Genetic diversity of *Macaca fascicularis* (Cercopithecidae) from Penang, Malaysia as inferred from mitochondrial control region segment

J.J. Rovie-Ryan[1], M. T. Abdullah[2,3], F. T. Sitam[1], S. G. Tan[4], Z. Z. Zainuddin[5], M. M. Basir[1,6], Z. Z. Abidin[1,6], C. Keliang[1], A. Denel[1], E. Joeneh[1] and **F. M. Ali[1]**

[1]Outbreak Response Team, Wildlife Genetic Resource Bank Laboratory, Ex-Situ Conservation Division, Department of Wildlife and National Parks Peninsular Malaysia, Kuala Lumpur, Malaysia.
[2]Kenyir Ecosystem Research Centre, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia..
[3]Department of Zoology, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, Malaysia.
[4]Dept. of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, Malaysia.
[5]Borneo Rhino Alliance, Sabah, Malaysia.
[6]Deputy Director General I and II Office, DWNP, KM 10 Cheras Road, 56100 Kuala Lumpur.

Corresponding author: J.J. Rovie-Ryan, email: *jeffrine@wildlife.gov.my*

Abstrak

Keanekaragaman genetic kera ekor panjang (*Macaca fascicularis*) dari Penang, termasuk Pulau Jerejak dan daratan utama Negeri Penang Malaysia telah dianalisis dengan menggunakan 1.042 bp control region (CR) segment DNA mitochondria (mtDNA). Dua puluh haplotipe menunjukkan adanya satu haplotipe tunggal yang sama antara daratan utama dan pulau, hal ini menandakan bahwa ini merupakan genetic yang diterima dari daratan. Dibandingkan dengan penelitian sebelumnya yang dilakukan berdasarkan CR, semua haplotipe dari Penang merupakan gambaran baru dan tidak ada yang sama dengan populasi *M. fascicularis* lainnya di wilayah ini. Adanya satu deletion mutasi unik pada contoh dari penang (Kelompok I dan II) bisa menjadi indicator yang baik untuk upaya konservasi keunikan genetic dan mungkin bisa dikelola sebagai satu unit pengelolaan. Sebuah ringkasan pohon filogenetik (NJ, MP, ML dan Bayesian) mendukung pengelompokan monofiletik dari *M. fascicularis* seperti digambarkan pada penelitian penelitian sebelumnya. Pemisahan topologi dari haplotype Penang kedalam tiga kelompok utama secara umum berhubungan dengan distribusi geografis mereka. Penelitian ini juga mencatat bahawa haplotipe Penang memiliki garis keturunan dari wilayah continental yang telah terpisah dari garis keturunan insular sekitar 1.04 juta tahun yang lalu. Penelitian ini juga menunjukkan bahwa CR dari mtDNA sangat baik digunakan untuk mengkuantifikasi keanekaragaman genetic intraspesifik pada *M. fascicularis*.

Abtract

The genetic diversity of the long-tailed macaques (*Macaca fascicularis*) from Penang, Malaysia, including Jerejak Island and the mainland area of the state of Penang, Seberang Perai were examined using 1,042 bp control region (CR) segment of the mitochondrial DNA (mtDNA). Twenty haplotypes were described with a single haplotype sharing between the mainland and the island which suggests that it is a remnant of the genetic makeup from the mainland. Compared to previous studies based on the CR, all the Penang haplotypes are newly described with none shared with the other regional populations of *M. fascicularis*. A single deletion mutation unique to the Penang samples (Groups I and II) could be a good indicator for the conservation of the genetic uniqueness and possibly should be managed as a management unit (MU). A summarised phylogenetic tree (NJ, MP, ML and Bayesian) supports the monophyletic clustering of the *M. fascicularis* as described in previous studies. The topology separates the Penang haplotypes into three major groups, which generally corresponds to their geographical distribution. We also noted that the Penang haplotypes are of the continental lineage which separated from the insular lineage at around 1.04 mya. Finally, we showed that the CR of the mtDNA is powerful and suitable for the quantification of intraspecific diversity in *M. fascicularis*.

Keywords: *Macaca fascicularis*, Penang Island, phylogenetics, hypervariable segments I and II

Introduction

The long-tailed macaque, *Macaca fascicularis*, also known as the cynomolgus macaques, are widely distributed in nature and occupies vast areas of mainland southeast Asia (Thailand, Cambodia, Vietnam, Laos, Myanmar, Peninsular Malaysia and Singapore) and the Greater and Lesser Sunda Islands (Indonesia, Brunei, and the Malaysian Borneo) and the Philippines (Eudey, 2008; Fooden, 1995). They can be found almost everywhere especially at low elevations preferring the mangrove and swampy forests, river banks, and

Received 26th February 2014; Revision accepted 2nd May, 2014

seashores (Eudey, 2008). In Peninsular Malaysia they are very common and populate areas in sympatry with the human settlements [Department of Wildlife and National Parks (DWNP), 2006]. In Penang particularly, *M. fascicularis* is distributed throughout the islands (Penang Island and smaller adjacent islands including Jerejak Island) and the mainland portion of Seberang Perai (DWNP, 2006; Karimullah & Shahrul, 2011).

Historically, Peninsular Malaysia was part of the Sundaland which was heavily influenced by events during the Quaternary Period (Pliocene and Pleistocene) (Voris, 2000). During the Pleistocene, periods of intermittent glacial caused the fluctuation of sea levels and at its maximum fell by 120 m below present-day levels and landmasses (Malay Peninsula, Borneo, Sumatra, Java, Bali, Palawan, the Mentawai Islands, and the smaller intervening islands), which are currently separated were joined and formed Sundaland (Bird et al., 2005; Harrison et al., 2006; Sathimurthy & Voris, 2006; Voris, 2000). Consequently, Penang Island which is currently situated about four km off the coast of Seberang Perai and separated by the narrow Penang Straits at a maximum depth of 20 m (Asadpour et al., 2011) would experience repeated connection to the mainland during Pleistocene, thus limiting faunal interchanges.

