ZOOLOGICAL RESEARCH

Multilocus phylogeny suggests a distinct species status for the Nepal population of Assam macaques (*Macaca assamensis*): implications for evolution and conservation

Laxman Khanal[1,2,*], Mukesh Kumar Chalise[1,3], Peng-Fei Fan[4], Randall C. Kyes[5], Xue-Long Jiang[2,*]

[1] *Central Department of Zoology*, *Institute of Science and Technology*, *Tribhuvan University*, *Kathmandu* 44613, *Nepal*

[2] *State Key Laboratory of Genetic Resources and Evolution*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650203, *China*

[3] *Nepal Biodiversity Research Society (NEBORS)*, *Lalitpur* 23513, *Nepal*

[4] *School of Life Sciences*, *Sun Yat-sen University*, *Guangzhou*, *Guangdong* 510275, *China*

[5] *Department of Psychology*, *Global Health*, *and Anthropology*, *Center for Global Field Study*, *and Washington National Primate Research Center*, *University of Washington*, *Seattle* 98195, *USA*

ABSTRACT

Phylogenetic relationships within the *sinica*-group of macaques based on morphological, behavioral, and molecular characteristics have remained controversial. The Nepal population of Assam macaques (*Macaca assamensis*) (NPAM), the westernmost population of the species, is morphologically distinct but has never been used in phylogenetic analyses. Here, the phylogenetic relationship of NPAM with other congeners was tested using multiple mitochondrial and Y-chromosomal loci. The divergence times and evolutionary genetic distances among macaques were also estimated. Results revealed two major mitochondrial DNA clades of macaques under the *sinica*-group: the first clade included *M. thibetana*, *M. sinica*, and eastern subspecies of Assam macaque (*M. assamensis assamensis*); the second clade included *M. radiata* together with species from the eastern and central Himalaya, namely, *M. leucogenys*, *M. munzala*, and NPAM. Among the second-clade species, NPAM was the first to diverge from the other members of the clade around 1.9 million years ago. Our results revealed that NPAM is phylogenetically distinct from the eastern Assam macaques and closer to other species and hence may represent a separate species. Because of its phylogenetic distinctiveness, isolated distribution, and small population size, the Nepal population of

Received: 26 September 2020; Accepted: 05 January 2021; Online: 06 January 2021

Foundation items: This study was supported by the Chinese Academy of Sciences-World Academy of Sciences (CAS-TWAS) President's PhD Fellowship Program to L.K., Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0501), and National Key Research and Development Plan (#2017YFC0505202) China. R.C.K. was supported in part by the Office of Research Infrastructure Programs (ORIP) of the National Institutes of Health through Grant Number P51OD010425 to the WaNPRC. L.K. was supported by the Rufford Small Grants (17666-1) of Rufford Foundation, UK partly for fieldworks.

*Corresponding authors, E-mail: lkhanal@cdztu.edu.np; jiangxl@mail.kiz.ac.cn

DOI: 10.24272/j.issn.2095-8137.2020.279

sinica-group macaques warrants detailed taxonomic revision and high conservation priority.

Keywords: Himalaya; Macaques; Paraphyletic; *sinica*-group; Taxonomy

INTRODUCTION

Macaques (genus *Macaca*) are represented by 23 extant species (Fan et al., 2017) with wide distribution across East, Southeast, and South Asia, as well as limited areas in northwestern Africa. They are considered one among few successful taxa in terms of speciation and distribution (Evans et al., 2020; Hayasaka et al., 1996; Thierry et al., 2004). Despite many studies on macaques, several evolutionary questions remain (Li et al., 2009). Macaques are believed to have entered Eurasia via northeast Africa, followed by divergence of the Asian lineage into three or four species groups less than three million years ago (mya) (Roos et al., 2019; Tosi et al., 2003). Macaque species have been variously separated into several species groups; for example, Fooden (1976) defined four species groups based on genital morphology, whereas Groves (2001) classified macaques into six groups. Most recently, the genus *Macaca* has been divided into seven species groups based primarily on genital morphology, geographical distribution patterns, behavior, and genetics (Fan et al., 2017; Liedigk et al., 2015; Roos et al., 2019). Although there is a consensus at present on the seven species groups of macaques, phylogenetic relationships among different species and species groups are not yet conclusive.

Among the seven macaque species groups, the *sinica*-group is polytypic and represented by six extant species, i.e., *M. assamensis*, *M. thibetana*, *M. sinica*, *M. radiata*, *M. munzala*, and *M. leucogenys*. The group was initially defined based on male genitalia, with a subacute and sagittate-shaped glans in dorsal view (Fooden, 1976). Former molecular phylogenetic studies described the *sinica*-group as a monophyletic assemblage of parapatric species, possibly derived from *M. fascicularis*-like ancestors (Tosi et al., 2003). Recent description of two new species, i.e., *M. munzala* (Sinha et al., 2005) and *M. leucogenys* (Li et al., 2015), has enriched the group composition. However, the taxonomic position of *M. munzala* has been questioned by some who argue that the morphological variations used to describe this species may be merely adaptive traits to colder climates (Biswas et al., 2011). Similarly, the circular glans of *M. leucogenys* (Li et al., 2015) is not sufficiently common for basic taxonomic studies of the *sinica*-group, thus introducing further questions. Additionally, consideration of a potential new subspecies of *M. assamensis* from Nepal (Molur et al., 2003), which still lacks detailed phylogenetic assessment, has contributed in the continued taxonomic confusion regarding this species group.

The *sinica*-group of macaques exhibits moderately fragmented distribution in South and Southeast Asia (Fan et al., 2017). The species in southern India and Sri Lanka include *M. radiata* and *M. sinica*; the Himalayan range harbors *M. munzala*, *M. leucogenys*, and *M. assamensis pelops* (western subspecies); the Southeast Asian region and China support *M. assamensis assamensis* (eastern subspecies) and *M. thibetana* (Fan et al., 2017; Khanal et al., 2018b). They are mostly parapatric, with few exceptions, e.g., *M. munzala*, *M. leucogenys*, *M. assamensis*, and *M. thibetana* populations from southeastern Tibet and Arunachal areas. However, taxonomic ambiguities exist among those sympatric populations (Biswas et al., 2011; Chakraborty et al., 2007; Fan et al., 2017; Li et al., 2015; Sinha et al., 2005).

