American Journal of Primatology 77:901–910 (2015)

RESEARCH ARTICLE

Full-Length *Numt* Analysis Provides Evidence for Hybridization Between the Asian Colobine Genera *Trachypithecus* and *Semnopithecus*

BOSHI WANG[1,2,3], XUMING ZHOU[2], FANGLEI SHI[2], ZHIJIN LIU[2], CHRISTIAN ROOS[4], PAUL A. GARBER[5], MING LI[2], AND HUIJUAN PAN[1]*

[1]*College of Nature Conservation, Beijing Forestry University, Beijing, China*
[2]*Key laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Chaoyang, Beijing, China*
[3]*University of Chinese Academy of Sciences, Beijing, China*
[4]*Primate Genetics Laboratory, German Primate Center, Göttingen, Germany*
[5]*Department of Anthropology and Program in Ecology and Evolutionary Biology, University of Illinois, Urbana, Illinois, USA*

The phylogenetic position of the genus *Semnopithecus*is unresolved because of topological incongruence when inferred using different molecular markers. Although some studies proposed hybridization between the genera *Semnopithecus* and *Trachypithecus* to explain the discordance, no conclusive evidence for hybridization has been identified. To address this issue, we used DNA walking and long-range PCR to describe a nuclear mitochondrial DNA (*Numt*) segment present in *Trachypithecus pileatus* which extends over more than 15 kb, and represents approximately 92% of the entire mitochondrial genome. We assessed the presence of this *Numt* in 16 other colobine species, including four species of the genus *Trachypithecus*, six species of the genus *Semnopithecus*, and representative species of six other genera belonging to the subfamily Colobinae. We failed to detect a *Numt* sequence in any of the other colobine species except for *T. shortridgei*, which is closely related to *T. pileatus*. The sister relationship of this *Numt* within the genus *Semnopithecus* suggests that it was derived from the mt genome of the genus *Semnopithecus* and invaded the nuclear genome of *T. pileatus* by unidirectional introgression hybridization. These results offer the most conclusive evidence for the existence of hybridization between *Semnopithecus* and *Trachypithecus*. Am. J. Primatol. 77:901–910, 2015.

Key words: *Numt*; *Trachypithecus*; Asian colobines; hybridization

INTRODUCTION

The Old World monkey subfamily Colobinae represents a diverse clade of more than 50 species grouped into ten genera [Brandon-Jones et al., 2004; Groves, 2001]. These "leaf-eating" primates occupy a wide range of forest and woodland habitats across Africa and southern and southeastern Asia. Among the ten genera, the genera *Trachypithecus* and *Semnopithecu*s are widely distributed and taxonomically diverse. Based on fur coloration, behavior, and ecology, *Trachypithecus* was traditionally divided into five species groups, including *T. obscurus*,*T. francoisi*, *T. cristatus*, *T. pileatus*, and *T. vetulus* [Groves, 2001; Rowe, 1996]. The genus *Trachypithecus* is found to range from mainland southeast Asia to the Sundaland. [Kay & Davies, 1994; Mittermeier et al., 2013]. In contrast to *Trachypithecus*, the genus *Semnopithecus* dispersed throughout the Indian subcontinent, overlapping geographically with *T. pileatus* in Bhutan, Bangladesh, and northeast India (Fig. 1). Traditionally, *Semnopithecus* was considered to be a single species, *S. entellus*, but with as many as 14–16 subspecies [Napier & Napier, 1967; Pocock, 1928; Roonwal & Mohnot, 1977]. Nowadays, most authors tend to elevate some subspecies to species level, and described nine distinct *Semnopithecus* species, including *S. schistaceus*, *S. entellus*, *S. ajax*, *S. hector*, *S.*

Contract grant sponsor: Natural Science Foundation of China; contract grant numbers: 31372185, 31130061; contract grant sponsor: Hubei Province Key Laboratory of Conservation Biology of Shennongjia Golden Monkey; contract grant sponsor: State Forestry Administration of China

*Correspondence to: Huijuan Pan, College of Nature Conservation, Beijing Forestry University, Beijing 100083, China. Email: phjjanine2013@gmail.com

Received 13 December 2014; revised 17 March 2015; revision accepted 23 March 2015

DOI: 10.1002/ajp.22419
Published online 22 April 2015 in Wiley Online Library (wileyonlinelibrary.com).