In Malaysia, very few genetic studies had been conducted on *M. fascicularis*. Most studies were conducted to investigate their conflict with humans (DWNP, 2006), association with zoonotic diseases (Cox-Singh & Singh, 2008; Thayaparan et al., 2013, 2014), distribution (DWNP, 2006; DWNP, unpublished data; Karimullah & Shahrul, 2011), and behaviour (Ling, 2006; Ping, 2003; Shuan, 2006). In other regional populations of *M. fascicularis* (Indochinese, Indonesian, Philippines, Singapore, and Mauritius) however, numerous genetic studies have been conducted using mitochondrial DNA (mtDNA) markers (Blancher et al., 2008; Harihara et al., 1988; Kawamoto et al., 2008; Lawler et al., 1995; Perwitasari-Farajallah et al., 1999; Perwitasari-Farajallah et al., 2001; Schillaci et al., 2011; Shiina et al., 2010; Smith et al., 2007; Tosi & Coke, 2007; Tosi et al., 2002). In this study, we employ the mtDNA control region (CR) as part of a major research initiative by the DWNP to comprehensively examine the population genetics, phylogeography, and the diseases associated with *M. fascicularis* in Malaysia (both from Peninsular Malaysia and from Sarawak and Sabah states on Borneo Island). In summary, this pilot study is designed to achieve these objectives; (1) to examine the genetic diversity of the Penang *M. fascicularis* and (2) to investigate the efficiency of the mtDNA control region in assessing the genetic diversity of *M. fascicularis* in Malaysia.

Methods

Sample Collection

Sampling was conducted by the Outbreak Response Team (ORT) of the DWNP on conflict long-tailed macaques as part of a Wildlife Disease Surveillance Programme (WDSP) launched by DWNP in 2011 to monitor the emergence of zoonotic diseases in wildlife species. Figure 1 shows the sampling sites (A-L) while Table 1 provides the details of each of the samples used in this study. All samples were kept at the Wildlife Genetic Resource Bank (WGRB) Laboratory, DWNP.

DNA Extraction, PCR Amplification and Sequencing

Total genomic DNA was extracted from 46 samples either from blood (preserved in lysis buffer) or from liver samples (see Table 1) using the QIAamp DNeasy Blood and Tissue Kit (QIAGEN Ag., Germany) protocol for blood and tissue samples as provided by the manufacturer. A pair of oligonucleotides; WGRB/MFCR/F15978 (5'–ACCACCAACACCCAAAGCTGGC–3') and WGRB/MFCR/R580(5'-TCAGTGTCTTGCTTTGGGTGGGT-3'), were designed using the program Primer3 (Rozen & Skaletsky, 2000) as a plug-in in the Geneious v5.6 (Drummond et al., 2012) to cover the complete length of the CR segment. Amplifications were carried out in an Arktik Thermal Cycler (Thermo Scientific, USA), using a 15 µl reaction volume consisting of 0.5 µl of DNA template (~15–20 ng), 0.2 µl (0.13 µM) of each primer and 14.5 µl of GoTaq® Colorless Master Mix (Promega, USA). Amplification was done using the following PCR profile: a preliminary denaturation at 98°C for 2 min followed by 30 cycles of 95°C for 30 sec, 69°C for 30 sec and 72°C for 40 sec. This was followed by a final extension period of 72°C for 3 min before the samples were cooled to 4°C. Cycle sequencing on both primers were done on an ABI PRISM®377 DNA Sequencer by a sequencing service provider (1st Base Laboratories Sdn. Bhd., Malaysia).

Sequence Analysis

Multiple alignments of the sequences were done and ambiguous flanking regions were identified and removed

Table 1. Detailed information on each *M. fascicularis* samples collected from the Penang State.

No	Sample Label	Location/GPS of capture	Sex	Map	Sample type**	Haplotype Designation	Haplotype Frequency	GenBank Acc. No.
1	WDSP/11/0037	Sg. Kecil, Seberang Perai, Penang	F	L	BLB	20	0.130	JX113316
2	WDSP/11/0038	Sg. Kecil, Seberang Perai, Penang	F	L	BLB	20	0.130	JX113317
3	WDSP/11/0039	Kuala Juru, Seberang Perai, Penang	M	J	BLB	9	0.022	JX113318
4	WDSP/11/0040	Kuala Juru, Seberang Perai, Penang	M	J	BLB	5	0.022	JX113319
5	WDSP/11/0041	Permatang Kriang, Seberang Perai, Penang	M	I	BLB	4	0.043	JX113320
6	WDSP/11/0042	Permatang Kriang, Seberang Perai, Penang	M	I	BLB	19	0.022	JX113321
7	WDSP/11/0043	Ladang Byram, Seberang Perai, Penang	F	K	BLB	20	0.130	JX113322
8	WDSP/11/0044	Ladang Byram, Seberang Perai, Penang	F	K	BLB	20	0.130	JX113323
9	WDSP/11/0045	Ladang Byram, Seberang Perai, Penang	F	K	BLB	20	0.130	JX113324
10	WDSP/11/0046	Ladang Byram, Seberang Perai, Penang	F	K	BLB	20	0.130	JX113325
11	WDSP/11/0047	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	12	0.065	JX113326
12	WDSP/11/0048	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	12	0.065	JX113327
13	WDSP/11/0049	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	11	0.022	JX113328
14	WDSP/11/0050	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	10	0.087	JX113329
15	WDSP/11/0051	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	10	0.087	JX113330
16	WDSP/11/0052	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	12	0.065	JX113331
17	WDSP/11/0053	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	10	0.087	JX113332
18	WDSP/11/0054	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	10	0.087	JX113333
19	WDSP/11/0055	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	15	0.022	JX113334
20	WDSP/11/0056	Jerejak Rainforest Resort, Jerejak Island	F	H	BLB	13	0.022	JX113335
21	WDSP/11/0057	Jerejak Rainforest Resort, Jerejak Island	M	H	BLB	16	0.022	JX113336
22	WDSP/11/0058	Sg. Baru, Teluk Kumbar, Penang Island	M	G	BLB	17	0.043	JX113337
23	WDSP/11/0059	Sg. Baru, Teluk Kumbar, Penang Island	F	G	BLB	17	0.043	JX113338
24	WDSP/11/0060	Surau MK 2, Tg. Bungah, Penang Island	M	B	BLB	6	0.022	JX113339
25	WDSP/11/0061	Batu Feringghi, Penang Island	M	A	BLB	14	0.022	JX113340
26	WDSP/11/0062	Permatang Pasir, Balik Pulau, Penang Island	M	E	BLB	18	0.022	JX113341
27	WDSP/11/0063	Surau MK 2, Tg. Bungah, Penang Island	M	B	BLB	7	0.109	JX113342
28	WDSP/11/0064	Surau MK 2, Tg. Bungah, Penang Island	M	B	BLB	7	0.109	JX113343
29	WDSP/11/0065	Surau MK 2, Tg. Bungah, Penang Island	M	B	BLB	7	0.109	JX113344
30	WDSP/11/0066	Surau MK 2, Tg. Bungah, Penang Island	F	B	BLB	7	0.109	JX113345
31	WDSP/11/0067	Surau MK 2, Tg. Bungah, Penang Island	F	B	BLB	8	0.022	JX113346
32	WDSP/11/0068	Bkt. Gambir, Gelugor, Penang Island	M	F	BLB	1	0.217	JX113347
33	WDSP/11/0069	Bkt. Gambir, Gelugor, Penang Island	M	F	BLB	1	0.217	JX113348
34	WDSP/11/0070	Bkt. Gambir, Gelugor, Penang Island	M	F	BLB	1	0.217	JX113349
35	WDSP/11/0071	Permatang Pasir, Balik Pulau, Penang Island	M	E	BLB	3	0.043	JX113350
36	WDSP/11/0072	Bkt. Gambir, Gelugor, Penang Island	F	F	BLB	2	0.022	JX113351
37	WDSP/11/0073	Bkt. Gambir, Gelugor, Penang Island	M	F	BLB	3	0.043	JX113352
38	WDSP/11/0074	Jln. Perak, Penang Island	M	D	BLB	7	0.109	JX113353
39	ZMW486	Tanjung Tokong, Penang Island	F	C	L	1	0.217	JX113354
40	ZMW487	Tanjung Tokong, Penang Island	M	C	L	1	0.217	JX113355
41	ZMW488	Tanjung Tokong, Penang Island	F	C	L	1	0.217	JX113356
42	ZMW489	Tanjung Tokong, Penang Island	F	C	L	1	0.217	JX113357
43	ZMW490	Tanjung Tokong, Penang Island	M	C	L	4	0.043	JX113358
44	ZMW491	Tanjung Tokong, Penang Island	M	C	L	1	0.217	JX113359
45	ZMW492	Tanjung Tokong, Penang Island	F	C	L	1	0.217	JX113360
46	ZMW493	Tanjung Tokong, Penang Island	M	C	L	1	0.217	JX113361