Phylogenetic relationships among extant taxa in the *sinica*-group and the biogeographic processes leading to their current distribution are not well understood or are disputed (Khanal et al., 2018b). One of the main reasons for these discrepancies is the extensive and complex geographical and cyclical environmental changes that repeatedly occurred during the Pleistocene in the Himalayan range. Of note, the last interglacial (LIG) and last glacial maximum (LGM) left imprints of large-scale climatic fluctuation effects on the distribution and diversity of macaque species (Roos et al., 2019). In addition, the complex geography of the Himalayan region and presence of physiographic barriers like rivers and mountains have isolated wildlife populations, causing distinct genetic structures and speciation (Khanal et al., 2018a, 2018c). Furthermore, molecular phylogenetic studies on macaques have produced incongruent results with the use of various markers and analytical tools. Although mitochondrial DNA (mtDNA) and Y-chromosomal genes (TSPY and SRY) have been used extensively for phylogenetic analyses of macaques, the selection of loci and sizes of amplicons have varied among studies, resulting in variation in phylogenetic tree topology (Chakraborty et al., 2007; Fan et al., 2017; Li et al., 2009; Li & Zhang, 2004, 2005; Morales & Melnick, 1998; Page et al., 1999; Tosi et al., 2000, 2003; Ziegler et al., 2007). For example, the phylogenetic tree of the *sinica*-group of macaques using mtDNA sequences, supports *M. radiata*, *M. munzala*, and *M. leucogenys* as a monophyletic clade, with *M. thibetana* and *M. a. assamensis* (eastern subspecies) forming another monophyletic clade; in contrast, Y-chromosomal loci result in different and poorly resolved tree topologies, depicting a monophyly for all *sinica*-group taxa with the inclusion of *M. arctoides* (Fan et al., 2017).

The Nepal population of Assam macaques (NPAM) is one of the least studied primate populations (Chalise, 2008). The westernmost global distribution of Assam macaques was considered central Nepal (Wada, 2005); however, several troops have been reported recently in far-western Nepal, almost 300 km west of the formerly known limit (Khanal et al., 2019). The sporadically distributed populations in fragmented habitats, isolated by physiographic barriers like rivers and mountains, display variations in both morphology and behavior at different latitudes and elevations (Chalise, 2005, 2013; Khanal et al., 2018a). They also differ in pelage color, morphometric indices, and distribution ranges, with the

nearest conspecific populations from adjacent countries like India and Bhutan (Molur et al., 2003). Considering these distinctions, NPAM has been suggested for distinct subspecies status and thus warrants further detailed taxonomic elaboration (Boonratana et al., 2020; Chalise, 2013). Despite being designated as the "*M. assamensis* Nepal Population" since 2003 (Molur et al., 2003), NPAM still lacks the molecular assessment needed to help resolve its obscure taxonomy. Therefore, there is need for a complete assessment of this macaque species and its congeners from eastern and central Himalaya to elucidate the taxonomy of the *sinica*-group. Therefore, using multiple mitochondrial and Y-chromosomal DNA loci, this study aimed to: (i) establish the phylogenetic relationship of NPAM with its congeners; (ii) estimate the divergence time of the study population from sister taxa; and (iii) evaluate the evolutionary distance between NPAM and other macaques. With the first ever inclusion of DNA sequences from the westernmost population of Assam macaques in phylogenetic analyses, we identified its evolutionary importance and recommend this species be given strong conservation priority.

MATERIALS AND METHODS

Sampling

A total of 277 fecal samples were collected non-invasively from 40 wild troops of Assam macaques covering the known distribution range of the species in Nepal (Figure 1). Sampling details are described in Khanal et al. (2018a). The fecal samples were placed in lysis buffer and stored at ambient temperature before extraction of genomic DNA.

DNA extraction, PCR amplification, and sequencing

Total genomic DNA was extracted from fecal samples using a QIAamp DNA Stool Mini Kit (Qiagen, Germany) following the manufacturer's protocols, with some modification. In brief, 2 mL sample tubes with lysis buffer were centrifuged at 12 000 r/min for 2 min at room temperature. After this, supernatants (600 μL) were removed for further processing. Propanol stored at –20 ºC was used instead of ethanol and final elution was performed using 75 μL of elution buffer. Genomic DNA was also extracted from four skin samples of *M. leucogenys* from Tibet used in the first phylogenetic analysis of the species (Fan et al., 2017) with a QIAamp DNA Blood and Tissue Kit (Qiagen, Germany) following the manufacturer's protocols.

Four mitochondrial loci, cytochrome b (cyt *b*), control region/D-loop (hypervariable region 1 [HVR1] and hypervariable region 2 [HVR2]), and 16S rRNA were used to trace maternal genealogy. Primer pairs (Table 1) under respective annealing temperatures were used with 1.0 μL of template DNA in a 20 μL PCR mixture (95 ºC for 5 min initial denaturation; 35 cycles of denaturation at 94 ºC for 30 s, 45 s annealing, and 1 min extension at 72 ºC; followed by final extension for 10 min at 72 ºC). PCR amplification of respective loci for *M. leucogenys* was performed in similar cycles with 0.5 μL of genomic DNA as the template.

Paternal genealogy and migration patterns of males were

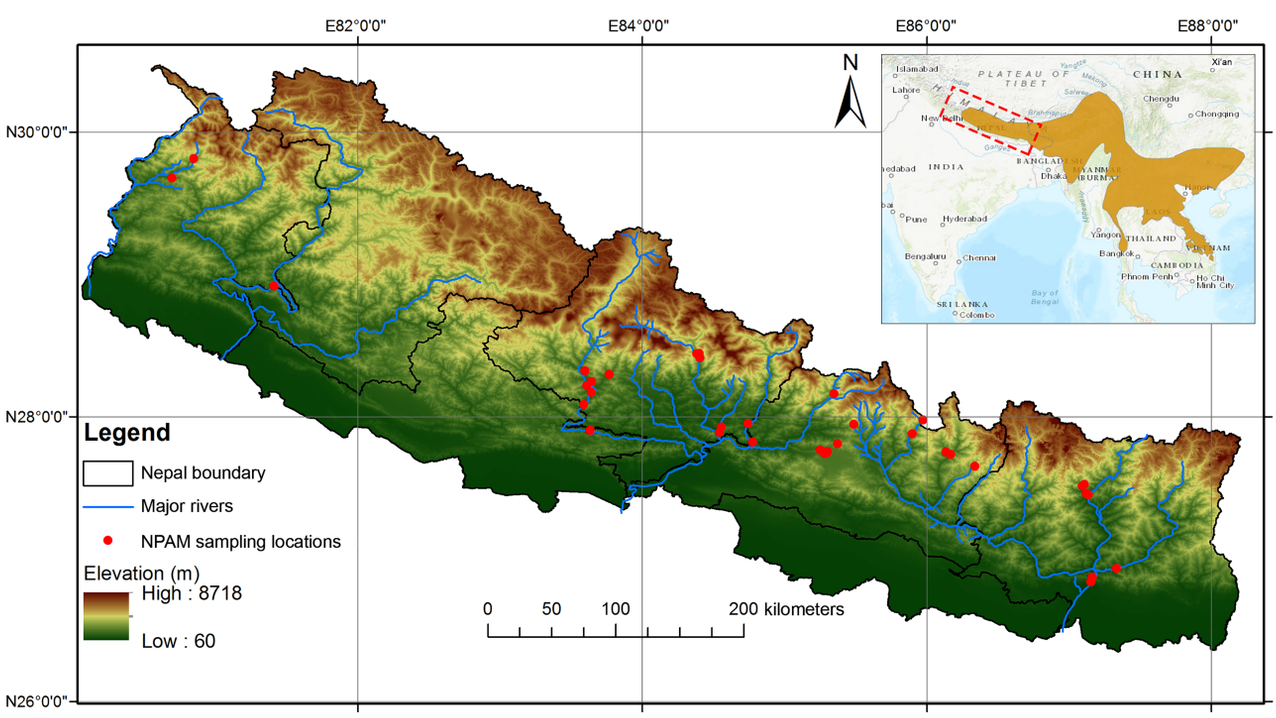

Figure 1 Map of Nepal showing NPAM sampling locations

Inset shows distribution range of Assam macaques, adapted from IUCN Red List 2020 (Boonratana et al., 2020).

