Zoological Studies **60**:50 (2021)
doi:10.6620/ZS.2021.60-50

Zoological Studies

Open Access

Taxonomy, Evolutionary and Dispersal Events of Pig-Tailed Macaque, *Macaca nemestrina* (Linnaeus, 1766) in Southeast Asia with Description of a New Subspecies, *Macaca nemestrina perakensis* in Malaysia

Muhammad Abu Bakar Abdul-Latiff[1,2,3], and Badrul Munir Md-Zain[1,*]

[1]*Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600, Bangi, Selangor, Malaysia. *Correspondence: E-mail: abgbadd1966@yahoo.com (Md-Zain)*
E-mail: abdullatiff02727@gmail.com or latiff@uthm.edu.my (Abdul-Latiff)
[2]*Oasis Integrated Group (OIG), Institute for Integrated Engineering (I²E), Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, Malaysia.*
[3]*Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia (Pagoh Campus), 84600, Muar, Johor, Malaysia.*

Received 23 June 2020 / Accepted 6 June 2021 / Published 8 October 2021
Communicated by Teng-Chiu Lin

The pig-tailed macaque, *Macaca nemestrina*, which is distributed in Peninsular Malaysia, Thailand, Borneo, and Sumatra, has been the subject of unstable and changing taxonomic entity in the *M. nemestrina* group. This species is involved with a human-macaque conflict in Malaysia and at the same time played an important role in the ethnozoological culture of Malaysian. Even so, comprehensive phylogenetic, population genetics and biogeographical analysis of *M. nemestrina* in Malaysia are non-existent after decades of intensive research on the genus itself. Thus, we conducted the first comprehensive genetic study of *M. nemestrina* in Malaysia, based on three mitochondrial loci—Cytochrome *b* (567 bp), D-loop (398 bp), and *COI* (577 bp)—from 27 individuals representing Malaysia, plus an additional 26 sequences of Southeast Asian macaques from Genbank. Comparative biogeographical analysis in this study supports the positions of *M. nemestrina* in *M. nemestrina* groups as opposed to the *silenus* or Sulawesi groups. Results from this study also indicate that Bornean populations are the first extant lineages to separate from the other examined lineages of *M. nemestrina*, *M. leonina*, *M. pagensis*, and *M. siberu* in Southeast Asia. Molecular clock analysis suggested that *M. nemestrina* arrived in the Malay Peninsula about 0.32 million years ago (MYA). Our results indicate that the population of pig-tailed macaque from Perak (west Peninsular Malaysia) differs genetically based on all phylogenetic and population genetic analyses. Morphologically, Perak's pig-tailed macaque shows brighter coloration than *M. n. nemestrina*. Thus, we proposed a new subspecies for Perak's pig-tailed macaque as *Macaca nemestrina perakensis* distributed in the state of Perak, Peninsular Malaysia. This research helps resolve the taxonomic position and population genetics of pig-tailed macaque in Malaysia, which contribute directly to conservation and management of the species in Malaysia.

Key words: Pig-tailed macaque, *Macaca nemestrina*, Cercopithecinae, Phylogeography, Taxonomy.

Citation: Abdul-Latiff MAB, Md-Zain BM. 2021. Taxonomy, evolutionary and dispersal events of pig-tailed macaque, *Macaca nemestrina* (Linnaeus, 1766) in southeast asia with description of a new subspecies, *Macaca nemestrina perakensis* in Malaysia. Zool Stud **60**:50. doi:10.6620/ZS.2021.60-50.

BACKGROUND

Macaca is the only genus representing the subfamily Cercopithecinae in Asia and shows the largest radiation history across the continent, being the most successful non-human primate (Liedigk et al. 2014). The systematics of the genus itself has been controversial, as primatologists have rearranged the species classification numerously on the basis of different approaches. Fooden (1976), for instance, leans toward a four species group classification, namely *silenus-sylvanus*, *sinica*, *fascicularis*, and *arctoides*, based on anatomical structures of sexual organs. Delson (1980) proposed a different four species group, Groves (2001) proposed six groups, and Zinner et al. (2013) proposed seven groups, all under which positions of *M. nemestrina* remain unresolved. In this study, we aim to answer two main issues involving Southeast Asian pig-tailed macaque's classification: (a) the systematic level at which all the pig-tailed macaques should be separated (species or subspecies) and (b) the barrier separating *M. nemestrina* and *M. leonina*, and whether hybridization occurs between their populations.

Fooden (1975) classified *M. nemestrina* into three subspecies: *M. nemestrina nemestrina* (Southern/Sundaland pig-tailed macaque), *M. nemestrina leonina* (Northern/Indochinese pig-tailed macaque) and *M. nemestrina pagensis* (Mentawai macaques). Albrecht (1980) supported these classifications based on regression analysis on skull and body size to latitude, as did Rosenblum et al. (1997), who used 2.3 Kb of mitochondrial DNA (mtDNA) fragments. Whereas Groves (1997 2001), Abegg and Thierry (2002) and Evans et al. (2003) supported the classifications of Mentawai macaque at the species level as *Macaca pagensis*. Roos et al. (2003) validated the paraphyletic relationships of *M. pagensis* in Mentawai islands, therefore nominating a new species, *M. siberu*, which showed closer relationships to *M. nemestrina* in Sumatra. Hence, taxonomically, there are four species of macaques (*M. nemestrina*, *M. leonina*, *M. siberu*, and *M. pagensis*) that were once classified as *M. nemestrina* across Southeast Asia. Generally, all four species occur in primary, secondary, and disturbed areas (Groves 2001, Malaivijitnond et al. 2012, Roos et al. 2003).

Macaca nemestrina is distributed in Malaysia (including Peninsular Malaysia and Malaysian Borneo), Brunei, Indonesia (Bangka, Kalimantan Borneo, and Sumatra), and southern peninsular Thailand (Groves 2001, Roos et al. 2014). *Macaca leonina*, conversely, occurs in eastern Bangladesh, Cambodia, southern China, northeastern India, Lao PDR, Myanmar, Thailand, and central and southern Vietnam (Groves 2001, Malaivijitnond et al. 2012).

Gippoliti (2001) classified *M. leonina* and *M. nemestrina* at the species level, based on morphological evidence of female sexual swellings. Likewise, Malaivijitnond et al. (2012) reported a distinctive morphological characteristic between these two species. Roos et al. (2007) also supported this classification by studying the Cyt *b*, 12S rRNA, and 16S rRNA sequences. However, whether the limits of *M. leonina* distributions were the Isthmus of Kra (10°30'N) (Brandon-Jones et al. 2004; Groves 2001; Nadler et al. 2007) or south of the Isthmus Kra in Surat Thani-Krabi (Fooden 1975) remains unknown.

While a few revisions have been done on the taxonomy and biogeography of *M. siberu*, *M. leonina*, and *M. pagensis*, almost none have been performed *M. nemestrina* in Malaysia. Molecular systematics of *M. nemestrina* in Malaysia have never been extensively studied, with a lack of data about population genetics and biogeography. These issues have severe conservation implications in Malaysia. *Macaca nemestrina* is the second in line in human-wildlife conflicts in Malaysia, after the widely distributed long-tailed macaque, *M. fascicularis* (DWNP 2012). Assuming only one taxonomic entity exists, the Departments of Wildlide and National Parks Peninsular Malaysia (DWNP) have been translocating *M. nemestrina* to resolve the human-macaque conflict in the country (DWNP 2012). Studying population genetics and identifying the unique gene pool of pig-tailed macaque's population in Malaysia is important for population translocation and to avoid disrupting any unique lineages present in Malaysia. Furthermore, *M. nemestrina* also plays a crucial role in the culture of Malaysian people living in the rural area, as they were raised as pets, domesticated, and used to pluck coconuts, fruits, and vegetables (Ruslin et al. 2017). These events were recorded as early as 1926 (Harrison 1926) during the British-Malaya era. The aftermath of this culture is the growing pet trade, especially among the people using the macaques to pluck fruits for economic gains instead of simple recreations.

To better understand the taxonomy of *M. nemestrina*, we carried out the first comprehensive analysis on the molecular systematics of *M. nemestrina* in Malaysia. The objective of this study is to review the taxonomy, population genetics and biogeography of *M. nemestrina* in Malaysia and its relationships to other Southeast Asian pig-tailed macaques using Cytochrome *c* Oxidase Subunit I (*COI*), Cytochrome *b* (Cyt *b*) and D-loop sequences of mitochondrial DNA (mtDNA). The mtDNA is widely used as a genetic marker because of its high mutation rate, maternal inheritance, and high frequency in the cells that make it easier to amplify (Zainudin et al. 2010). The non-coding portion D-loop

is a highly variable region and frequently used for intraspecific studies (Md-Zain et al. 2019). The protein-coding Cyt *b* and *COI* regions are frequently used for species delimitation and phylogenetic studies (Hebert et al. 2003; Md-Zain et al. 2018). Previous studies proved that mtDNA is an ideal genetic marker to solve taxonomic problems, phylogenetic relationships, and population genetics of primates (Abdul-Latiff et al. 2014a b 2017a b 2019; Aifat et al. 2016a b; Md-Zain 2001; Md-Zain et al. 2008 2010a b; Vun et al. 2011).

MATERIALS AND METHODS

Ethical Statement

All non-invasive genetic samples collected in this study were in compliance with guidelines of *Principles for the Ethical Treatment of Non-Human Primates* by the American Society of Primatologist. Feces and folicles were collected from macaques kept in captivity by people living in rural areas as this is permitted by law in Malaysia under the Wildlife Protection Act (Act 716). However, we reject the notion of chaining pig-tailed macaques by humans, and appropriate advices were given to the owner over the issues.

Samples collection, DNA extraction, PCR and sequencing

We collected 27 biological samples of *M. nemestrina* originating from the states of Kelantan, Johor, Perak, Pahang, Kedah representing the Malay Peninsula, and Sarawak representing Borneo (Table 1). These samples were collected on-field with assistance from a DWNP officer. Follicle samples were collected simply by using tweezer to pluck the hairs with assistance from the owner of the macaque, as most macaques are aggressive toward strangers. Fecal samples of the wild population were collected by following the wild population of pig-tailed macaque and waiting for them to drop their feces. These samples were stored in 90% ethanol for DNA extraction purposes. A total of 26 additional sequences (Table 1) representing *M. nemestrina*, *M. leonina*, *M. siberu*, and *M. pagensis* from various regions in Southeast Asia obtained from GenBank were also used for comparative analysis purposes. However, these analyses were restricted to the Cyt *b* locus alone, as no sufficient data exist for other loci in GenBank. In addition, two sequences of *M. fascicularis* (sequenced in this study) and *M. sylvanus* (NC002764) (Table 1) were used as the outgroup and calibration point for molecular clock estimation.

DNA extractions were carried out by using the extraction kit following procedures recommended by the supplier. Feces samples were extracted using DNA innuPREP Stool DNA Kit (Analytik Jena, Germany), whereas DNA innuPREP Forensic Kit (Analytik Jena, Germany) were used for follicle samples. We amplified three mtDNA markers: the D-loop region, the *COI* and the Cyt *b* through a polymerase chain reaction (PCR) using a Mastercycler Nexus (Eppendorf North America, Inc.). PCR reactions were generated using Phusion Flash High-Fidelity PCR Master Mix (Finnzymes, OY), which exhibits high accuracy (proofreading DNA polymerase with a fidelity of 25X *Taq polymerase*), extreme speed (extension times of 15 s/kb or less), and a very high yield in a reduced length of time. Three species-locus-specific primers adapted from Abdul-Latiff et al. (2017b) were used in order to ensure the purity of sequences obtained, and to avoid amplifying Numts. PCRs were employed with parameters as follows: initial denaturation for 10 s at 98°C, followed by 30 cycles of denaturation for 1 s at 98°C, annealing for 30 s at X °C (annealing temperature for different primers as suggested by Abdul-Latiff et al. (2017b), extension for 15 s at 72°C, and a final extension stage for 1 min at 72°C. PCR performance and product sizes were verified on 1% agarose gels. Quality PCR products were then purified using innuPREP DOUBLEpure (Analytik Jena, Germany) and were sent to 1st Base Sdn Bhd (Malaysia) for sequencing purposes.

Data Analyses

Sequences obtained were edited using Bioedit Sequence Alignment Editor v7.2.3 (Hall 1999). Subsequently, these sequences were validated using sequence similarity searches (GenBank BLASTn). MEGA7 (Tamura et al. 2013) was then used to perform multiple alignment using the ClustalW algorithm. MEGA7 (Tamura et al. 2013) and DnaSP v4.0 (Rozas et al. 2003) were mainly used in sequence analysis to reveal pairwise genetic distance, single nucleotide polymorphisms, net nucleotide divergence (Da), and nucleotide diversity (π). Genetic segregation at the sequence and population levels—including nucleotide subdivision (Nst), population subdivision (F_{ST}), and number of migrants per generation (Nm)—were investigated using DnaSP v4.0 (Rozas et al. 2003).