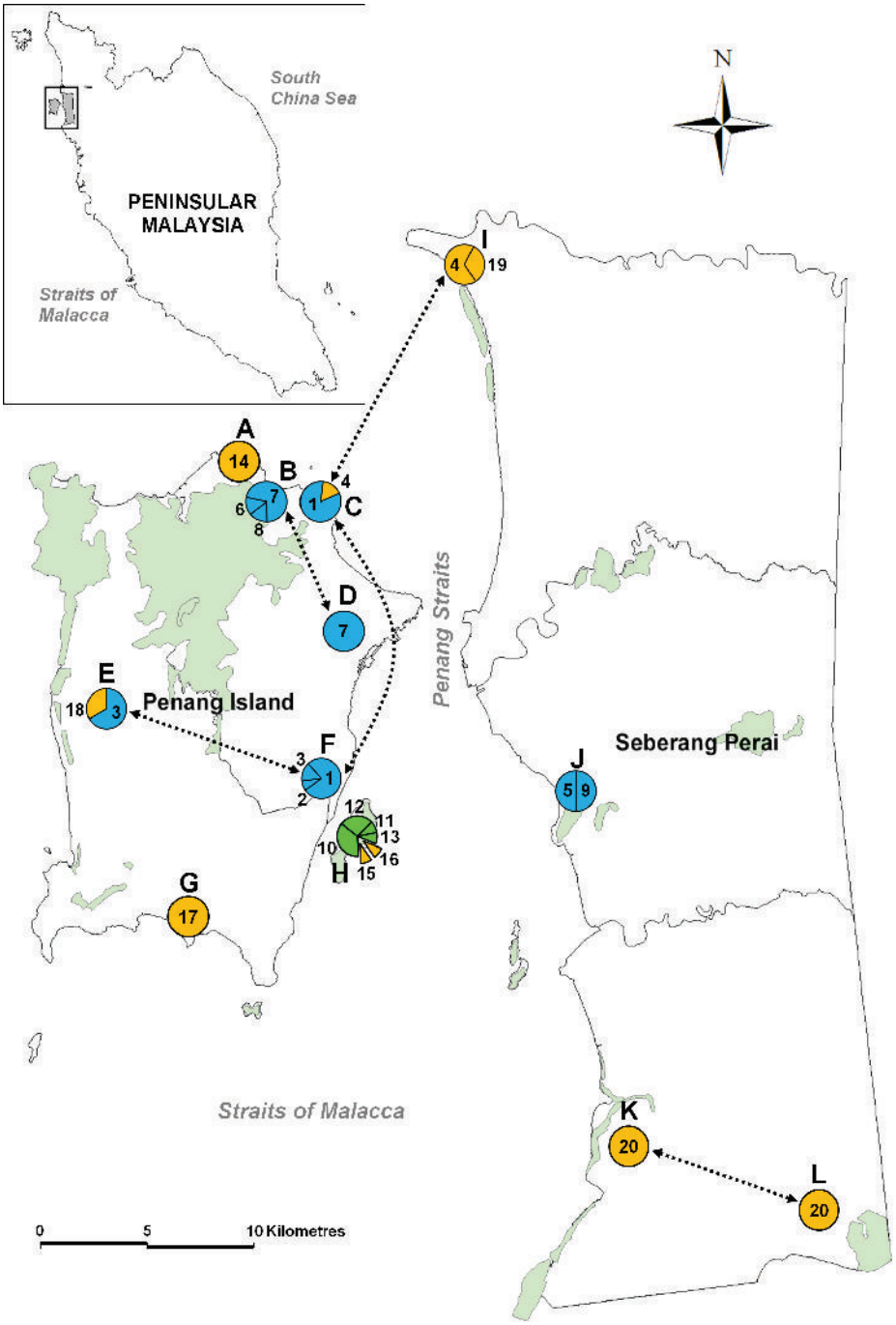

Figure 1. The map showing the sampling locations and haplotype designation of *M. fascicularis* in the Penang Island, Jerejak Island, and Seberang Perai (mainland). Each letter (A-L) represents a sampling location while the numbers (1-20) represents the assigned haplotype number (for details see Table 1). Dotted arrows indicate haplotypes sharing between the locations.

from the analysis by using the program Geneious v5.6. Sequence characterisations (variable sites, conserved sites and parsimony-informative sites) were done using Mega v5 (Tamura et al., 2011). Standard genetic diversity indices including the number of haplotypes (H), haplotype diversity (h), and nucleotide diversity (π) (Nei, 1987) were calculated in DnaSP v5 (Librado et al., 2009).