Table 1 Primer pairs and PCR conditions for amplification of studied loci

S.N.	Locus	Primer pairs (Forward/Reverse)		Ann. temp.	Length (bp)	References
		Name	Sequence (5′–3′)			
1.	cyt *b*	CYTF	AACCATCGTTGTACTTCAAC	56	1 140	Khanal et al., 2018a
		CYTR	TCTGGTTTACAAGGCCAGTG			Khanal et al., 2018a
2	D-loop	LqqF	TCCTAGGGCAATCAGAAAGAAAG	58	1 090	Li & Zhang, 2004
		Saru5	GGCCAGGACCAAGCCTATTT			Hayasaka et al., 1991
3	16S rRNA	284F-EX	GGATTAGATACCCCACTATGCTTG	58	1 397	Tosi et al., 2003
		384R-EX	GCTACCTTTGCACRGTCAGGGTACCG			Tosi et al., 2003
4	TSPY1	TSPYA	AGCCAGGAAGGCCTTTTCTCG	60	2 203	Kim et al., 1996
		TSR1012	TGTCACCTGTGACGTTCACGA			Bunlungsup et al., 2016
	TSPY2	TSF566	AGGTCATTCATGGATGCAGAT	64		Bunlungsup et al., 2016
		TSR1676	CCACAGTTATAACCTGCTTTGC			Bunlungsup et al., 2016
	TSPY3	TSF1383	AATCCCCTGCAATACTACAGGAGG	64		Bunlungsup et al., 2016
		TSPY5R	CTGTGCATAAGACCATGCTGAG			Tosi et al., 2000
5	SRY	SW2	CTTGAGAATACATTGTCAGGG	56	764	Whitfield et al., 1993
		SW3B	AGGTCTTTGTAGCCAATGTTACCCG			Whitfield et al., 1993

traced using two nuclear loci linked with the non-recombining region of the Y-chromosome: i.e., TSPY (testis specific protein Y-chromosome) and SRY (sex determining region Y-chromosome) genes. The TSPY gene (2 220 bp) was amplified using three primer pairs and 2.0 μL of template DNA in a 20 μL PCR mixture under respective annealing temperatures (Table 1). The SRY (770 bp) locus was amplified using 2.0 μL of genomic DNA and primer pairs (Table 1). The PCR protocols for both Y-chromosomal loci were run for 45 cycles (95 ºC for 5 min initial denaturation; 45 cycles of 45 s denaturation at 94 ºC, 45 s annealing at respective temperatures, and 75 s extension at 72 ºC; followed by final extension for 10 min at 72 ºC).

The amplicons were tested by electrophoresis in 1.2% agarose gels. Direct sequencing of both DNA strands of successful amplicons was performed using a Big Dye Terminator Cycle Sequencing Kit (Invitrogen, USA) and resolved in an ABI Prism 3700 Automated Sequencer (Applied Biosystems, USA).

Precautions were taken during the entire process of sample collection, DNA extraction, and PCR amplification to avoid contamination: (i) two sterilized cotton swabs were packed separately and used for each sample collection to help avoid human contact; (ii) tightly closed individual sample tubes were packed separately in polythene bags; (iii) fecal samples located at least 2 m apart on the ground were considered to be from separate individuals for collection; (iv) DNA extraction and PCR amplification steps were conducted at a separate location where no other lab work was ongoing; (v) only six samples were processed in one batch; and (vi) every PCR amplification batch included one ‘negative control’ test to ensure a contamination-free process.

The obtained amplicons, which were the intended loci, were confirmed not to be exogenous contaminants or nuclear insertions of mitochondrial DNA (numts) because: (i) macaque-specific pairs of primers that successively failed to amplify high-quality human DNA were used for each locus; (ii) a single PCR amplicon of expected size was detected in all individuals; and (iii) no highly variant sequences were detected. Moreover, fecal samples are richer in mtDNA molecules than in nuclear DNA, and the natural degradation of fecal DNA makes nuclear copies even scarcer, reducing the likelihood of amplifying numts. For the amplification of TSPY and SRY loci, 2.0 μL of template DNA was used for each 20 μL PCR mixture. This produced amplicons without abnormal sequences, which were then confirmed as macaque DNA sequences of the respective loci based on BLAST search of the GenBank database.

Sequence alignment

Sequences were successfully obtained from 208 fecal samples for mtDNA and 32 samples for Y-chromosomal loci from NPAM. Population genetic analysis of DNA sequences from NPAM with the same sampling effort did not reveal a distinct structure (Khanal et al., 2018a). Therefore, a subset of five mtDNA sequences representing the diverse sampling localities of NPAM were used for downstream analysis. All 32 sequences of the Y-chromosomal loci were monomorphic; hence, a single sequence was used for phylogenetic investigation. The HVR1 (492 bp) and HVR2 (478 bp) regions were extracted from the D-loop sequences (1 090 bp) of NPAM to make them comparable to sequences of other macaques in GenBank. The DNA sequences were assembled using DNASTAR Lasergene v.7.1 and aligned using CLUSTAL W (Thompson et al., 1994) implemented in MEGA X (Kumar et al., 2018). The sequences of the respective genes belonging to the eastern subspecies of Assam macaque and other macaque species were accessed from GenBank (Supplementary Tables S1, S2). The sequences for the mitochondrial loci of each macaque species (except *M. munzala*) were retrieved from whole mitochondrial genome (mitogenome) sequences. For mtDNA of *M. munzala* and Y-chromosomal loci of all macaques, sequences were retrieved

from GenBank representing species from the same or close sampling localities. The corresponding sequences of *Papio hamadryas* for each locus under study were downloaded from GenBank and used as the outgroup for phylogenetic analyses. Each locus was then aligned individually using MUSCLE (Edgar, 2004) in MEGA X (Kumar et al., 2018).

Phylogenetic analyses

The phylogenetic relationships of NPAM were tested using mtDNA (cyt *b*, HVR1, HVR2, and 16S rRNA) and Y-chromosomal (TSPY and SRY) loci separately. Phylogenetic analyses were performed using the maximum-likelihood (ML) algorithm in RAxML v.8.2.11 (Stamatakis, 2014) and Bayesian inference (BI) in BEAST v.1.10 (Suchard et al., 2018).

ML analysis was conducted with 1 000 bootstrap replicates using the partition scheme determined by the GTR+Γ model in PartitionFinder v.1.0.1 (Lanfear et al., 2012). For BI analyses, the best partitioning schemes and mutational models were estimated using 12 evolutionary models (TrNef, SYM, F81, HKY, TrN, GTR, TrNef+Γ, SYM+Γ, F81+Γ, HKY+Γ, TrN+Γ, and GTR+Γ) in PartitionFinder v.1.0.1 (Lanfear et al., 2012) based on Bayesian information criterion (BIC) (Supplementary Table S3). BI analyses were performed with two independent Markov Chain Monte Carlo (MCMC) chains, each with two million generations. Samples were collected every 1 000 generations, discarding the first four of the total generations as burn in. A <0.01 deviation of split frequencies between the two independent runs ensured the convergence of the two MCMC chains. Convergence was further assessed using TRACER v.1.6 (Rambaut et al., 2014), with an effective sample size (ESS) >200.

