PHYLOGEOGRAPHY OF LONG-TAILED MACAQUE, *MACACA FASCICULARIS AUREA* I. GEOFFROY [1831] FROM MON AND KAYIN STATES, MYANMAR

Phyu Pyar Tin[1], Kazunari Matsudaira[2], Aye Mi San[3]

Abstract

Macaca fascicularis aurea (Burmese long-tailed macaque) is one of the ten subspecies of the long-tailed macaques and is distributed along the Andaman seacoast. In Mon and Kayin States, Myanmar, the macaques inhabit some isolated limestone mountains. To uncover the phylogeography of *M. fascicularis aurea* in Mon and Kayin States and its relationship with those of other areas, the hypervariable segment 1 of mitochondrial DNA (mtDNA) was analysed using non-invasive faecal samples. Thirty-one sequences of *M. fascicularis aurea* were analysed; eight sequences from six populations in Mon State, seven sequences from three populations in Kayin State, four sequences from three populations at the Mergui Archipelago, and eleven sequences from three populations in the Thai Andaman seacoast. Both phylogenetic tree and haplotype network analyses revealed the presence of two mtDNA groups representing geographical differences: one group consisted of the populations from Mon and Kayin States, namely the mainland clade, and the other group consisted of the populations from the Mergui Archipelago and Thai Andaman seacoast, namely the coastal-island clade. Divergence time estimations suggested that *M. fascicularis aurea* initially diverged into the two clades one million years ago. Among the mainland clade, the divergence time of the most recent common ancestor was estimated to be four hundred thousand years ago. Three populations of Kayin State showed the same mtDNA haplotype and suggested a close maternal relationship. No spatial tendency was observed among the populations of Mon State. This study confirmed the close maternal genetic relationship among *M. fascicularis aurea* in Mon and Kayin States.

Keywords mitochondrial DNA, hypervariable segment 1, phylogenetic relationships

Introduction

Macaca fascicularis is one of the most geographically widespread and abundant non-human primate species in Southeast Asia, including Thailand, Indonesia, Malaysia, the Philippines, etc. The species is found in a wide geographic area that encompasses continental and insular populations (Fooden, 1995). *M. fascicularis aurea* (Burmese long-tailed macaque) is one of ten *M. fascicularis* subspecies and is distributed along the Bay of Bengal and the Andaman seacoast, including the Mergui Archipelago (Fooden, 1995; San and Hamada, 2011, Bunlungsup *et al*., 2016). This subspecies has been studied intensively in their use of stone tools for foraging, which is unique among Old World monkeys (Malaivijitnond *et al*., 2007; Gumert *et al*., 2009).

The genetic uniqueness of this subspecies among *M. fascicularis* has been revealed (Bunlungsup *et al*., 2016; Matsudaira *et al*., 2018; Osada *et al*., 2021; Padphone *et al*., 2021). The first phylogenetic study of *M. fascicularis aurea* revealed that *M. fascicularis aurea* was genetically distinct from *M. fascicularis fascicularis* both in the mitochondrial and Y chromosome phylogeny (Bunlunsup *et al*., 2016). Subsequently, a whole mitochondrial DNA (mtDNA) sequence analysis suggested ancient mtDNA introgression from a sinica- species group member to *M. fascicularis aurea* (Matsudaira *et al*., 2018). Whole- genome sequence analysis confirmed ancient introgression from the sinica-species group to *M. fascicularis aurea* but not to *M. fascicularis fascicularis* in the nuclear genome (Osada *et al*., 2021). Restriction Site Associated DNA Sequencing (RADseq) focusing on four *M. fascicularis aurea*

[1] Department of Zoology, Mawlamyine University
[2] Division of Southeast Asian Area Studies, Graduate School of Asian and African Area Studies, Kyoto University.
[3] Department of Zoology, University of Yangon

populations together with *M. fascicularis fascicularis* and *M. mulatta* populations revealed some extent of introgression, or gene flow, from *M. fascicularis aurea* to *M. fascicularis fascicularis* populations near the subspecies contact zone, but less or negligible amount of introgression from *M. fascicularis fascicularis* to *M. fascicularis aurea* (Padphone et al., 2021).