The mtDNA CR contains two segments; hypervariable segment I (HVI) and II (HVII). Apparently, there has been no comprehensive study utilizing the complete CR in *M. fascicularis*. However, several authors worked on either one of the segments. Thus we split our dataset to compare our sequences with the available sequences of the HVI (Chu et al., 2007; Smith et al., 2007), and the HVII (Blancher et al., 2008; Kawamoto et al., 2008; Shiina et al., 2010).

To analyse the genetic structuring between the mainland and the island as well as among the other regional populations, an estimate of population subdivision (FST) (Hudson et al., 1992) was calculated using DnaSP v5. Apart from that, the genetic distances between populations were also calculated by using the Kimura two-parameter model (Kimura, 1980) as performed using Mega v5.

Phylogenetic Trees and Estimation of Divergence Time

To infer the phylogenetic relationships, haplotypes data were used to generate the trees. Trees were constructed by using the neighbour-joining (NJ), maximum parsimony (MP), and maximum likelihood (ML) methods as implemented in Mega v5 and also the Bayesian method by using MrBayes (Huelsenbeck & Ronquist, 2001) as a plug-in in the Geneious program. A median-joining network was also constructed using the Network program (Bandelt et al., 1999). NJ was performed by using the Kimura 2-parameter distance model (Kimura, 1980) with pairwise deletion option while the MP analysis was done by using the Close-Neighbour-Interchange (CNI) option. The best-fit substitution model for ML was calculated in Mega v5 and the tree was constructed based on the HKY+G model using the Nearest-Neighbour-Interchange (NNI) option. The Bayesian analysis was performed by using the default settings with the HKY85 model (Hasegawa et al., 1985). Four heated Markov chains were initiated from random trees and sub-sampled at every 200 cycles. Delson (1996) proposed an early separation occurring at around 5.5 million years ago (mya) between the African and the Asian Macaca lineage. Therefore, for the complete CR dataset, members from both lineages were included in the analysis; the Asian lineage was represented by *M. mulatta* (AY612638) and *M. thibetana* (EU294187), while the African lineage was represented by *M. sylvanus* (AJ309865). All trees were rooted with the outgroup species of the tribe Papionini, *Papio hamdryas* (Y18001). To assess the robustness of the trees, bootstrapping (Felsenstein, 1985) with 10,000 replicates were conducted on all the NJ, MP, and ML trees.

To compare our dataset with the other regional populations of *M. fascicularis*, we constructed the ML tree, as described above. However, due to computational limitations, we ran the analysis with 100 bootstrapping replicates. The trees were rooted with sequences of *M. cyclopis* (HVI-DQ143986, HVII- AB261600), *M. fuscata* (HVI-AJ419862, HVII- AB261557), and *M. mulatta* from China (HVI-DQ373357, HVII-AB261958) and India (HVI-DQ373369, HVII-AB245416). By using the estimate of 1.6 million years ago (mya) for the divergence time between *M. mulatta* and *M. fascicularis* (Purvis, 1995; Hayasaka, 1996; Blancher et al., 2008), we re-calibrated the branch length of the ML tree in an attempt to date the divergence time of a major bifurcation event in the course of their evolution.

Results

Genetic Diversity

An alignment of 1,042 bp in length was produced from the 46 individual samples including the other *Macaca* species and outgroup sequences (N= 51) obtained from the GenBank. The sequences were later registered with the GenBank and were given accession numbers from JX113316-JX113361 (see Table 1).

The nucleotide diversity (π) for the complete CR dataset was 0.012. Although the pooled island samples (Penang and Jerejak Island) consisted of 36 individuals, the π of 0.009 was observed to be lower than the mainland π of 0.013 which was represented by only 10 individuals. In total 20 haplotypes were detected using the complete dataset with a haplotype diversity (h) of 0.921. Considering the indels, the mainland had 5 haplotypes (h= 0.667) while the island had 16 (h= 0.895). A single haplotype sharing between the mainland and the island was observed (Hap4; Fig. 1).

No	Population	N	Seg. Length	CS	VS	PIS	π	H	h
1	Jerejak Island	11	536	514	23	22	0.01445	4	0.764
2	Penang Island	25	536	504	33	21	0.0135	9	0.793
3	Seberang Perai[1]	10	536	501	36	21	0.02276	5	0.667
4	Malaysia*	149	536-9	378	163	147	0.064	104	0.99
5	Indonesia	70	537-9	390	152	131	0.075	55	0.99
6	Philippines	83	537-40	430	110	64	0.023	34	0.81
7	Vietnam	89	538-40	441	99	85	0.039	48	0.97
8	Mauritius	68	537	523	14	10	0.001	8	0.27
	Total	**505**	**546**	**345**	**198**	**173**	**0.074**	**246**	**0.98**

Seg. Length= Sequence Length, CS= Conserved sites, VS= Variable sites, PIS= Parsimony informative sites, π= Nucleotide diversity (Nei, 1987), H= No. of haplotypes, h= Haplotype diversity (Nei, 1987).
[1]Mainland

Table 2. Comparison among the different regional populations of *M. fascicularis* using the HVI segment. Number of haplotypes (H) and haplotype diversity (h) were calculated by considering sites with alignment gaps.

No	Population	1	2	3	4	5	6	7	8
1	Jerejak Island		2.0	3.4	6.1	7.5	9.6	5.3	13.5
2	Penang Island	0.3		3.0	5.9	7.3	9.7	4.9	13.7
3	Seberang Perai	0.4	0.4		6.2	7.4	9.7	6.0	12.9
4	Malaysia*	0.4	0.4	0.4		7.7	8.8	7.6	12.2
5	Indonesia	0.4	0.4	0.4	0.0		8.5	8.6	11.2
6	Philippines	0.8	0.8	0.8	0.5	0.4		9.8	8.9
7	Vietnam	0.6	0.6	0.6	0.3	0.3	0.7		12.4
8	Mauritius	0.9	0.9	0.9	0.7	0.6	0.9	0.8	

Table 3. Estimates of population subdivision (FST) (below the diagonal) and pairwise distances (%, above the diagonal) among the geographical populations of *M. fascicularis* analysed.

