RAFFLES BULLETIN OF ZOOLOGY **65**: 3–19
Date of publication: 3 February 2017
http://zoobank.org/urn:lsid:zoobank.org:pub:E995D01D-EF65-4A41-9B11-A68E0E6E781A

Evolutionary relationships of *Macaca fascicularis fascicularis* (Raffles 1821) (Primates: Cercopithecidae) from Singapore revealed by Bayesian analysis of mitochondrial DNA sequences

Michael A. Schillaci[1*], Amy R. Klegarth[2,7], William M. Switzer[3], Milena R. Shattuck[4], Benjamin P. Y-H. Lee[5], Hope Hollocher[6]

Abstract. Long-tailed macaques (*Macaca fascicularis*) have a wide geographic distribution across mainland and insular Southeast Asia. The evolutionary history of long-tailed macaques has been examined extensively through comparison of phenotypic variation and by phylogenetic analyses of molecular genetic data. Nonetheless, the complex evolutionary history of *M. fascicularis* throughout Southeast Asia is not fully understood. For the present study, we performed a Bayesian phylogenetic analysis of *M. fascicularis* mitochondrial 12S/tRNA-val/16S sequences to examine the evolutionary relationships of the long-tailed macaques from Singapore. More generally, we hoped to gain a better understanding of the evolutionary history of long-tailed macaques throughout Southeast Asia. We used previously archived sequences in GenBank and new sequences from Singapore (n=34) and Bali, Indonesia (n=2) in a Bayesian phylogenetic framework to co-infer evolutionary histories and divergence dates. Our results revealed two large clades, one composed of haplotypes primarily from Sundaic islands populations, and the second primarily from continental populations. These two larger clades comprise four primary regional clades. All three haplotypes from Singapore form a well-supported subclade within a larger peninsular clade. A median-joining network of haplotypes mirrored the results from the phylogenetic analyses. We found divergence dates that were largely consistent with previous studies using complete mitochondrial genomes. Based on an assessment of phylogenetic relationships, the pattern of estimated divergence dates, and the available fossil record, we suggest that the evolutionary history of *M. fascicularis* likely included multiple dispersal events.

Key words. evolution, dispersal, Southeast Asia, mtDNA, phylogeography

INTRODUCTION

The evolutionary history of macaques (Genus *Macaca*) has been a subject of considerable research due to the wide geographic distribution of the genus, its morphological and taxonomic diversity, and its seemingly complex dispersal history within Asia. The genus comprises four commonly recognised lineages or species groups: the *fascicularis* group, the *silenus-sylvanus* group, the *sinica* group, and the *arctoides* group (Fooden, 1976). One group in particular, the *fascicularis* group, exhibits a particularly complex and enigmatic history that is not yet completely understood. The *fascicularis* group is composed of four species: 1) *M. mulatta*, 2) *M. fuscata*, 3) *M. cyclopis*, and 4) *M. fascicularis*. The long-tailed macaque, *M. fascicularis*, has a wide geographic distribution across Southeast Asia, including most of the Indochinese mainland, the Indonesian and Malaysian archipelagos, and much of the Philippine Islands. Up to 10 subspecies of *M. fascicularis* have been recognised (Fooden, 1995). These subspecies can be broadly divided into two groups. The first group is composed of subspecies with lighter pelage distributed across continental populations, including the Indochinese mainland, and the shallow-water Sundaic islands of the Sunda Shelf such as Borneo, Sumatra and Java, and a second group made up of subspecies with darker pelage distributed among deep-water islands such as Nicobar, Simeulue, Lasia, Maratua, and Palawan (Fooden, 1995). As summarised by Abegg & Thierry (2002), based largely on work by Fooden (1995), it is thought that the darker pelage subspecies of the deep-water islands represent the initial dispersal of *M. fascicularis*, probably by rafting during the early to mid-Pleistocene. The lighter-pelage taxa, such as *M. fascicularis fascicularis*, are believed to

[1]Department of Anthropology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON, M1C1A4, Canada; Email: schillaci@utsc.utoronto.ca (*corresponding author)

[2]Department of Biological Sciences, University of Notre Dame, 100 Galvin Life Sciences Center, Notre Dame, IN 46556, USA; Email: amy.klegarth@gmail.com

[3]Laboratory Branch, Division of HIV/AIDS Prevention, National Center for HIV, Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, 30329 USA; Email: bis3@cdc.gov

[4]Center for the Study of Human Origins, Department of Anthropology, New York University, New York, NY, 10003, USA; Email: milenashattuck@gmail.com

[5]National Parks Board, 1 Cluny Road, Singapore 259569; and Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, United Kingdom; Email: BPYL2@kent.ac.uk

[6]Department of Biological Sciences, University of Notre Dame, 100 Galvin Life Sciences Center, Notre Dame, IN 46556, USA; Email: Hope.Hollocher.1@nd.edu

[7]Department of Anthropology, 230 Raitt Hall, University of Washington, Seattle, WA 98105, USA.

ISSN 2345-7600 (electronic) | ISSN 0217-2445 (print)

have subsequently colonised the Indochinese mainland and shallow water islands by land during glacial maxima when the sea levels were low and the Sunda Shelf was exposed (Fooden, 1995; Abegg & Thierry, 2002).

In addition to phenotypic characteristics such as pelage color, mitochondrial DNA (mtDNA) variation has been used widely to investigate the evolutionary history of *M. fascicularis*, including making inferences regarding taxonomy, phylogeography, hybridisation/introgression, and dispersal histories using comparative phylogenetics and haplotype analyses (e.g., Harihara et al., 1988; Tosi et al., 2002, 2003; Tosi & Coke, 2007; Smith et al., 2007, 2014; Blancher et al., 2008; Shiina et al., 2010; Abdul-Latiff et al., 2014a, b; Liedigk et al., 2015; Bunlungsup et al., 2016a, b). A number of studies have also attempted to determine the geographic origins of the Mauritius macaques (Lawler et al., 1995; Tosi & Coke, 2007; Kawamoto et al., 2008; Blancher et al., 2008; Smith et al., 2014; Badhan et al., 2015). Other studies have focused more on regional mainland (Abdul-Latiff et al., 2014a; Bunlungsup et al., 2016b) or insular populations (Kawamoto et al., 1984; Perwitasari-Farajallah et al., 2001; Schillaci et al., 2011). At least one study has used mtDNA variation to determine the geographic origins of individual long-tailed macaques from captive research colonies (Stevison & Kohn, 2008).

Generally, there are some important commonalities to the major findings of the research cited above. First, regardless of the mtDNA region analysed, two main phylogenetic clades are apparent: one composed of primarily mainland, or continental populations, and the other composed primarily of insular populations from the Sundaic islands that formed when the Sunda Shelf was submerged by rising sea levels after the last glacial maximum about 18 thousand years ago (kya) (Abegg & Thierry, 2002). Haplotypes from Sumatra, along with a single haplotype from Thailand, however, are the exception to this phylogeographic pattern (see results and discussion below). There is no universal consensus regarding the geographic origin of these two clades. Harrison et al. (2006 p. 356) suggested that the molecular data presented by Tosi et al. (2003) indicating *M. fascicularis* diverged from the other members of the fascicularis species group at approximately 2.2–2.5 million years ago (mya) implies *M. fascicularis* probably entered the Sunda islands from the Indochinese mainland at the end of the Pretiglian cold phase at 2.3 mya, or during the Eburonian cold phase starting at around 2.0 mya. This scenario would mean that the progenitor of both the continental and Sundaic islands clades originated on what is now the Indochinese mainland (cf., Tosi et al., 2003; Tosi & Coke, 2007; Blancher et al., 2008; Abdul-Latiff et al., 2014b). The dispersal south along the exposed Sunda Shelf may have occurred in pulses coincident with climatic changes, with subsequent rafting to smaller shallow- and deep-water islands (also see Fooden, 1995; Abegg & Thierry, 2002). Other researchers, however, have suggested that insular Southeast Asia, specifically Indonesia, is the most likely origin of the long-tailed macaque (Delson, 1980; Smith et al., 2007, 2014). Subsequent to his 1995 publication, Fooden (2006) implied that because the most derived members of the *M. fascicularis* species group are found north of the current geographic range of *M. fascicularis* on the Indochinese mainland, and the earliest fossils attributed to *M. fascicularis* (or a close relative) are found on the Indonesian island of Java, dispersal of long-tailed macaques likely occurred south to north, after an initial dispersal to the deep-water islands (cf. Fooden, 1995).

More recently Liedigk et al. (2015) discussed the possibility that the Indonesian island of Sumatra may be the geographic origin of *M. fascicularis fascicularis* given that, in their study of full mitochondrial sequences, Sumatran haplotypes were positioned within both the Sundaic islands (haplotypes from southern Sumatra, Bangka) and continental (northern Sumatra haplotypes) clades. Similar results have since been reported by Bunlungsup et al. (2016a). A previous study by Tosi & Coke (2007) placed Sumatran mtDNA haplotypes solely within the Sundaic islands clade, but assigned Sumatran *Y*-chromosomal lineages to both continental and Sundaic islands clades. The authors of that study suggested this pattern reflected immigration since the last glacial maximum by Indochinese mainland *Y*-lineages, and that it is possible that the Indochinese mainland *Y*-lineage is in the process of replacing the native (Sundaic islands) lineage. A subsequent study of *Y*-chromosome markers found two primary clades, a continental Indochinese clade that included *M. mulatta* and *M. fascicularis* sequences from populations north of the Isthmus of Kra, and another Sundaic clade which included *M. fascicularis* sequences from insular and peninsular populations south of the Isthmus of Kra (Bunlungsup et al., 2016a).

The second commonality among studies of mtDNA variation is that island populations have reduced genetic variation, likely the product of genetic drift, including founder effect (Kawamoto et al., 1984; Perwitasari-Farajallah et al., 2001; Smith et al., 2007). Recent research on the Mauritius macaques suggests the founders of this population were from Indonesia, specifically, either Java (Lawler et al., 1995; Kawamoto et al., 2007; also see Kondo et al., 1993; Sussman & Tattersall, 1981) or multiple islands, including Java and Sumatra (Tosi & Coke, 2007; Bonhomme et al., 2008; Stevison & Kohn, 2008).

The current study examines the genetic variation of the long-tailed macaques (*Macaca fascicularis fascicularis*) from Singapore, and the evolutionary relationships of these primates across their geographic range through detailed phylogenetic analyses of the 12S/tRNA-val/16S segment of the mitochondrial genome. We used sequences previously archived in GenBank and new sequences from Singapore (n=34) and Bali, Indonesia (n=2) (Fig. 1). In addition to this specific goal, we hoped to gain a better understanding of the evolutionary history of long-tailed macaques throughout Southeast Asia. We chose the 12S/tRNA-val/16S segment because of its demonstrated utility for inferring the evolutionary history of *Macaca* (Tosi et al., 2002, 2003; Tosi & Coke, 2007; Stevison & Kohn, 2008) and because there is an adequate number and geographic coverage (with provenance) of sequences available in GenBank.