It is interesting how this genetic uniqueness has been formed. For example, when and where did *M. fascicularis aurea* diverge from *M. fascicularis fascicularis*? When and where did *M. fascicularis aurea* meet the sinica-species group and hybridise? Many aspects of the evolutionary history of *M. fascicularis aurea* were still not clear. This was mainly due to the limited number of populations analysed in the previous studies. Most of the *M. fascicularis aurea* populations previously analysed were those of the islands of the Mergui Archipelago and Thai Andaman seacoast. Only one population of Bayin Nyi Naung, in Kayin State was analysed as a representative of the "mainland" area, and not near the coast. In this study, to improve our knowledge of the evolution of *M. fascicularis aurea*, the phylogeography of the nine populations of *M. fascicularis aurea* from the mainland area, Mon and Kayin States, were studied for the following objectives: to investigate the phylogeographic relationship of *M. fascicularis aurea* in Mon and Kayin States and to uncover the divergence times of *M. fascicularis aurea* living in mainland Myanmar.

Materials and Methods

Study sites

Six populations in Mon State (Kyauk Taung [KT], Mein Ma Hlein [MMH], Eaint Phet Taung [EPT], Kyauk Tha Lone Taung [KTL], Kaylar Tha [KLT] and Kha Yone Gu [KYG]) and three populations in Kayin State (Kaw Goon [KG], Ya Thae Pyan [YTP] and Bayin Nyi Naung [BYNN]) were investigated. (Table 1 and Figure 1)

Study period

The present study was conducted from 2018 to 2020.

Sample collection and DNA Extraction

Faecal samples were collected from the nine study populations. The epithelial cells of the digestive tract were collected by the tip of a cotton bud from faecal samples. The cotton bud was rolled on the surface of faecal samples and then dipped in the 1 ml of lysis solution in a microcentrifuge tube (Hayaishi and Kawamoto, 2006). This process was conducted three or more times until the lysis solution in the tube converted to dingy colour. The tubes were kept in a zip bag and stored at room temperature until DNA extraction. DNA extraction from the faeces was performed using the QIAamp DNA Stool Mini Kit (Qiagen, Germany). DNA extraction was carried out according to the manufacturer's protocol with some modifications. At first, the supernatant of the lysis solution with the faecal sample was transferred to a 2 ml microcentrifuge tube. The total volume of faecal lysis solution was adjusted to 1.4 ml by adding buffer ASL and vortex mixed well, which was different from the manufacturer's protocol. At the final step, 200 µl of Buffer AE was incubated on the QIAamp membrane for 20 min to maximize the amount of retrieved DNA. The remaining steps of the extraction were carried out according to the manufacturer's procedures.

PCR and sequencing

The fragment of mtDNA, hypervariable segment 1 (HVS1) was PCR amplified using the primers MFA-DF2: AGCATGATATTCCGTCCACTCAG and MFA-DR2: GGTGATAGACC TGTGATCCATCG. PCR amplifications were carried out in a 25 µl mixture, consisting of 13.85 µl of deionized water, 2.5 µl of 10× PCR Buffer II for AmpliTag Gold, 2.5 µl of 2mM each dNTP, 2.0 µl of 25mM $MgCL_2$, 1.0 µl of 100mg/ml BSA (Sigma Aldrich, USA), 0.5 µl each of forward and reverse primers, 0.15 µl of AmpliTaq Gold (Thermo Scientific, USA), and 2 µl of DNA template. The PCR cycle conditions were as follows; the initial denaturation at 94°C for 9 min, followed by 40 cycles of denaturation at 94°C for 30 sec, annealing at 60°C for 30 min, and extension at 72°C for 1 min and the final extension at 72ºC for 10 min. PCR products were separated by 1% agarose-TAE gel electrophoresis and visualized by ethidium bromide staining and UV transillumination. The purification of the PCR products was conducted by using FavorPrep^{TM} PCR Purification Kit (Favorgen, Taiwan). The nucleic acid concentration and purity of each purified PCR product were measured by using a Nanodrop Spectrophotometric machine (Thermo Scientific, USA).

PCR direct sequencing was conducted by using, Big-Dye Terminator v3.1 Cycle Sequencing Kit (Thermo Scientific, USA) with the same PCR primers. The electrophoresis was conducted by an ABI 3500 Genetic Analyzer (Thermo Scientific, USA).