Table 2 compared the genetic diversity indices of the Penang populations to the other regional populations based on the HVI segment while Table 3 compared the estimates of population subdivision (FST) as well as the genetic distances. Separately, our finding revealed that the HVI segment contained more variations when compared to the HVII segment (see Appendix 1 for comparison between the HVI and HVII segments).

Phylogeography

The phylogenetic trees constructed using the NJ, MP (CI= 0.76, RI= 0.70), ML (–lnL= 3738.65), and Bayesian methods produced similar topologies and thus we summarised them using the ML tree as shown in Fig. 2A while Fig. 2B represents the median-joining haplotype network. Overall, the tree showed the monophyletic grouping of *M. fascicularis*. The tree topology and the haplotype network generally separated the 46 samples into three major groups: (1) Group I which consisted of three subgroups, (2) Group II which was formed entirely by the Jerejak Island samples, and (3) Group III which consisted of a mixture of haplotypes from the mainland and the island samples. In Group I, subgroup I-1 consisted of haplotypes from the north-eastern part of the island (sites C, E, and F), subgroup I-2 consisted of haplotypes from the mainland (site J) which was adjacent to the island, while subgroup I-3 contained of a mixture of haplotypes from the rest of the island (sites A, B, D, and G). Furthermore, the single deletion mutation observed as mentioned earlier delineated Groups I and II from Group III.

Both the HVI and HVII segment produced similar topologies (data not shown). A broader comparison of the Penang haplotypes with that of the other regional populations revealed that they were clustered within the continental lineage (Fig. 3). Additionally, in Figure 3, by calibrating node A which represents the bifurcation between *M. mulatta* and *M. fascicularis*, an estimated date of 1.04 mya at node B was obtained, which represents the last common ancestor between the continental and insular lineage.

Figure 2. A. Phylogenetic relationships of the *M. fascicularis* haplotypes from Penang as represented by the ML tree. Values above the branches represent bootstrap confidence levels (10,000 replications) for NJ, MP and ML respectively. Values below the branches represent the Bayesian posterior probability. **B.** Median-joining haplotype network of the Penang *M. fascicularis* that formed the 20 haplotypes. Mutational steps are indicated in dotted circles (if not indicated, mutational step is equal to one; link lengths are proportional to mutational steps). The median vectors that represent hypothetical intermediates or missing (unsampled) haplotypes are shown in squares. The areas of the circles are proportional to the haplotype frequency.

Discussion

Several studies on *M. fascicularis* have showed differences in variations between the HVI and HVII of the CR (Smith et al., 2007; Blancher et al., 2008; Kawamoto et al., 2008; Shiina et al., 2010). Similarly, our findings also revealed that the HVI harboured more variations compared to the HVII. Thus, we recommend that future studies on the CR should concentrate on the HVI.

Previous studies investigating the genetic variations and diversity of *M. fascicularis* either examined the HVI (Smith et al., 2007) or HVII (Blancher et al., 2008; Kawamoto et al., 2008; Shiina et al., 2010). Thus, only a handful of the complete CR sequences were available in Genbank for comparison. Smith et al. (2007) used the HVI on the Malaysian, Indonesian, Mauritius, Philippines, and Vietnamese *M. fascicularis* while Shiina et al. (2010) worked on three Indochinese subpopulations including the populations from Indonesia and the Philippines. Our finding revealed that all the haplotypes detected from Penang represent newly described haplotypes with none shared with the other *M. fascicularis* regional populations. Additionally, none of the haplotypes matched the samples (149 samples of unknown localities from Malaysia) used by Smith et al. (2007). Therefore, we would assume that the samples used by Smith et al. (2007) did not originate from

the Penang population. However, the small number of samples used in this study could prevent us from detecting any possible haplotypes sharing.

The π within the island for all three datasets (complete CR, HVI, and HVII) showed lower genetic diversity when compared to the mainland. This condition is very similar with the other island populations of Philippines and Mauritius, which could suggest a bottleneck or a small founder size following colonisation from the mainland (Blancher et al., 2008; Kawamoto et al., 2008). Nevertheless, the small sample number could also provide bias to the diversity observed.

The distribution pattern of the haplotypes in this study strongly correspond to their geographical distributions (Fig. 1). Macaca in particular displays an extreme level of sex-biased dispersal (Melnick & Hoelzer, 1992) where only males disperse from one population to another while the females remain sedentary in nature. This will lead to geographically structured mitochondrial haplotypes which shows structuring of populations according to historical cladogenic events (Melnick et al., 1993; Tosi & Coke, 2007). In this study, we observed moderate structuring between the Penang mainland and the two island populations (FST of 0.43 and 0.37 respectively for Jerejak and Penang Island) as well as to the other neighbouring populations (FST of 0.37 for both the other Malaysian and Indonesian populations). The genetic distances between the mainland to the island populations are however low as compared to the other populations (Table 3). The narrow Penang Strait would act as a contemporary barrier to the gene flow between the mainland and the island apart from the sedentary nature of the females. The historical connectivity and the close proximity would explain the low genetic distances between the mainland and the island populations.

On the other hand, the single haplotype sharing (Hap4) observed between the mainland and island could be explained as a remnant of the ancestral genetic makeup from the earlier colonisation from the mainland. The emergence of past land bridge connections permitted gene flow between the mainland and island populations. This is supported by our recent work on the Y-chromosomal gene flow of the males using the same samples set (Rovie-Ryan et al., 2013) where similar haplotypes sharing were observed. However in the past, translocations of *M. fascicularis* by the authorities from the island to the adjacent mainland as a response to the increasing human-macaque conflict due to the urbanisation (S. Elagupillay, pers. comm.) could also explain the haplotype sharing.