For estimation of divergence time of the maternal lineage of macaques, a time-tree was constructed in BEAST using the concatenated mtDNA genes (3 526 bp). The estimated uncorrelated lognormal relaxed clock was selected as the clock model, and the Yule process and piecewise and constant root model were used as tree priors with 2×10^7 MCMC generations and sampling every 10 000 generations. Nodal calibrations were performed using the previously defined *Papio-Macaca* split at 8 mya and most recent common ancestor of macaques at 5.5 mya (Tosi et al., 2003).

Kimura's two-parameter (K2P) model (Kimura, 1980) of genetic distance is the most effective model for assessing interspecific and intraspecific genetic distances at low levels of divergence (Nei & Kumar, 2000; Hebert et al., 2003). Therefore, pairwise evolutionary distances between *sinica*-group species were calculated with the K2P model of nucleotide substitution using cyt *b* sequences in MEGA X (Kumar et al., 2018).

RESULTS

Mitochondrial gene trees

Both the ML (Figure 2) and BI (Figure 3) trees based on 3 526 bp long mtDNA sequences revealed two distinct clades within the *sinica*-group of macaques. Results supported NPAM, *M. munzala* from India, *M. leucogenys* from Tibet, and *M. radiata* from southern India forming a monophyletic clade (Clade-2) with relatively high nodal support. The eastern subspecies of Assam macaque (*M. a. assamensis*), *M. sinica*, and *M. thibetana* formed another clade (Clade-1). NPAM, which is currently believed to be the westernmost subspecies of Assam macaques (*M. a. pelops*), belonged to Clade-2, whereas the eastern subspecies (*M. a. assamensis*) from southern China fitted into Clade-1.

Y-chromosomal gene trees

The phylogenetic tree based on Y-chromosomal genes (TSPY and SRY; 2 967 bp) resulted in four major clades of macaque species. All six known species under the *sinica*-group, together with NPAM and *M. arctoides*, remained in the same clade (Figure 4). The intraspecific genetic variations within NPAM and the interspecific genetic variations among different species of macaques were very low in Y-chromosomal genes. This low variation resulted in a phylogenetic tree with relatively low nodal support.

Divergence time estimation among macaques

The divergence time estimation results suggested that the time of the most recent common ancestor for the *sinica*-group of macaques was 2.95 mya, when the ancestral stock of Clade-1 (*M. a. assamensis*, *M. sinica*, and *M. thibetana*) and Clade-2 (NPAM, *M. leucogenys*, *M. munzala*, and *M. radiata*) diverged (Figure 5). NPAM was the first to diverge from the other Clade-2 species stock about 1.9 mya. Furthermore, *M. radiata* diverged from the ancestral stock about 1.8 mya, whereas *M. leucogenys* and *M. munzala* split about 1.5 mya.

Evolutionary distance among macaques

The genetic distances estimated by the K2P model using cyt *b* sequences (Table 2) showed very low (0.039) genetic distance between NPAM and *M. munzala* from north-eastern India, followed by *M. radiata* (0.055) and *M. leucogenys* (0.069). The genetic distance between NPAM and *M. a. assamensis* from Yunnan, China, was much higher (0.093).

DISCUSSION

In this study, we examined the phylogenetic position of the Nepal population of Assam macaques (NPAM) based on phylogenetic analyses of extant macaques using multiple maternally inherited mitochondrial and paternally inherited Y-chromosomal loci. Our study is novel as it includes: (i) ~6.5 kb long DNA sequences from multiple genes; (ii) additional gene segments from *M. leucogenys* not used in previous phylogenetic analyses (Fan et al., 2017); and (iii) for the first time, sequences from the previously unsampled westernmost population of Assam macaques from the Nepal Himalaya. The increase in taxon sampling, larger size of DNA sequences, and broader geographic coverage provide better insight into the evolution of the *sinica*-group of macaques.

Phylogenetic distinctiveness of NPAM

The phylogenetic assignment of macaques has remained

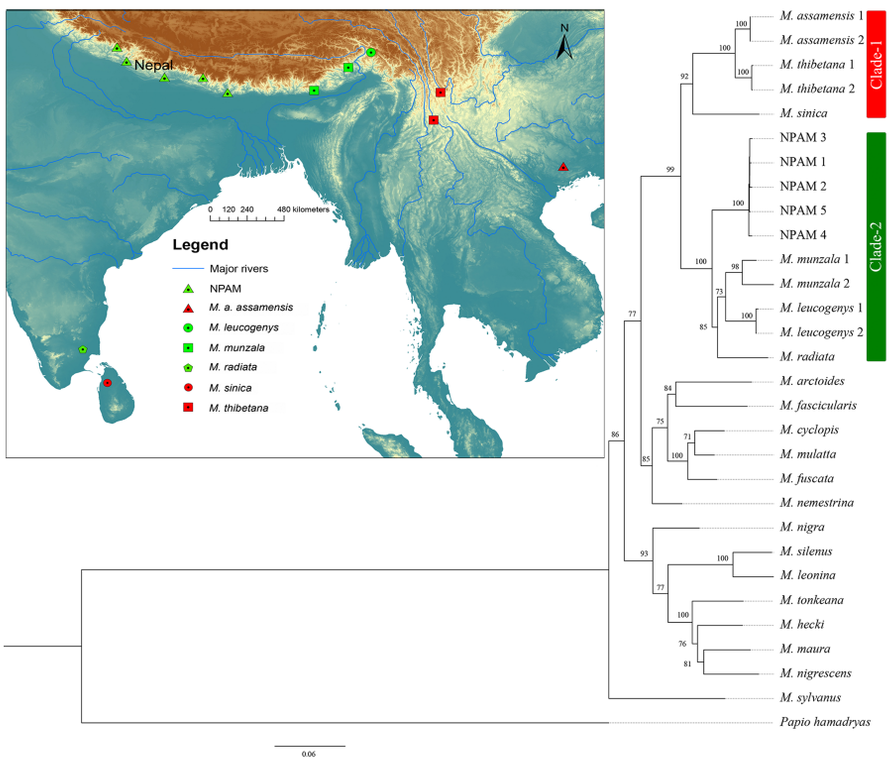

Figure 2 ML gene tree of concatenated mtDNA (3 526 bp) among *sinica*-group macaques

Node values represent RAxML percentage bootstrap probability. Inset shows geographical distribution of *sinica*-group species; source: *M. leucogenys* (Fan et al., 2017); *M. munzala* (Biswas et al., 2011; Chakraborty et al., 2014; Sarania et al., 2017), *M. a. assamensis* (Minge et al., 2016; Zhou et al., 2011); NPAM (Khanal et al., 2018a, 2019).