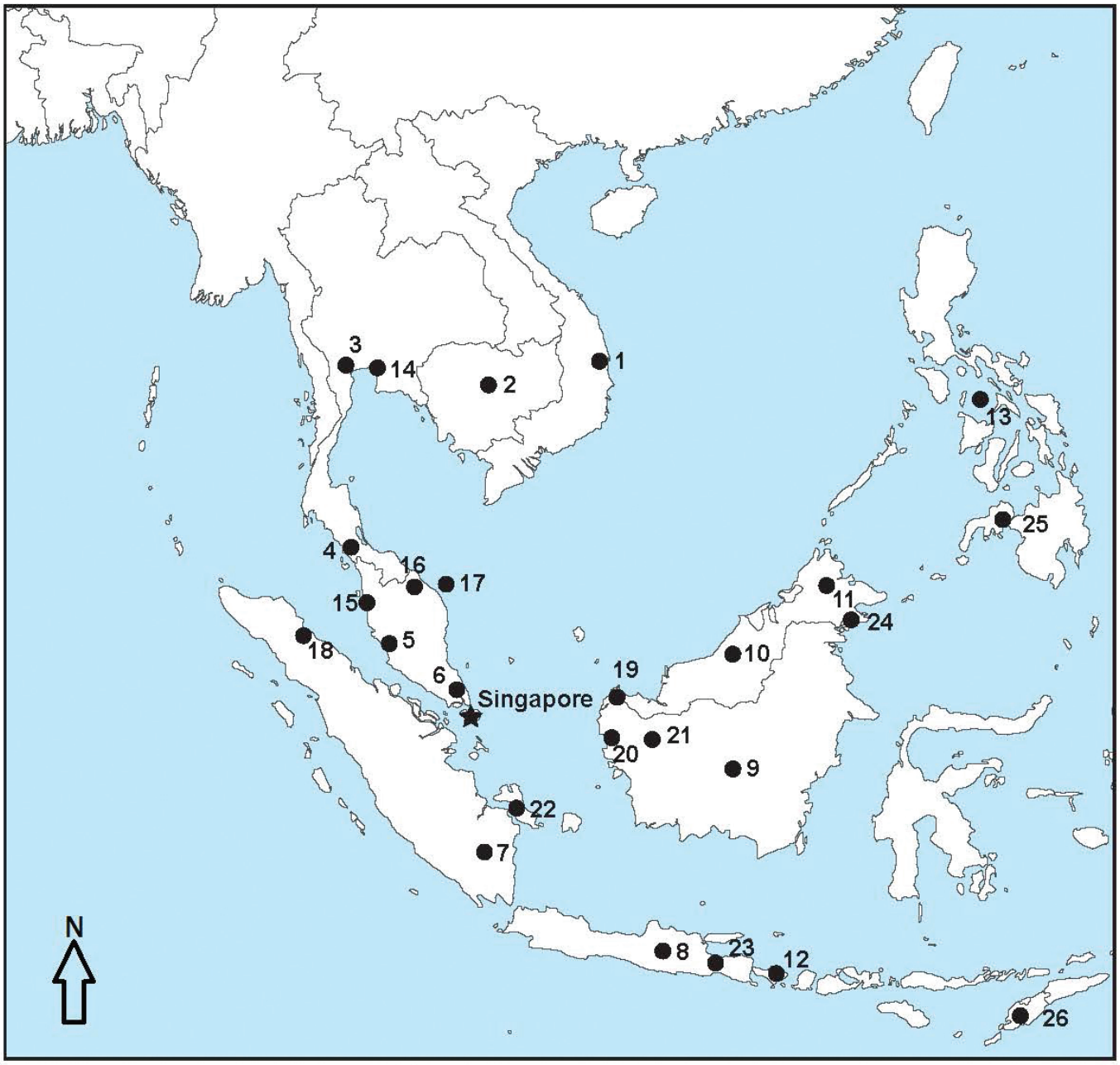

Fig 1. Map of Southeast Asia showing the approximate location of the new (Singapore and Bali) and GenBank sequences included in the study. Numbers correspond to the following locations (haplotype IDs in parentheses): ★, Singapore (Sing1–3) ; 1, Vietnam (Viet1 & 2); 2, Cambodia (Camb1 & 2); 3, Thailand (Thai1); 4, Thailand (Thai2); 5, Malaysia (Selangor1 & 2); 6, Malaysia (Johor); 7, south Sumatra, Indonesia (Sumatra1 & 2, Java1); 8, Java (Java1); 9, Kalimantan, Borneo (Borneo3); 10, Sarawak, Borneo (Borneo1); 11, Sepilok, Borneo (Borneo2); 12, Bali, Indonesia (Bali1 & 2); 13, Sibuyan, Philippines (Phil1); 14, Bangkok, Thailand (Thai3 & 4); 15, Malaysia (W. Malay); 16, Malaysia (E. Malay2); 17 Malaysia (E. Malay1);18, north Sumatra (Sumatra3–6, 9); 19, west Borneo, Indonesia (Borneo9); 20, west Borneo, Indonesia (Borneo4–7); 21, central Borneo, Indonesia (Borneo4 & 6); 22, Bangka, south Sumatra (Sumatra 7 & 8); 23, Java, Indonesia (Java2 & 3); 24, northeast Borneo, Indonesia (Borneo8); 25, Mindanao, Philippines (Phil2); 26, Timor (Timor). Several Borneo haplotypes appear in multiple locations.

In conjunction with the phylogenetic analysis, divergence dates are inferred through Bayesian analysis of the 12S/tRNA-val/16S sequences. We also estimate the nucleotide and haplotype diversity of various *M. f. fascicularis* clades identified by the phylogenetic analysis. The results of our analyses point to a complex evolutionary history for long-tailed macaques, which likely included multiple, or even episodic, dispersal events.

METHODS

Sample collection. Blood samples were collected from a total of 34 long-tailed macaques (*M. fascicularis fascicularis*) trapped at six different locations in 2005, 2011, and 2012 within Singapore's Central Catchment (CCNR) and Bukit Timah (BTNR) Nature Reserves (Table 1, Fig. 2). In addition, two *M. fascicularis* blood samples collected in Bali, Indonesia in 2006 were included in the analysis. Monkeys were caught with a portable aluminum cage-like trap covered with a nylon net (see Schillaci et al., 2007). Once trapped, monkeys were hand injected with <5 mg kg^{-1} of Telazol® (tiletamine/zolazepam) (Fort Dodge Laboratories, Fort Dodge IA, 50501, USA) to achieve anaesthesia. Following universal precautions, blood was drawn from the femoral vein and transferred to Whatman® FTA® cards. The macaques were monitored closely during anaesthesia and recovery. All macaques were released after data collection at the same location they were trapped. None of the monkeys was injured during any of the procedures. All trapping and sample collection protocols used for this study were approved by the institutional animal care and use committees of the University of Washington (Protocol no. 3143-03), University of Toronto (Protocol no. 20005356), and the University of Notre Dame (Protocol no. 14-020, 14-05-1835). All research reported in this manuscript adhered to the legal requirements of Singapore, and permits were obtained from National Parks Board Singapore to conduct the research.

Mitochondrial DNA extraction, amplification and sequencing. Total genomic DNA was extracted from dried blood spots on Whatman® FTA® cards using the QiAmp DNA mini-kit dried blood spot protocol (Qiagen). Following the protocol of Tosi et al. (2002), a 1,549-bp segment region encompassing the 3' end of the 12S ribosomal gene (500-bp), tRNA-val (70-bp), the 5' end of the 16S ribosomal gene (930-bp), and flanking regions was amplified and sequenced in duplicate (i.e., two sequences each for the forward and reverse sequences) for a total of 36 macaques (34 from Singapore and 2 from Bali, Indonesia). A subset of the 36 samples that did not sequence cleanly with previously published primer pairs (n=15) was amplified and sequenced in duplicate using the newly designed primers 12S16S_FB (5'-AGCCAAGACCCAAACTGGGATTAG-3') and 12S16S_RA (5'-ACAGTTAAACCCTCGTGGTGCCTT-3'). These 15 samples were sequenced using the newly designed primer pairs with overlapping coverage from both old and new primer sets. PCR reactions were prepared in 30 ul volumes and used 2ul template DNA, 1× buffer, 45 mM $MgCl_2$, 2 mM dNTP, 4.0 uM of each primer, 1 unit 100× BSA, and 0.2 units Taq polymerase (Invitrogen). Amplification was performed with the following thermocycling parameters for both primer pairs: initial denaturation at 94°C for 5 minutes followed by 35 cycles of 94°C for 45 seconds, 60°C for 45 seconds, and 72°C for 45 seconds, followed by a one-time extension period of 72°C for 20 minutes and then the products were held at 4°C.

Sequencing was conducted at the DNA Analysis Facility at Yale University (New Haven, CT). Samples were sequenced in both the forward and reverse directions using the PCR primers and BigDye chemistry (Applied Biosystems), and products were analyzed using a 3730xl DNA Genetic Analyzer (Applied Biosystems). Sequences were trimmed and assembled automatically by Sequencher 3.8 (Gene Codes Corp.) using default parameters and manually checked against the chromatograms. Nuclear mitochondrial DNA insertions (numts) are commonly associated with disruptions to the open reading frame (Lopez et al., 1994; Buhay, 2009). As such, numts were screened using NCBI's Open Reading Frame Finder (Wheeler et al., 2003) to confirm that all generated sequences had open reading frames. All of the 36 *M. f. fascicularis* sequences we generated had a minimum of 2× coverage, as well as a verified open reading frame. The maximum composite likelihood estimate of the transition/transversion ratio for these sequences was approximately 13.5:1. The nucleotide frequencies were: A = 36.73%, T/U = 20.12%, C = 25.93%, and G = 17.22%.

Phylogenetic analysis. 12S/tRNA-val/16S sequences obtained from the macaque blood samples (n=36) were combined with macaque and baboon sequences obtained from GenBank (n=71), resulting in a final data set of 107 sequences (Table 1). The sequences obtained from GenBank were derived primarily from Tosi et al. (2002, 2003) and Liedigk et al. (2015). Nucleotide alignments were prepared using MAFFT (Multiple Alignment using Fast Fourier Transform) (Katoh et al., 2002) as implemented in the program Geneious v8.1.6. The quality of the alignment for phylogenetic and Bayesian inference was confirmed with likelihood mapping in the program TreePuzzle (Schmidt et al., 2002). There was a minimum of 1,312 positions in the final dataset. 53 unique haplotypes, 45 of which were *M. fascicularis* haplotypes, were identified from the 107 sequences (Appendix) using the ARLEQUIN computer software package (version 3.5.1.2). Sequences representing new haplotypes from Singapore (KY075767, KY075768, KY075769) and Bali (KY075770, KY075771) described in this study were submitted to Genbank. Subsequent phylogenetic analyses were restricted to these 53 unique haplotypes.

The best fitting distance model of nucleotide substitution for the final alignment was inferred using the maximum likelihood method with goodness of fit measured by the Bayesian information criterion (BIC) using jModelTest 2 (Darriba et al., 2012). Bayesian phylogenies were inferred for the 12S/tRNA-val/16S sequences using the program BEAST v2.1.3 (Bouckaert et al., 2014) with a Birth-Death speciation tree prior, and the best fitting nucleotide substitution model (HYK+G+I, BIC= 9890.41). The null hypothesis of equal

Table 1. Sample information including sample and haplotype identifications (IDs), GenBank accession numbers, taxon, country of origin, map location IDs appearing in Figure 1, and sample locations.