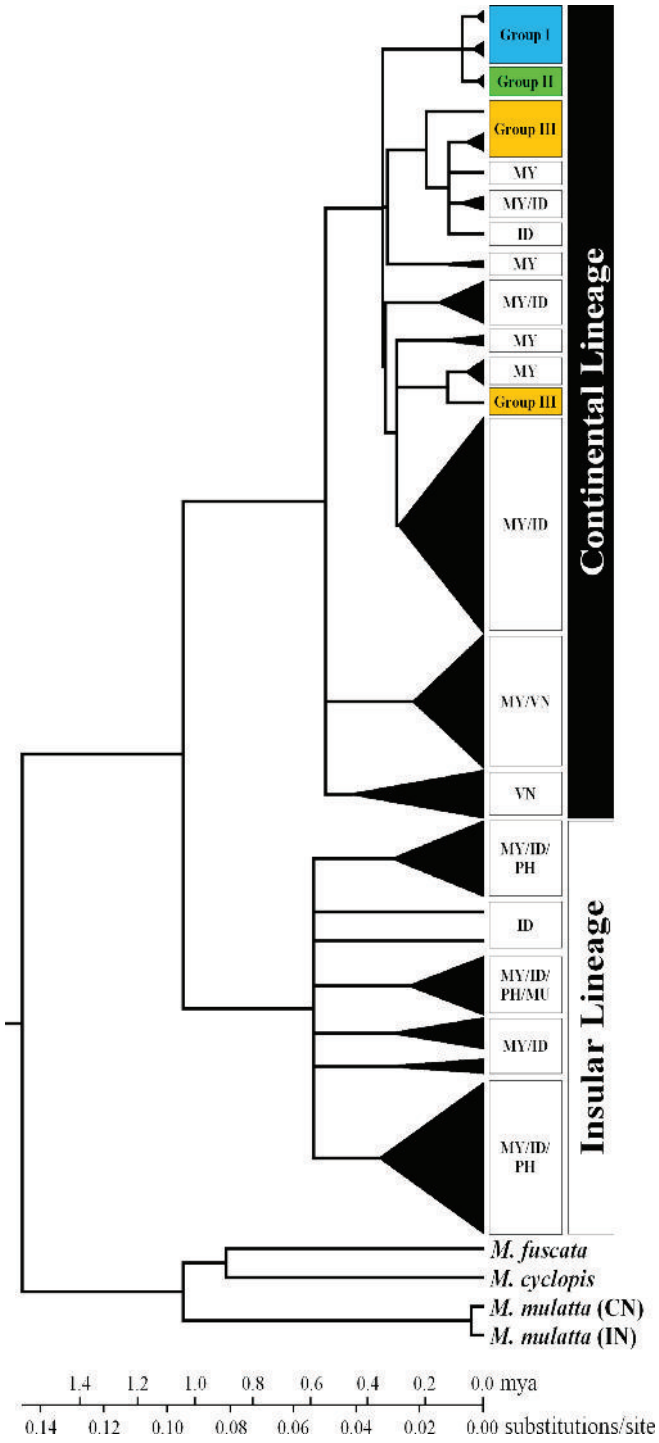

Figure 3. ML tree the HVI segment (-lnL= 8267.40) constructed using the HKY+G model (Hasegawa et al. 1985) with 100 bootstrapping. Note that the Penang haplotypes (Group I, II, and III) are clustered within the continental lineage. Following the ISO 3166-1 alpha-2 codes for the names of countries, MY codes for Malaysia, ID for Indonesia, VN for Vietnam, PH for Philippines, MU for Mauritius, CN for China, and IN for India. No major discrepancy was observed between the HVI and HVII trees (tree not shown).

The single indel observed at np 217 distinguished the three major groupings of the Penang *M. fascicularis*. The insertion mutation represents the haplotypes within Group III while the deletion mutation represents the haplotypes for Groups I and II as also seen in the phylogenetic tree and the haplotype network (Fig. 2). Based on the basal positioning in the phylogeny tree, we suggest that the haplotypes of Group III represents the ancestral form of *M. fascicularis* in Penang. On a wider scale, all the sequences from the other regional populations of *M. fascicularis* using the HVI dataset also displayed the insertion mutation. Therefore, the deletion mutation observed in Group I and II is unique only to the Penang *M. fascicularis*. Due to this uniqueness, we propose that the Penang *M. fascicularis* particularly from the island should be managed as a management unit (MU). MUs are a second category of population subsets that are demographically distinct and are much smaller than the evolutionary significant units (ESUs) (Moritz, 1994). In Peninsular Malaysia, several strategies of population management has been used by the DWNP to control the human-Macaca conflicts including on-going population inventories, public awareness programmes, translocation, culling (DWNP, 2006), and recently experimenting on chemical castration (Karuppannan et al., 2013). The findings from this study would therefore provide policy makers with information on which priority areas or populations to preserve.

The monophyletic clustering of the *M. fascicularis* as shown in the phylogenetic trees (Fig. 2 and 3) were in agreement with previous molecular studies using mtDNA data (Blancher et al., 2008; Hayasaka et al., 1996; Tosi et al., 2002). Comparisons with the other regional populations revealed that the Penang samples were clustered within the continental lineage (Fig. 3). The dichotomy between the continental and the insular lineage were first discovered by Harihara et al. (1988) and later confirmed by Tosi et al. (2002, 2003), Tosi and Coke (2007), and Blancher et al. (2008). Based on our calibration, we propose that by the middle Pleistocene, around 1.04 mya (Fig. 3), the separation of the continental and the insular lineage of *M. fascicularis* occurred. This estimate is close to that proposed by Tosi et al. (2003) and Blancher et al. (2008) at ~1.2 mya.

The findings of this study revealed that genetic diversity and divergence of *M. fascicularis* were lower in the island as compared to the mainland which could suggest a bottleneck or small founder size following colonisation from the mainland. The shared haplotype provided evidence on the historical connectivity between the mainland and island. The unique haplotypes observed in the Penang samples (especially the island population) would be a good indicator to conserve the genetic uniqueness of *M. fascicularis* in Penang. These findings could be used in management and conservation strategies especially in the population control, and would also be beneficial to other studies for example on the co-evolution of primates with vectors of diseases associated with primates (like evolution of simian malaria parasite) and biogeographical history of the Penang Island. Finally, we also discovered that the CR of the mtDNA is a powerful region to be used in addressing the genetic structuring of *M. fascicularis*. We acknowledge that the small sample numbers could be a source of bias in our findings. Therefore, for a more robust and geographically wide analysis, more samples should be collected from Penang and from a wider range of localities covering the entire Peninsular Malaysia, Sabah and Sarawak. Also, to infer current gene flow, further studies utilizing the Y-chromosome and microsatellite markers are currently being done.