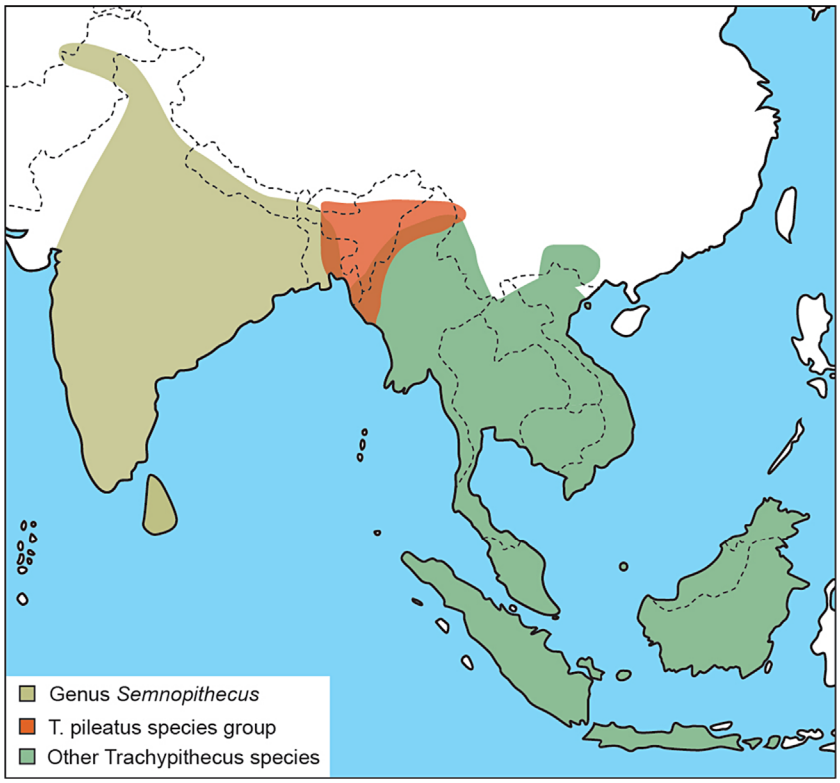

Figure 1. Distribution of the genus *Semnopithecus* and *Trachypithecus*.

hypoleucos, *S. priam*, *S. dussumieri*, *S. vetulus*, and *S. johnii* [Groves, 2001; Mittermeier et al., 2013; Nag et al., 2011]. Among these, the ranges of *S. schistaceus*, *S. entellus*, and *S. hector* overlap with *T. pileatus*.

Recent advances in DNA sequencing technology and phylogenetic analysis afford new opportunities to examine primate evolutionary histories, and these data have challenged traditional morphology-based taxonomic classifications [Packer et al., 2009]. Based on an analysis of the lysozyme gene, Messier & Stewart [1997] suggested that the *T. vetulus* species group, comprising the Sri Lankan *T. vetulus* and the south Indian *T. johnii*, was more closely related to the genus *Semnopithecus* than the genus *Trachypithecus*. This also was supported using results from retroposon integrations, nuclear DNA, mtDNA, and karyotype data [Bigoni et al., 2003; Karanth et al., 2008; Osterholz et al., 2008]. Each of these results also indicated that the *T. vetulus* species group should be reclassified and placed in the genus *Semnopithecus.* [Mittermeier et al., 2013; Perelman et al., 2011]. However, the status of the *T. pileatus* group, including *T. pileatus, T. geei*, and *T. shortridgei*, remains unclear. Based on mitochondrial DNA sequences, the *T. pileatus* group is best assigned to *Semnopithecus.* In contrast, nuclear data aligns the *T. pileatus* group within *Trachypithecus* [Karanth et al., 2008; Osterholz et al., 2008]. It also has been suggested that the *T. pileatus* group might have undergone ancestral hybridization with *Semnopithecus*.