Sample ID/GenBank No.	Haplotype ID	Taxon	Country	Map ID	Region/Location
1) SM1 (KY075769)	Sing1	*Macaca fascicularis fascicularis*	Singapore	★	BTNR Visitor Center
2) SM2	Sing1	*M. f. fascicularis*	Singapore	★	BTNR Visitor Center
3) SM3	Sing1	*M. f. fascicularis*	Singapore	★	BTNR Visitor Center
4) SM4	Sing1	*M. f. fascicularis*	Singapore	★	BTNR Visitor Center
5) SM5	Sing1	*M. f. fascicularis*	Singapore	★	BTNR Visitor Center
6) AGU	Sing1	*M. f. fascicularis*	Singapore	★	BTNR Bike Path
7) SAM	Sing2	*M. f. fascicularis*	Singapore	★	BTNR Visitor Center
8) SM6 (KY075767)	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Rifle Range Road
9) SM7	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Rifle Range Road
10) SM8	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Rifle Range Road
11) SM9 (KY075768)	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Rifle Range Road
12) SM10	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Rifle Range Road
13) SM11	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Peirce Res.
14) SM12	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Upper Peirce Res.
15) SM14	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Lower Peirce Res.
16) SM15	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Lower Peirce Res.
17) SM16	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Lower Peirce Res.
18) SM19	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Bukit Kalang
19) SM20	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Bukit Kalang
20) SM21	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Bukit Kalang
21) SM22	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Bukit Kalang
22) SM24	Sing 3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
23) SM25	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
24) SM26	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
25) SM27	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
26) SM28	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
27) SM30	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
28) SM31	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
29) SM32	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
30) DJ	Sing3	*M. f. fascicularis*	Singapore	★	CCNR Upper Seletar
31) SM34	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Lower Peirce Res.
32) SM35	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Lower Peirce Res.
33) SM36	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Lower Peirce Res.
34) SM37	Sing2	*M. f. fascicularis*	Singapore	★	CCNR Lower Peirce Res.
35) AF424963	Selangor1	*M. f. fascicularis*	Malaysia	5	Peninsular (Selangor)
36) AF424964	Selangor1	*M. f. fascicularis*	Malaysia	5	Peninsular (Selangor)
37) KM851017	Selangor2	*M. f. fascicularis*	Malaysia	5	Peninsular (Selangor)
38) AF424965	Johor	*M. f. fascicularis*	Malaysia	6	Peninsular (Johor)
39) KM851015	Johor	*M. f. fascicularis*	Malaysia	6	Peninsular (Johor)
40) KM851012	E. Malay1	*M. f. fascicularis*	Malaysia	17	Peninsular (Pulau Redang)
41) KM851016	E. Malay2	*M. f. fascicularis*	Malaysia	16	Peninsular (East)
42) KM851018	W. Malay	*M. f. fascicularis*	Malaysia	15	Peninsular (West)
43) AF424958	Viet1	*M. f. fascicularis*	Vietnam	1	
44) AF424961	Viet2	*M. f. fascicularis*	Vietnam	1	
45) AF424955	Camb1	*M. f. fascicularis*	Cambodia	2	
46) AF424956	Camb2	*M. f. fascicularis*	Cambodia	2	
47) AF424954	Thai1	*M. f. fascicularis*	Thailand	3	Tum Chompol
48) AF424962	Thai2	*M. f. fascicularis*	Thailand	4	Songkhla
49) KM851024	Thai3	*M. f. fascicularis*	Thailand	14	Bangkok
50) KM851036	Thai4	*M. f. fascicularis*	Thailand	14	Bangkok
51) PCM1 (KY075770)	Bali1	*M. f. fascicularis*	Indonesia	12	Bali
52) PCM3 (KY075771)	Bali2	*M. f. fascicularis*	Indonesia	12	Bali
53) DQ832630	Java1	*M. f. fascicularis*	Indonesia	7	Sumatra, south

Sample ID/GenBank No.	Haplotype ID	Taxon	Country	Map ID	Region/Location
54) AF424969	Java1	*M. f. fascicularis*	Indonesia	8	Java
55) KM851020	Java2	*M. f. fascicularis*	Indonesia	23	Java
56) KM851021	Java2	*M. f. fascicularis*	Indonesia	23	Java
57) KM851028	Java3	*M. f. fascicularis*	Indonesia	23	Java
58) KM851029	Java3	*M. f. fascicularis*	Indonesia	23	Java
59) KM851031	Java2	*M. f. fascicularis*	Indonesia	23	Java
60) KM851019	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
61) KM851025	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
62) KM851026	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
63) KM851027	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
64) KM851030	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
65) KM851032	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
66) KM851033	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
67) KM851037	Timor	*M. f. fascicularis*	Indonesia	26	Timor, west
68) DQ832629	Sumatra1	*M. f. fascicularis*	Indonesia	7	Sumatra, south
69) DQ832628	Sumatra1	*M. f. fascicularis*	Indonesia	7	Sumatra, south
70) DQ832631	Sumatra2	*M. f. fascicularis*	Indonesia	7	Sumatra, south
71) KM850998	Sumatra3	*M. f. fascicularis*	Indonesia	18	Sumatra, north
72) KM850999	Sumatra4	*M. f. fascicularis*	Indonesia	18	Sumatra, north
73) KM851010	Sumatra5	*M. f. fascicularis*	Indonesia	18	Sumatra, north
74) KM851011	Sumatra5	*M. f. fascicularis*	Indonesia	18	Sumatra, north
75) KM851022	Sumatra6	*M. f. fascicularis*	Indonesia	18	Sumatra, north
76) KM851023	Sumatra7	*M. f. fascicularis*	Indonesia	22	Sumatra, south, Bangka
77) KM851034	Sumatra8	*M. f. fascicularis*	Indonesia	22	Sumatra, south, Bangka
78) KM851035	Sumatra9	*M. f. fascicularis*	Indonesia	18	Sumatra, north
79) AF424967	Borneo1	*M. f. fascicularis*	Malaysia	10	Borneo, Sarawak
80) AF424968	Borneo2	*M. f. fascicularis*	Malaysia	11	Borneo, Sepilok
81) AF424966	Borneo3	*M. f. fascicularis*	Indonesia	9	Borneo, Kalimantan
82) KM851004	Borneo4	*M. f. fascicularis*	Indonesia	20	Borneo, west coast
83) KM851002	Borneo4	*M. f. fascicularis*	Indonesia	21	Borneo, west
84) KM851008	Borneo4	*M. f. fascicularis*	Indonesia	21	Borneo, west
85) KM851003	Borneo5	*M. f. fascicularis*	Indonesia	20	Borneo, west coast
86) KM851005	Borneo6	*M. f. fascicularis*	Indonesia	20	Borneo, west coast
87) KM851006	Borneo6	*M. f. fascicularis*	Indonesia	21	Borneo, west
88) KM851007	Borneo6	*M. f. fascicularis*	Indonesia	20	Borneo, west coast
89) KM851009	Borneo7	*M. f. fascicularis*	Indonesia	20	Borneo, west coast
90) KM851013	Borneo8	*M. f. fascicularis*	Malaysia	24	Borneo, northeast
91) KM851014	Borneo9	*M. f. fascicularis*	Indonesia	19	Borneo, northwest
92) DQ832618	Mauritius1	*M. f. fascicularis*	Mauritius		
93) DQ832619	Mauritius1	*M. f. fascicularis*	Mauritius		
94) DQ832620	Mauritius1	*M. f. fascicularis*	Mauritius		
95) KM851000	Mauritius2	*M. f. fascicularis*	Mauritius		
96) AF424970	Phil1	*M. f. philippensis*	Philippines	13	Sibuyan Island
97) KM851001	Phil2	*M. f. fascicularis*	Philippines	25	Mindanao, west
98) AF424949	Mmul1	*M. mulatta*	Burma		
99) AF424950	Mmul2	*M. mulatta*	China		Southeast
100) AF424951	Mmul3	*M. mulatta*	India		North
101) AF424952	Mmul3	*M. mulatta*	India		North
102) AF424971	Msyl1	*M. sylvanus*	Africa		Northwest
103) AF424972	Msyl1	*M. sylvanus*	Africa		Northwest
104) AF424973	Msyl2	*M. sylvanus*	Africa		Algeria
105) AF424974		*Papio hamadryas*	Africa		Eastern
106) JX946201		*Papio hamadryas*	Africa		Eastern
107) JX946196		*Papio anubis*	Africa		Eastern

BTNR, Bukit Timah Nature Reserve; CCNR, Central Catchment Nature Reserve

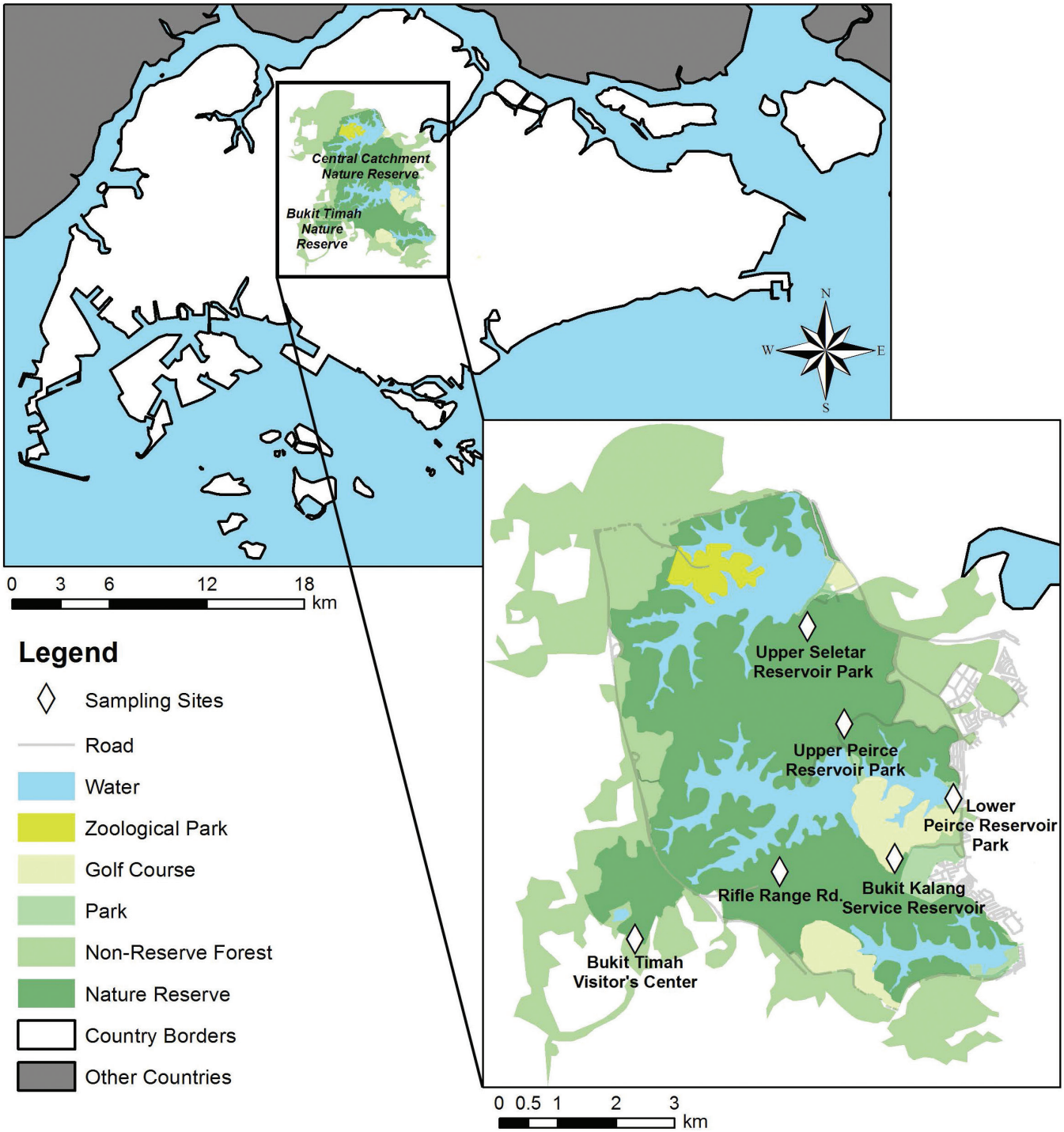

Fig 2. Map of central Singapore showing the sampling locations in the Bukit Timah (BTNR) and Central Catchment (CCNR) Nature Reserves. Map created using ArcGIS® (ESRI® 2015).

evolutionary rates (with HYK+G+I model) was tested in MEGA6 (Tamura et al., 2013) and rejected at the α=0.05 level. Therefore, divergence dates (time to most recent common ancestor, TMRCA) and substitution rates were co-inferred using the uncorrelated lognormal relaxed clock model. The clock was calibrated using a strong uniform prior of 5.3–5.9 mya marking the dispersal of macaques out of Africa (Alba et al., 2014). Two 12S/tRNA-val/16S sequences from *Papio hamadryas* available at GenBank (AF424974, JX946201), and one from *Papio anubis* (JX946196), were used as the outgroup for the analyses. The model was run twice, each time with one billion Markov chain Monte Carlo (MCMC) iterations. The outputs from both runs were combined using LogCombiner in the BEAST v2.1.3 package (Bouckaert et al., 2014). Convergence of the chain sampling was checked in the program Tracer for effective sample sizes (ESS). All ESS values of the combined runs were >418. Trees were saved every 100,000 generations and the tree with the maximum product of the posterior clade probabilities (maximum clade credibility tree) was chosen from the posterior distribution of 10,001 sampled trees after burning in the first 1,000 sampled trees with the

Table 2. Estimated time in millions of years to most recent common ancestor (i.e., node) and node support (posterior probability) from Bayesian analysis of the 12S/tRNA-val/16S mtDNA fragment. Node identifications (Node ID) correspond to those depicted in the Bayesian phylogeny in Figure 3.

