT CTC TCT CTC TCT CA	54	570-1300	4	4	0.5785 ± 0.0839	100
UBC 824	TCT CTC TCT CTC TCT CG	55	760-1517	3	3	0.5874 ± 0.0113	100
UBC 825	ACA CAC ACA CAC ACA CT	58	600-1210	6	5	0.5366 ± 0.1131	83.3
UBC 840	GAG AGA GAG AGA GAG AYT	54	400-1160	11	10	0.5697 ± 0.1131	90.9
UBC 841	GAG AGA GAG AGA GAG AYC	54	490-1020	8	8	0.5859 ± 0.3514	100
UBC 842	GAG AGA GAG AGA GAG AYG	56	600-1100	5	5	0.5868 ± 0.2301	100
UBC 844	CTC TCT CTC TCT CTC TRC	56	1000-1600	3	3	0.6763 ± 0.0000	100
UBC 845	CTC TCT CTC TCT CTC TRG	56	485-1210	2	2	0.6228 ± 0.1426	100
UBC 855	ACA CAC ACA CAC ACA CYT	60	380-1350	9	8	0.5559 ± 0.0377	88.8
UBC 856	ACA CAC ACA CAC ACA CY	60	620-1600	8	6	0.4398 ± 0.1262	75.0
UBC 868	GAA GAA GAA GAA GAA GAA	52	620-760	3	2	0.4024 ± 0.1712	66.6
UBC 881	GGG TGG GGT GGG GTG	60	810-1340	7	7	0.6624 ± 0.0549	100
Total / Mean	–	–	627.8-1240.1	109 / 5.7	95 / 5.0	0.5404 ± 0.1303	87.6

T_a: Annealing temperature, TL: total number of loci, and TPL: total number of polymorphic loci, Y: C/T, and R: A/G.