Population demographical history was analyzed using Tajima's test of neutrality, *D* (Tajima 1989), Fu and Li's *D** and *F** (Fu and Li 1993), and Fu's Fs (Fu 1997). All these parameters were estimated using DnaSP v4.0 (Rozas et al. 2003). The number of haplotypes, haplotype diversity, and segregation sites present in the population were determined in DnaSP v4.0 (Rozas et al. 2003). The haplotypes were then analyzed in Arlequin

Table 1. Details on samples and sequences used in this study. Label on sample code represent domesticated (2) or wild (1) population

No.	Code	Species	Origin	Source
1	[1]ALMND18	*M. nemestrina*	Tanah Merah, Kelantan	Abdul-Latiff
2	[1]ALMND22	*M. nemestrina*	Tanah Merah, Kelantan	Abdul-Latiff
3	[1]ALMND24	*M. nemestrina*	Tanah Merah, Kelantan	Abdul-Latiff
4	[1]ALMND26	*M. nemestrina*	Tanah Merah, Kelantan	Abdul-Latiff
5	[2]ALMNJ275	*M. nemestrina*	Paloh, Johor	Abdul-Latiff
6	[2]ALMNJ277	*M. nemestrina*	Paloh, Johor	Abdul-Latiff
7	[2]ALMNJ278	*M. nemestrina*	Paloh, Johor	Abdul-Latiff
8	[2]ALMNJ279	*M. nemestrina*	Paloh, Johor	Abdul-Latiff
9	[2]ALMNJ280	*M. nemestrina*	Paloh, Johor	Abdul-Latiff
10	[2]ALMNA353	*M. nemestrina*	Kuala Gula, Perak	Abdul-Latiff
11	[2]ALMNA355	*M. nemestrina*	Kuala Gula, Perak	Abdul-Latiff
12	[1]ALMNA356	*M. nemestrina*	Selama, Perak	Abdul-Latiff
13	[1]ALMNA361	*M. nemestrina*	Selama, Perak	Abdul-Latiff
14	[1]ALMNA366	*M. nemestrina*	Selama, Perak	Abdul-Latiff
15	[1]ALMNA367	*M. nemestrina*	Selama, Perak	Abdul-Latiff
16	[1]ALMNA368	*M. nemestrina*	Selama, Perak	Abdul-Latiff
17	[2]ALMNC229	*M. nemestrina*	Jerantut, Pahang	Abdul-Latiff
18	[2]ALMNC230	*M. nemestrina*	Jerantut, Pahang	Abdul-Latiff
19	[1]ALMNK103	*M. nemestrina*	Baling, Kedah	Abdul-Latiff
20	[1]ALMNK104	*M. nemestrina*	Baling, Kedah	Abdul-Latiff
21	[1]ALMNK125	*M. nemestrina*	Baling, Kedah	Abdul-Latiff
22	[1]ALMNK126	*M. nemestrina*	Baling, Kedah	Abdul-Latiff
23	[2]AAMNQ1	*M. nemestrina*	Lundu, Sarawak	Ahmad Ampeng
24	[2]AAMNQ2	*M. nemestrina*	Lundu, Sarawak	Ahmad Ampeng
25	[2]AAMNQ287	*M. nemestrina*	Lundu, Sarawak	Ahmad Ampeng
26	[2]AAMNQ289	*M. nemestrina*	Lundu, Sarawak	Ahmad Ampeng
27	[2]AAMNQ342	*M. nemestrina*	Lundu, Sarawak	Ahmad Ampeng
28	AY151101	*M. nemestrina*	Sumatra	Roos et al. (2003)
29	AY151100	*M. nemestrina*	Sumatra	
30	AY151099	*M. nemestrina*	Sumatra	
31	AY151095	*M. nemestrina*	Sumatra	
32	AY151096	*M. nemestrina*	Sumatra	
33	HM071134	*M. nemestrina*	Borneo	Evans et al. (2010)
34	HM071133	*M. nemestrina*	Borneo	
35	HM071132	*M. nemestrina*	Borneo	
36	HM071131	*M. nemestrina*	Borneo	
37	HM071130	*M. nemestrina*	Borneo	
38	DQ355484	*M. nemestrina*	Borneo	Ziegler et al. (2007)
39	DQ355483	*M. nemestrina*	Borneo	
40	AY151110	*M. siberu*	Pulau Siberu, Indonesia	Roos et al. (2003)
41	AY151111	*M. siberu*	Pulau Siberu, Indonesia	
42	AY151112	*M. siberu*	Pulau Siberu, Indonesia	
43	AY151113	*M. siberu*	Pulau Siberu, Indonesia	
44	HM071136	*M. leonina*	Thailand	Evans et al. (2010)
45	DQ355487	*M. leonina*	Laos	Ziegler et al. (2007)
46	DQ355488	*M. leonina*	Laos	
47	DQ355489	*M. leonina*	Vietnam	
48	AY151127	*M. pagensis*	Pulau Pagai, Indonesia	Roos et al. (2003)
49	AY151128	*M. pagensis*	Pulau Pagai, Indonesia	
50	AY151129	*M. pagensis*	Pulau Pagai, Indonesia	
51	AY151119	*M. pagensis*	Pulau Sipora, Indonesia	
52	AY151120	*M. pagensis*	Pulau Sipora, Indonesia	
53	AY151121	*M. pagensis*	Pulau Sipora, Indonesia	
54	NC002764	*M. sylvanus*	-	Arnason et al. (2000)

v3.1 (Excoffier et al. 2005) to investigate population expansion events in mismatch distribution (Rogers and Harpending 1992) using pairwise differences in the sequences along with the sudden expansion model (Slatkin and Hudson 1991; Rogers and Harpending 1992) and spatial expansion model (Ray et al. 2003; Excoffier et al. 2005). Furthermore, the relationships among haplotypes obtained were reconstructed by assuming that at any given site, two randomly drawn haplotypes were unlikely to have arisen from more than one mutational step (Alexandrino et al. 2002). These relationships were determined in the form of minimum-spanning network (MSN) drawn in Network 4.6.1.2.

Three criteria were used to build the phylogenetic tree in this study: distance-based criteria producing Neighbor-Joining tree (NJ), character-based criteria producing Maximum Parsimony tree (MP), and Bayesian Inference (BI) phylogenetic tree. Both NJ and MP trees were analyzed in MEGA6 (Tamura et al. 2013) and PAUP v4.0b10 (Swofford 2002). The NJ phylogenetic tree was reconstructed using the Kimura-2-Parameter algorithm, and 1,000 bootstrap replications were performed. Tree bisection and reconnection (TBR) algorithms, heuristic searching method, 1,000 random stepwise additions, and 1,000 bootstrap replications were applied to find the best MP tree using the 50% consensus majority rule. The best substitution models for all loci were determined using Modeltest v3.7 (Posada and Crandall 1998) by means of Akaike information criterion (AIC) requirements (Table 2). This information was then used in MrBayes v3.1 (Huelsenbeck and Ronquist 2001) to reconstruct a BI tree. Metropolis-coupled Markov chain Monte Carlo (MCMC) was run with 10 million generations, and the tree was sampled every 1,000 generations. Split-frequencies probabilities (*P*) of 0.003199 (D-loop), 0.002585 (*COI*) and 0.003059 (Cyt *b*) were obtained. The first 10% of the trees obtained in the analysis were discarded as burn-in (1,000 trees discarded from a total of 10,000 trees). A majority-rule consensus for the remaining trees was constructed, and the posterior probabilities (*PP*)were summarized for each branch.

The molecular clock phylogenetic tree revealing divergence times of *M. nemestrina* genealogy were estimated using BEAUti v1.7.5 and BEAST v1.7.5 (Drummond et al. 2012). Two datasets and an ingroup dataset comprising all samples—excluding *M. sylvanus* and the outgroup sample containing only *M. sylvanus*—were built in BEAUti v1.7.5. *Macaca sylvanus* which was estimated to diverge as early as 5.5 MYA based on fossil data (Delson 1996; Rook et al. 2004) was used as the calibration point to estimate the molecular divergence of *M. nemestrina* and its lineages. The Unweighted Pair Group Method with Arithmetic Mean approach was used to identify the initial topology of the tree. Subsequently, the Birth-Death speciation model (Gernhard 2008a b) was employed to reconstruct the phylogenetic tree, as it assumes the birth and death of a lineage occurring at constant rate and is independent from external factors. The uncorrelated lognormal relaxed-clock model (Drummond et al. 2006) was used to estimate the substitution rate for all nodes in the tree with uniform priors on the mean (0,10), standard deviation (0, 10), and ucld.mean at uniform rate (0.033, upper limit = 1, lower limit = 0). The MCMC tree was run for 10 million generations, and trees were sampled

Table 2. Nucleotide Substitution models determined by Modeltest v3.7 (Posada and Crandall 1998) by means of Akaike information criterion (AIC) requirements

Parameter	D-loop	*COI*	Cyt *b*
Best Nucleotide Substitution Model	TrN+I+G	HKY+G	TrN+I+G
Proportion of invariable sites (I)	0.4362	0.0000	0.5323
Rates	Gamma	Gamma	Gamma
Variable Sites (G)	0.4473	0.5283	0.9077
Base Frequency:			
Adenosine	0.3108	0.2518	0.2657
Cytosine	0.3104	0.3077	0.3379
Guanine	0.1311	0.1617	0.1353
Thymine	0.2478	0.2788	0.2611
Nucleotide Substitution Model:			
A-C	1.0000	0.5302	1.0000
A-G	25.7608	26.7354	55.8441
A-T	1.0000	1.6887	1.0000
C-G	1.0000	1.1433	1.0000
C-T	15.2268	20.6501	27.6800
G-T	1.0000	0.0000	1.0000

every 1,000 generations with 10% of the initial tree discarded as burn-in. Tracer version 1.5 was used to assess the estimated sample size (ESS) from the log files produced by BEAST. After 10 million generations, the ESS of all parameters (posterior, prior, likelihood, ucld.mean, etc.) well exceeded 200, suggesting that the MCMC steps were more than adequate. Maximum-clade-credibility tree were then built using TreeAnnotator v1.7.5 (Rambaut and Drummond 2013).

RESULTS

Sequence Analysis and Genetic Distance

We obtained reliable sequences of D-loop, *COI* and Cyt *b* fragments of ~500, ~650, and ~750 bp, respectively, from all analysed samples. All sequences were blasted in GenBank BLASTn to validate the targeted loci and species with minimum query coverage and max ident of 99% and an E value = 0. Both D-loop and *COI* sequences were aligned with *M. fascicularis* and *M. sylvanus* alone, while Cyt *b* sequences were aligned with additional taxa, namely *M. leonina*, *M. siberu*, and *M. pagensis*. Final alignments of the sequences produced 398 bp, 577 bp, and 567 bp for D-loop, *COI,* and Cyt *b*, respectively. D-loop region sequences showed the highest percentage of parsimony informative characters with 17.6%, followed by Cyt *b* (11.3%) and *COI* (9.5%).

The pairwise genetic distance (Tables 3, 4) revealed unresolved relationships among populations in the Malay Peninsula, in the states of Kelantan, Johor, Pahang, and Kedah. The genetic distance between these populations ranged from 0.018–0.027 (D-loop), 0.002–0.004 (*COI*), and 0.002–0.005 (Cyt *b*). However, Perak's population hinted at a more distant relationship compared to other states with genetic distances as high as 0.064, 0.011, and 0.010 for D-loop, *COI*, and Cyt *b* sequences, respectively. Borneo's population indicated a clear separation from the Peninsula's population, with a genetic distance as much as 0.128 (D-loop), 0.105 (*COI*), and 0.115 (Cyt *b*). A far higher notation of genetic distances were observed between Borneo's population and Perak's population: 0.142 and 0.119 for D-loop and Cyt *b* sequences, respectively.

Phylogenetic and Molecular Clock Estimation

Three strongly supported monophyletic clades

Table 3. Genetic distance using Kimura-2-Parameter of *M. nemestrina* population in Malaysia based on D-loop and *COI* (in parentheses)

	1	2	3	4	5	6
[1] Kelantan						
[2] Johor	0.026 (0.004)					
[3] Pahang	0.024 (0.004)	0.027 (0.003)				
[4] Kedah	0.026 (0.002)	0.018 (0.002)	0.027 (0.002)			
[5] Perak	0.062 (0.011)	0.064 (0.011)	0.066 (0.007)	0.062 (0.009)		
[6] Sarawak	0.128 (0.105)	0.121 (0.100)	0.128 (0.100)	0.125 (0.102)	0.142 (0.096)	- (-)

Table 4. Genetic distance (Cyt *b*) of *M. nemestrina* population in Malaysia as compared to other *Macaca* species based on Kimura-2-Parameter

	1	2	3	4	5	6	7	8	9	10	11
[1] Kelantan											
[2] Johor	0.002										
[3] Pahang	0.005	0.004									
[4] Kedah	0.004	0.002	0.005								
[5] Perak	0.010	0.009	0.008	0.010							
[6] Sarawak	0.115	0.113	0.113	0.113	0.119						
[7] Sumatra	0.013	0.012	0.012	0.013	0.013	0.113					
[8] Borneo	0.108	0.106	0.106	0.106	0.112	0.035	0.104				
[9] *M. leonina*	0.053	0.051	0.051	0.053	0.053	0.116	0.049	0.112			
[10] *M. siberu*	0.038	0.037	0.037	0.035	0.036	0.107	0.035	0.102	0.043		
[11] *M. pagensis*	0.062	0.061	0.061	0.063	0.059	0.104	0.062	0.109	0.074	0.065	-

were detected in congruent within all three NJ (Fig. 1), MP (Fig. 2), and BI (Fig. 3) phylogenetic trees namely Borneo's clade, Perak's clade (clade A), and the Peninsula's clade (clade B). Clade B was designated as the Peninsula's clade because it encompasses all populations in the Malay Peninsula, excluding Perak's population. Borneo's clade diverged earlier than the Peninsula's population and was supported by a minimum of 98% of bootstrap values and 0.99 posterior probabilities (*PP*) across all particular phylogenetic trees. In Cyt *b* phylogenetic trees, the sequences of Borneo's population extracted from GenBank were also consistently grouped together in all trees.

The monophyletic state of the Malay Peninsula's population (including Perak) was consistently supported with no less than 84% of bootstrap values and 0.99 *PP*. Nevertheless, the early divergence of Clade A was not anticipated by any means. This clade was supported unconditionally with perfect score of 1.00 *PP* and a minimum of 82% of bootstrap values. Conversely, Clade B—encompassing from Kedah, Pahang, Kelantan, and Johor—showed an unresolved relationship among the populations, proving a much closer relationship between each other than to Perak's populations.

The time-scaled phylogenetic tree (Fig. 4) indicated that *M. nemestrina* originating from Borneo experienced the earliest divergence across all pig-tailed macaque in Southeast Asia, estimated at approximately 4.57 MYA. Before the subsequent diversification of the *M. nemestrina* lineage occurred, *M. pagensis* was observed to have formed its species group at approximately 2.51 MYA. *Macaca siberu* and *M. leonina* diverged from *M. nemestrina* lineages at only 1.67 MYA, and they diverged from each other at approximately 1.06 MYA. Sumatran pig-tailed macaque finally diverged from the Malay Peninsula's population at approximately 0.51 MYA. Perak's population subsequently separated from the rest of the Malay Peninsula's populations as early as 0.32 MYA.

Population Genetic Analyses

The distant relationship between the Borneo and Malay Peninsula's populations detected in previous analyses was maintained as indicated by population genetic analyses. The highest π and Da values were detected between these two populations as high as 0.08 (D-loop), 0.05 (*COI*), and 0.059 (Cyt *b*) for π and 0.1076 (D-loop), 0.0920 (*COI*), and 0.105 (Cyt *b*) for Da. In addition, F_{ST} and Nst were also the highest between these populations with 0.8647 (D-loop), 0.9950 (*COI*), and 0.990 (Cyt *b*) for F_{ST} and 0.8720 (D-loop), 0.9953 (*COI*), and 0.990 (Cyt *b*) for Nst. The lowest number of migrants between these populations further strengthen the almost cut-off gene flow relationship between these populations with as little as 0.18 (D-loop), 0.01 (*COI*), and 0.01 (Cyt *b*) for Nm (Tables 5, 6).