Node ID/Clade	Posterior Probability	Node Age (95% HPD)[1]	Liedigk et al. (2015) Node Age (95% HPD)[2]	Tosi et al. (2003) Mean (± 0.5 SD)[3]
(A) *M. fascicularis–M. mulatta*	1.0	3.24 (2.26–4.17)	3.42 (2.83–4.01)	2.5 (±0.3)[4]
(B) *M. fascicularis*	1.0	2.06 (1.37–2.90)	1.70 (1.36–2.04)	1.2 (±0.2)
(C) *M. mulatta*	1.0	1.59 (0.91–2.37)	1.67 (1.19–2.15)	1.2 (±0.4)[5]
(D) Sundaic Islands and a single haplotype from Thailand[6]	0.50	1.77 (1.14–2.55)		
(E) Peninsular Malay/N. Sumatra Indochinese mainland	1.0	1.11 (0.69–1.66)	0.96 (0.78–1.16)	
(F) Sundaic Islands	1.0	1.04 (0.64–1.55)	0.93 (0.74–1.12)	
(G) Peninsular Malay/N. Sumatra	1.0	0.80 (0.48–1.21)	0.70 (0.55–0.84)	
(H) Philippines/Borneo/S. Sumatra	0.92	0.75 (0.43–1.19)		
(I) Java/Bali/Timor/Mauritius	0.61	0.88 (0.50–1.32)		
(J) Singapore/N. Sumatra/S. Malaysia	1.0	0.41 (0.23–0.65)		
(K) Borneo/S. Sumatra/ Philippines	0.90	0.62 (0.35–0.97)	0.61 (0.47–0.75)	
(L) Borneo/Philippines	1.0	0.36 (0.18–0.59)	0.33 (0.26–0.41)	
(M) Mainland Indochina	0.78	0.81 (0.40–1.29)		
(N) Singapore	1.0	0.27 (0.11–0.47)		
(O) Borneo/S. Sumatra	0.39	0.52 (0.26–0.83)		
(P) North Sumatra	0.24	0.23 (0.12–0.40)		

[1]95% highest posterior density (HPD) range in millions of years presented in parentheses.
[2]HPD, highest posterior density range in millions of years before present.
[3]Time to most recent common ancestor (TMRCA) in millions of years before present. Values in parentheses indicate the range of divergence values within one half standard deviation of the mean (Tosi et al., 2003 p. 431)
[4]Includes *M. arctoides* (see Tosi et al., 2003, Fig 5)
[5]Includes *M. cyclopis* and *M. fuscata* (see Tosi et al., 2003, Fig 5)
[6]Haplotype Thai2 from Songkhla Thailand (AF424962)

program TreeAnnotator in the BEAST v2.1.3 package. Trees were viewed in FigTree version 1.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Analysis of mtDNA sequence variation and haplotype diversity. Using the complete data set composed of all sequences, the nucleotide diversity (π) for each major clade identified in the phylogenetic analysis was calculated as the number of nucleotide differences per site averaged over all sequence pairs (Tamura et al., 2013). The neutral mutation model was examined using Tajima's test of neutrality (Tajima, 1989) for each of the major clades identified in the phylogenetic analysis. Following Smith et al. (2014), we interpreted statistically significant negative values of Tajima's test statistic D as indicative of population growth, with D values of zero reflecting constant population size. Haplotype frequencies and diversity (h) were estimated using the program ARLEQUIN (version 3.5.1.2). To visualise the overall relationships among haplotypes, we created a haplotype network using the median joining method (Bandelt et al., 1999) in Network v4.6 (fluxus-engineering.com), with transversions weighted three times as much as transitions. Following Smith et al. (2014) and Hasan et al. (2014), we included the *M. sylvanus* haplotypes as a means of rooting the median joining network.

RESULTS

The 97 *M. fascicularis* sequences used in our analysis comprised a minimum 1,312 positions, of which 1,195 were conserved, and 118 (9.87%) were variable. Of the variable sites, 86 (72.9%) were parsimony informative. The phylogenetic tree inferred from the Bayesian analysis (Fig. 3) exhibited two primary *M. fascicularis* clades: one well-supported clade with a posterior probability (PP) of 1 corresponding to the Malay peninsula, north Sumatra and Indochinese mainland haplotypes, i.e., the continental clade (Clade E), and the other a poorly supported clade (PP=0.50) consisting of haplotypes from the Sundaic islands of Indonesia and Malaysia including Borneo, Java, Bali, south Sumatra, and Timor (Clade D), in addition to a single continental haplotype from Songkhla Thailand (Thai2). A similar phylogenetic position for this same haplotype was presented by Stevison & Kohn (2008), and research presented by Tosi and colleagues (Tosi et al., 2002, 2003, 2007), from which the Songkhla Thailand sequence (AF424962) was derived, gave low bootstrap values indicative of uncertainty for the position of this haplotype in their phylogenetic trees.

The two primary clades revealed in our analysis can be subdivided into a total of four regional clades: 1) peninsular Malay/north Sumatra (Clade G), 2) Indochinese mainland (Clade M), 3) Java/Bali/Timor (Clade I), and 4) Philippines/

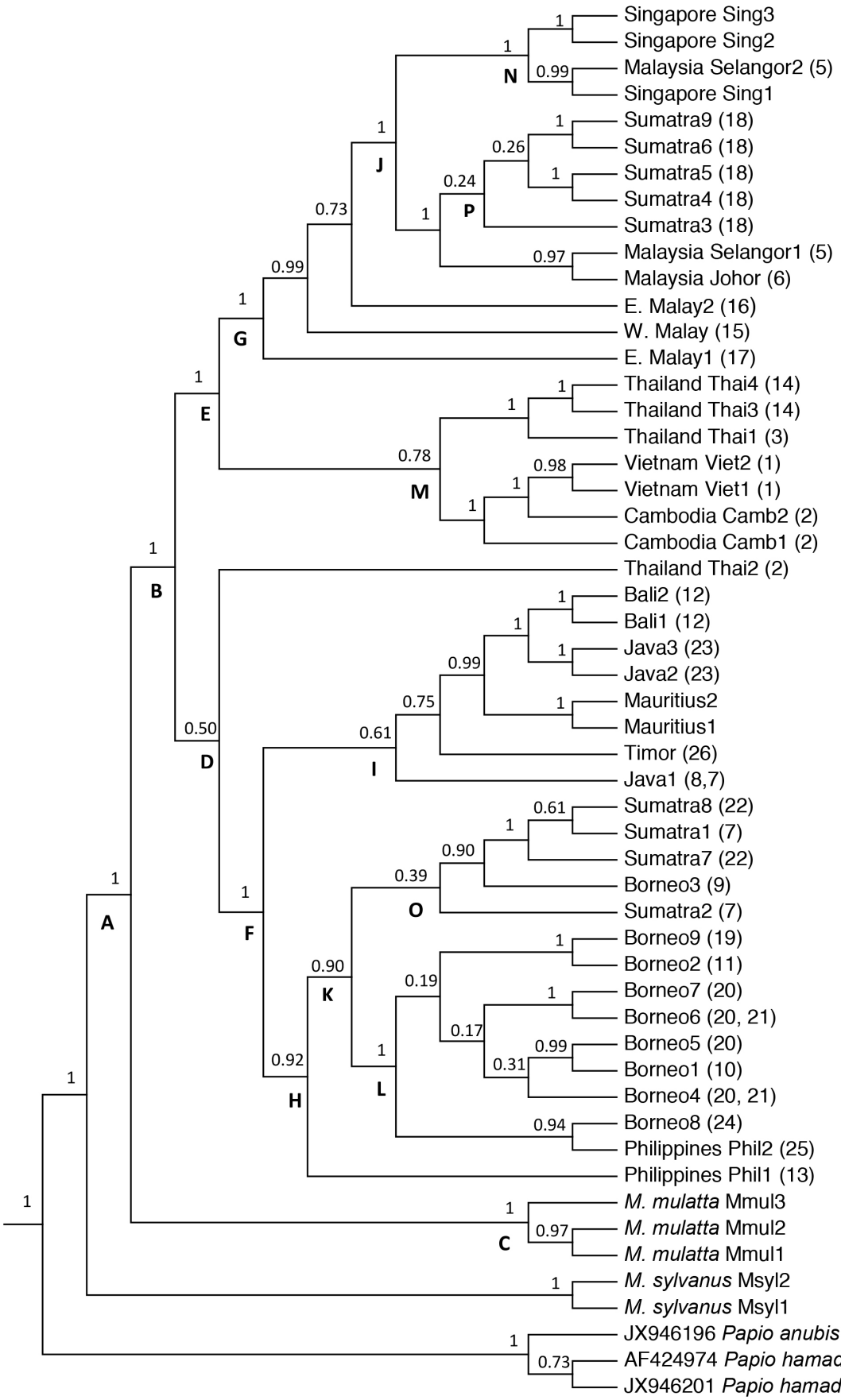

Fig 3. Phylogenetic tree topology from Bayesian inference of 12S/tRNA-val/16S mtDNA sequences using a Birth-Death speciation tree prior, and HYK+G+I nucleotide substitution model in BEAST v2.1.3. Lettered identifications for clades are presented below the branches at major nodes. Posterior probabilities are displayed above the branches at nodes. Numbers in parentheses appearing with haplotype identifications are presented in Table 1, and correspond to numbered locations presented on the Figure 1 map. The Singapore haplotypes form two phylogenetic subgroupings, one from the Bukit Timah Nature Reserve (Sing1) and the other from the Central Catchment Nature Reserve (Sing2 & Sing3).

Table 3. Long-tailed macaque 12S/tRNA-val/16S sequence and haplotype diversity measures. Node identifications (Node ID) correspond to those depicted in the Bayesian phylogeny in Figure 3.