Acknowledgements

We would like to thank the Department of Wildlife and National Parks (DWNP), Malaysia, for providing the facilities, equipments, and personnels to conduct this pilot study. Also thanks to the State Director of DWNP Penang, Mr. Jamalun Nasir Ibrahim and his staff for their assistance. This project was funded by the DWNP, and partly by the Proboscis Genome Research grant awarded to M. T. Abdullah and colleagues. The sampling work for ZMW coded samples was supported in part by the USAID Emerging Pandemic Threat Program – PREDICT Project and by the Skoll Foundation and Google Incorporated through Global Viral. We would also like to thank the following individuals for their dedication and help during sampling: Tang Teong Kim, Osman Bakar and Rozeta Hj. Zahran. The samplings conducted complied with the rules, regulations and ethical standards in the treatment of the animals as laid down by all the relevant wildlife authorities in Malaysia as well as the guidelines by the Institutional Animal Care and Use Committee (IACUC), University of California, Davis, USA as adopted by the PREDICT Project. We also acknowledge the comments by two anonymous reviewers who helped to improve the earlier version of this manuscript.

References

Asadpour, R., Lim, H.S., Alashloo, M.M., Shekafti, S.A. and S. Moussavi (2011). Application of THEOS Imagery to study chlorophyll-a at the Strait of Penang Island, Malaysia. *International Proceedings of Chemical, Biological and Environmental Engineering* **17**: 117-122.

Bandelt, H-J., Forster, P. and A. Röhl (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37-48.

Bird, M.I., Taylor, D. and C. Hunt (2005). Palaeoenvironments of insular Southeast Asia during the Last Glacial Period: a savanna corridor in Sundaland? *Quaternary Science Reviews* **24**: 2228-2242.

Blancher, A., Bonhomme, M., Crouau-Roy, B., Terao, K., Kitano, T. and N. Saitou (2008). Mitochondrial DNA sequence phylogeny of 4 populations of the widely distributed cynomolgus macaque (*Macaca fascicularis fascicularis*). *Journal of Heredity* **99**: 254–264.

Chu, J-H., Lin, Y-S. and H- Y. Wu (2007). Evolution and dispersal of three closely related macaque species, *Macaca mulatta*, *M. cyclopis*, and *M. fuscata*, in eastern Asia. *Molecular Phylogenetics and Evolution* **43**: 418–429.

Cox-Singh, J. and B. Singh (2008). Knowlesi malaria: newly emergent and of public health importance? *Trends in Parasitology* **24**: 406–410.

Delson, E. (1996). The oldest monkeys in Asia. In International Symposium: Evolution of Asian Primates, (ed O. Takenaka), pp. 40. Primate Research Institute, Freude & Kyoto University, Inuyama, Aichi, Japan.

Drummond, A.J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T. and A. Wilson (2012). Geneious version 5.6. Http://www.geneious.com.

DWNP. (2006). Pelan Pengurusan Kera (Macaca fascicularis) Bermasalah di Semenanjung Malaysia (English translation: Management Plan for Problematic Macaques (*Macaca fascicularis*) in Peninsular Malaysia). Department of Wildlife and National Parks, Kuala Lumpur, Malaysia.

Eudey, A.A. (2008). The crab-eating macaque (*Macaca fascicularis*): widespread and rapidly declining. *Primate Conservation* **23**: 129–132.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Fooden, J. (1995). Systematic review of Southeast Asian long- tail macaques, *Macaca fascicularis* (Raffles, [1821]). *Fieldiana Zoology* new series **81**: 1–206.

Harihara, S., Saitou, N., Hirai, M., Aoto, N., Terao, K., Cho, F., Honjo, S. and K. Omoto (1988). Differentiation of mitochondrial DNA types in *Macaca fascicularis*. *Primates* **29**: 117–127.

Harrison, T., Krigbaum, J. and J. Manser (2006). Primate biogeography and ecology on the Sunda Shelf Islands: A Paleontological and Zooarchaeological Perspective. In Primate Biogeography, (eds S.M. Lehman & J.G. Fleagle), pp. 331–372. Springer, New York, USA.

Hasegawa, M., Kishino, H. and T. Yano (1985). Dating of human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**: 160–174.

Hayasaka, K., Fujii, K. and S. Horai (1996). Molecular phylogeny of macaques: implications of nucleotide sequences from an 896-base pair region of mitochondrial DNA. *Molecular Biology and Evolution* **13**: 1044–1053.

Hudson, R.R., Boos, D.D. and N.L. Kaplan (1992). A statistical test for detecting population subdivision. *Molecular Biology and Evolution* **9**: 138–151.

Huelsenbeck, J.P. and F. Ronquist (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754– 755.

Karimullah and A. Shahrul (2011). Condition and population size of *Macaca fascicularis* (long-tailed macaque). *Journal of Cell and Animal Biology* **5**: 41–46.

Karuppannan K.V., Saaban, S., Firdaus Ariff, A.R. and A.R. Mustapa (2013). Non-surgical castration in controlling long tailed macaque (*Macaca fascicularis*) population by Department of Wildlife and National Parks (DWNP). *Malaysian Journal of Veterinary Research* **4(1)**: 33–36.

Kawamoto, Y., Kawamoto, S., Matsubayashi, K., Nozawa, K., Watanabe, T., Stanley, M-A. and D. Perwitasari-Farajallah (2008). Genetic diversity of the longtail macaques (*Macaca fascicularis*) on the island of Mauritius: an assessment of nuclear and mitochondrial DNA polymoprhisms. *Journal of Medical Primatology* **37**: 45–54.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotides sequence. *Journal of Molecular Evolution* **16**: 111–120.

Lawler, S.H., Sussman, R.W. and L.L. Taylor (1995). Mitochondrial DNA of the Mauritian macaques (*Macaca fascicularis*): an example of the founder effect. *American Journal of Physical Anthropology* **96**: 133–141.

Librado, P. and J. Rozas (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Ling, T.S. (2006). Behavioral study of *Macaca fascicularis* and interaction with humans at Taman Botani, Pulau Pinang. B.Sc. thesis, Universiti Kebangsaan Malaysia, Selangor, Malaysia.

Melnick, D.J. and G.A. Hoelzer (1992). Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. *International Journal of Primatology* **13**: 379–393.