Natural hybridization is regarded as an important evolutionary mechanism that results in an admixture of previously isolated gene pools [Stebbins, 1959; Zinner et al., 2011]. The genes flow between species may accelerate adaptation, facilitate ecological diversity, as well as drive speciation processes. In primates, hybridization has been reported mainly between subspecies and species, but has also been detected between genera [Detwiler et al., 2005; Dunbar & Dunbar, 1974; Won & Hey, 2005]. In the past, most of the reported hybridization in primates were ongoing events, which were evaluated by the approaches of field observation [Bynum et al., 1997; Dunbar & Dunbar, 1974; Nagel, 1973]. Recently, with the application of molecular methods, past hybridization between two ancestral, divergent lineages can also be detected [Arnold & Meyer, 2006; Burrell et al., 2009; Roos et al., 2011]. The incongruence between phylogenies based on different molecular markers is often an evidential footprint of ancestral hybridization events. For example, the hybrid speciation hypothesis in *Macaca arctoides* accounts for not only the incongruity between nuclear and

mitochondrial data, but also its inexplicable reproductive morphologies, including a unique glans penis and baculum structure in males and a reciprocal vaginal and exocervix morphology in females [Tosi et al., 2000].

The *T. pileatus* group was once considered as a typical example of hybrid speciation based on the incongruity between nuclear and mitochondrial data until two studies suggest that mitochondrial sequences from *T. pileatus* are nuclear mitochondrial pseudogenes. This is based on *T. pileatus*' complete mitochondrial genome sequence, and the fact that its phylogenetic position based on complete mtDNA was also consistent with that based on nuclear data [Wang et al., 2012; Shi et al., unpublished data]. Nuclear mitochondrial-like sequences, which transfer from mitochondrial DNA to the nuclear genome, are referred to as *Numts* [Lopez et al., 1994]. *Numts* have been detected in more than 80 species of eukaryotes, and exhibit different degrees of homology to their mitochondrial counterparts and variation in size [Bensasson et al., 2001]. *Numt*s can be thought of as "molecular fossils" which, after inserting into nuclear genomes, evolve much more slowly than their mitochondrial counterparts. Thus, the analysis of *Numt*s mistakenly considered as organelle mtDNA can confound phylogenetic and population genetic analyses because of their slow evolutionary rate and biparental mode of inheritance compared with authentic organelle mtDNA [Smith et al., 1992; Zhang & Hewitt, 1996b].

Although the misclassification of *T. pileatus* caused by the *Numt* has been corrected, it still remains an open question why a *Semnopithecus* mitochondrial-like *Numt* exists in the *T. pileatus* nuclear genome. In the present study, our aims are to elucidate the characteristics of this *Numt* in *T. pileatus*, and trace its origin based on the phylogenetic and coalescence analyses comparing four *Trachypithecus* species and other colobine genera.

METHODS

Sample Collection and DNA Extraction

For this study, we collected samples of blood and muscle tissue from 17 colobine species and compared them to 20 sample sequences from other studies (Table I). All sample collections were carried out in compliance with the relevant institutions and laws of China, and this research adhered to the American Society of Primatologists principles for the ethical treatment of primates. Muscle was stored in 95% ethanol. Blood samples were collected while trapping