Node ID/Clade	N	n	S	Ps	π	D	Number of haplotypes	h (±SE)
(B) *M. f. fascicularis*[1]	96	1312	117	0.08917	0.01782	0.08689	44	0.9447 (0.013)
(E) Peninsular Malay/N. Sumatra Indochinese mainland	55	1312	61	0.04649	0.00615	-1.36165	21	0.8593 (0.034)
(D) Sundaic Islands w/ Thai2[2]	41	1312	68	0.05183	0.01009	-0.60024	23	0.9476 (0.021)
(F) Sundaic Islands[2]	40	1312	51	0.03887	0.00953	0.15396	22	0.9449 (0.021)
(G) Peninsular Malay/N. Sumatra	48	1312	42	0.03201	0.00389	-1.58408	14	0.8147 (0.039)
(I) Java/Bali/Timor/Mauritius	21	1312	31	0.02363	0.00872	1.27245	8	0.8286 (0.065)
(H) Philippines/Borneo/S. Sumatra	20	1312	34	0.02592	0.00519	-1.142819	15	0.9632 (0.028)
(K) Philippines/Borneo/S. Sumatra	19	1312	29	0.02210	0.00486	-0.91716	14	0.9591 (0.031)
(L) Borneo/ Philippines	13	1312	12	0.00915	0.00272	-0.31977	9	0.9231 (0.057)
(M) Mainland Indochina	7	1312	21	0.01600	0.00729	0.65651	7	1.0000 (0.076)
(N) Singapore	35	1312	7	0.00534	0.00158	0.627591	4	0.6538 (0.048)
(O) Borneo/S. Sumatra	6	1313	17	0.01295	0.00447	-1.315156	5	0.9333 (0.122)
(P) North Sumatra	6	1312	9	0.00686	0.00285	-0.314715	5	0.9333 (0.122)

N, number of sequences; n, number of positions after gap stripping; S, number of segregating sites; Ps=S/n; π, nucleotide diversity; D, Tajima's D statistic; h, haplotype diversity.
[1]Does not include the *M. f. philippensis* haplotype (Phil1) from Sibuyan Island (AF424970)
[2]Does not include the *M. f. philippensis* haplotype (Phil1) from Sibuyan Island (AF424970)

Borneo/south Sumatra (Clade H). The haplotypes from Mauritius fell within the Java/Bali/Timor clade. Although strong statistical support was observed for the continental clade (Clade E) in the Bayesian tree, with a posterior probability of 1, there was no statistical support for the large Sundaic islands clade (Clade D; PP=0.50). Exclusion of the single continental haplotype from Songkhla Thailand (Thai2) results in a well-supported (PP=1) clade containing haplotypes solely from the Sundaic islands (Clade F). While the haplotypes from the Philippines, Borneo, and south Sumatra formed a well-supported (PP=0.92) clade (Clade H), the haplotypes from Java, Bali, Timor, and Mauritius comprised a weakly supported (PP=0.61) clade (Clade I). Although the two Philippine haplotypes, which represent *M. f. fascicularis* (Phil2) and *M. f. philippensis* (Phil1), were both positioned within the larger Philippines/Borneo/south Sumatra clade (Clade H), the *M. f. philippensis* haplotype was positioned as sister to the other haplotypes comprising the clade. The three haplotypes originating from Singapore, along with a single haplotype from the southwestern Malay Peninsula (Selangor2) formed a strongly supported (PP=1) subclade (Clade N) within the continental clade. Within this Singapore subclade, haplotypes from the Bukit Timah Nature Reserve (BTNR) (Sing1, KY075769) and the Central Catchment Nature Reserve (CCNR) (Sing2, KY075767 & Sing3, KY075768) formed separate phylogenetic groups (also see Schillaci et al., 2011).

Our Bayesian TMRCA estimates using an uncorrelated lognormal relaxed clock model are presented in Table 2. Our estimated TMRCA, or divergence, of *M. fascicularis* and *M. mulatta* is 3.24 mya, with a 95% highest probability density (HPD) range of (2.26–4.17 mya). The TMRCA of the *M. mulatta* clade is 1.59 mya (95% HPD 0.91–2.37 mya), and 2.06 mya (95% HPD 1.37–2.90 mya) for the *M. fascicularis* clade. The TMRCA for both the continental (Clade E; 1.11 mya, 95% HPD 0.69–1.66 mya) and Sundaic Islands (Clade F; 1.04 mya, 95% HPD 0.64–1.55 mya) clades is about 1 mya. The estimated TMRCA of the Singapore clade (Clade N) was 0.27 mya (95% HPD 0.11–0.47 mya).

The results of our analysis of genetic diversity are presented in Table 3. None of the Tajima's tests of neutrality was significant at the a=0.05 level assuming a beta distribution (Tajima, 1989), indicating the neutral mutation hypothesis can explain the observed polymorphisms within the 12S/tRNA-val/16S mtDNA fragment for all clades. Because none of the D values deviated significantly from zero, none of the clades was interpreted as experiencing population expansion. The Singapore subclade (Clade N) exhibited the lowest nucleotide (π) and haplotype (h) diversity of any clade. The larger Sundaic islands clade (Clade D) exhibited the greatest genetic diversity as measured by nucleotide diversity, but not haplotype diversity. The greatest haplotype diversity was observed for the Indochinese mainland clade (Clade M) and the Borneo/south Sumatra/Philippines clades (Clades H & K). It is important to note, however, that the sample size for the Indochinese mainland clade is small (n=7). The Borneo/Philippines clade (Clade L) exhibited lower nucleotide diversity than any other major regional sub-clade.

A total of 45 *M. fascicularis* haplotypes were identified, including three from Singapore. The median-joining network shows little reticulation among haplotypes, with haplotype clusters corresponding closely with the regional clades revealed by the Bayesian analysis (Fig. 4). The haplotypes from Singapore all form a distinct grouping along with a single peninsular haplotype (Selangor2). One of the Singapore haplotypes originates from the BTNR (Sing1, KY075769), while the other two originate from separate locations within

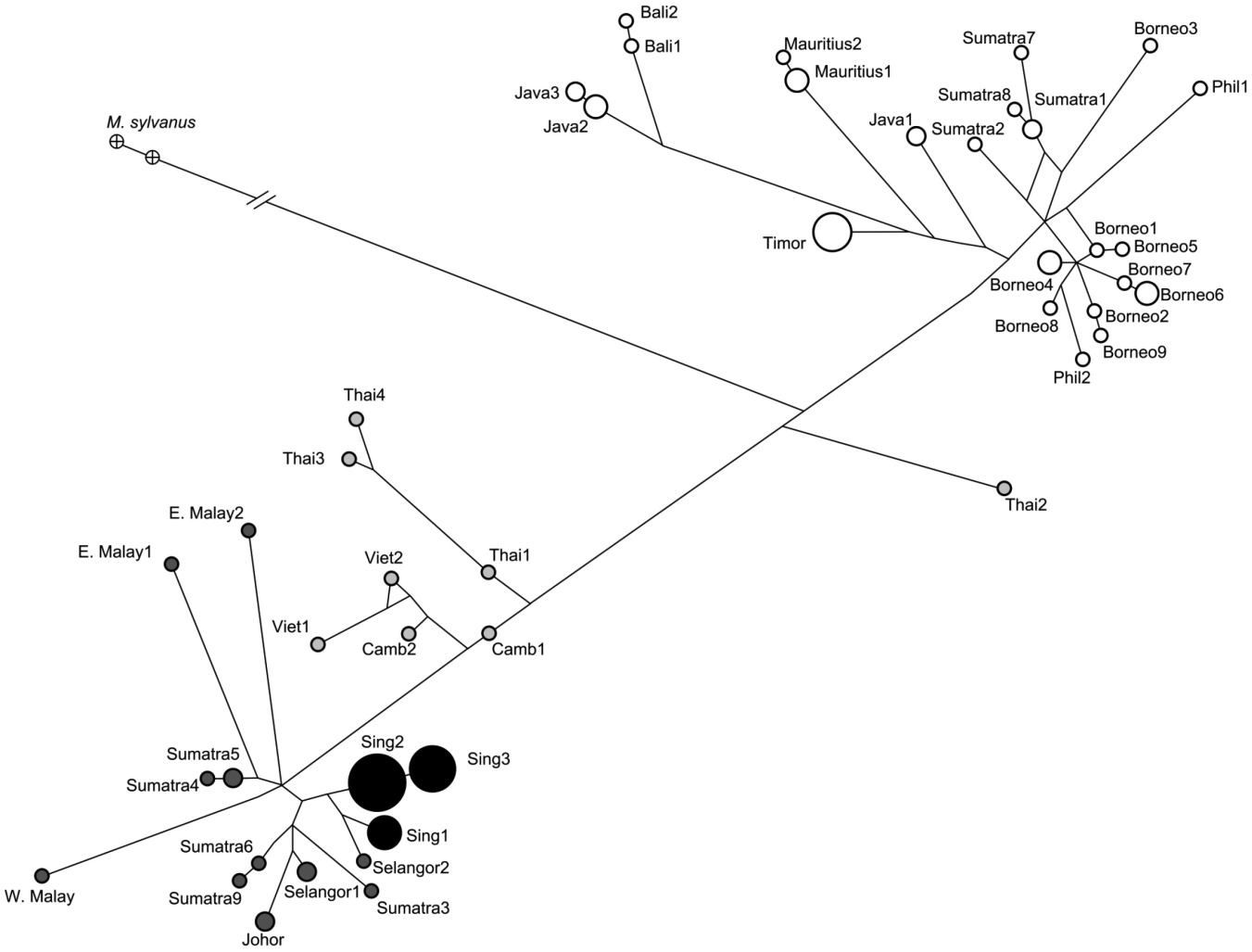

Fig 4. Median-joining haplotype network for *M. fascicularis*. The size of the circular nodes representing haplotypes is proportional to the number of sequences comprising the haplotype. Shading of circular nodes corresponds to general geographic groupings including Sundaic islands (white), mainland Indochina (gray), Malay Peninsula and northern Sumatra (dark gray), and Singapore (black). Haplotype identifications are presented in Table 1.

the CCNR (Sing2, KY075767; Sing3, KY075768). The haplotypes from the continental clade, including the three haplotypes from Singapore, form a separate haplotype group within the median-joining network, as do the haplotypes from the Sundaic islands. Haplotype groups corresponding to Java/Bali/Timor (Clade I) and the Borneo/south Sumatra/ Philippines phylogenetic (Clade K) clades comprise the larger Sundaic haplotype grouping. Only minor reticulation is observed within the Indochinese mainland and Borneo/ south Sumatra haplotype groupings. The *M. sylvanus* haplotypes used to root the median-joining network, along with the Thai2 haplotype from Songkhla Thailand, appear in an intermediate position within the network.

DISCUSSION

We performed a Bayesian phylogenetic analysis of 12S/ tRNA-val/16S mtDNA sequences to better understand the evolutionary history of *M. f. fascicularis*, particularly *M. f. fascicularis* from Singapore. The results from our analyses are consistent with findings reported previously using the same region of the mitochondrial genome (e.g., Tosi et al., 2003; Tosi & Coke, 2007; Stevison & Kohn, 2008), with two primary clades apparent, one composed of continental macaque sequences, and the other Sundaic island sequences. Our study contributes haplotypes from Singapore and Bali, thus providing additional resolution to phylogenetic reconstructions describing *M. fascicularis* evolutionary history. These two primary clades representing haplotypes from Sundaic islands and continental populations have also been observed by researchers utilising other segments of the *M. fascicularis* mtDNA genome (e.g., Blancher et al., 2008; Shiina et al., 2010; Abdul-Latiff et al., 2014b; Smith et al., 2014; Bunlungsup et al., 2016a), and the full mitogenome (Liedigk et al., 2015). Our Bayesian tree, though quite similar to that presented by Liedigk et al. (2015) based on full mitochondrial genomes, differs in several important ways. While the haplotypes from mainland Indochina (i.e., not from peninsular Malaysia) formed a monophyletic clade in our analysis, the non-peninsular mainland Southeast Asian (i.e., mainland Indochina) sequences in Liedigk et al.'s (2015) study did not form a monophyletic clade in their Bayesian tree. Also, the Timor, Java, and Mauritius sequences included in the Liedigk et al. (2015) study did not form a single clade, though the haplotypes derived from those sequences, in addition to others originating from Bali, Java and Mauritius,

did form a single clade in our analysis. Perhaps the biggest difference is the position of the statistically well supported Timor clade in the study by Liedigk et al. (2015). In their study the Timor clade was sister to all other Sundaic islands clades, including the sequence from the Philippines. In our study, the Timor haplotype was nested within the clade formed by haplotypes from Java, Bali and Mauritius.