Melnick, D.J., Hoelzer, G.A., Absher, R. and M.V. Ashley (1993). MtDNA diversity in rhesus monkeys reveals overestimates of divergence time and paraphyly with neighboring species. *Molecular Biology and Evolution* **10**: 282–295.

Moritz, C. (1994). Defining 'Evolutionary Significant Units' for conservation. *Trends in Ecology and Evolution* **9**: 373–375.

Nei, M. (1987). Molecular Evolutionary Genetics. Columbia University Press, New York, USA.

Perwitasari-Farajallah, D., Kawamoto, Y. and B. Suryobroto (1999). Variation in blood proteins and mitochondrial DNA within and between local populations of longtail macaques, *Macaca fascicularis* on the Island of Java, Indonesia. *Primates* **40(4)**: 581–595.

Perwitasari-Farajallah, D., Kawamoto, Y. and R. C. Kyes (2001). Genetic characterization of long-tailed macaques (*Macaca fascicularis*) on Tabuan Island, Indonesia. *Primates* **42(2)**: 141–152.

Ping, K.O. (2003). Behavioral study of *Macaca fascicularis* and interaction with humans at Taman Belia, Pulau Pinang. B.Sc. thesis, Universiti Kebangsaan Malaysia, Bangi, Malaysia.

Purvis A. (1995). A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society B: Biological Sciences* **348**: 405–421.

Rovie-Ryan, J.J., Abdullah, M.T., Sitam, F.T., Abidin, Z.Z. and S.G. Tan (2013). Y-chromosomal gene flow of *Macaca fascicularis* (Cercopithecidae) between the insular and mainland peninsula of Penang state, Malaysia. *Journal of Science and Technology in the Tropics* **9**: 113-126.

Rozen, S. and H. Skaletsky (2000). Primer3 on the WWW for general users and for biologist programmers. In Bioinformatics Methods and Protocols: Methods in Molecular Biology (eds. S. Krawetz & S. Misener), pp. 365–386. Humana Press, Totowa, USA.

Sathimurthy, E. and H.K. Voris (2006). Maps of Holocene sea level transgression and submerged lakes on the Sunda Shelf. *Natural History Journal of Chulalongkorn University* **Supplement 2**: 1–44.

Schillaci M.A., Saravia, S., Lee, B.P.Y.-H. and C. Matheson (2011). Preliminary report on mitochondrial DNA variation in *Macaca fascicularis* from Singapore. *The Raffles Bulletin of Zoology* **59(1)**: 101–108.

Shiina, T., K. Tanaka, Y. Katsuyama, K. Otabe, K. Sakamoto, M. Kurata, M. Nomura, H. Yamanaka, H. Nakagawa, H. Inoko and M. Ota, (2010). Mitochondrial DNA diversity among three subpopulations of cynomolgus macaques (*Macaca fascicularis*) originating from the Indochinese region. *Experimental Animal* **59**: 567–578.

Shuan, H.H. (2006). A study on aggressive and submissive behaviour in *Macaca fascicularis* at Botanical Garden, Penang. B.Sc. thesis, Universiti Kebangsaan Malaysia, Selangor, Malaysia.

Smith, D.G., McDonough, J.W. and D.A. George (2007). Mitochondrial DNA variation within and among regional populations of longtail macaques (*Macaca fascicularis*) in relation to other species of the fascicularis group of macaques. *American Journal of Primatology* **69**: 182–198.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and S. Kumar (2011). Mega version 5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731–2739.

Thayaparan, S., Robertson, I., Amran, F., Su'ut, L. and M.T. Abdullah (2013). Serological prevalence of leptospiral infection in wildlife in Sarawak, Malaysia. *Borneo Journal of Resource Science and Technology* **2(2)**: 79–82.

Thayaparan, S., Robertson, I.D., Fairuz, A., Suut, L. and M.T. Abdullah (2014). Leptospirosis, an emerging zoonotic disease in Malaysia. *Malaysian Journal of Pathology* **35(2)**: 125–134.

Tosi, A.J. and C.S. Coke (2007). Comparative phylogenetics offer new insights into the biogeographic history of *Macaca fascicularis* and the origin of the Mauritian macaques. *Molecular Phylogenetics and Evolution* **42**: 498–504.

Tosi, A.J., Morales, J.C. and D.J. Melnick (2002). Y-Chromosome and mitochondrial markers in *Macaca fascicularis* indicate introgression with Indochinese *M. mulatta* and a biogeographic barrier in the Isthmus of Kra. *International Journal of Primatology* **23**: 161–178.

Tosi, A.J., Morales, J.C. and D.J. Melnick (2003). Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution* **57**: 1419–1435.

Voris, H.K. (2000). Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography* **27(5)**: 1153–1167.

Appendix 1. Comparison between the HVI and HVII of the mtDNA CR from different regional populations of *M. fascicularis*. Number of haplotypes (H) and haplotype diversity (h) were calculated by considering sites with alignment gaps.

		HVI[a] (546 bp)								HVII[b] (466 bp)							
No	Population	N	Seg. Length	CS	VS	PIS	π	H	h	N	Seg. Length	CS	VS	PIS	π	H	h
1	Penang[1]	46	536-7	490	47	40	0.021	17	0.91	46	461	451	10	6	0.003	9	0.64
2	Malaysia	149	536-9	378	163	147	0.064	104	0.99								
3	Indonesia	70	537-9	390	152	131	0.075	55	0.99	65	438-61	384	82	64	0.033	56	0.99
4	Philippines	83	537-40	430	110	64	0.023	34	0.81	14	438-58	451	13	8	0.010	7	0.87
5	Vietnam	89	538-40	441	99	85	0.039	48	0.97	23	461-2	408	55	35	0.027	22	1.00
6	Cambodia									48	441-62	392	71	35	0.021	42	0.99
7	Mauritius	68	537	523	14	10	0.001	8	0.27	6	439-59	461	1	1	0.001	4	0.87
	Total	511	546	346	200	176	0.072	261	0.98	202	466	332	134	93	0.05	140	0.98

Seg. Length= Sequence Length, CS= Conserved sites, VS= Variable sites, PIS= Parsimony informative sites, π= Nucleotide diversity (Nei, 1987), H= No. of haplotypes, h= Haplotype diversity (Nei, 1987).
[a]Smith et al. (2007)
[b]Shiina et al. (2010), Kawamoto et al. (2008), and Blancher et al. (2008)