Data analysis and phylogenetic tree construction

The mtDNA sequences of *M. fascicularis aurea* determined in this study (14 individuals) were analysed together with sequences determined in previous studies: 16 mtDNA sequences of *M. fascicularis aurea* (Accession number: LC093210 to LC093225) from the southernmost part of Myanmar and the Southwestern part of Thailand (Bunlungsup *et al.*, 2016) and one mtDNA sequence of *M. sylvanus* (Barbary macaque) (Accession number: NC_002764) as the outgroup. These sequences were downloaded from GenBank. The sequences were aligned by MUSCLE implemented in MEGA X (Kumar *et al*., 2018), and the insertion and deletion sites were removed, which resulted in the dataset of 31 sequences consisting of 675 nucleotide sites.

Phylogenetic trees were constructed using a distance-based method, the Neighbour-Joining (NJ; Saitou and Nei, 1987), and the maximum likelihood method (ML; Felsenstein, 1981) using MEGA X. In both the analyses, the Tamura-Nei model (Tamura and Nei, 1993) was used as the substitution model. The reliability of the phylogenetic relationship was assessed by 1,000 bootstrap replicates.

Haplotype network analysis of mtDNA

A total of 25 mtDNA sequences, nine sequences from the nine present study populations and the rest of those from the previous study populations were analysed by haplotype network analysis to uncover the geographical relationship between the haplotypes of Mon and Kayin States. A minimum-spanning network (Chen and Morris, 2014) of the mtDNA sequences was constructed by using PopART 1.7 (Leigh and Bryant, 2015).

Divergence time estimation

The divergence times of mtDNA were estimated by the Bayesian framework using BEAST2 (v2.6.3) (Boukaert *et al*., 2014). For this analysis, the gamma site model and Hasegawa-Kishino-Yano (HKY) substitution model were used. The relaxed clock log normal model was used as the clock model and the Coalescent Bayesian Skyline model was used for the tree prior. Markov Chain Monte Carlo (MCMC) runs were conducted by using the chain length of 30,000,000 generations and logging every 1000 generations. Sample distributions of multiple independent replicated runs were combined with LogCombiner 2.6.3 and summarized

(25% burn-in) by TreeAnnotator 2.6.3 (both programs are part of the BEAST2 package). For calibration priors of divergence times, the divergence between *M. sylvanus* and *M. fascicularis aurea* was assumed as the normal distribution divergence of 5.0 million years ago (MYA) (95% lower and upper limits: 5.5 - 6.5 MYA) (Alba *et al.*, 2014; Roos *et al.*, 2019). The phylogenetic tree with the divergence times was visualized in FigTree v1.4.4.

(https://github.com/rambaut/figtree/releases).

Table 1 Study sites

No.	State	Name of location	Coordinate
1.	Mon	Kyauk Taung (KT)	16º 49' N, 97º 35' E
2.		Eaint Phet Taung (EPT)	16º 30' N, 97º 76' E
3.		Mein Ma Hlein Taung (MMH)	16º 30' N, 97º 49' E
4.		Ka Yone cave (KY)	16º 31' N, 97º 42' E
5.		Kay Lar Tha (KLT)	17º 13' N, 97º 05' E
6.		Kalar Kyauk Tha Lone (KTL)	16º 19' N, 97º 42' E
7.	Kayin	Kaw Goon cave (KG)	16º 49' N, 97º 35' E
8.		Ya Thae Pyan cave (YTP)	16º 44' N, 97º 29' E
9.		Bayin Nyi Naung mountain (BYNN)	16º 58' N, 97º 29' E