TABLE I. Species Used in This Study

Family	Subfamily	Genus	Species	GenBank Accession No.	*Numt*
Cercopithecidae	*Colobinae*	*Trachypithecus*	*T. pileatus*	KF680163[a]	+
			T. shortridgei	KP834334	+
			T. francoisi	NC_023970[b]	_
			T. obscurus	NC_006900[b]	_
			T. cristatus	NC_023971[b]	_
		Semnopithecus	*S. entellus*	NC_008215[b]	_
			S. vetulus	NC_019582[b]	_
			S. johnii	NC_019583[b]	_
			S. hector	n.d.	_
			S. dussumieri	n.d.	_
			S. priam	n.d.	_
		Presbytis	*P. melalophos*	NC_008217b[b]	_
		Rhinopithecus	*R. roxellana*	NC_008218[b]	_
		Pygathrix	*P. nemaeus*	NC_008220[b]	_
		Nasalis	*N. larvatus*	NC_008216[a]	_
		Simias	*S. concolor*	NC_020667[a]	_
		Colobus	*C. guereza*	NC_006901[a]	_
		Piliocolobus	*P. badius*	NC_008219[a]	_
		Procolobus	*P. verus*	NC_020666[a]	_
	Cercopithecinae	*Macaca*	*M. mulatta*	NC_005943[a]	_
		Theropithecus	*T. gelada*	NC_019802[a]	_
		Papio	*P. papio*	NC_020009[a]	_
Hominidae		*Homo*	*H. sapiens*	X93334[b]	_
		Pan	*P. troglodytes*	D38113[b]	_

+/–, the presence/absence of the *Numt*; n.d., data.
[a]From [Wang et al., 2012]
[b]From [Peng et al., 2009]

individuals for physical examination, and were stored in a refrigerator at −80 °C. For each samples, high molecular weight cellular DNA was isolated from blood or frozen tissues by QIAamp DNA Mini Kit (Qiagen Inc., Valencia, CA) following the protocols in the kit. Extracted DNA was diluted 20 times with doubly distilled water and stored at −20 °C.

Amplification and Sequencing

DNA walking experiments were performed to locate the *Numt* in the *T. pileatus* genome. The genomic DNA sequence flanking the *Numt* integration site was identified using DNA Walking SpeedUp premix kit-II (Seegene, Seoul, South Korea) according to the manufacturer's protocol. To get the target sequence, we first amplified the *Numt* region from the *T. pileatus* sample using primers described previously [Osterholz et al., 2008]. The obtained 573 bp mitochondrial pseudogene sequence was aligned to the *T. pileatus* mitochondrial genome. The target-specific primers (TSPs) were designed within the region that differentiated them from mitochondrial DNA. Two sets of TSP primers were used to extend the unknown region of this *Numt* from both upstream and downstream locations. The products of the third round of DNA walking PCR were cut out of 1.5% agarose gel containing EtBr and purified using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI). They were then cloned directly into the pMD18-T Vector (TaKaRa, Dalian, China). The cloned plasmids were sequenced and analyzed to determine whether the obtained walling products reached the genomic region flanking the *Numt*. The procedures mentioned above were repeated until we obtained one side of the flanking region. The sequences of the flanking region were analyzed by comparison to the *Macaca mulatta* genome database at GenBank by using BLAST (NCBI) to locate the integration site and deduce the other side of the flanking sequences. Then a pair of primers (LL/LR) was designed within the flanking region to amplify the whole *Numt* sequences by long-range PCR (sequences available upon request). A combination of the primer walking approach was employed to sequence the complete *Numt* sequence. With this pair of primers, we screened for the presence versus absence of the *Numt* in 15 other colobine species (detailed in Table I).

Long-range PCR was used to amplify the complete mitochondrial DNA genomes of *T. pileatus* and *T. shortridgei* following the Expand Long Template PCR system protocol (TaKaRa), which is an effective method that minimizes the possibility of amplifying nuclear mitochondrial pseudogenes [Thalmann et al., 2004]. We designed two sets of primers to amplify two overlapping segments (each 9–10 kb in size) that together cover the entire mitochondrial DNA genome. The reactions were carried out in a total volume of 50 μl, containing 2.0 μl of genomic DNA, 2 μl of each primer (10 μM), 25 μl of Premix Taq™ (LA Taq™ version 2.0, TaKaRa), and 19 μl of distilled water. The thermal cycle profile consisted of an initial denaturation at 95 °C for 3 min, and then 35 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 10 min. In the final cycle, extension was carried out at 72 °C for 30 min. We used 24 walking primers to sequence these overlapping segments and obtain the complete mitochondrial genome.