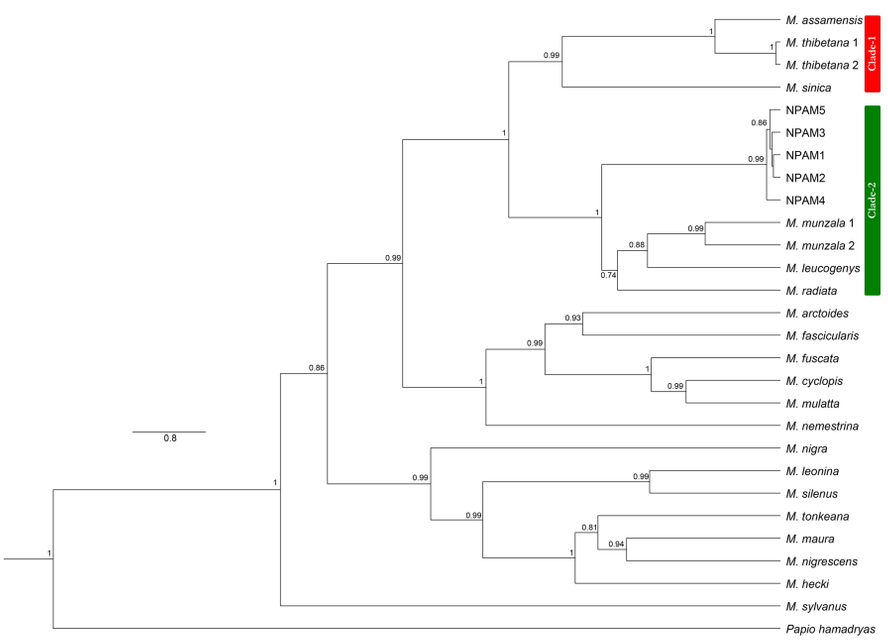

Figure 3 BI tree of concatenated mtDNA (3 526 bp) among macaques

Node numbers are Bayesian posterior probability values.

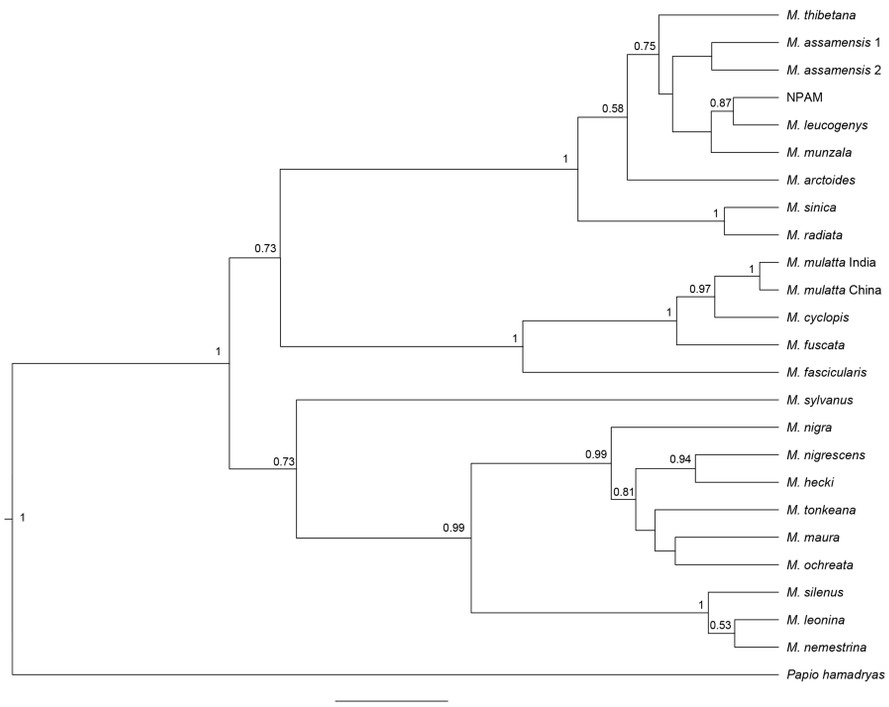

Figure 4 BI tree of concatenated Y-chromosomal genes (2 967 bp)

Node numbers are Bayesian posterior probability values.

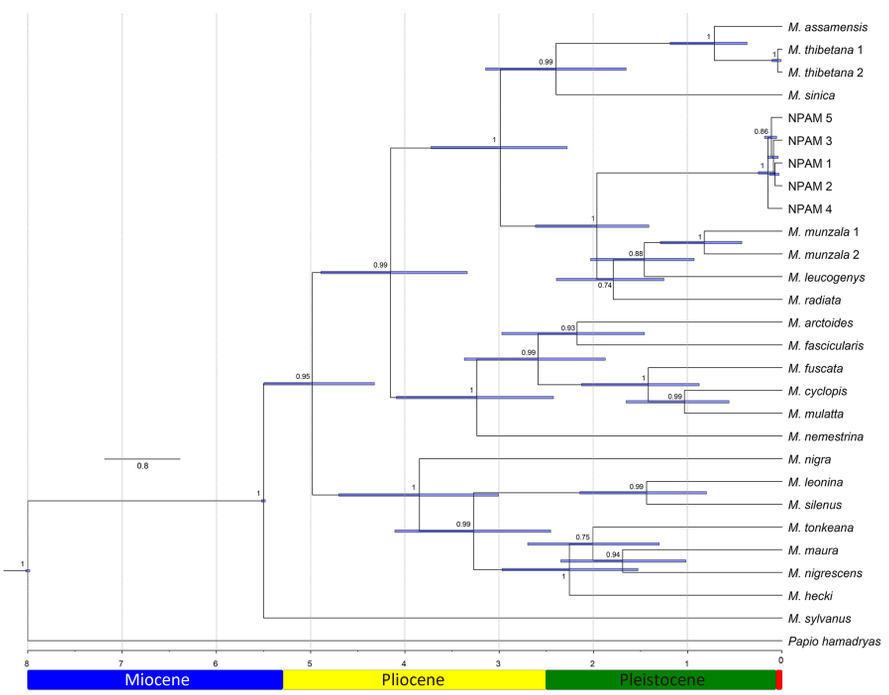

Figure 5 Ultrametric tree showing divergence time estimates for macaques based on concatenated mtDNA sequences (3 526 bp) using BEAST

Blue bars indicate 95% highest posterior densities of divergence times and time scale below represents million years ago before present; nodal values represent Bayesian posterior probabilities.