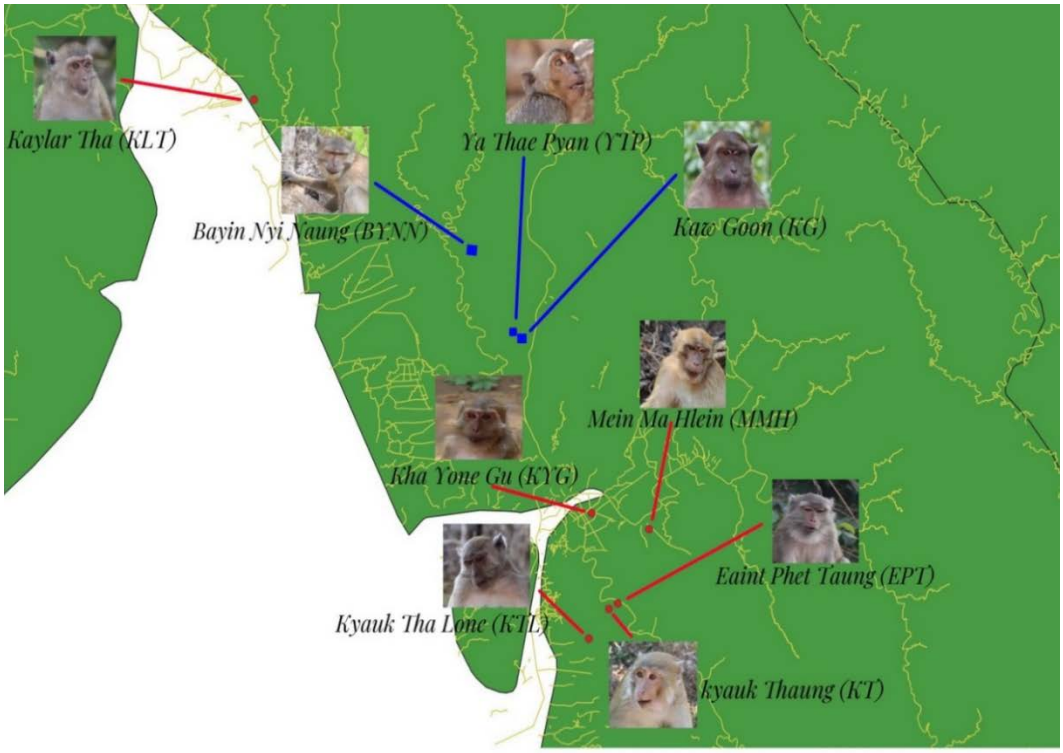

Figure 1 Locations of the study sites of *M. fascicularis aurea* from Mon and Kayin States.

Red cycles indicate Mon State populations (n= 6) and blue quadrangles indicate Kayin State populations (n= 3). Yellow lines indicate river and canals.

Results

Phylogenetic relationship of *M. fascicularis aurea* from Mon and Kayin States

In this study, a total of 14 sequences were determined for the *M. fascicularis aurea* samples collected in the nine populations (one to two samples per population). By combining these sequences with the downloaded sequences of *M. fascicularis aurea* (n = 16) and *M. sylvanus* (n = 1), a total of 31 sequences were analysed. After the alignment, some poorly aligned positions were removed and the final alignment comprised a total of 675 bp.

Both ML and NJ trees showed the same tree topology (Fig 2 and 3). The mtDNA sequences of *M. fascicularis aurea* populations were separated into two main clades, namely the coastal-island clade and the mainland clade. The coastal-island clade comprised Thai Andaman seacoast populations (Piak Nam Yai Island [PNY], Mangrove Forest Research Center [MFRC] and Wat Paknam Pracharangsarith [WPN]) and Mergui Archipelago populations (Lampi Island [LPI], Zadetkyi [ZDK] and Jarlan Island [JLI]). The mainland clade, on the other hand, was the assemblage of the populations in Mon and Kayin States. The three study populations of Kayin State (YTP, KG and BYNN) clustered tightly and separated from the other populations of Mon State, which was supported by moderate and high bootstrap values in the ML (68) and NJ (89) trees.

Among the Mon State populations, MMH was tightly clustered with the Kayin State populations with high bootstrap values both in ML (87) and in NJ (95) trees. In addition, the three populations, KTL, KT and KLT, clustered each other with high supporting values in the ML (97) and NJ (96) trees.

In the coastal-island clade, the three populations, JLI, LPI and ZDK from the Mergui Archipelago, were clustered with the Thai populations (PNY, MFRC and WPN).

Haplotype Network of *M. fascicularis aurea*

The haplotype network is shown in Fig 4. The 10 sequences of the mainland region showed seven haplotypes. The remaining 15 sequences were from the previous study (Bunlungsup *et al.*, 2016) and consisted of seven sequences from Mergui Archipelago populations and eight sequences from Thai Andaman seacoast populations. These 15 sequences showed seven haplotypes and each haplotype consisted of one to four sequences.

In the haplotype network, relatively large nucleotide differences between the haplotypes from the mainland clade and the coastal-island clade were observed. 33 substitutions were observed between BYNN and WPN1298, and 38 substitutions were observed between EPT and PNY2005. On the other hand, only one to 21 substitutions were observed between the mainland haplotypes, and five to 12 substitutions were observed between the coastal-island haplotypes. This large difference between the two regions was consistent with the phylogenetic trees that showed the two clades among *M. fascicularis aurea*.