We found a peninsular clade that included macaque haplotypes from northern Sumatra, the Malay Peninsula and Singapore, but not haplotypes originating from north of the Isthmus of Kra, at locations in Thailand, Vietnam, and Cambodia. Liedigk et al. (2015) also reported a peninsular clade composed of sequences from Malaysia and northern Sumatra. In some sense this finding supports recent assertions based on analyses of mtDNA D-loop HVSII sequences (Abdul-Latiff et al., 2014b) that *M. fascicularis* from the Malay Peninsula south of the Isthmus of Kra form a monophyletic clade relative to other Indochinese mainland populations; however, our results indicate that this peninsular clade also includes mitochondrial haplotypes from north Sumatra (also see Liedigk et al., 2015). Similar to our findings, and those presented by Liedigk et al. (2015), results from the recent study by Bunlungsup et al. (2016a) using a 835 base pair fragment of mtDNA including the hypervariable region I (HVSI), tRNA proline, tRNA threonine, and part of the cytochrome b gene also revealed a peninsular clade that included haplotypes from northern Sumatra. Parenthetically, their results also indicated that the phylogeography of peninsular *M. fascicularis* may be more complicated than previously thought. Bunlungsup et al. (2016a) found that the *M. fascicularis* populations immediately south of the Isthmus of Kra formed a monophyletic clade positioned sister to a larger continental clade composed of Indochinese and peninsular subclades. Whether this clade or others are present just north of the Isthmus of Kra was not determined in that study. Nevertheless, the close relationship between peninsular and Sumatran populations of *M. f. fascicularis* described here and by Liedigk et al. (2015) and Bunlungsup et al. (2016a) seems to be consistent with the biogeographic histories of many vertebrate species. A recent phylogeographic study of mtDNA sequences by Leonard et al. (2015) found that for a strong majority of 28 non-migratory, forest-dependent vertebrates, populations on the Malay Peninsula and Sumatra formed a clade sister to populations from Borneo. Interestingly, using the results presented by Tosi et al. (2007), the authors of the study indicated that *M. fascicularis* was a notable exception to this phylogeographic pattern. The results presented here, and by Liedigk et al. (2015), based on new sequences are in fact consistent with the phylogeographic pattern described by Leonard et al. (2015).

Our Bayesian TMRCA estimates were similar to those presented by Liedigk et al. (2015) using complete mitogenomes, but were somewhat older than estimates presented by Tosi et al. (2003) based on likelihood branch lengths and a molecular clock model using the same 12S/tRNA-val/16S mtDNA fragment and a similar African/Asian macaque calibration date. Our estimates, like those of Liedigk et al. (2015), were between about 0.90 and 0.40 million years older than the three comparable estimates presented by Tosi et al. (2003), who did not include sequences from the Malay Peninsula due to observed rate heterogeneity. It is important to note that Tosi et al. (2003 p. 1432) suggest that their reported divergence dates might be slightly younger than the true dates due to the calibration used in their methodology. Nonetheless, when standard deviations or 95% HPD ranges are considered, all three studies yielded broadly comparable estimates. For example, our estimated 95% HPD for the *M. fascicularis*-*M. mulatta* divergence (2.26–4.17 mya) encompasses the estimates for this same divergence by both Liedigk et al. (3.42 mya) and Tosi et al., (2003) (2.5 mya). Similarly, both our estimate and that presented by Liedigk et al. (2015) closely resemble recent estimates by Pozzi et al. (2014) for the *M. fascicularis*-*M. mulatta* divergence, based on a Bayesian analysis of the complete mitochondrial genome (i.e., 3.44 mya, 95% HPD 2.75–4.21 mya).

Our TMRCA estimate for the *M. fascicularis* clade (2.06 mya; 95% HPD, 1.37–2.90 mya) falls just outside the 95% HPD presented by Liedigk et al. (1.36–2.04 mya), though our 95% HPD encompasses theirs. The ages for the Sundaic (Clade F) and continental (Clade E) clades span a period of less than 100 thousand years of each other in both Liedigk et al. (2015) and our study, with both clades dating to about 1 mya. The peninsular Malay/north Sumatra clade (Clade G) dates to 0.80 mya (95% HPD 0.48–1.21 mya) in our study, and 0.70 mya (95% HPD 0.55–0.84 mya) in Liedigk et al. (2015). The mainland Indochinese clade (Clade M), which was not monophyletic in Liedigk et al. (2015), dated to 0.81 mya (95% HPD, 0.40–1.29 mya). Similarly, our monophyletic Java/Bali/Timor clade (Clade I) dated to 0.88 mya (95% HPD, 0.50–1.32 mya). The Philppines/Borneo/south Sumatra clade dates to about 0.62 mya in both studies. Relevant to this study, the Singapore clade (Clade N), which includes a single haplotype from the Sepang district within the state of Selangor on the Malay Peninsula, dates to about 0.27 mya (95% HPD, 0.11–0.47 mya). Collectively, seven of eight of our estimates for the ages of clades (i.e., internal nodes) fall within the 95% HPDs presented by Liedigk et al. (2015). Together, these comparisons suggest that our Bayesian estimates for the internal nodes within the larger *M. fascicularis* clade that were not included in previous studies are likely reasonable.

Interestingly, the reported divergence dates generated using various portions of the mitochondrial control region, and various different analytical methods, seem to be considerably more recent than those based on the complete mitogenome or 12S/tRNA-val/16S fragment. This may be a consequence of the higher substitution rate observed for the control region relative to the rest of the mitochondrial genome, including the 12S/tRNA-val/16S fragment (see Ingman et al., 2000). Unlike the 12S/tRNA-val/16S fragment utilised in our analyses, the hypervariable control region is non-coding and thus has a different mutational pattern that could affect inference of evolutionary divergence dates and histories, especially in more recently diverged species (Ho et al., 2005; Knaus et al., 2011). For example, Abdul-Latiff et al. (2014b) and

Bunlungsup et al. (2016a) reported dates for the divergence of *M. fascicularis* and *M. mulatta* that are between about 1.34–1.53 mya more recent than ours (i.e., 3.24 mya), and fall outside our 95% HPD for that node (2.26–4.17 mya). Abdul-Latiff et al. (2014b) estimate a very recent divergence of the continental and Sundaic subclades (0.910 mya), while Bunlungsup et al. (2016a) and Blancher et al. (2008) provide estimates of 1.19 and 1.2 mya, similar to that presented by Tosi et al. (2003; 1.2 mya). Both of these latter estimates are roughly 0.87 million years more recent than that provided by our study (2.06 mya), and neither fall within our 95% HPD for that node (1.37–2.90 mya). The estimate presented by Liedigk et al. (2015) for the divergence of the continental and Sundaic clades based on complete mitogenomes was 1.70 mya (95% HPD, 1.36–2.04 mya), closer to ours, with both estimates exhibiting overlapping 95% HPDs. Similarly, the TMRCA estimates of the Sundaic islands macaque clade by Abdul-Latiff et al. (2014b) (0.46 mya), Blancher et al. (2008) (~0.62 mya), and Bunlungsup et al. (2016a) (~0.65 mya) are all more recent than ours (1.04 mya, 95% HPD 0.64–1.55 mya) and Liedigk et al.'s (2015) (0.93 mya, HPD 0.74–1.12 mya). Although the estimates for the TMRCA of the continental clade by Abdul-Latiff et al. (2014b; 0.56 mya) and Blancher et al. (2008) (0.55 mya) are younger than our estimate (0.80 mya), all fall inside our 95% HPD interval (0.40–1.29 mya), but not that of Liedigk et al. (2015) (0.96 mya, 95% HPD 0.78–1.16 mya). Bunlungsup et al.'s (2016a) estimate of 1.07 mya (95% HPD 0.50–1.76 mya) for TMRCA for all continental sequences, including those of a newly described peninsular subclade, is close to ours and Liedigk et al. (2015). Our inferred TMRCA of the peninsular haplotypes south of the Isthmus of Kra (0.80 mya, 95% HPD 0.48–1.21 mya) is also considerably older than the estimate provided by Abdul-Latiff et al. (2014b) (0.33 mya), but is similar to the estimated TMRCA presented by Liedigk et al. (2015) (0.70 mya, HPD 0.55–0.84 mya). As mentioned earlier, Bunlungsup et al. (2016a) indicated that the phylogeographic picture for *M. fascicularis* just south of the Isthmus of Kra may be more complicated than previously thought. Bunlungsup et al. (2016a) did not find a single monophyletic peninsular clade composed of sequences originating from *M. fascicularis* south of the Isthmus of Kra, but instead found two—one clade composed of sequences from northern Sumatra and from the Malay Peninsula well south of the isthmus, which they named the Sundaic Andaman Sea Coast clade, and another clade composed of sequences obtained from macaques just south of the isthmus, which they named the Sundaic Thai Gulf clade. This newly described Sundaic Thai Gulf clade was estimated by Bunlungsup et al. (2016a) to have a TMRCA of 0.49 mya, while the TMRCA of their Sundaic Andaman Sea Coast clade was estimated to be 0.60 mya. Bunlungsup et al.'s (2016a) estimate for the TMRCA of the Sundaic Andaman Sea Coast clade is younger than our estimate for our peninsular clade but falls within our 95% HPD interval (0.48–1.21 mya), as well as within the interval presented by Liedigk et al. (2015) (0.55–0.84 mya).

Singapore lineages. Our analysis revealed a well-supported (PP=1) subclade (Clade N) comprising three haplotypes from Singapore and one from the western Malay Peninsula. This largely Singapore subclade split from a clade of north Sumatran, Johor and Selangor haplotypes around 0.41 mya (95% HPD, 0.23–0.65 mya). The low genetic (π) and haplotype diversity (h) for the Singapore clade has been noted before (Schillaci et al., 2011) and is likely reflecting the smaller population size of the island, along with isolation from migrating mainland females, and possibly founder effect associated either with the original populating of the Singapore region, or genetic drift acting on the small population that was isolated on the island after sea levels rose as the consequence of melting ice sheets after the last glacial maximum, eventually cutting Singapore off from the mainland between 8–10 kya (Bird et al., 2007).

Similar to previous results (Schillaci et al., 2011), we found two primary groupings within the Singapore subclade, one from BTNR and the other from CCNR, which was composed of two haplotypes. The make-up of sequences comprising haplotypes within the CCNR subclade did not strictly reflect geography within the reserve. We estimated the age of the split between the BTNR and CCNR haplotypes to be around 0.27 mya (95% HPD, 0.11–0.47 mya). The early age of this node does not support earlier suggestions by Schillaci et al. (2011) that the split between the BTNR and CCNR haplotypes in Singapore might be attributed to the construction of a major highway in 1986 that now separates the two nature reserves. In that report, Schillaci et al. (2011 p. 106) suggested that the smaller BTNR group (n≈125; see Sha et al., 2009) might be a subset of a larger original CCNR population (n≈902; Sha et al., 2009) that may have become cut off by forest fragmentation associated with the artificial modification of the landscape surrounding the Bukit Timah area. In particular, the construction of the Bukit Timah Expressway may have resulted in some form of random drift such as founder effect for the BTNR group. Under this scenario, random drift, in conjunction with isolation from any possible migrating females, resulted in two differentiated mtDNA sub-populations. The results we present here, however, suggest that divergence of the haplotype lineages that make up the Singapore clade occurred much earlier than the human occupation of Singapore, and given the presence of a haplotype from Selangor, Malaysia, at least one of these divergences could have occurred outside present-day Singapore, likely somewhere on the Malay Peninsula south of the Isthmus of Kra.