The Peninsula's population (excluding Perak's population) revealed unresolved segregations between each other and does not exhibit any separation or segregation paradigm. The statistical values obtained for π, Da, Nst, F_{ST}, and Nm between the populations of Kelantan, Pahang, Kedah, and Johor were deemed as not significant enough to support any relationship between them. Perak's population by contrast implies that they are somewhat separated from the rest of the Peninsula's populations. Nucleotides diversity and divergence between Perak's population and the Peninsula's population were significant enough to support this separation as high as 0.0365 (D-loop), 0.0038 (*COI*), and 0.006 (Cyt *b*) for π, and 0.0543 (D-loop), 0.0104 (*COI*), and 0.010 (Cyt *b*) for Da. Nucleotide and population subdivision between these populations were also in congruent with π and Da exhibiting F_{ST} as high as 0.9028 (D-loop), 0.9730 (*COI*), and 0.958 (Cyt *b*) and Nst as much as 0.9061 (D-loop), 0.9731 (*COI*), and 0.959 (Cyt *b*). Another startling fact was the small number of migrants between these populations with 0.25 (D-loop), 0.01 (*COI*), and 0.08 (Cyt *b*), suggesting some sort of barriers to gene flow.

A total of 29 (Cyt *b*), 19 (D-loop), and 10 (*COI*) haplotypes were detected across *M. nemestrina* populations in Malaysia (Tables 7, 8, 9). Borneo's population indicated the presence of 4 (D-loop), 2 (*COI*), and 1 (Cyt *b*) haplotypes with a wide range of haplotype diversity—0.900 ± 0.161 (D-loop), 0.533 ± 0.172 (*COI*), and 0.031 ± 0.175 (Cyt *b*) (Tables 10, 11). Perak's population showed 6 (D-loop) and 2 (*COI* and Cyt *b*) haplotypes with haplotype diversities as high as 0.952 ± 0.096 (D-loop), 0.333 ± 0.046 (*COI*), and 0.476 ± 0.171 (Cyt *b*). The number of haplotypes and diversity across the remaining Peninsula's populations was consistently low, except for Kelantan's population.

MSNs were generated using the haplotypes obtained to visualize the relationship between the population of *M. nemestrina* in Malaysia and those of *M. siberu*, *M. pagensis*, and *M. leonina* (Fig. 5). Haplotypes present in Borneo's population displayed 38 (D-loop) and 48 (*COI*) mutational steps to the nearest Malay Peninsula's haplotype. Furthermore, within the MSN based on the Cyt *b* sequence, which also contained another sequences representing Borneo, the smallest mutational distance between haplotype of these two populations was 44 steps, altogether proving the length of the relationship detected throughout this study. The constant assimilations inside of the Peninsula's population revealed in this study was again proven in the MSN analysis, showing only 2 (Cyt *b*), 3 (*COI*), and

9 (D-loop) mutational steps between their haplotypes. Peninsula's population also showed 4 mutational steps compared to Sumatran *M. nemestrina* populations in MSN for Cyt *b* sequences. The peculiar relationships between Perak's population and Peninsula's population similarly hinted at a population separation with 18 (D-loop) and 4 (Cyt *b* and *COI*) mutational steps between their haplotypes. Interestingly, Perak's haplotypes were only separated by 2 mutational steps from Sumatran's haplotypes, much closer than the Peninsula's haplotypes. Although *M. pagensis* were the earliest to diverge from *M. leonina*, *M. siberu*, and *M.*

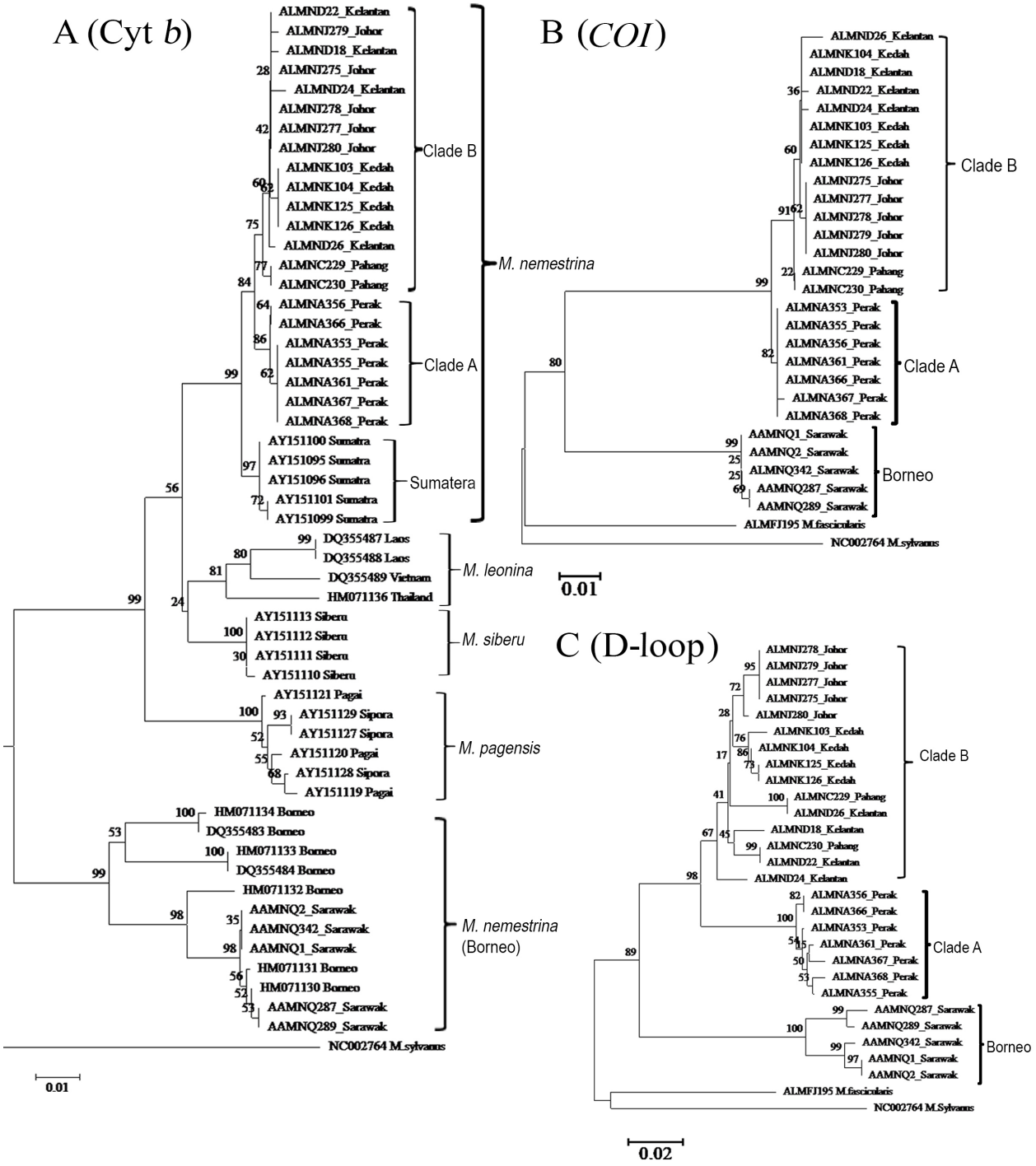

Fig. 1. Neighbor-Joining phylogenetic tree reconstructed based on Kimura-2-Parameter for three loci (Cyt *b*, D-loop and *COI*) with 1000 bootstrap replications. Bootstrap values indicated on the nodes.

nemestrina in all phylogenetic trees, the MSNs revealed that *M. pagensis* showed a more distant relationship to Borneo's population as compared to *M. siberu* and *M. leonina*. The MSNs also revealed a much closer relationship between *M. leonina* (24 steps) and *M. siberu* (18 steps) in Sumatran's population as compared to *M. pagensis* (30 steps).

The demographical history of *M. nemestrina* was investigated through four different indices: D, F_S, F*, and D* (Table 12). Positive values for D in Perak, Borneo, Sumatra, and Borneo populations suggested an excess of rare polymorphisms present in the population,

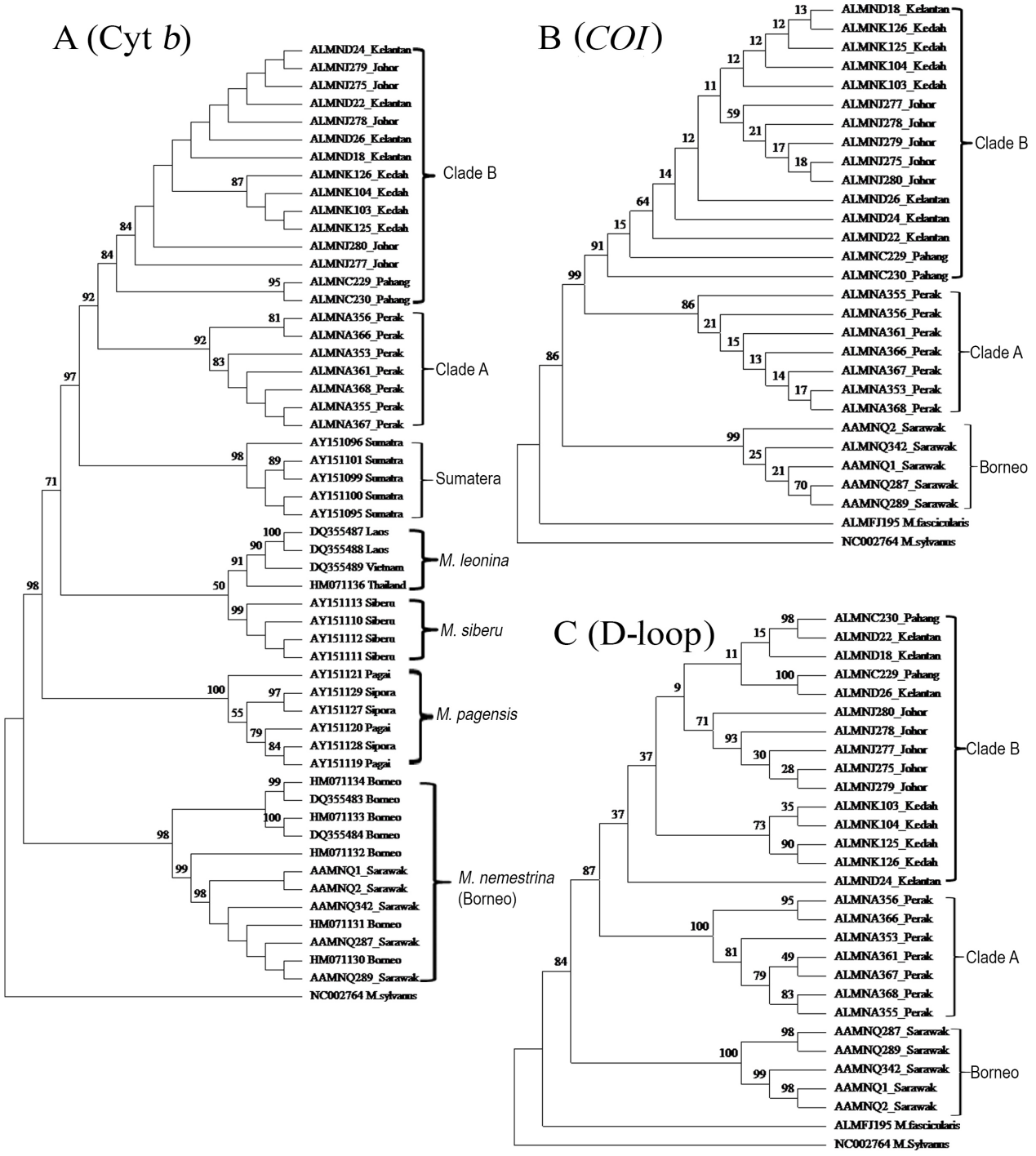

Fig. 2. Maximum Parsimony phylogenetic tree reconstructed for three loci (Cyt *b*, D-loop and *COI*) based on tree bisection and reconnection (TBR) algorithms, heuristic searching method, 1,000 random stepwise additions and 1,000 bootstrap replications. Bootstrap values indicated on the nodes.

indicating a positive selection, while negative values in other populations pointed to an excess of high-frequency variants, indicating balancing selection (Oleksyk et al. 2010). However, all populations failed to reject the null hypothesis at $P > 0.10$. The presence of a unique haplotype indicated by negative values of F_S, that is Kelantan (-1.87) and Perak (-0.003), were acquired. This result could determine the uniqueness of Perak's gene pool parallel with other results projected throughout the analyses. However, a lack of significance at $P > 0.10$ could not determine whether this was due to population expansion events or genetic hitchhiking (Fu 1997). Both D* and F* were negative for all *M. nemestrina* populations in the Malay Peninsula, except for Perak's population (Cyt *b*). These indicates the presence of new haplotypes in the population (Fu and Li 1993), although these haplotypes were not significant at $P > 0.10$.

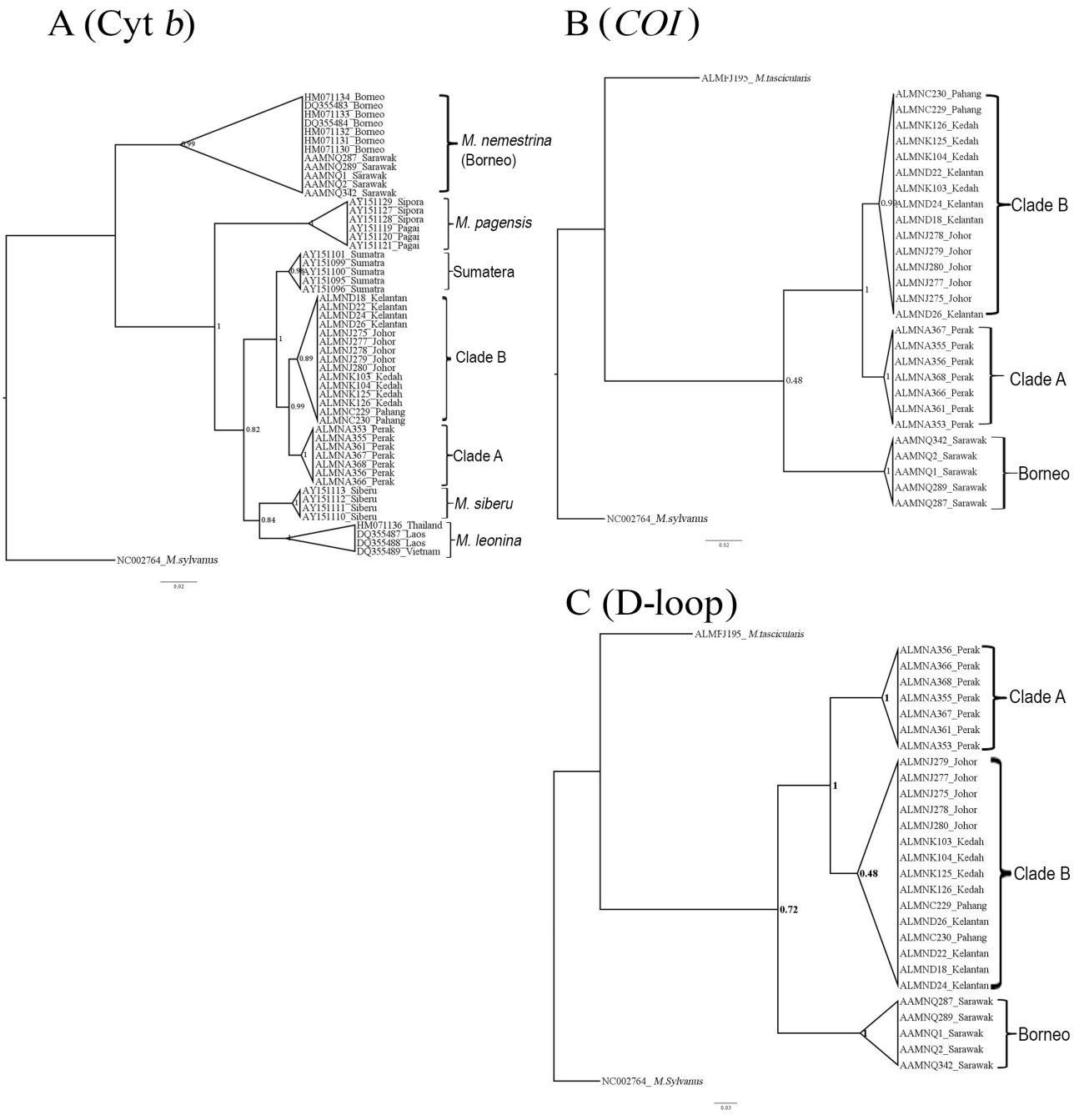

Fig. 3. Bayesian inference of the 50% majority rule consensus tree for three loci (Cyt *b*, D-loop and *COI*) with bayesian posterior probability (*PP*) accordingly indicated on the branch.