Among the mainland region, the sequences from the three populations of Kayin State (BYNN, KG, YTP) showed the same haplotype. In addition, one sequence from the previous study (Mfa_BNT 1447, Bunlungsup *et al*., 2016), which was obtained from a sample that had been collected at Bayin Nyi Naung Mountain in 2007 (Aye Mi San, unpublished data), showed only one nucleotide difference from the BYNN sequence collected at the same location in 2019. Furthermore, the haplotype of MMH was the closest to the haplotype from the Kayin State populations.

Among the mainland region, no clear relationship between the haplotype network and the distribution of populations was observed, except that Kayin State populations showed the same haplotype.

Among the coastal-island region, no tendency was observed. The number of substitutions between haplotypes was similar to that observed among the mainland populations. The haplotypes of island populations relatively diverged from those of the coastal populations, and were more largely different from those of the mainland populations.

Divergence times of *Macaca fascicularis aurea*

The divergence time of the mtDNA of *M. fascicularis aurea* and that of *M. sylvanus* was estimated to be 4.74 MYA (95% high posterior density credibility interval [HPD CI] = 3.27–6.37). Among *M. fascicularis aurea*, the divergence between the two clades, *i.e.*, the mainland clade and coastal island clade was estimated to have occurred 1.0 MYA (95% HPD CI= 0.40 – 1.85). Furthermore, the divergence within the mainland clade was estimated to have started about 0.40 MYA (95% HPD CI = 0.14 – 0.76). On the other hand, the divergence within the coastal-island clade was estimated to have started around 0.31MYA (95% HPD CI = 0.10 – 0.59). Within the mainland clade, EPT and KYG formed a cluster and the cluster was estimated to have split from the other cluster at 0.3 MYA (95% HPD CI = 0.14 – 0.76). KLT, KT and KTL populations of Mon State separated from the remaining cluster at 0.25 MYA (95% HPD CI = 0.09 – 0.5). The divergence time between the three populations of Kayin State and MMH population of Mon State was estimated to be 0.09 MYA (95% HPD CI = 0.02 – 0.22) (Table 2 and Fig 5).

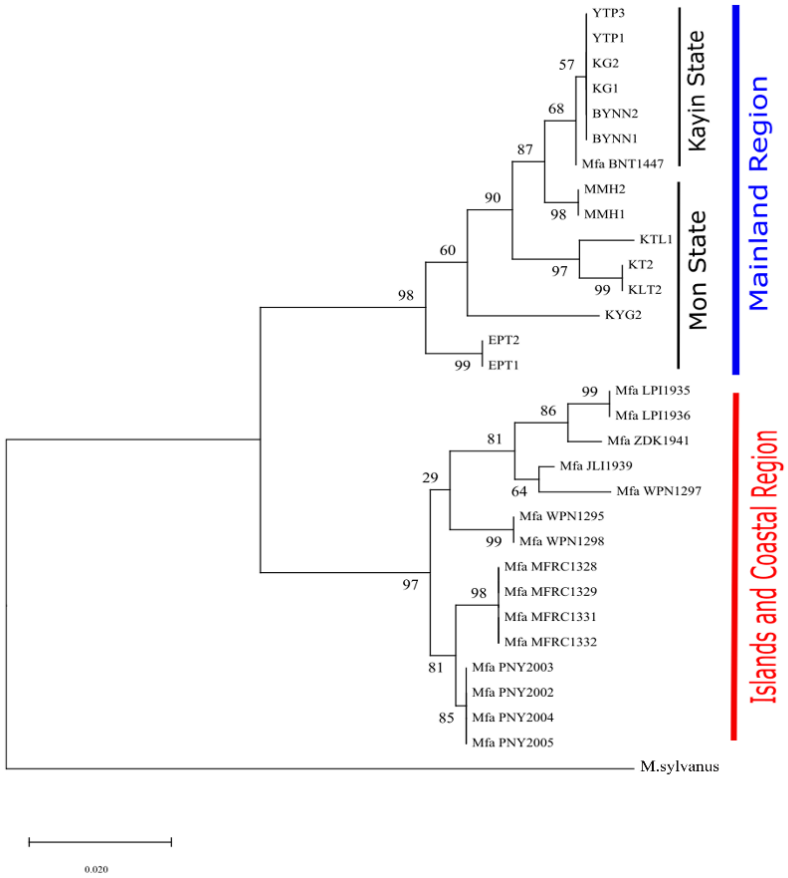

Figure 2 Maximum-likelihood phylogenetic tree of *M. fascicularis aurea* on mtDNA (675 bp) sequences