Sumatran lineages and the Sundaic islands clade. Within the Sundaic islands clade were three subclades, one composed of eight haplotypes from Java, Bali, Mauritius and Timor (Clade I), one composed of eight haplotypes from Borneo and one from the Philippines (Clade L), and the other composed of four haplotypes from south Sumatra and a single haplotype from Borneo (Clade O). While the largely Borneo subclade was well supported in the Bayesian analysis (PP=1), the largely south Sumatra subclade was not (PP=0.39). The node (Clade K) representing the divergence between these two subclades is well supported (PP=0.90) and dates to 0.62 mya (95% HPD 0.35–0.97 mya). In comparison, the poorly supported (PP=0.24) north Sumatra clade (Clade P)

nested within the larger peninsular Malay/north Sumatra clade (Clade G), diverged from two peninsular haplotypes (Johor and Selangor1) around 0.29 mya (95% HPD 0.14–0.46 mya). Given the more recent inferred origin of the northern Sumatra subclade, it seems reasonable to suggest that the south Sumatran haplotypes represent the earlier residents of the island and that the northern haplotypes are the product of a more recent dispersal. In support of this suggestion, the south Sumatran clade exhibits greater genetic diversity (south, π=0.0045, h=0.933 vs. north, π=0.0029, h=0.933) (see discussion in Liedigk et al., 2015). This scenario is a plausible explanation for the presence of peninsular *Y*-chromosomal lineages in north Sumatra described by Tosi & Coke (2007).

Origins of Mauritius macaques. Our analyses also yielded results relevant to the origin of the Mauritius macaques. Similar to previous analyses of the 12S/tRNA-val/16S fragment by Tosi & Coke (2007), and Stevison & Kohn (2008), our results placed the macaque haplotypes from Mauritius within the clade that included haplotypes from Java. In their assessment of the origins of the Mauritius macaques, Tosi & Coke (2007 p. 502) suggested that because the Mauritius sequence grouped with their Sundaic Java/Sumatra clade in an analysis of the 12S/tRNA-val/16S fragment, but with the continental clade in their analysis of *Y*-chromosomal DNA, the Mauritius macaques are either derived solely from Sumatran lineages, or from a mixture of continental males and insular females. Both our Bayesian and haplotype network analyses grouped those same Mauritius sequences, with the addition of a second Mauritius haplotype, within our Java/Bali/Timor clade. In our analyses, the two Mauritius haplotypes form a well-supported (PP=1) sister clade to Javan and Bali haplotypes not included in the Tosi & Coke (2007) study, but not with the haplotype (Java1 in Figure 3) that they used. Notably, the haplotype used by Tosi & Coke (2007) is shared by two individuals, one from Java and one from south Sumatra. The fact that this Javan haplotype, which includes a Sumatran sequence, does not group with our south Sumatran clade may be an indicator of uncertain ancestry or provenance. Although Tosi & Coke (2007) indicate that their Sumatran samples were collected from a colony of animals that descended from populations in Palembang and Lampung (south Sumatra), they do not mention the precise location of that colony. In a recent study of mtDNA D-loop sequences by Abdul-Latiff et al. (2014b), the Mauritius macaques grouped with one sequence from Java and one other sequence of unknown provenance from Indonesia, but not with several sequences from Bali and Lombok. In the Bayesian analysis presented by Liedigk et al. (2015), a single complete mitochondrial sequence from a Mauritius macaque forms a poorly supported clade with sequences from Borneo and south Sumatra macaques, which did not share a common ancestor with sequences from Timor and Java. However, their analysis did not include sequences from Bali. Together, these findings suggest Java, rather than Sumatra, might be the geographic origin of the Mauritius macaques. Additional analysis of sequences originating from the region of Java, south Sumatra, and Bali, however, is needed before this hypothesis can be substantiated.

Evolutionary history of *M. fascicularis.* The results of our analyses do not provide a straightforward pattern of evolutionary relationships or clade divergence timing that would support a clear inference regarding the dispersal history of *M. fascicularis*. If there was a southern dispersal from a northern point of origin, or homeland, for long-tailed macaques, then we would expect a more recent date for the TMRCA for the Sundaic islands clade. Similarly, if there was a northern dispersal from a southern homeland, we would expect a more recent TMRCA for the continental mainland clade. Our results, like those of Liedigk et al. (2015), indicate that the divergence of these clades from each other occurred around 2 mya, and that TMRCAs for the Sundaic islands and continental mainland clades are essentially the same, occurring around 1 mya, thus offering no temporal information regarding the sequence of dispersal. Other potential phylogenetic indicators of dispersal history are also ambiguous. For example, the Java/Bali/Timor/Mauritius clade (Clade I) is the oldest of the regional subclades with a TMRCA of 0.88 mya (95% HPD 0.51–1.32 mya), while also exhibiting the highest nucleotide diversity (π=0.0087) of the regional subclades. Although a regional population with a high level of diversity combined with the earliest TMRCA might be indicative of a species homeland, there is no north-south geographic patterning to these two attributes across the regional subclades. The Indochinese mainland clade (Clade M) exhibits a TMRCA (0.81 mya, 95% HPD 0.40–1.29 mya) almost as old as the Java/Bali/Timor/Mauritius clade (Clade I; 0.88 mya, 95% HPD 0.50–1.32 mya), and a similarly high nucleotide diversity (π=0.0073).

The internal branching pattern and estimated divergence dates discussed above likely reflect a complex dispersal history involving multiple dispersals (see Abegg & Thierry 2002). Such complexity is indicated in the present study by the intermediate and unresolved position of the haplotype Thai2 (AF424962) from Songkhla Thailand just south of the Isthmus of Kra (and Khlong Marui Fault) within both our Bayesian phylogenetic tree and the median-joining haplotype network. A reviewer of this manuscript suggested that the intermediate and unresolved position of the Thai2 haplotype within our Bayesian phylogenetic tree, and its position within the median-joining haplotype network, might indicate that this sequence represents an old or relict haplotype that is derived from insular Southeast Asia (Sundaic islands), and that the Peninsular Malay/N. Sumatra clade (Clade G) reflects a later re-expansion from the mainland back to Sundaic Islands. However, given its provenance, the Thai2 haplotype from Songkhla Thailand might belong to the Sundaic Thai Gulf clade recently described by Bunlungsup et al. (2016a). If the Thai2 haplotype does belong to any of the mitochondrial lineages represented by the Sundaic Thai Gulf clade presented by Bunlungsup et al. (2016a), it is likely not derived from insular Southeast Asia (Sundaic islands) given the inclusion of the Sundaic Thai Gulf clade within their larger moderately supported (PP=0.86) continental clade. Nonetheless, this clade, including our Thai2 haplotype, may provide a clue to the dispersal history of *M. f. fascicularis*. If the Sundaic Thai Gulf clade presented by Bunlungsup et al. (2016a) represents the relicts of one of two daughter

lineages resulting from the initial divergence of proto-*M. fascicularis* into continental and Sundaic islands clades, then it seems likely this divergence occurred south of about 9°–10°N, the approximate latitude of the Khlong Marui Fault within the Isthmus of Kra.

The fossil record for *Macaca* sheds little light on the complex dispersal history of *M. fascicularis* suggested by phylogenetic analyses of genetic sequence data. The earliest fossils firmly attributed to *M. fascicularis* have been found in Java, Indonesia within the main bone bed ("Hauptknochenschicht") of Trinil during excavations by Eugene Dubois in the 1890s (Stremme, 1911; Hooijer, 1962; Fooden, 1995, Table 33). The maximum age of the main bone bed at Trinil was recently estimated by Joordens et al. (2015) to be 0.546 ± 0.10 mya using $^{40}Ar/^{39}Ar$ dating. More recently, fossils dating to around 2.2–0.8 mya ascribed to *Macaca* cf. *fascicularis* have been found from cave deposits in southern China (Takai et al., 2014), but this designation was made on size variability of the upper and lower third molars, not morphology. If the fossil macaques described from China are in fact the progenitors of present-day *M. fascicularis* (proto-*M. fascicularis*), given they are considerably older than the fossil *M. fascicularis* described from Java, the notion of an initial dispersal of proto-*M. fascicularis* from the Indochinese mainland south to the Sunda Shelf would seem more plausible. A comparative analysis of the dental morphology of the *Macaca* cf. *fascicularis* fossil remains from China are needed before such an assertion can be made. Nonetheless, fossil evidence notwithstanding, the internal branching pattern in our phylogeny does not seem to support the notion of a southern Chinese/northern Indochinese origin for *M. fascicularis*. Rather than appearing as the sister clade to the peninsular and Sundaic clades, the mainland Indochinese clade (Clade M), composed of sequences from Vietnam, Cambodia, and Thailand, is nested within the larger continental clade (Clade E), and shares an ancestor with peninsular haplotypes. Unfortunately, it appears that the fossil record for this species is not complete enough at this time to resolve these questions. More paleontological work is needed to identify the geographic origins of the progenitors of *M. fascicularis*.

We propose a possible dispersal history in which, after the divergence of proto-*M. fascicularis* from proto-*M. mulatta/M. arctoides* somewhere within southern China/ northern Indochina around 3.0–3.5 mya, there was a southern expansion of proto-*M. fascicularis* to about 9°–10°N—the approximate latitude of the Khlong Marui Fault—that resulted in the colonisation of southern Indochina and the northern Sunda Shelf. Parenthetically, there might also have been at this time a western expansion into what is now Myanmar of proto-*M. fascicularis*. Such an expansion would have eventually resulted in the emergence of *M. fascicularis aurea*. Additional analyses of mtDNA variation, however, are needed to assess this proposition (see Bunlungsup et al., 2016b). Subsequently, about 1.7–2.0 mya there was continued dispersal to the south from the northern Sunda shelf (i.e., Khlong Marui Fault region), resulting in the colonisation of what would become the Sundaic islands, including what is now Borneo, southern Sumatra, the Philippines, Java, and Bali. This resulted in the initial divergence of the *M. f. fascicularis* clade. The continental and Sundaic islands lineages then diversified at around 1.0 mya, resulting in regional subclades such as the Java/Bali/Timor/Mauritius, Borneo/south Sumatra/ Philippines, Indochinese mainland and Peninsular Malay/N. Sumatra subclades described in the present study, or the Indochinese, Sundaic Andaman Sea Coast, and Vietnamese subclades recently described by Bunlungsup et al. (2016a). The diversification of the continental clade included an expansion north, resulting in the mainland Indochinese clade, as well as a southern expansion into the southern Malay Peninsula and northern Sumatra by the peninsular clade. This scenario could account for our observed internal branching pattern within the continental clade, as well as that described by Bunlungsup et al. (2016a) (including an eastern expansion resulting in the Vietnamese subclade) and others (cf., Liedigk et al., 2015). It also reflects the fact that the greatest species diversity within the fascicularis species group is found to the north within a region encompassing southern China/ northern Indochina. Other researchers (cf., Fooden, 1995; Abegg & Thierry, 2002; Tosi & Coke, 2007; Blancher et al., 2008; Abdul-Latiff et al., 2014; Liedigk et al., 2015) have suggested broadly comparable scenarios involving multiple dispersals based on morphology or phylogenetic analysis of mtDNA variation. Smith et al. (2014), however, have recently proposed a Sundaic homeland, possibly Sumatra, for *M. fascicularis* based on their analysis of nuclear STR loci and mtDNA variation. Clearly, additional paleontological evidence and a greater geographic sampling of genetic sequences are needed before a more complete description of the evolutionary history of *M. fascicularis* can be made.