Mismatch distribution analyses (Fig. 6) with their respective parameters are listed in tables 13, 14, and 15. The variance in the mismatch distribution was too small for several populations; thus, no demographic parameters could be estimated. Though varied outcome can be interpreted from the distributions, Perak's population distributions were largely noticeable among the rest. This distribution exhibited a bimodal distribution (D-loop), which indicates either two population expansions or the presence of two or more mixed populations that subsequently expanded (Jalil et al. 2008). The latter is more unlikely as all the analyses pointed to a population separation, and no shared haplotypes were detected. Unimodal distributions were also identified for Perak's population (*COI* and Cyt *b*), which hinted at recent population expansion in the past (Rogers and Harpending 1992; Slatkin and Hudson 1991).

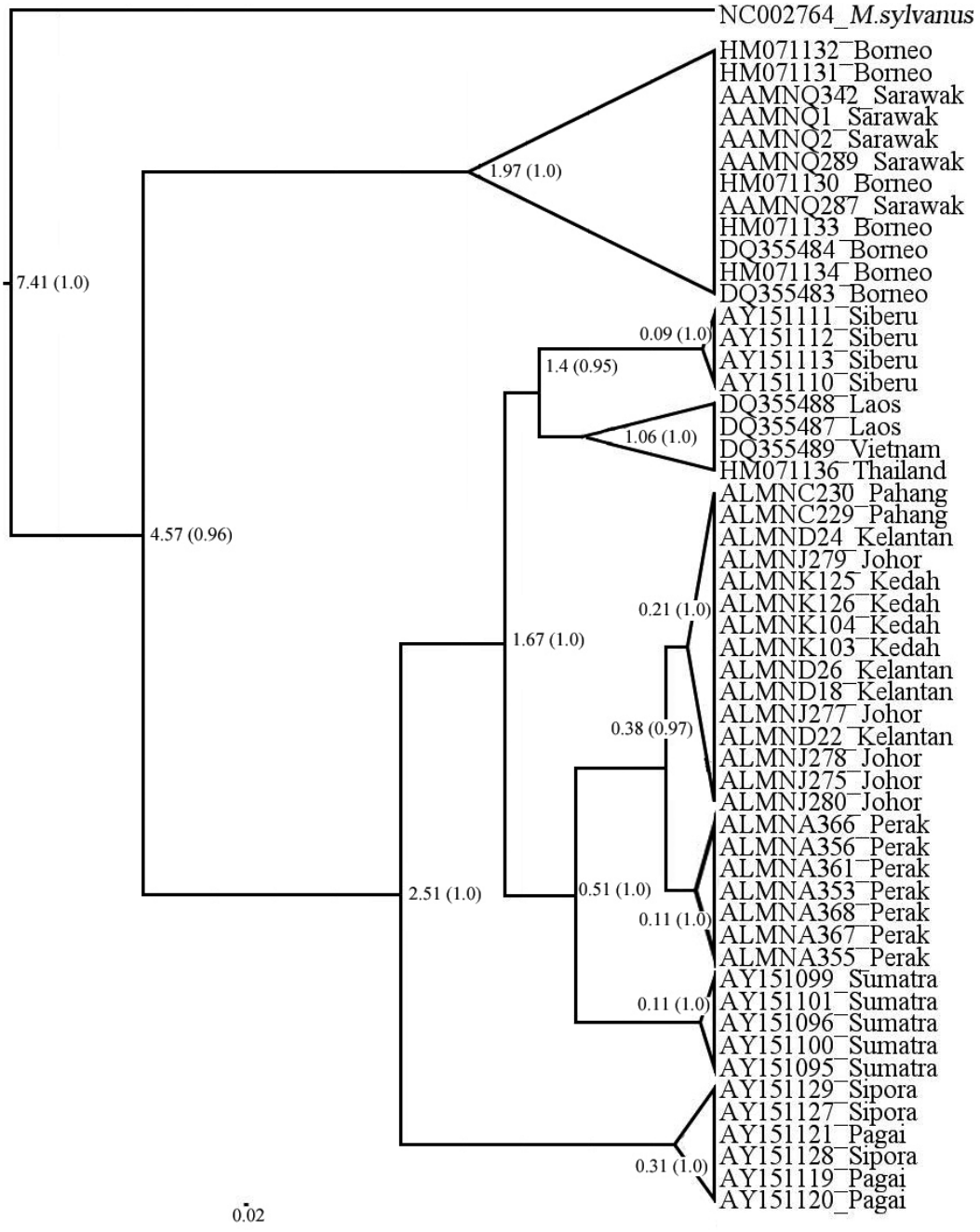

Fig. 4. Molecular divergence tree of the Cyt *b* sequence of Southeast Asia's *M. nemestrina*, *M. pagensis*, *M. siberu*, and *M. leonina* with the numbers on the nodes representing divergence time in millions of years (MYA) and *PP* in parentheses indicating the confidence interval.

DISCUSSION

Bornean pig-tailed macaque populations are the first extant lineages to separate from other examined lineages

Borneo's pig-tailed macaques demonstrate the most distant relationships of all pig-tailed macaques, including *Macaca siberu*, *M. pagensis*, and *M. leonina*. This result raised many questions over the taxonomic positions of Borneo's populations with respect to other Southeast Asian pig-tailed macaques. The result obtained in this study was not the first to report the peculiarity surrounding *M. nemestrina* of Borneo's lineages. Evans et al. (1999) found a paraphyletic relationship between Bornean *M. nemestrina* and *M. tonkeana* (Sulawesi macaque), while the study by Tosi et al. (2000) reveals incongruence of mtDNA and Y chromosome markers as windows onto species phylogeny. Further investigation by Tosi et al. (2003) successfully revealed three distinct mitochondrial linegaes of *M. nemestrina*, namely the Isthmus Kra group, Sumatra-Mentawai cluster, and Bornean aggregate. Evans et al. (2003), in his review of Sulawesi macaque dispersal theory, tried to stitch together all previous data by suggesting that the tendency of speciation is higher in small isolated populations (*i.e.*, Sulawesi and Mentawai). Evans et al. (2003) also acknowledge the inconsistencies of nuclear data with mtDNA markers and justified it with the male migration mediated gene flow.

With an abundance of data and discussion on the origins and dispersal of Sulawesi macaques, in this article, we focus on pig-tailed macaques alone. The results in this study suggest that populations of *M. nemestrina* in Borneo embody the oldest genetic lineages among the species groups, including *M. pagensis*, *M. leonina*, and *M. siberu*. The ancestral states of *M. nemestrina* were discussed at length by Tosi et al. (2003), suggesting that ancestral *M. nemestrina* exhibited *M. silenus*-like morphology and only attained their current form after separation from extant *M. silenus*. While many would argue how Bornean populations were the oldest genetic stock, this idea is contradicting with natural radiation and dispersal mechanism of other species such as *M. fascicularis* (Abdul-Latiff et al. 2014a b). Over the past 30 million years, climatic changes in the Sunda Shelf (Abegg and Thierry 2002) has been the main hypothesis in justifying the presence and absence of particular primate taxa in the region. The sophisticated events of active tectonics, rising and falling of sea levels, along with drastic climatic changes in Southeast Asia (Holloway and Hall 1998; Meijaard 2003; Voris 2000) ultimately influenced the small isolated populations in the Malay Peninsula, Sumatra, and Sulawesi. However, populations of *M. nemestrina* in Borneo, the world's third largest island (743,000 km^2) (Meiri et al. 2008), may be able to withstand this evolutionary pressure. Populations of pig-tailed macaques here may well be large enough to buffer against founder and bottleneck effects as they exhibit a larger effective population size (Evans et al. 2003). As a result, they became the genetic stocks of all

Table 5. Measures of nucleotide diversity (π), net nucleotide divergence (Da), nucleotide subdivision (Nst), estimate of population subdivision (F_{ST}) and gene flow (number of migrants, Nm) between population of *M. nemestrina* in Malaysia based on D-loop and *COI* (in parantheses). Kelantan, Johor, Pahang, Kedah and Perak represent Malay Peninsula's population, while Sarawak represent Borneo's population

Population	π	Da	F_{ST}	Nst	Nm
Kelantan-Johor	0.0198 (0.0029)	0.0083 (0.0017)	0.3333 (0.4444)	0.3315 (0.4442)	1.00 (0.63)
Kelantan-Pahang	0.0268 (0.0038)	0.0082 (0.0017)	0.3514 (0.4444)	0.3512 (0.4442)	1.92 (0.31)
Kelantan-Kedah	0.0221 (0.0022)	0.0074 (0.0000)	0.2917 (0.3570)	0.2908 (0.3257)	1.21 (0.83)
Kelantan-Perak	0.0365 (0.0067)	0.0391 (0.0087)	0.6677 (0.7792)	0.6752 (0.7802)	0.25 (0.14)
Kelantan-Sarawak	0.0757 (0.0513)	0.0857 (0.0920)	0.7519 (0.9722)	0.7661 (0.9739)	0.16 (0.01)
Johor-Pahang	0.0154 (0.0017)	0.0081 (0.0035)	0.3107 (0.3926)	0.3087 (0.3595)	1.11 (0.36)
Johor-Kedah	0.0115 (0.0010)	0.0133 (0.0017)	0.7578 (0.3265)	0.7594 (0.3327)	0.16 (0.41)
Johor-Perak	0.0351 (0.0060)	0.0543 (0.0104)	0.9028 (0.9730)	0.9061 (0.9731)	0.05 (0.01)
Johor-Sarawak	0.0667 (0.0497)	0.0936 (0.0902)	0.8647 (0.9949)	0.8720 (0.9952)	0.08 (0.00)
Pahang-Kedah	0.0185 (0.0009)	0.0074 (0.0017)	0.2778 (0.3648)	0.2771 (0.3626)	1.30 (0.53)
Pahang-Perak	0.0303 (0.0034)	0.0419 (0.0069)	0.6686 (0.9600)	0.6766 (0.9602)	0.25 (0.02)
Pahang-Sarawak	0.0682 (0.0394)	0.0841 (0.0902)	0.7400 (0.9949)	0.7547 (0.9952)	0.18 (0.00)
Kedah-Perak	0.0335 (0.0050)	0.0511 (0.0087)	0.8786 (0.9677)	0.8826 (0.9679)	0.07 (0.02)
Kedah-Sarawak	0.0704 (0.0496)	0.0961 (0.0920)	0.8583 (0.9950)	0.8664 (0.9953)	0.08 (0.00)
Perak-Sarawak	0.0730 (0.0481)	0.1076 (0.0868)	0.8604 (0.9914)	0.8698 (0.9919)	0.08 (0.00)

Table 6. Measures of nucleotide diversity (π), net nucleotide divergence (Da), nucleotide subdivision (Nst), estimate of population subdivision (F_{ST}) and gene flow (number of migrants, Nm) between population of *M. nemestrina* in Malaysia and other species of *Macaca* based on Cyt *b* sequences. Kelantan, Johor, Pahang, Kedah and Perak represent Malay Peninsula's population, while Sarawak represent Borneo's population