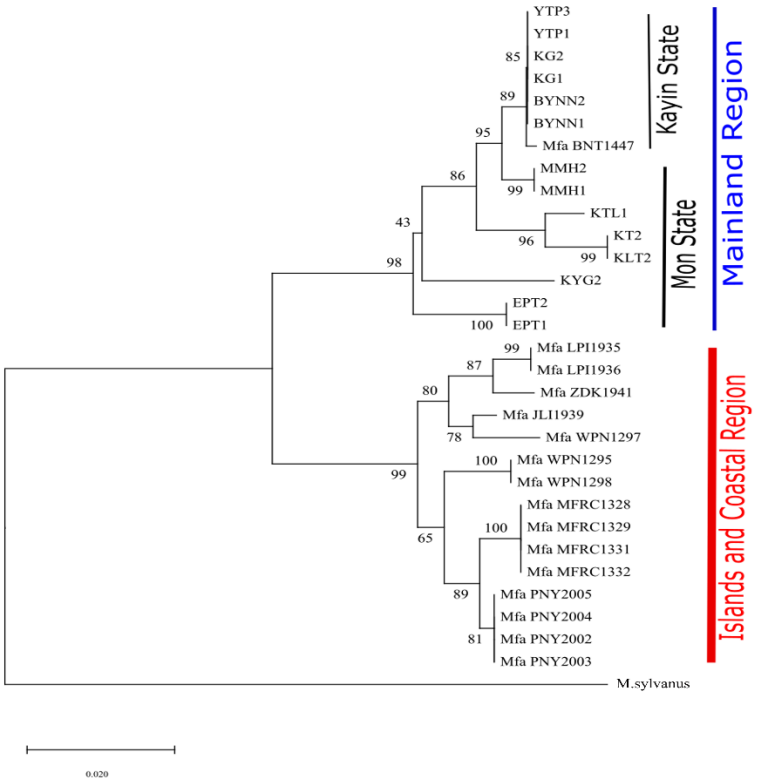

Figure 3 Neighbour-joining phylogenetic tree of *M. fascicularis aurea* on mtDNA (675 bp) sequences

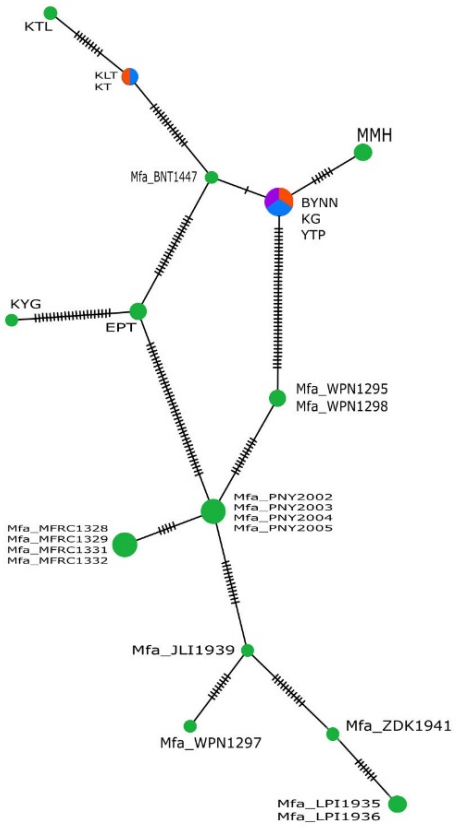

Figure 4 The Minimum-spanning network of mtDNA (675 bp). The short bars on each edge indicate substitutions