Table 2 Estimates of evolutionary divergence over sequence pairs between *sinica*-group macaques using K2P model with cyt *b* sequences

Species	1	2	3	4	5	6	7
1. NPAM		0.01	0.013	0.016	0.012	0.016	0.016
2. *M. munzala*	0.039		0.013	0.018	0.012	0.017	0.017
3. *M. leucogenys*	0.069	0.073		0.016	0.013	0.016	0.016
4. *M. assamensis*	0.093	0.111	0.090		0.014	0.015	0.008
5. *M. radiata*	0.055	0.058	0.061	0.078		0.016	0.014
6. *M. sinica*	0.093	0.099	0.102	0.081	0.096		0.013
7. *M. thibetana*	0.090	0.102	0.087	0.026	0.075	0.067	

Evolutionary distances (below diagonal) and standard error estimates (above diagonal).

complicated due to the relatively recent evolution of the taxa and incongruent methods and sources of data used for analyses, such as pelage color, relative tail lengths, facial features, penile morphology, female reproductive anatomy, copulatory behaviors, and molecular markers. Furthermore, controversies exist around the taxonomic assignment of macaque species within the different species groups based on competing morphological (Fooden, 1979), paleontological (Delson, 1980), and genetic evidence (Fa, 1989; Li & Zhang, 2005). Multiple authorities have attempted to resolve the phylogenetic relationships among macaques and have proposed various phylogenetic positions of the taxa and number of species groups (Deinard & Smith, 2001; Fan et al., 2017; Fooden, 1979; Hayasaka et al., 1996; Li and Zhang, 2005; Morales & Melnick, 1998; Roos et al., 2019; Tosi et al., 2000, 2003). Mitogenomic analyses have greatly improved our taxonomic understanding of macaques (Evans et al., 2020; Liedigk et al., 2015; Matsudaira et al., 2018; Roos et al., 2019). However, the use of different DNA markers and unrepresentative spatial and taxon coverage have introduced further complications.

The phylogenetic analysis of the *sinica*-group of macaques has remained controversial for several reasons, such as inclusion of *M. arctoides* in the group, intraspecific paraphyly of some species (Li & Zhang, 2005), and existence of morphotypes in many species. Most molecular phylogenetic analyses of macaques show lower nodal/bootstrap support for monophyly of the *sinica*-group. For example, Hayasaka et al. (1996) reported 30%–58% bootstrap support for monophyly of the *sinica*-group. Similarly, Li & Zhang (2005) and Chakraborty et al. (2007) reported values of 55% and 75%, respectively. Low bootstrap values do not provide strong support for the associated analyses, as only bootstrap values ≥70% indicate sufficiently resolved topologies (Hillis & Bull, 1993). In the current study, the ML and BI analyses showed almost perfect nodal support (99% bootstrap percentage in ML and 1.00 Bayesian posterior probability in BI) for the monophyly of the overall *sinica*-group members and two reciprocally monophyletic clades (Clade-1 and Clade-2) based on mtDNA analysis. NPAM formed a monophyletic clade with *M. munzala*, *M. leucogenys*, and *M. radiata* with strong nodal support (100% for ML and 1.00 for BI). Interestingly, the eastern subspecies *M. a. assamensis* exhibited monophyly with *M. thibetana* with perfect nodal support in both analyses. This high nodal support might be associated with the use of longer DNA sequences from multiple genes.

In the last several decades, two new species, i.e., *M. munzala* (Sinha et al., 2005) and *M. leucogenys* (Li et al., 2015), have been added to the *sinica*-group. The species status of *M. munzala* has been questioned and was considered part of a population of Assam macaques found in the gap between the two subspecies ranges (Biswas et al., 2011). The morphological variations, such as darker pelage, larger body size, and shorter tails, were considered adaptive modifications for survival in the colder climates found at higher elevations and latitudes (Biswas et al., 2011). The white-cheeked macaque (*M. leucogenys*) originally described from southeastern Tibet (Li et al., 2015) also bears adaptive features like that of *M. munzala* (Biswas et al., 2011). In addition, *M. leucogenys* differs from other *sinica*-group species in having a rounded glans penis, in contrast to the most distinctive feature of the group, i.e., sagittate-shaped glans penis (Fan et al., 2017; Li et al., 2015). Phylogenetic analyses of *M. munzala* (Chakraborty et al., 2007) and *M. leucogenys* (Fan et al., 2017) lacked comparison with congeneric or conspecific populations from neighboring Himalayan countries, including Nepal. In addition to these ambiguities, NPAM has been considered as a potential new subspecies due to its morphological variations compared to conspecific populations (Boonratana et al., 2020; Chalise, 2005, 2013; Molur et al., 2003) but has not been included in previous phylogenetic analyses.

Use of DNA markers of longer length should result in more convincing phylogenetic inferences for the complex relationships among Himalayan macaques. In the current study, multiple mtDNA and Y-chromosomal loci were used to infer the phylogenetic relationships among the *sinica*-group of macaques. Our analyses revealed two major facts about the phylogeny of these macaques: (i) monophyly of NPAM, *M. munzala*, *M. leucogenys*, and *M. radiata* clade; and (ii) polyphyly of *M. assamensis*. NPAM (currently believed to be a western subspecies *M. a. pelops*) belonged to Clade-2, whereas the eastern subspecies (*M. a. assamensis*) belonged to Clade-1. These findings are further supported by the lower

level of genetic differences between NPAM and *M. munzala* (0.039) and *M. leucogenys* (0.069) but much higher evolutionary distance (0.093) with *M. a. assamensis*. This evidence strongly supports the possibility of a distinct species status for NPAM.

Divergence and evolutionary significance of NPAM

Estimations of divergence times among *sinica*-group members vary considerably among studies. Chakraborty et al. (2007) estimated divergence of *M. munzala* from its ancestral stock 0.48 mya, whereas Fan et al. (2017) estimated it to be 1.17 mya. Furthermore, Fan et al. (2017) estimated divergence of *M. leucogenys* from other species of the *sinica*-group around 2.51 mya, with the species being the first to diverge from common ancestral stock of the group. However, Tosi et al. (2003) estimated the last common ancestor for the *sinica*-group of macaques to be 1.5±0.6 mya based on Y-chromosomal genes and 1.7 mya based on mtDNA. Therefore, the incongruences in divergence times among these studies may be associated with the markers being used; the divergence time of *M. munzala* appears to be underestimated while that of *M. leucogenys* appears to be overestimated. This study revealed that NPAM diverged ~1.9 mya from the stock of *M. leucogenys*, *M. munzala*, and *M. radiata*. Furthermore, *M. radiata* appeared to diverge from the rest of the monophyletic clade 1.8 mya; and finally, *M. munzala* and *M. leucogenys* split around 1.5 mya. These results are congruent with the geographical distribution of the taxa in eastern and central Himalaya, supporting that closely related species with proximal distribution may have diverged more recently. Paleontological analysis has revealed that the *sinica*-group species dispersed from Lufeng within the southeast corner of the Qinghai-Tibet Plateau and Mt. Hengduan during the upper Pliocene (Li et al., 2020). After divergence, NPAM, the westernmost population of the species, may have been the first to disperse westward along the southern flank of the Himalaya, acquiring distinct genetic characters (Khanal et al., 2018a). The divergence time estimates using molecular data obtained in this study are corroborated by the fossil-based analysis of Li et al. (2020) and phylogeographical analysis of Khanal et al. (2018a). The phylogenetic distinctiveness of NPAM with *M. a. assamensis* could be the consequence of its westward dispersal, in contrast to the eastern subspecies from the point of divergence.