Population	π	Da	F_{ST}	Nst	Nm
Kelantan-Johor	0.002	0.000	0.000	0.000	0.63
Kelantan-Pahang	0.004	0.004	0.667	0.667	0.13
Kelantan-Kedah	0.003	0.002	0.500	0.500	0.25
Kelantan-Perak	0.006	0.007	0.769	0.770	0.08
Kelantan-Sarawak	0.059	0.100	0.973	0.975	0.01
Kelantan-Sumatra	0.008	0.011	0.824	0.825	0.05
Kelantan-Borneo	0.065	0.077	0.782	0.791	0.07
Kelantan-*M. leonina*	0.037	0.022	0.689	0.692	0.11
Kelantan-*M. siberu*	0.034	0.053	0.940	0.941	0.02
Kelantan-*M. pagensis*	0.034	0.053	0.903	0.906	0.03
Johor-Pahang	0.002	0.004	0.909	0.909	0.03
Johor-Kedah	0.001	0.002	0.833	0.833	0.05
Johor-Perak	0.005	0.008	0.911	0.911	0.02
Johor-Sarawak	0.057	0.101	0.986	0.987	0.00
Johor- Sumatra	0.007	0.011	0.924	0.924	0.02
Johor- Borneo	0.064	0.077	0.794	0.801	0.06
Johor-*M. leonina*	0.035	0.035	0.709	0.711	0.10
Johor-*M. siberu*	0.020	0.035	0.978	0.979	0.01
Johor-*M. pagensis*	0.034	0.053	0.927	0.929	0.02
Pahang-Kedah	0.003	0.005	1.000	1.000	0.00
Pahang-Perak	0.004	0.008	0.949	0.950	0.01
Pahang-Sarawak	0.005	0.101	0.990	0.990	0.00
Pahang- Sumatra	0.006	0.011	0.953	0.953	0.01
Pahang- Borneo	0.060	0.077	0.797	0.804	0.06
Pahang-*M. leonina*	0.035	0.035	0.714	0.716	0.10
Pahang-*M. siberu*	0.019	0.035	0.988	0.988	0.00
Pahang-*M. pagensis*	0.029	0.053	0.932	0.935	0.02
Kedah-Perak	0.005	0.010	0.958	0.959	0.01
Kedah-Sarawak	0.057	0.101	0.990	0.990	0.00
Kedah- Sumatra	0.008	0.013	0.959	0.960	0.01
Kedah- Borneo	0.064	0.077	0.797	0.804	0.06
Kedah-*M. leonina*	0.036	0.037	0.724	0.726	0.10
Kedah-*M. siberu*	0.020	0.034	0.987	0.987	0.00
Kedah-*M. pagensis*	0.034	0.055	0.934	0.937	0.02
Perak-Sarawak	0.057	0.105	0.986	0.987	0.00
Perak- Sumatra	0.008	0.012	0.924	0.925	0.02
Perak- Borneo	0.064	0.081	0.802	0.810	0.06
Perak-*M. leonina*	0.035	0.036	0.717	0.720	0.10
Perak-*M. siberu*	0.018	0.034	0.975	0.976	0.01
Perak-*M. pagensis*	0.032	0.052	0.923	0.926	0.02
Sarawak-Sumatra	0.058	0.101	0.985	0.986	0.00
Sarawak- Borneo	0.030	0.013	0.379	0.380	0.41
Sarawak-*M. leonina*	0.064	0.090	0.856	0.863	0.04
Sarawak-*M. siberu*	0.055	0.096	0.985	0.986	0.00
Sarawak-*M. pagensis*	0.054	0.089	0.948	0.951	0.01
Sumatra-Borneo	0.063	0.075	0.788	0.795	0.07
Sumatra-*M. leonina*	0.034	0.032	0.689	0.691	0.11
Borneo-*M. leonina*	0.069	0.068	0.668	0.682	0.12
M. leonina-M. siberu	0.032	0.027	0.654	0.656	0.13
M. leonina-M.pagensis	0.046	0.051	0.742	0.749	0.09
M. pagensis-M. siberu	0.035	0.057	0.929	0.932	0.02

Table 7. Segregating sites (123 bp) in the 567-bp segment of Cyt *b* sequences defining 29 haplotypes and their distribution across *M. nemestrina* population in Malaysia and other *Macaca* species groups. Haplotype 1, 17-18 – *M. leonina*; 2-7 – *M. nemestrina* (Borneo); 8-9 – *M. siberu*; 10-11 – *M. nemesterina* (Sumatra); 12-16 – *M. pagensis*; 19,21,22 – Kelantan; 20,23 – Johor; 24 – Pahang; 25 – Kedah; 26-27 – Perak; 28-29 – Sarawak

Haplotype	Sites
[	1111111 1111111111 1112222222 2222222222 2222222333 3333333333 3333333334 4444444444 4444444445 5555555555 555]
[	13333334 4444556677 9990001223 4555557788 8890001112 2334444556 6677889000 1224455677 7888889990 0012333334 4456677891 1122444555 666]
[	3920347892 6789180735 3460581068 1034591706 7851480343 5470369134 6706251039 5475847702 6014786790 2523245781 4965817001 3625039258 247]
#Hap_1	CAACTAGCCC ACTCAGCTTA AGTATCTTTT GCAATTCTTT GCTCTCATAG CCTACTTAAC TCAACTAATA TAAAGCATCA GTTATATATA CCAATGCACC TTCACCACGC GTACCCCATC GCG
#Hap_2	...T.G.... GT...ATC.G ..CTCTC... ..G.C..C.C ATC....CG. ...G.CCTG. .TG..CC.CG .GG.AA.C.. ..CG.GCG.. .T.G..T.T. CC........ AC..T.T... ..A
#Hap_3	...T.G...T .T...ATC.. ..CTCTC... AT..C..C.C ATC..T.CG. .T..TCCTG. .TC.TCCG.G .GG..ATC.. ..C...CG.. .T.GC.T.T. CCTG...... A...T.T... ..A
#Hap_4	...T.G.... GT...AT... ..CT.TCCC. ..G..C.C.. ATC..T.CG. .T.GT.CTG. .T..ACCG.G .GG..A..T. .CCG..CG.. .T.GC.T... CC...T.A.. A...TTTGC. A.A
#Hap_5	...T.G.... GT...ATC.. ..CT.TCC.. A.G..C.C.. ATC..T.CG. .T.GTCCTG. .T..ACCG.. .CG..A..T. ..CG..CG.. TT.GC.T... CC...TG... A.G.T.TG.. A.A
#Hap_6	...T.G.... GT...ATC.. ..CT.TCC.. A.G..C.C.. ATC..T.CG. .T.GTCCTG. .T..ACCG.G .CG..A..T. A.CG..CG.. TT.GC.T... CC...TG... A.G.T.TG.. A.A
#Hap_7	...T.G.... GT...ATC.G ..CTCTC... ..G.C..C.C ATC....CG. ...G.CCTG. .TG..CC.CG .GGGAA.C.. ..CG.GCG.. .T.G..T.T. CC........ AC..T.T... ..A
#Hap_8	.G........ .....A.... ...T..C..C ..G.....C. ..C....C.. ....GG. ..G.T...CG ...G...... ..C..... ....G.. .C....G.A. A......... ..A
#Hap_9	.G........ .....A.... ...T..C..C ..G.....C. ..C....C.. T......GG. ..G.T...CG ...G...... ..C..... ....G.. .C....G.A. A......... ..A
#Hap_10	.G...G.... .....A.... ..CT..C.C. C.G..... ..C...GC.A ..C.....G. ..G.T....G ...G...C.G ..CC..... ....G.T .C.G..... A......... ...
#Hap_11	.G...G.... .....A.... ..CT..C.C. C.G..... ..C....C.A ..C.....G. ..G.T....G ...G...C.G ..CC..... ....G.T .C.G..... A......... ...
#Hap_12	TG..CGA... .T..GA.... G.CTC.C... A.G..CTC.. ..C....C.. .TC...C.G. .TG.T....G ..GG..... ..CGC..... T.G.....T .C..TT.TA. A........T ATA
#Hap_13	TG..CGA... .T..GA.... G.CT..C... A.G..CTC.. ..C....C.. .TC...C... .TGGT....G ...GA..... ..C.C..... T........T .C..TT.TA. A........T ATA
#Hap_14	TG..CGA... .T..GA.... G.CT..C... A.G..CTC.. ..C....C.. .TC...C.G. .TG.T....G ..GG..... ..CGC..... T........T .C..TT.TA. A........T A.A
#Hap_15	TG..CGA... .T..GA.... G.CT..C... A.G..CTC.. ..C.C..C.. .TC...C.G. .TG.T....G ...GA..... ..CGC..... T........T .C..TT.TA. A........T ATA
#Hap_16	TG..CGA... .T..GA.... G.CTC.C... A.G..CTC.. ..C....C.. .TC...C... .TGGT....G ...GA..... ..CGC..... T........T .C..TT.TA. A........T ATA
#Hap_17	.G..C....T ..C..A.C.. ...T..C... ...G.C.... .....C.. ....G.T....G ...GA..... ....G.T....G .C........ AC...T.... A..
#Hap_18	.G........ ..C..A.C.. ...T..C... A......... .....C.. ....T C.G.T....G ...G...... ...G....CC ...G.A...T .C.......T AA........ A..
#Hap_19	..G..G.... .....A.... .ACT..C.C. C.G..... ..C....C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G..... A......... ...
#Hap_20	.....G.... .....A.... .ACT..C.C. C.G..... ..C....C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G..... A......... ...
#Hap_21	.....G.... .....A..C. .ACT..C.C. C.G..... ..C....C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G..... A..T..... ...
#Hap_22	.....G.... .....A.... .ACT..C.C. C.G..... ..C....C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G....A. A......... ...
#Hap_23	.....G.T.. .....A.... .ACT..C.C. C.G..... ..C....C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G..... A......... ...
#Hap_24	.....G.... .....A.... ..CT..C.C. C.G..... ..CT...C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G..... A......... ...
#Hap_25	.....G.... .....A.... .ACT..C.C. C.G..... ..C....C.. ..C....GG. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G..... A......... ...
#Hap_26	.G...G..T. ...T.A.... ..CT..C.C. C.G..... ..C....C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G....A. A......... ...
#Hap_27	.G...G..T. ...T.A.... ..CT..C.C. C.G..... ..C....C.. ..C.....G. ..G.T....G C..G.....G ..CCC..... ....G.T .C.G..... A......... ...
#Hap_28	...T.G.... GT...ATC.. ..CT.TCC.. A.G..C.C.. ATC..T.CG. .T..TCCTG. .T..ACCG.. .CG..A..T. ..CG..CG.. TT.GC.T... CC...TG... A.G.T.TG.. A.A
#Hap_29	...T.G.... GT...ATC.. ..CT.TCC.. A.G..C.C.. ATC..T.CG. .T.GTCCTG. .T..ACCG.. .CG..A..T. A.CG..CG.. TT.GC.T... CC...TG... A.G.T.TG.. A.A

Table 8. Segregating sites (79 bp) in the 398-bp segment of D-loop sequences defining 19 haplotypes and their distribution across the *M. nemestrina* population in Malaysia

Haplotype	Sites
[	11111111 1111111111 1111111111 1111111111 2222222222 2222222222 222233333]
[	5567778888 9900001223 3333334455 5556666667 7777888899 0011111122 2333455666 789911145]
[	4812390189 8914891580 1234671212 3784567891 4689123679 4701237801 5089117345 204807803]
#Hap_1	TCGAACCAAC CACCGACGAC CCTCAATTAC AAATCTCTTC ATGGCACTAC AGGTTCCCGT CGGGCATCCT TTCCGGCAG
#Hap_2	....T...... .....GA... ....C. ..A..G.... .C.CCT.... .AA....... .C.......
#Hap_3	....T...... .T....A... .G........ ..A..G.... .C..C..... .AA..G.... .C..A....
#Hap_4	.....T.... ...TA..... TTC...A... ....A..G.... .C...T.... .AA....... CC...A...
#Hap_5	.....T..G. ..TT...... ......A... ....A..G.... .C...T...C ..A....... CC.......
#Hap_6	.....T.... ..TT...... ......A... ....A..G.... .C...T...C .AA....... .C.......
#Hap_7	...G.T.... ..TT...... ......A... ....A..G.... .C...T...C .AA....... .C.......
#Hap_8	.....T.... ...T...A.. ......A... G......... ..A..G.... .C........ .AA....... .C.....G.
#Hap_9	.....T.... ...T...A.. ......A... ....A..G.... .C...T.... ..A....... .C.....G.
#Hap_10	.TAGC..... ..T.AG.A.. .TCAGGACGT ..G.TCAC.A TC.TTGT.G. .AA.CTTTA. T.A...CTTC CCT...TG.
#Hap_11	.TAGC..... ..TTAG.A.. .TCAGGACGT ..G.TCAC.A TC.T.GT.G. GAA.CTTTA. T.A...CTTC CCT....G.
#Hap_12	.TA.C..... .GT.AGTAG. ..CAGGACGT G....CAC.A TC.T.GT... .AA.CTTTA. TAA.T.CTTC CC.....G.
#Hap_13	.TA.C.T... ..T.AGTAG. ...AGGACGT G....CAC.A TC.T.GT... .AA.CTTTA. TAA.T.CTTC CC.....GA
#Hap_14	A....TT..T T..TA..... .TC..GA.G. ...CT..C.. ..A...TC.. .C..C....C ..AA...... .C...A...
#Hap_15	A....T...T T..TA..A.. .TC..GA.G. ...CT..C.. ..A...TC.T .C..C.T..C ..AA...... .C...A...
#Hap_16	A....T.G.T T..TA..... .TC..GA.G. ...CT..C.. ..A...TC.T .C..C.T..C ..AA...... .C..AA...
#Hap_17	A....T...T T..TA....T .TC..GA.G. ...CT..C.. ..A...TC.T .C..C.T..C ..AA...... .C.T.A...
#Hap_18	A....T...T T..TA..... .TC..GA.G. ...CT..C.. ..A...TC.T .C..C....C ..AA...... .C...A...
#Hap_19	A....T...T T..TA..... .TC..GA... ...CT..C.. ..A...TC.T .C..C.T..C ..AA...... .C.T.A...

Table 9. Segregating sites (61 bp) in the 398-bp segment of *COI* sequences defining 19 haplotypes and their distribution across the *M. nemestrina* population in Malaysia

Haplotype	Sites
	1111111111 1122222222 2222223333 3333333444 4445555555 5
	2233456689 1233445588 8901233445 5689990012 3445689147 8890033446 6
	0358445995 0545032625 8483406280 1113562539 2489832753 2956969080 6
#Hap_1	CTACTCTGAT TGAAACCCTT ATAATCCAGT GGGGTATTGG ATGTATAGAA TACTGGATTC T
#Hap_2	.... T......... ....
#Hap_3	TCGTCTCAGC .AG.GTTTCC GC..CTTGA. AAAA..CC.A .CACGCGATG CGTCAAGCCT C
#Hap_4	TCGTCTCAGC .AG.GTTTCC GC..CTTGA. AAAA.GCC.A .CACGCGATG CGTCAAGCCT C
#Hap_5	TCGTCTCAGC .AGGGTTTCC GC..CTTGA. AAAA.GCC.A .CACGCGATG CGTCAAGCCT C
#Hap_6	TCGTCTCAGC .AG.GTTTCC GC..CTTGA. AAAA.GCCCA .CACGCGATG CGTCAAGCCT C
#Hap_7	TCGTCTCAGC .AG.GTTTCC GCG.CTTGAG AAAAGGCC.A .CACGCGATG CGTCAAGCCT C
#Hap_8	TCGTCTCAGC .AG.GTTTCC GC..CTT.A. AAAA.GCC.A .CACGCGATG CGTCAAGCCT C
#Hap_9	TCGTCTCAGC CAG.GTTTCC G...CTTGA. AAAA..C..A .CACG.GATG CGTCAAGCCT C
#Hap_10	TCGTCTCAGC CAG.GTTTCC G..GCTTGA. AAAA..C..A .CACG.GATG CGTCAAGCCT C

Table 10. Summary statistic of haplotypes from D-loop dan *COI* (parantheses) sequence of *M. nemestrina* population in Malaysia

Population	N	H	S	Hd†	π†	K
Kelantan	4	4	23	1.000 ± 0.177	0.030 ± 0.006	12.00
		(4)	(5)	(1.000 ± 0.177)	(0.004 ± 0.000)	(2.50)
Johor	5	2	3	0.400 ± 0.237	0.003 ± 0.000	1.20
		(1)	(0)	(0.000 ± 0.000)	(0.000 ± 0.000)	(-)
Pahang	2	2	13	1.000 ± 0.500	0.033 ± 0.000	13.00
		(1)	(0)	(0.000 ± 0.000)	(0.000 ± 0.000)	(-)
Kedah	4	3	4	0.833 ± 0.222	0.005 ± 0.000	2.17
		(1)	(0)	(0.000 ± 0.000)	(0.000 ± 0.000)	(-)
Perak	7	6	9	0.952 ± 0.096	0.009 ± 0.000	3.43
		(2)	(1)	(0.333 ± 0.046)	(0.0005 ± 0.00)	(0.33)
Sarawak	5	4	19	0.900 ± 0.161	0.026 ± 0.006	10.40
		(2)	(1)	(0.533 ± 0.172)	(0.0009 ± 0.00)	(0.53)

N, number of sequences analyzed; H, number of haplotypes; S, number of segregating sites; Hd, haplotype diversity; π, nucleotide diversity; K, average number of nucleotide differences.