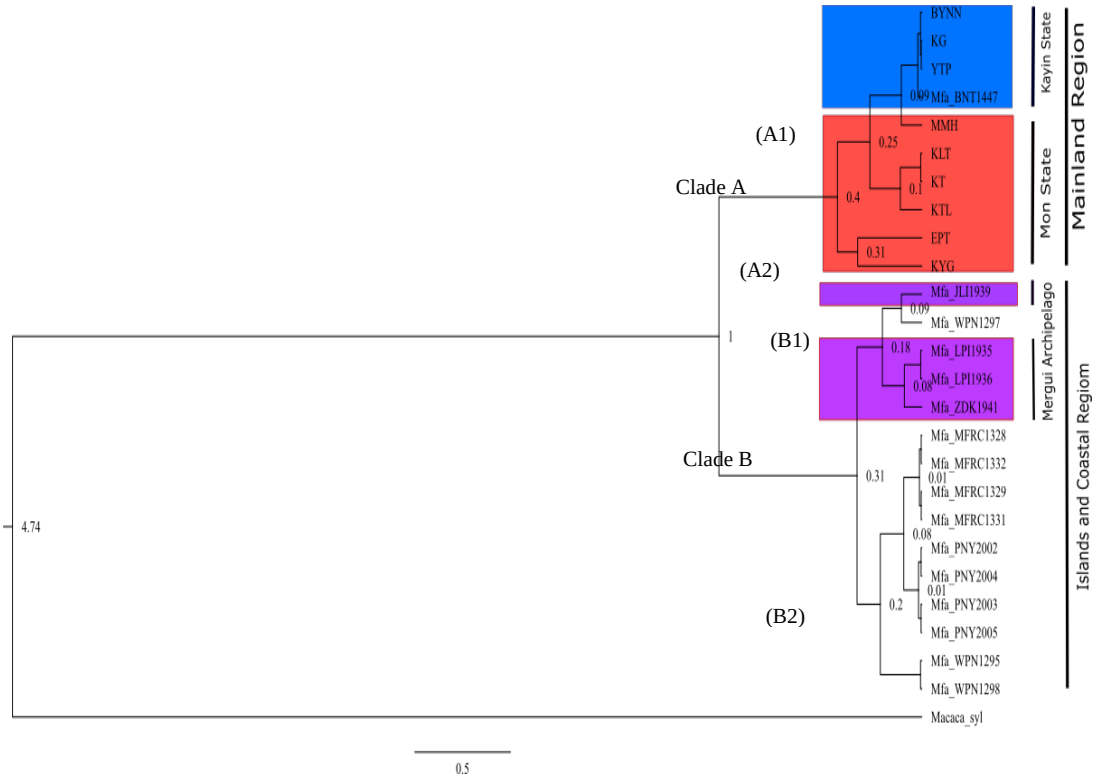

Figure 5 Phylogenetic tree with the estimated divergence times of partial mtDNA (675 bp)

Table 2 Divergence time of the mtDNA (in million years ago)

Divergence	Height median (MYA)	95% Credibility interval
Sylvanus- Asian macaques	4.74	3.27 - 6.37
Mainland-coastal region	1.00	0.40 – 1.85
The mainland clade (A1 and A2)	0.40	0.14 – 0.76
The mainland clade (Mon-Kayin State)	0.09	0.02 – 0.22
The coastal-island clade (B1 and B2)	0.31	0.10 – 0.59

Discussion

This study confirmed the presence of two clades among the mtDNA phylogeny of *M. fascicularis aurea*, one representing the populations from the mainland region and the other representing the populations from the coastal-island region, which was consistent with the mtDNA phylogeny shown by Bunlungsup *et al.*, 2016. In the previous study, only one sequence from the mainland region (BYNN) was implemented in the analysis, and thus the presence of the mainland clade and its geographical extent were not clear. In this study, all nine populations from Mon and Kayin States formed one clade, and thus maternally close relationship among the Mon and Kayin populations was uncovered.

Among the mainland clade, samples from the three populations in Kayin State (BYNN, KG, YTP) showed the same mtDNA haplotype (except that one previous published sequence

showed one nucleotide difference from the common haplotype) and thus genetically distinct from the populations of Mon State. These three populations are located between two major rivers, Donthami and Thanlwin, which might prevent the migration of monkeys between Mon and Kayin States as natural barriers.

Among the populations from Mon State, no clear relationship between the geographic distributions and the phylogenetic relationships of mtDNA was observed. This might be because of the absence of major geographic barriers among the Mon State populations and/or low resolution in the use of HVS1 of mtDNA.

The estimated divergence time of mtDNA between the mainland clade, and the coastal-island clade was about 1.00 MYA which was consistent with the age estimated by the whole mitochondrial genomes (Matsudaira *et al*., 2018). The divergence time estimates of the mtDNA sequence analysis indicated that the divergence of the mainland clade started earlier (0.40 MYA) than the coastal-island region (0.31 MYA). Further study including more populations from both areas is required to confirm this observation.