In conclusion, consistent with previous molecular studies, our analysis revealed two primary clades, one composed of *M. fascicularis* mtDNA sequences originating from the Sundaic islands, and the other of sequences originating primarily from continental populations. Both clades exhibit ancestral nodes dating to about one million years ago. Unlike most other studies, however, our primarily continental clade also included sequences from north Sumatra (also see Liedigk et al., 2015; Bunlungsup et al., 2016). Our evaluation of the fossil record, phylogenetic branching patterns, the timing of ancestral nodes (i.e., TMRCAs), and genetic diversity, suggests a complex evolutionary history for this species involving multiple dispersals is likely. However, in order to make more complete inferences regarding the dispersal history of *M. fascicularis*, additional sequences from the same region of mtDNA as those already available in GenBank, such as 12S/tRNA-val/16S, D-loop, or complete mtDNA genomes, are needed. Future research should sample *M. fascicularis* populations from the deep-water islands such as the Nicobar islands, Simeulue, and Maratua, as well as from the shallow-water small islands of the Sundaic archipelago. Additional sampling from continental populations including those from Bangladesh and Myanmar, from peninsular populations just north of the Isthmus of Kra (see Bunlungsup et al., 2016a),

and from insular populations such as Java, and the Lesser Sunda Islands east of Wallace's Line, including Lombok and Sumbawa, could also be informative.

ACKNOWLEDGEMENTS

We are grateful to L. Jones-Engel (University of Washington) for her participation in this research, and to S. Chan (NParks Singapore) for facilitating field work in Singapore. We thank S. Oon, W. Ng, and J. Castillo Garcia for their help with data collection, M. Reid for his help with obtaining sequences from GenBank, and D. Nikitovic for her help with graphics. We also thank E. Delson for his input regarding fossil calibration. Partial funding for this project was provided by the University of Toronto Scarborough (awarded to MAS). We are grateful to the anonymous reviewers of this paper for their constructive comments and suggestions. Use of trade names is for identification only and does not imply endorsement by the U.S. Department of Health and Human Services, the Public Health Service, or the Centers for Disease Control and Prevention. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention, or any of the authors' affiliated institutions. None of the authors has a conflict of interest.

LITERATURE CITED

Abdul-Latiff MAB, Ruslin F, Fui VV, Abu M-H, Rovie-Ryan JJ, Abdul-Patah P, Lakim M, Roos C, Yaakop S & Md-Zain BM (2014a) Phylogenetic relationships of Malaysia's long-tailed macaques, *Macaca fascicularis*, based on cytochrome b sequences. Zookeys, 407: 121–139.

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 Research International Volume 2014, Article ID 897682, 18 pp.

Abegg C & Thierry B (2002) Macaque evolution and dispersal in insular South-east Asia. Biological Journal of the Linnean Society, 75: 555–576.

Alba DM, Delson E, Carnevale G, Colombero S, Delfino M, Giuntelli P, Pavia M & Pavia G (2014) First joint record of *Mesopithecus* and cf. *Macaca* in the Miocene of Europe. Journal of Human Evolution, 67: 1–18.

Badhan A, Eichstaedt CA, Almond NM, Knapp LA & Rose NJ (2015) Analysis of full-length mitochondrial DNA D-loop sequences from *Macaca fascicularis* of different geographical origin reveals novel haplotypes. Journal of Medical Primatology, 44: 125–136.

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

Bird MI, Fifield LK, Teh TS, Chang CH, Shirlaw N & Lambeck K (2007) An inflection in the rate of early mid-Holocene eustatic sea-level rise: a new sea-level curve from Singapore. Estuarine, Coastal and Shelf Science, 71: 523–536.

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

Bonhomme M, Blancher A, Cuartero S, Chikhi L & Crouau-Roy B (2008) Origin and number of founders in an introduced insular primate: estimation from nuclear genetic data. Molecular Ecology, 17: 1009–1019.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A & Drummond AJ (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Computational Biology, 10(4), e1003537.

Buhay JE (2009) "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. Journal of Crustacean Biology, 29: 96–110.

Bunlungsup S, Imai H, Hamada Y, Matsudaira K & Malaivijitnond S (2016a) Mitochondrial DNA and two Y-chromosome genes of common long-tailed macaques (*Macaca fascicularis fascicularis*) throughout Thailand and vicinity. American Journal of Primatology, (in press) doi 10.1002/ajp22596.

Bunlungsup S, Imai H, Hamada Y, Gumert MD, San AM & Malaivijitnond S (2016b) Morphological characteristics and genetic diversity of Burmese long-tailed Macaques (*Macaca fascicularis aurea*). American Journal of Primatology, 78: 441–455.

Darriba D, Taboada GL, Doallo R & Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nature Methods, 9: 772.

Delson E (1980) Fossil macaques, phyletic relationships and a scenario of deployment. In: Lindberg DG (ed.) The Macaques: Studies in ecology, behavior and evolution. van Nostrand Rheinhold, New York. Pp. 10–30.

ESRI® (2015) ArcGIS Desktop: Release 10.3. Redlands, CA: Environmental Systems Research Institute. http://www.esri.com/software/arcgis/arcgisonline. (Accessed 16 November 2016).

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

Fooden J (1995) Systematic review of southeast Asian longtail macaques, *Macaca fascicularis* (Raffles, 1821). Fieldiana Zoology, 81: 1–206.

Fooden J (2006) Comparative review of fascicularis-group species of macaques (Primates: *Macaca*). Fieldiana Zoology, 107: 1–43.

Hasan MK, Feeroz MM, Jones-Engel L, Engel GA, Kanthaswamy S & Smith DG (2014) Diversity and molecular phylogeny of mitochondrial DNA of rhesus macaques (*Macaca mulatta*) in Bangladesh. American Journal of Primatology, 76: 1094–1104.

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

Harrison T, Krigbaum J & Manser J (2006) Primate biogeography and ecology on the Sunda Shelf Islands: a paleontological and zooarchaeological perspective. In: Lehman SM & Fleagle JG (eds.) Primate Biogeography. Springer, New York. Pp. 331–372.

Ho SY, Phillips MJ, Cooper A & Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular Biology and Evolution, 22(7): 1561–1568.

Hooijer DA (1962) Quaternary langurs and macaques from the Malay archipelago. Zoologische Verhandelingen Uitgegeven door het Rijksmuseum van Natuurlijke Historie te Leiden, 55: 1–64.

Ingman M, Kaessmann H, Pääbo S & Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. Nature, 408: 708–713.

Joordens JCA, d'Errico F, Wesselingh FP, Munro S, de Vos J, Wallinga J, Ankjærgaard C, Reimann T, Wijbrans JR, Kuiper KF, Mücher HJ, Coqueugniot H, Prié V, Joosten I, van Os B, Schulp AS, Panuel M, van der Haas V, Lustenhouwer W, Reijmer JJG & Roebroeks W (2015) *Homo erectus* at Trinil on Java used shells for tool production and engraving. Nature, 518: 228–231.

Katoh K, Misawa K, Kuma K-I & Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research, 30: 3059–3066.

Kawamoto Y, Ischak TM & Supriatna J (1984) Genetic variations within and between troops of the crab-eating macaque (*Macaca fascicularis*) on Sumatra, Java, Bali, Lombok and Sumbawa, Indonesia. Primates, 25: 131–159.

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

Kondo M, Kawamoto Y, Nozawa K, Matsubayashi K, Watanabe T, Griffiths O & Stanley M-A (1993) Population genetics of crab-eating macaques (*Macaca fascicularis*) on the island of Mauritius. American Journal of Primatology, 29: 167–182.

Knaus BJ, Cronn R, Liston A, Pilgrim K & Schwartz M K (2011) Mitochondrial genome sequences illuminate maternal lineages of conservation concern in a rare carnivore. BMC Ecology, 11(1): 10.

Lawler SH, Sussman RW & Taylor LL (1995) Mitochondrial DNA of the Mauritian macaques (*Macaca fascicularis*): an example of the founder effect. American Journal of Physical Anthropology, 96: 133–141.

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(1): 222.

Leonard JA, den Tex RJ, Hawkins MTR, Muñoz-Fuentes V, Thorington R & Maldonado JE (2015) Phylogeography of vertebrates on the Sunda Shelf: a multi-species comparison. Journal of Biogeography, 42: 871–879.

Lopez JV, Yuhki N, Masuda R, Modi W & O'Brien SJ (1994) Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. Journal of Molecular Evolution, 39: 174–190.

Perwitasari-Farajallah D, Kawamoto Y, Kyes RC, Lelana RPA & Sajuthi D (2001) Genetic characterization of long-tailed macaques (*Macaca fascicularis*) on Tabuan Island, Indonesia. Primates, 42: 141–152.

Pozzi L, Hodgson JA, Burrell AS, Sterner KN, Raaum RL & Disotell TR (2014) Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. Molecular Phylogenetics and Evolution, 75: 165–183.

Raffles TS (1821) XVII. Descriptive catalogue of a zoological collection, made on account of the honourable East India Company, in the island of Sumatra and its vicinity, under the direction of sir Thomas Stamford Raffles, Lieutenant-Governor of Fort Marlborough; with additional notices illustrative of the natural history of those countries. Transactions of the Linnean Society of London, 13(1): 239–274.

Schillaci MA, Jones-Engel L, Lee BPY-H, Fuentes A, Aggimarangsee N, Engel GA & Sutthipat T (2007) Morphology and somatometric growth of long-tailed macaques (*Macaca fascicularis fascicularis*) in Singapore. Biological Journal of the Linnean Society, 92: 675–694.

Schillaci MA, Saravia S, Lee BPY-H & Matheson C (2011) Preliminary report on mitochondrial DNA variation in *Macaca fascicularis* from Singapore. Raffles Bulletin of Zoology, 59: 101–108.

Schmidt HA, Strimmer K, Vingron M & von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics, 18: 502–504.

Sha JCM, Gumert MD, Lee BPY-H, Fuentes A, Rajathurai S, Chan S & Jones-Engel L (2009) Status of the long-tailed macaque *Macaca fascicularis* in Singapore and implications for management. Biodiversity and Conservation, 18: 2909–2926.

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

Smith DG, McDonough JW & George DA (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.

Smith DG, Ng J, George D, Trask JS, Houghton P, Singh B, Villano J & Kanthaswamy S (2014) A genetic comparison of two alleged subspecies of Philippine cynomolgus macaques. American Journal of Physical Anthropology, 155: 136–148.

Stevison LS & Kohn MH (2008) Determining genetic background in captive stocks of cynomolgus macaques (*Macaca fascicularis*). Journal of Medical Primatology, 37: 311–317.

Stremme H (1911) Die Säugetiere mit Ausnahme der proboscidier. In: Selenka L & Blanckenhorn MLP (eds.) Die Pithecanthropus-schichten auf Java. Wilhelm Engelmann, Leipzig. Pp. 82–150.

Sussman RW & Tattersall I (1981) Behavior and ecology of *Macaca fascicularis* in Mauritius: a preliminary study. Primates, 22: 192–205.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123: 585–595.

Takai, M, Zhang Y, Kono RT & Jin C (2014) Changes in the composition of the Pleistocene primate fauna in southern China. Quaternary International 354: 75–85. http://dx.doi.org/10.1016/j.quaint.2014.02.021. (Accessed 28 November 2016).

Tamura K, Stecher G, Peterson D, Filipski A & Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30: 2725–2729.

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

Tosi AJ, Morales JC & Melnick DJ (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 AJ, Morales JC & Melnick DJ (2003) Parental, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. Evolution, 57: 1419–1435.

Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA & Wagner L (2003) Database resources of the National Center for Biotechnology. Nucleic Acids Research, 31: 28–33.