Tree topology and divergence time estimations provided strong support for the distinct species status of NPAM. Based on recent molecular phylogenetic study, Fan et al. (2017) concluded that the white-cheeked macaque (*M. leucogenys*) is a valid species because of its genetic uniqueness in both mitochondrial DNA and Y-chromosomal genes. They further doubted probable interbreeding amongst Sino-Himalayan macaque species due to imperfect reproductive separation. Male-mediated gene flow among the Himalayan macaque species is still unidentified. Previous studies have also reported different topologies of gene trees for matrilineally inherited mtDNA and patrilineally inherited Y-chromosomal genes (Chakraborty et al., 2007; Fan et al., 2017; Roos et al., 2019; Tosi et al., 2003). This study also revealed that the Y-chromosomal genes exhibited very low polymorphism among members of the *sinica*-group, which might be associated with the lower rate of mutation of nuclear DNA than mtDNA (Melnick & Hoelzer, 1993; Tosi et al., 2003). In addition, the low level of genetic variation in the Y-chromosomal genes may be associated with the interspecific genetic introgression brought by male-biased dispersal because females in cercopithecine primates are philopatric. The taxonomic-phylogenetic problems connected with the species-level classification of the eastern and central Himalayan species (NPAM, *M. leucogenys* and *M. munzala*) should be evaluated as a possible evolutionary species complex under the process of speciation. Their adjacent distribution range in the Himalaya, morphological resemblance, Y-chromosomal gene homogeneity, and lesser pairwise evolutionary/genetic distance provide support for this hypothesis. Therefore, the taxonomic ambiguity among *M. munzala*, *M. leucogenys*, and NPAM needs to be resolved by detailed phylogenetic assessment using multiple mitochondrial and nuclear DNA loci together with morphological, behavioral, and distribution data, and should include samples from the type localities of *M. a. assamensis* in Assam, India, and *M. a. pelops* in western Arunachal.

Conservation implications

This study confirmed that NPAM is phylogenetically distinct. Population genetic analysis covering the entire distribution range within the Nepal territory has revealed low genetic diversity and shallow genetic structure among the troops in the fragmented landscape (Khanal et al., 2018a). The macaque population is distributed along a narrow elevational range of mid-hills (Chalise, 2013; Khanal et al., 2019) and it is a habitat specialist requiring broad-leaved riverine forest (Khanal et al., 2019). Late Quaternary climatic fluctuations left an imprint in the genetic structure of the population that suggests macaques experienced a range shift in the past (Khanal et al., 2018a) and ongoing anthropogenic climate change could have significant effects on the future survival and distribution of the population. Additionally, more than half of the population currently resides outside protected areas (Khanal et al., 2019) and incidents of human-macaque conflict, especially driven by crop-raiding, are high. Habitat loss and alteration as well as retaliatory killings continue to threaten this small population of Nepal Assam macaques. Therefore, given its low genetic diversity, phylogenetic distinctiveness, small population size, sporadic distribution, fragmented habitat, and on-going anthropogenic pressure, NPAM is recommended as an Evolutionarily Significant Unit (ESU) candidate (Moritz, 1994) requiring high conservation priority.

CONCLUSIONS

Phylogenetic inference of the *sinica*-group of macaques using multiple mtDNA loci concluded with strong nodal support that NPAM is phylogenetically closer to *M. munzala* and *M.*

leucogenys than to the eastern subspecies of Assam macaque (*M. a. assamensis*). Among the members of the monophyletic clade of *sinica*-group macaques from the Himalaya, NPAM was the first taxon to diverge (around 2.0 mya). Our results suggest that NPAM is likely a distinct species within the *sinica*-group of macaques. Therefore, detailed phylogenetic assessment of NPAM, including Himalayan species of the *sinica*-group of macaques, using mitogenomic and nuclear DNA with morphological, geographical, and behavioral data is warranted.

SCIENTIFIC FIELD SURVEY PERMISSION INFORMATION

Permission for fieldwork in Nepal was granted by Department of National Parks and Wildlife Conservation (072/73ECO20-483) and Department of Forest and Soil Conservation (2072 /073-943) of the Government of Nepal.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

L.K., M.K.C., and X.L.J. designed the study. L.K, M.K.C., and P.F.F. collected the samples and L.K. performed laboratory analysis. L.K. analyzed the data and wrote the manuscript with input from P.F.F., R.C.K., and X.L.J. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank the Department of National Parks and Wildlife Conservation and Department of Forest and Soil Conservation of the Government of Nepal for research permission, and our field assistants for their support in the field.

REFERENCES

Biswas J, Borah DK, Das A, Das J, Bhattacharjee PC, Mohnot SM, Horwich RH. 2011. The enigmatic Arunachal macaque: its biogeography, biology and taxonomy in northeastern India. *American Journal of Primatology*, **73**(5): 458−473.

Boonratana R, Chalise MK, Das J, Htun S, Timmins RJ. 2020(2020-07-12). The IUCN red list of threatened species. https://www.iucn.org/theme/species/our-work/iucn-red-list-threatened-species.

Bunlungsup S, Imai H, Hamada Y, Gumert MD, San AM, Malaivijitnond S. 2016. Morphological characteristics and genetic diversity of Burmese long-tailed Macaques (*Macaca fascicularis aurea*). *American Journal of Primatology*, **78**(4): 441−455.

Chakraborty D, Ramakrishnan U, Panor J, Mishra C, Sinha A. 2007. Phylogenetic relationships and morphometric affinities of the Arunachal macaque *Macaca munzala*, a newly described primate from Arunachal Pradesh, northeastern India. *Molecular Phylogenetics and Evolution*, **44**(2): 838−849.

Chakraborty D, Sinha A, Ramakrishnan U. 2014. Mixed fortunes: ancient expansion and recent decline in population size of a subtropical montane primate, the Arunachal macaque *Macaca munzala*. *PLoS One*, **9**(7): e97061.

Chalise MK. 2005. Characteristics of the Assamese monkey (*Macaca assamensis*) of Nepal. *American Journal of Primatology*, **66**: 195.

Chalise MK. 2008. Primate census in Kathmandu and west parts of Nepal. *Journal of natural History Museum*, **23**: 60−64.

Chalise MK. 2013. Fragmented primate population of Nepal. *In*: Marsh LK, Chapman CA. Primates in Fragments: Complexity and Resilience. New York: Springer, 329–356.

Deinard A, Smith DG. 2001. Phylogenetic relationships among the macaques: evidence from the nuclear locus *NRAMP1*. *Journal of Human Evolution*, **41**(1): 45−59.

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

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**(5): 1792−1797.

Evans BJ, Gansauge MT, Tocheri MW, Schillaci MA, Sutikna T, Jatmiko, et al. 2020. Mitogenomics of macaques (*Macaca*) across Wallace's Line in the context of modern human dispersals. *Journal of Human Evolution*, **146**: 102852.

Fa JE. 1989. The genus *Macaca*: a review of taxonomy and evolution. *Mammal Review*, **19**(2): 45−81.

Fan PF, Liu Y, Zhang ZC, Zhao C, Li C, Liu WL, et al. 2017. Phylogenetic position of the white-cheeked macaque (*Macaca leucogenys*), a newly described primate from southeastern Tibet. *Molecular Phylogenetics and Evolution*, **107**: 80−89.

Fooden J. 1976. Provisional classification and key to living species of macaques (Primates: *Macaca*). *Folia Primatologica*, **25**(2–3): 225−236.

Fooden J. 1979. Taxonomy and evolution of the Sinica group of macaques: I. Species and subspecies accounts of *Macaca sinica*. *Primates*, **20**(1): 109−140.