Table 11. Summary statistic of haplotypes from Cyt *b* sequence of *M. nemestrina* population in Malaysia and respective *Macaca* species

Population	N	H	S	Hd†	π†	K
Kelantan	4	4	4	1.000 ± 0.177	0.003 ± 0.000	2.00
Johor	5	2	1	0.400 ± 0.237	0.001 ± 0.000	0.40
Pahang	2	1	0	0.000 ± 0.000	0.000 ± 0.000	-
Kedah	4	1	0	0.000 ± 0.000	0.000 ± 0.000	-
Perak	7	2	1	0.476 ± 0.171	0.001 ± 0.000	0.47
Sarawak	5	1	1	0.031 ± 0.175	0.002 ± 0.001	1.20
Sumatra	5	2	2	0.600 ± 0.175	0.002 ± 0.001	1.20
Borneo	7	5	47	0.952 ± 0.000	0.039 ± 0.000	22.19
M. leonina	4	3	31	0.833 ± 0.222	0.030 ± 0.000	16.83
M. siberu	4	2	1	0.500 ± 0.265	0.001 ± 0.000	0.50
M. pagensis	6	5	9	0.933 ± 0.122	0.008 ± 0.000	4.40

N, number of sequences analyzed; H, number of haplotypes; S, number of segregating sites; Hd, haplotype diversity; π, nucleotide diversity; K, average number of nucleotide differences.

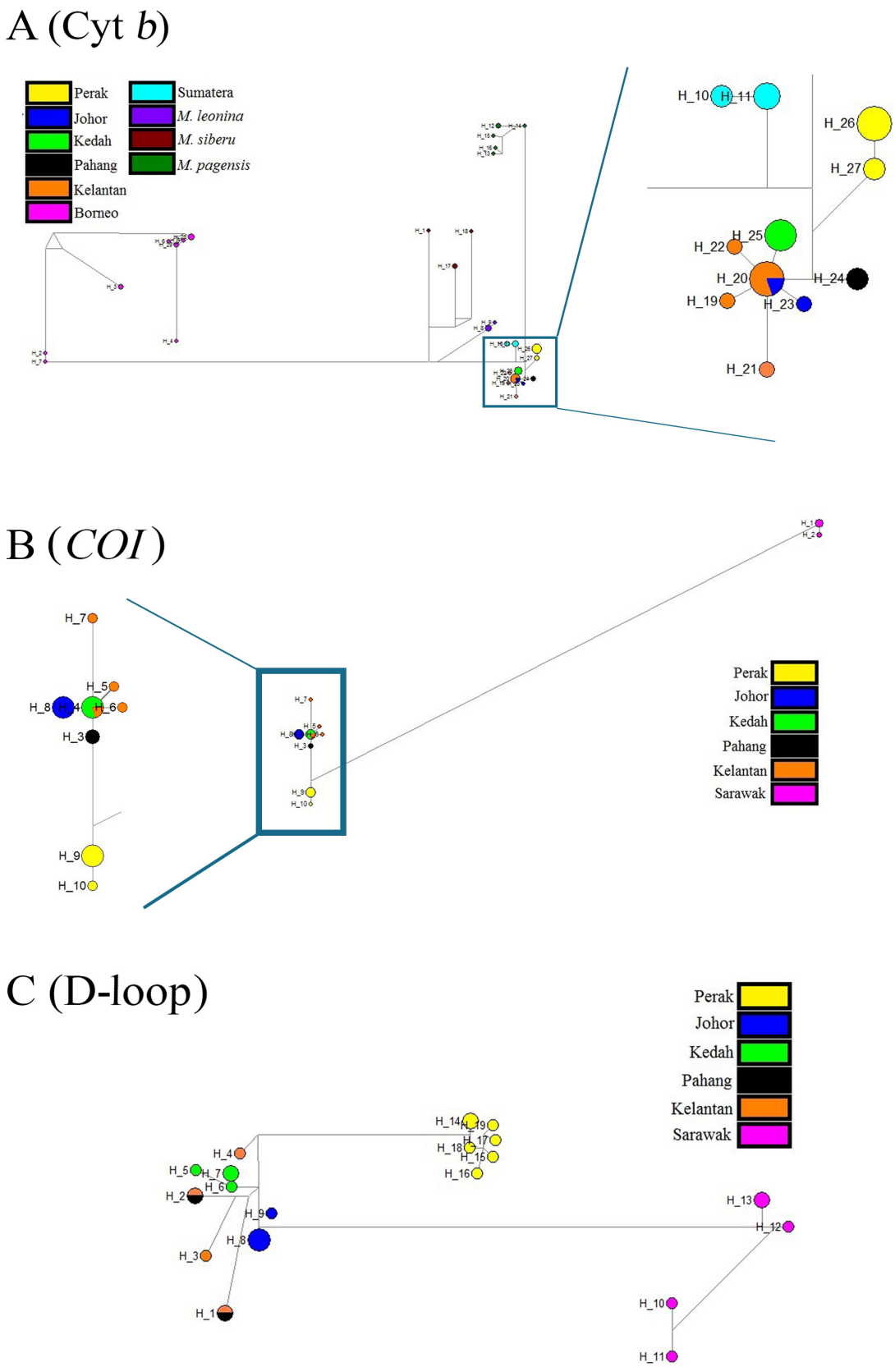

Fig. 5. Minimum Spanning Network (MSN) illustrating haplotype relationships among *M. nemestrina*'s populations based on Cyt *b*, D-loop and *COI* sequences.

extant *M. nemestrina* in Sumatra and Malay Peninsula, *M. leonina*, *M. siberu*, and *M. pagensis*.

Unfortunately, this resolution does come with a price: the contradicting positions of *M. nemestrina* in Sulawesi species groups or *silenus* species groups (Liedigk et al. 2015; Roos et al. 2014; Zinner et al. 2013). By acknowledging two genetic entities of *M. nemestrina* specifically in Borneo; and Sumatra, the Malay Peninsula, and Indochinese, this idea furthers complicates the matter. Ziegler et al. (2007) reported identical findings, which placed *M. nemestrina* from Borneo in Sulawesi groups, and *M. nemestrina* from Malay Peninsula, and Sumatra in *silenus* groups. Tosi et al. (2003) believes that the ancestor *M. silenus* were long gone from the Indian subcontinent and the extant *M. silenus* evolved from the genetic stocks represented by modern *M. nemestrina*. The classifications of species groups were meant to ease elucidations of the dispersal mechanism of particular taxa. However, to place *M. nemestrina* in any particular species groups seems impossible. Morphological data (Delson 1980; Fooden 1975) suggested that the ancestor of *M. nemestrina* arose from *M. silenus*. Genetic data however suggest that Sulawesi macaques were derived from extant lineages of *M. nemestrina* in Borneo, whereas *M. nemestrina* in the Malay Peninsula, *M. pagensis*, *M. siberu*, and *M. leonina* are more closely related to the *silenus* group (Abegg and Thierry 2002; Evans et al. 2003; Roos et al. 2003; Tosi et al. 2003; Ziegler et al. 2007). We believe that *M. nemestrina* should be placed under the *M. nemestrina* group (Groves 2001), based on sequential evolutionary events regardless of *M. nemestrina* in Borneo being the ancestral stocks to Sulawesi macaques or *M. silenus* in Asian Mainland.

Dispersal from Borneo to Sumatra, Mentawai Islands, Indochinese and the Malay Peninsula

Results in this study suggest separations of *M. nemestrina* in Sumatra and the Malay Peninsula from Bornean populations happened around 1.97 million years ago (MYA), in parallel with findings by Ziegler et al. (2007) around 2 MYA. Ancestral *M. pagensis* colonized the entire Mentawai group at approximately 2.51 MYA facilitated by the formation of land bridges as a consequence of the decrease in the sea level during the Pleistocene era as proposed by Roos et al. (2003).

Once more, our diversification estimations were concordant with Ziegler et al. (2007), suggesting the same colonization period, around 2.4–2.6 MYA. The connections between these two lands were disrupted by a rise of sea levels, thus isolating the populations in Sipora and Pagai islands (Roos et al. 2003). *Macaca siberu* arrived later in Siberut Islands around 1.4 MYA, and Roos et al. (2003) justified this as second dispersal events based on the same mechanism as *M. pagensis*.

Our findings are unique because we identify a close relationship between *M. leonina* in Indochinese regions with *M. siberu* supported by all phylogenetic trees, population genetic data, and haplotypes in MSN. It is possible that, before the second colonizations of Mentawai islands, populations from the same stock dispersed upwards toward Indochinese regions. The fluctuations in sea level, which led to the emersion of a huge continental shelf extending to marine areas estimated at around 200 meters in depth (Hutchinson 1989), may well be facilitating this dispersal events. Around these periods, the dispersal period of *M. nemestrina* to Java remains a mystery as there are no extant populations in Java anymore. However, Fooden (1975) assumed it is extinct in Java, results that were supported by findings of fossils of *M. nemestrina* in Java (Aimi 1981). Populations of *M. nemestrina* from Sumatra finally radiate through the Strait of Malacca to the Malay Peninsula around 0.32 MYA (Fig. 7), through formation of land bridges as a result of falling sea levels (Hutchinson 1989; Worldbath 2000).

Table 12. Statistical analysis for demographical history of *M. nemestrina* population based on Cyt *b*/D-loop/*COI* loci

Populasi	D	F_S	D*	F*
Kelantan	-0.78 / -0.447 / -0.797	-1.87/ 1.000 / 1.000	-0.780 /-0.447 / -0.797	-0.721 / -0.466 / -0.753
Johor	-0.817 /-1.048 / -	0.090 /0.546 / -	-0.817 / -1.048 / -	-0.772 / -1.052 / -
Pahang	-/-/-	-/1.000/-	-/-/-	-/-/-
Kedah	- / -0.065 / -	- / 0.251 / -	- / -0.065 / -	- / -0.060 / -
Perak	0.559 / -0.354 / -0.933	0.59 / 0.982 / -0.003	0.953 / -0.256 / -0.950	0.918 / -0.304 / -0.965
Sarawak	1.459 / 1.036 / 0.851	0.546 / 0.571 / 0.773	1.459 / 1.036 / 1.052	1.432 / 1.109 / 1.029
Sumatra	1.459 / - / -	0.546 / - / -	1.459 / - / -	1.432 / - / -
Borneo	0.633 / - / -	0.579 / - / -	0.957 / - / -	0.979 / - / -
M. leonina	-0.368 / - / -	0.284 / - / -	-0.270 / - / -	-0.302 / - / -
M. siberu	-0.612 / - / -	0.876 / - / -	-0.612 / - / -	-0.479 / - / -
M. pagensis	0.692 / - / -	0.918 / - / -	0.591 / - / -	0.660 / - / -

Taxonomic Positions of *M. n. nemestrina* and *M. n. perakensis* in Malaysia

In previous sections, we discussed the dispersal and evolutionary events leading to the formations of *M. leonina*, *M. siberu*, and *M. pagensis*. The phylogenetic relationships of *M. nemestrina* in Malaysia were more sophisticated. Assimilation of populations originating from the states of Johor, Kedah, Pahang, and Kelantan were proven in this study throughout all mtDNA markers and analyses. Populations from all these states formed a monophyletic clade with high bootstrap values and *PP*, along with insignificant population divisions. Differences as low as four mutational steps

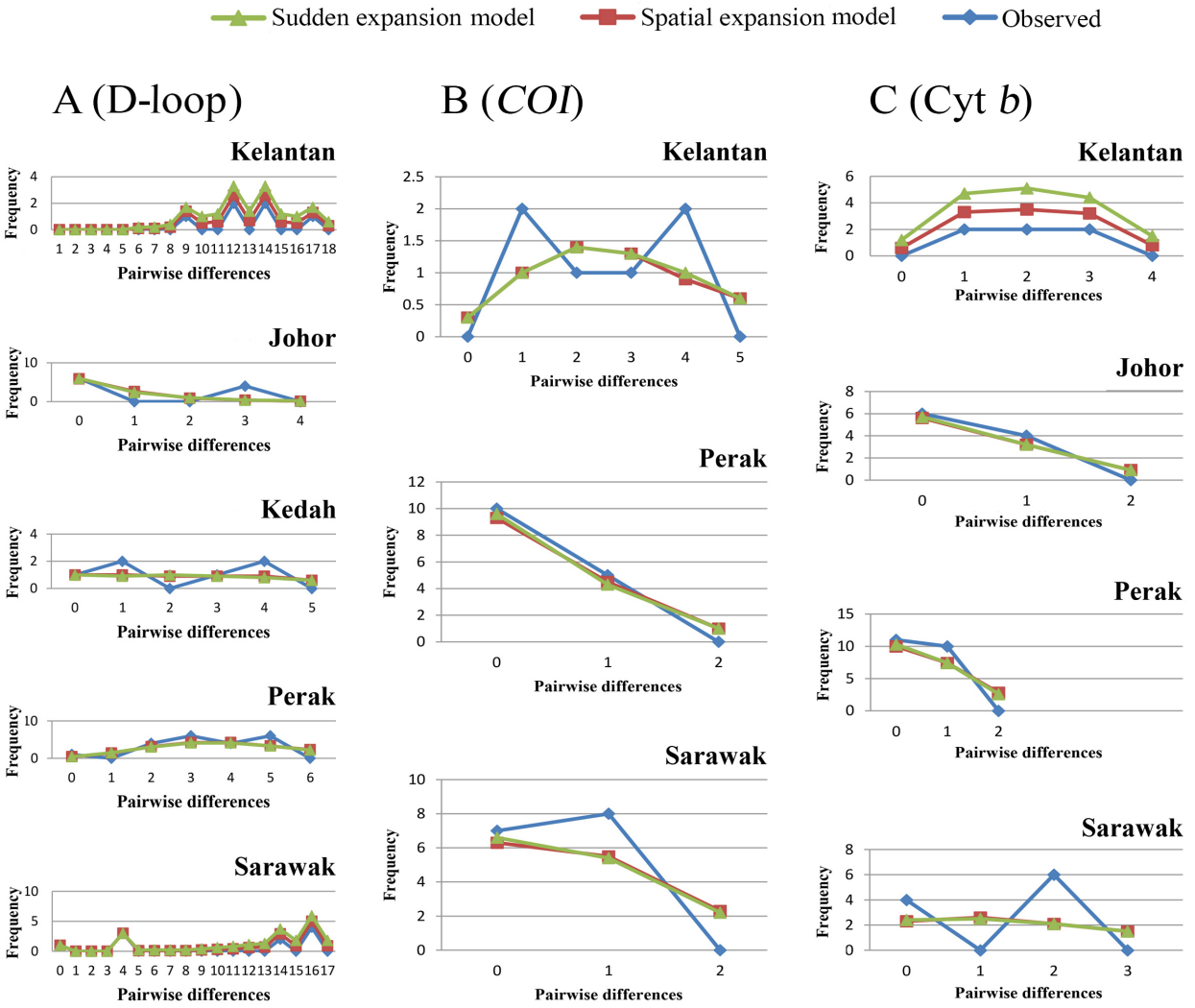

Fig. 6. Mismatch distribution of expected and observed frequencies of pairwise differences among Cyt *b*, D-loop and *COI* sequences for *M. nemestrina* population in Malaysia.