Still, there is a gap of unsampled populations between the northern part of Mon State sampled in the present study and the southern part of Mergui Archipelago sampled in the previous study. Further sampling of populations located between the two areas will uncover the border of the distribution of the two mtDNA clades. The border may reflect the maternal origin of *M. fascicularis aurea* where the mtDNA introgression from the sinica-species group occurred. In addition, there are some populations distributed along the Bay of Bengal which has not been studied. Further investigation of *M. fascicularis aurea* in the area should also be studied to delineate the scenario of the origin of *M. fascicularis aurea.*

Conclusion

The phylogenetic study of *M. fascicularis aurea* was conducted for the populations of Mon and Kayin States based on mtDNA. The phylogenetic analyses confirmed that the populations of Mon and Kayin States are maternally related to each other and form the mitochondrial clade, the mainland clade, which is distinct from that of the coastal-island populations. Furthermore, the estimated divergence time suggested that the divergence among the mainland populations may have started earlier than that of the coastal-island populations. Still, the exact origin of the *M. fascicularis aurea* is not clear. Further studies focusing on the unsampled region, i.e., the gap between the mainland and island-coastal regions, and the Bay of Bengal are expected to uncover more details of the evolutionary history of *M. fascicularis aurea*.

Acknowledgments

We would like to thank Dr Win Win Myint, Professor and Head (Retired) and Dr Aye Aye Myint, Professor and Head, Department of Zoology, Mawlamyine University for their permission to conduct this research work. We are also grateful to the monks and local people from the study areas for their good-humoured and sincere responses during our interview. We are most grateful to Wildlife Conservation Society (WCS) Myanmar for the research grant to carry out the present work.

References

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

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A. and Drummond, A. J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis, *Plos Computational Biology* 10:e1003537.

Bunlungsup, S., Imai, H., Hamada, Y., Gumert, M. D. and San, A. M. (2016) "Morphological characteristics and genetic diversity of Burmese long-tailed macaques (*Macaca fascicularis aurea)" American Journal of Primatology* 78:441-455.

Felsenstein, J. (1981) "Evolutionary trees from DNA sequences: a maximum likelihood approach" *Journal of Molecular Evolution* 17: 368-376.

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

Gumert, M. D., Kluck, M. and Malaivijitnond, S. (2009) "The physical characteristics and usage patterns of stone axe and pounding hammers used by long-tailed macaques in the Andaman Sea region of Thailand" *American Journal of Primatology* 71:594-608.

Hayaishi, S. and Kawamoto, Y. (2006) "Low genetic diversity and biased distribution of mitochondrial DNA haplotypes in the Japanese macaque (*Macaca fuscata yakui*) on Yakushima Island" *Primates* 47:158-164.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) "MEGA X: molecular evolutionary genetics analysis across computing platforms" *Molecular Biology and Evolution* 35:1547-1549.

Leigh, J. W. and Bryant, D. (2015) "POPART: full-feature software for haplotype network construction" *Methods in Ecology and Evolution* 6:1110-1116.

Malaivijitnond, S., Lekprayoon, C., Tandavanittj, N., Panha, S., Cheewatham, C. and Hamada, Y. (2007) "Stone-tool usage by Thai long-tailed macaques (*Macaca fascicularis*)" *American Journal of Primatology* 69:227-233.

Matsudaria, K., Hamada, Y., Bunlungsup, S., Ishida, T., San, A. M. and Malaivijitnond, S. (2018) "Whole mitochondrial genomic and Y-chromosomal phylogenies of Burmese long-tailed macaque (*Macaca fascicularis aurea*) suggest ancient hybridization between fascicularis and sinica species groups" *Journal of Heredity* 109*:*360-371.

Osada, N., Matsudaira, K., Hamada, Y. and Malaivijitnond, S. (2021) "Testing sex-biased admixture origin of macaque species using autosomal and X-chromosomal genomic sequences" *Genome Biology and Evolution* 13:evaa209.

Phadphon, P., Kanthaswamy, S., Oldt, R. F., Hamada, Y., Malaivijitnond, S. (2021) "Population structure of *Macaca fascicularis aurea*, and their genetic relationships with *M. f. fascicularis* and *M. mulatta* determined by 868 RADseq-derived autosomal SNPs—a consideration for biomedical research" *Journal of Medical Primatology* 51:33-44.

Roos, C., Kothe, M., Alba, D. M., Delson, E. and Zinner, D. (2019) "The radiation of macaques out of Africa: evidence from mitogenome divergence times and the fossil record" *Journal of Human Evolution* 133:114-132.

Saitou, N. and Nei, M. (1987) "The neighbor-joining method: a new method for reconstructing phylogenetic trees" *Molecular Biology and Evolution* 4:406-425.

Tamura, K. and Nei, M. (1993) "Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees" *Molecular Biology and Evolution* 10:512-526.