Groves CP. 2001. Primate Taxonomy. Washington DC: Smithsonian Institution Press.

Hayasaka K, Fujii K, Horai S. 1996. Molecular phylogeny of Macaques: implications of nucleotide sequences from an 896-base pair region of mitochondrial DNA. *Molecular Biology and Evolution*, **13**(7): 1044−1053.

Hayasaka K, Ishida T, Horai S. 1991. Heteroplasmy and polymorphism in the major noncoding region of Mitochondrial DNA in Japanese monkeys: association with tandemly repeated sequences. *Molecular Biology and Evolution*, **8**(4): 399−415.

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. *Proceedings of Royal Society Biological Sciences: Biological Sciences*, **270**(1512): 313−321.

Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, **42**(2): 182−192.

Khanal L, Chalise MK, He K, Acharya BK, Kawamoto Y, Jiang XL. 2018a. Mitochondrial DNA analyses and ecological niche modeling reveal post-LGM expansion of the Assam macaque (*Macaca assamensis*) in the foothills of Nepal Himalaya. *American Journal of Primatology*, **80**(3): e22748.

Khanal L, Chalise MK, Jiang XL. 2018b. Inconsistency in phylogeny of

sinica-group of macaques. *In*: Satellite International Symposium on Asian Primates-2018. Kathmandu: Nepal Biodiveristy Research Society, 48–49.

Khanal L, Chalise MK, Jiang XL. 2019. Distribution of the threatened Assamese Macaque *Macaca assamensis* (Mammalia: Primates: Cercopithecidae) population in Nepal. *Journal of Threatened Taxa*, **11**(1): 13047−13057.

Khanal L, Chalise MK, Wan T, Jiang XL. 2018c. Riverine barrier effects on population genetic structure of the Hanuman langur (*Semnopithecus entellus*) in the Nepal Himalaya. *BMC Evolutionary Biology*, **18**(1): 159.

Kim HS, Hirai H, Takenaka O. 1996. Molecular features of the TSPY gene of gibbons and Old World monkeys. *Chromosome Research*, **4**(7): 500−506.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**(2): 111−120.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*, **35**(6): 1547−1549.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**(6): 1695−1701.

Li BG, He G, Guo ST, Hou R, Huang K, Zhang P, et al. 2020. Macaques in China: evolutionary dispersion and subsequent development. *American Journal of Primatology*, **82**(7): e23142.

Li C, Zhao C, Fan PF. 2015. White-cheeked macaque (*Macaca leucogenys*): a new macaque species from Medog, southeastern Tibet. *American Journal of Primatology*, **77**(7): 753−766.

Li J, Han K, Xing JC, Kim HS, Rogers J, Ryder OA, et al. 2009. Phylogeny of the macaques (Cercopithecidae: *Macaca*) based on Alu elements. *Gene*, **448**(2): 242−249.

Li QQ, Zhang YP. 2004. A molecular phylogeny of *Macaca* based on mitochondrial control region sequences. *Zoological Research*, **25**(5): 385−390.

Li QQ, Zhang YP. 2005. Phylogenetic relationships of the macaques (Cercopithecidae: *Macaca*), inferred from mitochondrial DNA sequences. *Biochemical Genetics*, **43**(7–8): 375−386.

Liedigk R, Kolleck J, Böker KO, Meijaard E, Md-Zain BM, Abdul-Latiff MAB, et al. 2015. Mitogenomic phylogeny of the common long-tailed macaque (*Macaca fascicularis fascicularis*). *BMC Genomics*, **16**(1): 222.

Matsudaira K, Hamada Y, Bunlungsup S, Ishida T, San AM, 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**(4): 360−371.

Melnick DJ, Hoelzer GA. 1993. What is mtDNA good for in the study of primate evolution?. *Evolutionary Anthropology*, **2**(1): 2−10.

Minge C, Berghänel A, Schülke O, Ostner J. 2016. Patterns and consequences of male-infant relationships in wild assamese macaques (*Macaca assamensis*). *International Journal of Primatology*, **37**(3): 350−370.

Molur S, Brandon-Jones D, Dittus W, Eudey A, Kumar A, Singh M, et al. 2003. Status of South Asian primates: conservation assessment and management plan (C.A.M. P.). Coimbatore, India: Zoo Outreach Organization /CBSG-South Asia.

Morales JC, Melnick DJ. 1998. Phylogenetic relationships of the macaques (Cercopithecidae: *Macaca*), as revealed by high resolution restriction site mapping of mitochondrial ribosomal genes. *Journal of Human Evolution*, **34**(1): 1−23.

Moritz C. 1994. Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution*, **9**(10): 373−375.

Nei M, Kumar S. 2000. Molecular Evolution and Phylogenetics. New York: Oxford University Press, 333.

Page SL, Chiu CH, Goodman M. 1999. Molecular phylogeny of old world monkeys (Cercopithecidae) as inferred from γ-globin DNA sequences. *Molecular Phylogenetics and Evolution*, **13**(2): 348−359.

Rambaut A, Suchard M, Xie D, Drummond A. 2014(2020-04-25). Tracer v1.6. – Computer program and documentation. https://www.scirp.org/reference/referencespapers.aspx?referenceid=2565011.

Roos C, Kothe M, Alba DM, Delson E, 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.

Sarania B, Devi A, Kumar A, Sarma K, Gupta AK. 2017. Predictive distribution modeling and population status of the endangered *Macaca munzala* in Arunachal Pradesh, India. *American Journal of Primatology*, **79**(2): e22592.

Sinha A, Datta A, Madhusudan MD, Mishra C. 2005. *Macaca munzala*: a new species from western Arunachal Pradesh, Northeastern India. *International Journal of Primatology*, **26**(4): 977−989.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**(9): 1312−1313.

Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, **4**(1): vey016.

Thierry B, Singh M, Kaumanns W. 2004. Why macaque societies?. *In*: Thierry B, Singh M, Kaumanns W. Macaque Societies – A Model for the Study of Social Organization. Cambridge: Cambridge University Press.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**(22): 4673−4680.

Tosi AJ, Morales JC, Melnick DJ. 2000. Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history. *Molecular Phylogenetics and Evolution*, **17**(2): 133−144.

Tosi AJ, Morales JC, Melnick DJ. 2003. Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution*, **57**(6): 1419−1435.

Wada K. 2005. The distribution pattern of rhesus and Assamese monkeys in Nepal. *Primates*, **46**(2): 115−119.

Whitfield LS, Lovell-Badge R, Goodfellow PN. 1993. Rapid sequence evolution of the mammalian sex-determining gene SRY. *Nature*, **364**(6439): 713−715.

Zhou QH, Wei H, Huang ZH, Huang CM. 2011. Diet of the Assamese macaque *Macaca assamensis* in limestone habitats of Nonggang, China. *Current Zoology*, **57**(1): 18−25.

Ziegler T, Abegg C, Meijaard E, Perwitasari-Farajallah D, Walter L, Hodges JK, et al. 2007. Molecular phylogeny and evolutionary history of Southeast Asian macaques forming the *M. silenus* group. *Molecular Phylogenetics and Evolution*, **42**(3): 807−816.