Table 13. Mismatch distribution analysis based on D-loop sequences

Population	SSD		HRI	
Kelantan	0.170	P = 0.206	0.556	P = 0.202
Johor	0.198	P = 0.105	0.680	P = 0.416
Kedah	0.097	P = 0.358	0.306	P = 0.648
Perak	0.033	P = 0.338	0.147	P = 0.392
Sarawak	0.223	P = 0.017	0.590	P = 0.218

separating each haplotype from these populations further strengthen this inference. This phenomenon does support the classifications of *M. n. nemestrina* as proposed by Brandon-Jones et al. (2004), Groves (2001), Rosenblum et al. (1997), Roos et al. (2003), and Ziegler et al. (2007) as the subspecies became distributed throughout the Malay Peninsula.

Conversely, populations from the states of Perak exhibited the presence of a distinct form of pig-tailed macaque. Perak's populations remain differentiated from other populations separated throughout all analyses conducted on the basis of all three mtDNA markers. To strengthen these relationships, haplotypes from Perak's populations showed a minimum of 18 mutational steps to haplotypes from other regions in the Malay Peninsula. Our judgment was conservative at first because we assumed that this is due to the hybridized population between *M. nemestrina* and *M. leonina*. This is possible through natural introgression, since Royal-Belum State Park (forest reserve located in Northern

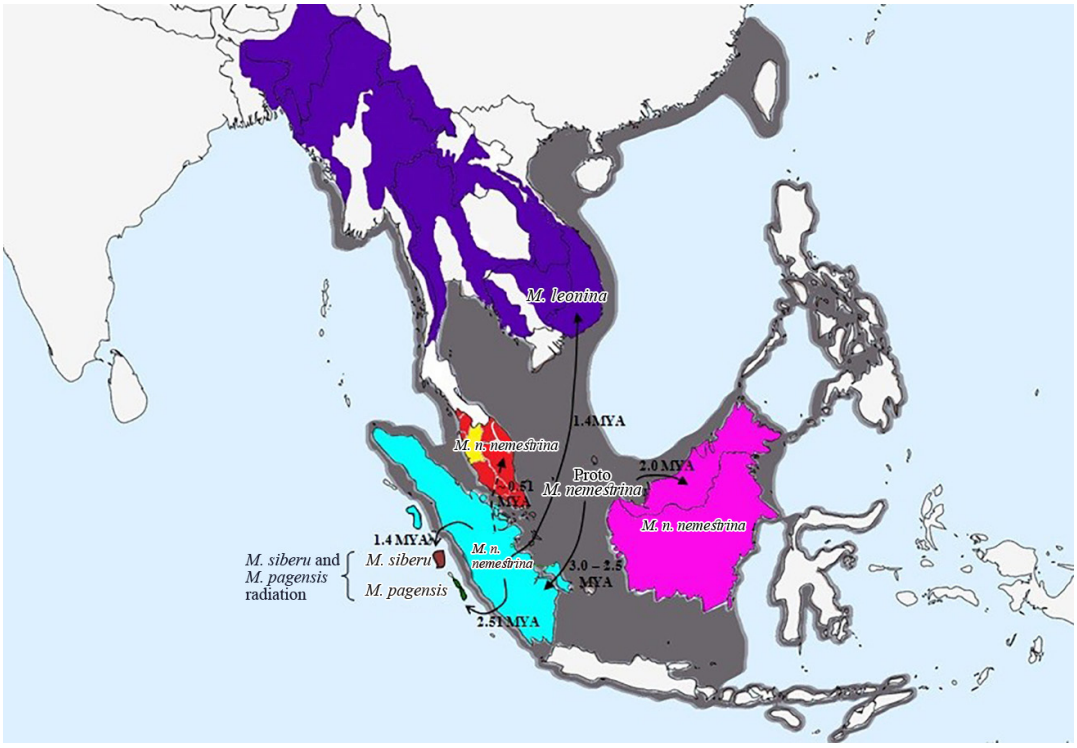

Fig. 7. Hypothetical biogeographical scenario proposed for dispersal events of *M. nemestrina* in Southeast Asia.

Table 14. Mismatch distribution analysis based on *COI* sequences

Populasi	SSD		HRI	
Kelantan	0.070	P = 0.438	0.278	P = 0.612
Perak	0.003	P = 0.759	0.222	P = 0.903
Sarawak	0.030	P = 0.353	0.289	P = 0.398

Table 15. Mismatch distribution analysis based on Cyt *b* sequences

Populasi	SSD		HRI	
Kelantan	0.042	P = 0.572	0.222	P = 0.795
Johor	0.007	P = 0.749	0.200	P = 0.941
Perak	0.017	P = 0.289	0.229	P = 0.600
Sarawak	0.233	P = 0.112	0.880	P = 0.137

Perak) is neighboring Bang Lang National Park of Thailand. This might be the entry point of *M. leonina* in the Malay Peninsula, thus supporting the introgression theory. Malaivijitnond et al. (2012) also suspected that this phenomenon happened toward the south of Isthmus Kra. However, results from the analyses, especially clearly depicted in MSN, pointed out that the minimum mutational steps between *M. leonina* and *M. nemestrina* in the Malay Peninsula were 18 mutational steps, as opposed to only six mutational steps to between Malayan-Sumatran populations.

Distinctions in the pelage and crown coloration on *M. nemestrina* and the length of the cheek whiskers, which were used as a diagnosis by Groves (2001) and Albrecht (1980), were applied to distinguish Perak's population from other populations in the Malay Peninsula. Groves (2001) diagnosis on *M. n. nemestrina* was blackish median dorsal contrasting with agouti brown of flanks, blackish crown, and short cheek whiskers. Conversely, *M. leonina* demonstrate agouti golden brown fur with no contrast between back and flanks with a brown crown and long cheek whiskers (Groves 2001). Perak's population did not demonstrate clear contrasting coloration on their median dorsal, with their flanks brighter-almost reddish coloration as compared to *M. n. nemestrina* (picture taken from neighboring state, Kedah Malaysia). Crown coloration, as shown in figure 8, clearly indicates that the crown of Perak's population is neither black (*M. n. nemestrina*) nor brown (*M. leonina*), but light brown. Lastly, the cheek whiskers of Perak's population surely were not as long and fully face-covering as *M. leonina*, but certainly longer than *M. n. nemestrina*. Using these three morphological criteria, Perak's population clearly shows distinct morphological characteristics as compared to *M. n. nemestrina* and *M. leonina*.

As result from our study, Perak's Pig-tailed macaques differ genetically from other *M. n. nemestrina* populations in Peninsular Malaysia, Borneo, and Sumatra including sister species such as *M. leonina*, *M. pagensis*, and *M. siberu*. Since Perak's pig-tailed population is not only distinguishable on the genetic level but also morphologically, Perak's pig-tailed macaques should be classified as a distinct taxon. While the differences are clear between Perak's population and the remaining *M. n. nemestrina* populations in Southeast Asia, genetic analyses prove the Perak's population still rightfully belong to *M. nemestrina* species classifications. Thus, we propose Perak's populations as a different subspecies of the nominate form *M. nemestrina*. Although several synonyms are available for pig-tailed macaques (Groves 2001), none is appropriate for the Perak's populations. Hence, we describe the Perak's pig-tailed population as a new subspecies.

***Macaca nemestrina perakensis* nov. subsp.**

(Fig. 8)

urn:lsid:zoobank.org:act:DC5A306C-1C3D-4F0A-A484-484EF3B8ADA9

Type Locality: Selama, Perak, West Malaysia.

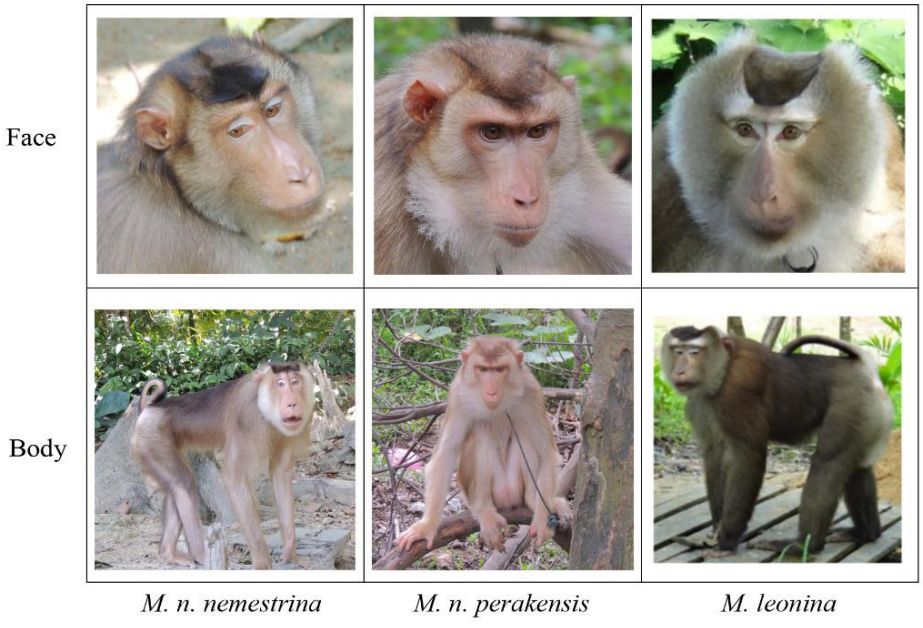

Fig. 8. Morphological differences between *M. n. nemestrina*, *M. n. perakensis* and *M. leonina* (Pictures of *M. leonina* adapted from Malaivijitnond et al. 2012).

Description: Tail well furred, carried in a backward arch with tip pointing down, less than half of head plus body length; median dorsal region yellowish brown with subtle contrasting reddish yellow flanks, light brown crown, whorled region broad in front; long cheek whisker; pale at base with white tips.

Distribution: We obtained samples for this study from district of Selama, Perak, Malaysia. Genetic analysis proved that it is not similar to a neighboring state, thus restricting the distribution of this new subspecies to the state of Perak only. However, we believe both *M. n. nemestrina* and *M. n. perakensis* are sympatrically distributed in Perak, and the exact range of *M. n. perakensis* is smaller than whole Perak.

Etymology: The subspecies in named after the type locality, the west Malaysian province of Perak.

CONCLUSIONS

Progenitors of present-day *Macaca nemestrina* are the survivors of Bornean ancestral lineages, which then evolved and radiated throughout Sumatra, Mentawai Islands, and Indochina, forming *M. pagensis*, *M. siberu*, *M. leonina*, and *M. nemestrina*. The complex and sophisticated diversifications events of these lineages happened around 4.57–0.21 MYA. The peculiarity of Borneo's populations leads us to support the positions of *M. nemestrina* in *M. nemestrina* groups as opposed to the *silenus* or Sulawesi groups. *Macaca nemestrina* arrived in the Malay Peninsula around 0.32 MYA, and consequently divided into two distinct lineages, *M. n. nemestrina* and *M. n. perakensis*, based on genetic and morphological data.

Acknowledgment: This work and the new species name (*Macaca nemestrina perakensis*) were registered with ZooBank under urn:lsid:zoobank.org:pub:5491 CA92-795F-42B6-B4AB-F4A9BDF98D92. We are grateful to YBhg. Dato' Abdul Kadir bin Abu Hashim, Director General of the Department of Wildlife and National Parks and Dr Pazil Abdul Patah who provided us with the necessary facilities and assistance. We also thank the Sarawak Forest Department (Dr Ahmad Ampeng) for providing us the genetic samples. The research was conducted under a research permit (JPHL&TN(IP):100-34/1.24 Jld 8). We acknowledged Mohd Faudzir Najmuddin and Mohd-Ridwan Abd Rahman for their contribution to improving the graphical figures. The authors acknowledge Universiti Kebangsaan Malaysia and Universiti Tun Hussein Onn Malaysia for providing the necessary funding, facilities, and assistance. This research was supported by GUP-2019-037, AP-2015-004, UKM-GUP-2017-087, Fundamental Research Grants Scheme FRGS/1/2018/WAB13/UTHM/03/2 by the Ministry of Education Malaysia, and UTHM-GPPS-H552.

Authors' contributions: MABAL conducted field sampling and laboratory works. MABAL wrote the manuscript, conducted all the molecular analyses. BMMZ involved in critical revision of intellectual content. All authors read and approved the final version of the manuscript.

Competing interests: MABAL and BMMZ declare that they have no conflict of interest.

Availability of data and materials: This manuscript does not contain any personal data beyond that belonging to the authors.

Consent for publication: The authors give their consent to publish.

Ethics approval consent to participate: Research methods reported in this manuscript adhered to the legal requirements of Malaysia and was approved by Department of Wildlife and National Parks (PERHILITAN), Peninsular Malaysia, KM10 Jalan Cheras, Kuala Lumpur, Malaysia under research permit (JPHL&TN(IP):100-6/1/14 Jld 2(40).

REFERENCES

Abdul-Latiff MAB, Ampeng A, Yaakop S, Md-Zain BM. 2014a. Phylogenetic relationships of Malaysia's pig-tailed macaque *Macaca nemestrina* based on D-loop region sequences. AIP Conf Proc **1614:**772–776. doi:10.1063/1.4895300.

Abdul-Latiff MAB, Ruslin F, Faiq H, Hairul MS, Rovie-Ryan JJ, Abdul-Patah P, Yaakop S, Md-Zain BM. 2014b. Continental monophyly and molecular divergence of Peninsular Malaysia's *Macaca fascicularis fascicularis*. Biomed Res Int **2014:**897682. doi:10.1155/2014/897682.

Abdul-Latiff MAB, Abdul-Patah P, Yaakop S, Md-Zain BM. 2017a. Aiding pest control management of long-tailed macaques (*Macaca fascicularis fascicularis*) in Malaysia by using molecular markers of mitochondrial DNA. AIP Conf Proc **1891(1):**020003. doi:10.1063/1.5005336.

Abdul-Latiff MAB, Aifat NR, Yaakop S, Md-Zain BM. 2017b. A noninvasive molecular approach: Exploiting species-locus-specific PCR primers in defeating numts and DNA cross-contamination of Cercopithecidae. J Anim Plant Sci **27:**1015–1023.

Abdul-Latiff MAB, Baharuddin H, Abdul-Patah P, Md-Zain BM. 2019. Is Malaysia's banded langur, *Presbytis femoralis femoralis*, actually *Presbytis neglectus neglectus*? Taxonomic revision with new insights on the radiation history of the *Presbytis* species group in Southeast Asia. Primates **60:**63–79. doi:10.1007/s10329-018-0699-y.

Abegg C, Thierry B. 2002. Macaque evolution and dispersal in insular

south-east Asia. Biol J Linn Soc **75:**555–576. doi:10.1046/j.1095-8312.2002.00045.x.

Aifat NR, Yaakop S, Md-Zain BM. 2016a. Optimization of partial Cyt *b* gene sequence from selected ancient *Presbytis* museum skin specimens. Malays Appl Biol **45:**93–96.

Aifat NR, Yaakop S, Md-Zain BM. 2016b. Ancient DNA analyses of museum specimens from selected *Presbytis* (Primate: Colobinae) based on partial Cyt *b* sequences. AIP Conf Proc **1784:**060024. doi:10.1063/1.4966862.

Aimi M. 1981. Fossil *Macaca nemestrina* (Linnaeus, 1766) from Java, Indonesia. Primates **22:**409–413.

Albrecht GH. 1980. Latitudinal, taxonomic, sexual, and insular determinants of size variation in pigtail macaques, *Macaca nemestrina*. Int J Primatol **1:**141–152.

Alexandrino J, Jarntzen JW, Ferrand N. 2002. Nested clade analysis & the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). J Hered **88:**66–74. doi:10.1038/sj.hdy.6800010.

Arnason U, Gullberg A, Burguete AS, Janice A. 2000. Molecular estimates of primate divergences and new hypotheses for primate dispersal and the origin of modern humans. Hereditas **133:**217–228. doi:10.1111/j.1601-5223.2000.00217.x.

Brandon-Jones D, Eudey AA, Geissmann T, Groves CP, Melnick DJ, Morales JC, Shelkelle M, Stewart CB. 2004. Asian primate classification. Int J Primatol **25:**97–164.

Delson E. 1980. Fossil macaques, phyletic relationships and a scenario of deployment. *In*: Lindburg DG (ed) The Macaques: Studies in Ecology, Behavior and Evolution. Van Nostrand Reinhold, New York, pp. 10–30.

Delson E. 1996. The oldest monkeys in Asia. International Symposium: Evolution of Asian primates. Primate Research Institute, Inuyama, Japan, pp. 40.

Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol **4:**e88. doi:10.1371/journal.pbio.0040088.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol **29:**1969–1973. doi:10.1093/molbev/mss075.

DWNP annual report. 2012. Laporan Tahunan. Jabatan Perlindungan Hidupan Liar dan Taman Negara. Kuala Lumpur, Malaysia.

Evans BJ, Morales JC, Supriatna J, Melnick DJ. 1999. Origin of the Sulawesi macaques (Cercopithecidae: *Macaca*) as suggested by mitochondrial DNA phylogeny. Biol J Linn Soc **66:**539–560. doi:10.1111/j.1095-8312.1999.tb01925.x.

Evans BJ, Supriatna J, Andayani N, Melnick DJ. 2003. Diversification of Sulawesi macaque monkeys: decoupled evolution of mitochondrial and autosomal DNA. Evolution **57:**1931–1946. doi:10.1554/02-476.

Evans BJ, Pin L, Melnick DJ, Wright SI. 2010. Sex-linked inheritance in macaque monkeys: implications for effective population size and dispersal to Sulawesi. Genetics **185:**923–937. doi:10.1534/genetics.110.116228.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online **1:**47–50. doi:10.1177/117693430500100003.

Fooden J. 1975. Taxonomy and evolution of liontail and pigtail macaques (Primates: Cercopithecidae). Field Zool **67:**1–169.

Fooden J. 1976. Provisional classification and key to living species of macaques (Primates: *Macaca*). Folia Primatol **25:**225–236.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics **147:**915–925. doi:10.1093/genetics/147.2.915.

Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. Genetics **133:**693–709. doi:10.1093/genetics/133.3.693.

Gernhard T. 2008a. New analytic results for speciation times in neutral models. Bull Math Biol **70:**1082–1097. doi:10.1007/s11538-007-9291-0.

Gernhard T. 2008b. The conditioned reconstructed process. J Theor Biol **253:**769–778. doi:10.1016/j.jtbi.2008.04.005.

Gippoliti S. 2001. Notes on the taxonomy of *Macaca nemestrina leonina* Blyth, 1863 (Primates: Cercopithecidae). Hystrix It J Mamm **12:**51–54. doi:10.4404/hystrix-12.1-4171.

Groves CP. 1997. Taxonomy and phylogeny of primates. Molecular biology and evolution of blood group and MHC antigens in primates. Springer, Berlin, Heidelberg, pp. 3–23.

Groves CP. 2001. Primate Taxonomy. Smithsonian Institution Press, Washington, USA.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser **41:**95–98.

Harrison JL. 1926. The Apes and Monkeys of Malaya. Malayan Museum Pamphlets **9:**1–10.

Hebert PD, Cywinska A, Ball SL. 2003. Biological identifications through DNA barcodes. Proc R Soc B **270:**313–321. doi:10.1098/rspb.2002.2218.

Holloway JD, Hall R. 1998. SE Asian geology and biogeography: an introduction. *In*: Hall R, Holloway JD (ed) Biogeography and Geological Evolution of SE Asia. Backhuys Publishers, Leiden.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754–755. doi:10.1093/bioinformatics/17.8.754.

Hutchinson CS. 1989. Geological Evolution of South-East Asia. Clarendon Press, Oxford, UK.

Jalil MF, Cable J, Sinyor J, Lackman-Ancrenaz I, Ancrenaz M, Bruford MW, Goossens B. 2008. Riverine effects on mitochondrial structure of Bornean orang-utans (*Pongo pygmaeus*) at two spatial scales. Mol Ecol **17:**2898–2909. doi:10.1111/j.1365-294x.2008.03793.x.

Liedigk R, Roos C, Brameier M, Zinner D. 2014. Mitogenomics of the Old-World monkey tribe Papionini. BMC Evol Biol **14:**176. doi:10.1186/s12862-014-0176-1.

Liedigk R, Kolleck J, Böker KO, Meijaard E, Md-Zain BM, Abdul-Latiff MAB, Ampeng A, Lakim M, Abdul-Patah P, Tosi AJ, Brameier M, Zinner D, Roos C. 2015. Mitogenomic phylogeny of the common long-tailed macaque (*Macaca fascicularis fascicularis*). BMC Genomics **16:**222. doi:10.1186/s12864-015-1437-0.

Malaivijitnond S, Arsaithamkul V, Tanaka H, Pomchote P, Jaroenporn S, Suryobroto B, Hamada Y. 2012. Boundary zone between northern and southern pig-tailed macaques and their morphological differences. Primates **53:**377–389. doi:10.1007/s10329-012-0316-4.

Md-Zain BM. 2001. Molecular systematic of the genus *Presbytis*. PhD Dissertation. Columbia University, New York.

Md-Zain BM, Morales JC, Hassan MN, Jasmi A, Melnick DJ. 2008. Is *Presbytis* a distinct monophyletic genus: inferences from mitochondrial DNA sequences. Asian Primates Journal **1:**26–36.

Md-Zain BM, Mohamad M, Ernie-Muneerah MA, Ampeng A, Jasmi A, Lakim M, Mahani MC. 2010a. Phylogenetic relationships of Malaysian monkeys, Cercopithecidae, based on mitochondrial cytochrome *c* sequences. Genet Mol Res **9:**1987–1996. doi:10.4238/vol9-4gmr942.

Md-Zain BM, Lee SJ, Lakim M, Ampeng A, Mahani MC. 2010b. Phylogenetic position of *Tarsius bancanus* based on partial cytochrome *b* DNA sequences. J Biol Sci **10(4):**348–354. doi:10.3923/jbs.2010.348.354.

Md-Zain BM, Abid-Kamal SNA, Aifat NR, Abdul-Latiff MAB, Mohd-Hashim A, Ampeng A, Yaakop S, Samat A. 2018. Molecular identification of shark fins in Malaysian Borneo's local markets. Biodiversitas **19:**1035–1043. doi:10.13057/biodiv/d190336.

Md-Zain BM, Abdul-Aziz A, Aifat NR, Mohd-Yusof NS, Zulkifli NA, Japning JRR, Rosli N, Yaakop S. 2019. Sequence variation data of the mitochondrial DNA D-loop region of the captive Malayan gaur (*Bos gaurus hubbacki*). Data in Brief **24:**103532. doi:10.1016/j.dib.2018.11.117.

Meijaard E. 2003. Mammals of South-East Asian islands and their late pleistocene environments. J Biogeogr **30:**1245–1257. doi:10.1046/j.1365-2699.2003.00890.x.

Meiri S, Meijaard E, Wich SA, Groves CP, Helgen KM. 2008. Mammals of Borneo-small size on a large island. J Biogeogr **35:**1087–1094. doi:10.1111/j.1365-2699.2008.01897.x.

Nadler T, Thanh VN Streicher U. 2007. Conservation status of Vietnamese primates. Vietnamese Journal of Primatology **1:**7–26.

Oleksyk TK, Smith MW, O'Brien SJ. 2010. Genome-wide scans for footprints of natural selection. Philos Trans Royal Soc B **365:**185–205. doi:10.1098/rstb.2009.0219.

Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817–818. doi:10.1093/bioinformatics/14.9.817.

Rambaut A, Drummond AJ. 2013. TreeAnnotator v1.7.5. Available at http://beast.bio.ed.ac.uk.

Ray N, Currat M, Excoffier L. 2003. Intra-deme molecular diversity in spatially expanding populations. Mol Biol Evol **20:**76–86. doi:10.1093/molbev/msg009.

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol **9:**552–569. doi:10.1093/oxfordjournals.molbev.a040727.

Rook L, Martínez-Navarro B, Howell FC. 2004. Occurrence of *Theropithecus* sp. in the Late Villafranchian of Southern Italy and implication for early pleistocene "out of Africa" dispersals. J Hum Evol **47:**267–277. doi:10.1016/j.jhevol.2004.08.001.

Roos C, Ziegler T, Keith Hodges J, Zischler H, Abegg C. 2003. Molecular phylogeny of Mentawai macaques: taxonomic and biogeographic implications. Mol Phylogenetics Evo **29:**139–150. doi:10.1016/s1055-7903(03)00076-9.

Roos C, Thanh VN, Walter L, Nadler T. 2007. Molecular systematics of Indochinese primates. Vietnamese Journal of Primatology **1:**41–53.

Roos C, Boonratana R, Supriatna J, Fellowes JR, Groves CP, Nash SD, Rylands B, Mittermeier RA. 2014. An updated taxonomy and conservation status review of Asian primates. Asian Primates Journal **4:**2–38.

Rosenblum LL, Supriatna J, Melnick DJ. 1997. Phylogeographic analysis of pigtail macaque populations (*Macaca nemestrina*) inferred from mitochondrial DNA. Am J Phys Anthropol **104:**35–45. doi:10.1002/(sici)1096-8644(199709)104:1<35::aid-ajpa3>3.0.co;2-c.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19:**2496–2497. doi:10.1093/bioinformatics/btg359.

Ruslin F, Azmi MA, Matsuda I, Amir R, Md-Zain BM. 2017. Monkey school: Training phases for coconut-picking macaques (*Macaca nemestrina*). Malayan Nat J **69:**301–306.

Slatkin M, Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics **129:**555–562. doi:10.1093/genetics/129.2.555.

Swofford DL. 2002. Phylogenetic analysis using parsimony and other methods version 4.0 beta version. Sinauer Associates, Sunderland.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:**585–595. doi:10.1093/genetics/123.3.585.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol **30:**2725–2729. doi:10.1093/molbev/mst197.

Tosi AJ, Morales JC, Melnick DJ. 2000. Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history. Mol Phylogenetics Evol **17:**133–144. doi:10.1006/mpev.2000.0834.

Tosi AJ, Morales JC, Melnick DJ. 2003. Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. Evolution **57:**1419–1435. doi:10.1111/j.0014-3820.2003.tb00349.x.

Voris HK. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. J Biogeogr **27:**1153–1167. doi:10.1046/j.1365-2699.2000.00489.x.

Vun VF, Mahani MC, Lakim M, Ampeng A, Md-Zain BM. 2011. Phylogenetic relationships of leaf monkeys (*Presbytis*; Colobinae) based on cytochrome *b* and 12S rRNA genes. Genet Mol Res **10:**368–381. doi:10.4238/vol10-1gmr1048.

Worldbath. 2000. Topographical and bathymetrical data from the National Geophysical. Data Center (NDGC) under the National Oceanic and Atmospheric Administration (NOAA). Available at http://ingrid.ldgo.columbia.edu/SOURCES/WORLDBATH/bath/.

Zainudin R, Shukor MN, Norhayati A, Md-Zain BM, Mustafa AR. 2010. Genetic structure of *Hylarana erythraea* (Amphibia: Anura: Ranidae) from Malaysia. Zool Stud **49(5):**688–702.

Ziegler T, Abegg C, Meijaard E, Perwitasari-Farajallah D, Walter L, Hodges JK, Roos C. 2007. Molecular phylogeny and evolutionary history of Southeast Asian macaques forming the *M. silenus* group. Mol Phylogenetics Evol **42:**807–816. doi:10.1016/j.ympev.2006.11.015.

Zinner D, Fickenscher G, Roos C. 2013. Family Cercopithecidae (Old World Monkeys). *In*: Mittermeier RA (ed) The Handbook of the Mammals of the World, Primates Vol. 3. Lynx Edicions, Barcelona, pp. 550–567.