Phylogenetic Analysis

The *Numt*-Tpi (*Numt* in *T. pileatus*) sequence was aligned with the *Numt*-Tsh (*Numt* in *T. shortridgei*) sequence, as well as the 17 Cercopithecidae mtDNA genomes available in GenBank in order to provide a phylogenetic context for the data. All datasets comprised 19 sequences, including five species of the genus *Trachypithecus* and three species of the genus *Semnopithecus*, and 11 sequences representing each of the other eight colobine genera (*Colobus*, *Piliocolobus*, *Procolobus*, *Presbytis*, *Rhinopithecus, Pygathrix*, *Nasalis*, *Simias*). The genus *Macaca* was used as an outgroup. The insertions, deletions, and inversions in *Numts* were adjusted and corrected by eye. Sequences were aligned using Muscle implemented in MEGA 6 (http://www.megasoftware.net) with the default settings [Tamura et al., 2013]. Estimates of evolutionary divergence between sequences for each lineage within the subfamily Colobinae were computed using the p-distance algorithm of the MEGA software package.

Because the Bayesian trees and Maximum Likelihood trees estimated from DNA sequences were the most accurate [Hall, 2004], two methods were conducted to infer phylogenetic relationships in this study: Bayesian inference (BI) and Maximum likelihood (ML). The Modeltest v3.7 [Posada & Crandall, 1998] and MrModeltest v2.2 [Nylander et al., 2004] programs were used to select the optimal nucleotide substitution model, with MrMTgui v1.0 interface [Nuin, 2008]. ML phylogenetic analysis was performed in the programs PAUP v.4b10 [Swofford, 2003], by running 1,000 replicates with a heuristic search incorporating the previously estimated parameters for GTR, invariant sites, and gamma values using the Akaike Information Criterion (AIC) of Modeltest. A Bayesian-based phylogenetic analysis was conducted with an MPI enabled MrBayes v3.1.2 [Altekar et al., 2004; Huelsenbeck et al., 2001; Ronquist & Huelsenbeck, 2003]. Four simultaneous runs were initiated, each using a random starting tree for 10 million generations sampled every 1,000 generations, and used four chains with the default heating temperature of 0.1. The first 25% of each run was discarded as burn-in. The posteriors from all four

runs were then combined to provide the consensus estimate of the coalescent-based species tree. Posteriors above 85% were considered as evidence for substantial support at a node.

Divergence Time Estimation

A Bayesian analysis implemented in BEAST v1.8.0 was used to estimate divergence times based on 21 Cercopithecidae mtDNA genomes. In addition to the colobine mtDNA genomes, five non-colobine catarrhine mtDNA genomes also were integrated to provide nodes temporally constrained by well-supported fossil records. These included three *cercopithecine* genera (*Papio*, *Macaca*, and *Theropithecus*), and two hominoid genera (*Homo*, *Pan*), which were defined as outgroups. The importance of a carefully designed calibration scheme in a molecular dating study cannot be overemphasized [Forest, 2009], so we selected three fossil constraints based on criteria for choosing appropriate calibration points [Ho & Phillips, 2009; Rutschmann et al., 2007], including the divergence of the Old World monkey and hominoid lineages at about 24–29–Ma [Zalmout et al., 2010], the split between *Theropithecus* and *Papio* at about 4 Ma [Leakey, 1993], and the divergence of the human and chimpanzee lineages at 6–7 Ma [Brunet et al., 2005; Vignaud et al., 2002]. Concatenating alignments of all 12 H-stranded protein-coding genes of the mtDNA genome was partitioned into two unlinked codon positions [(1 + 2), 3], and the Yule model was selected as tree prior and an uncorrelated relaxed lognormal molecular clock model was used to estimate rate variation along lineages [Drummond et al., 2006]. Four independent analyses of 50 million generations each with samples logged every 1,000 generations were run to ensure sampling of estimated sample size (ESS) values. Output from each run was imported into Tracer v1.6 [Rambaut et al., 2013] to determine burn-in value. Trees sampled from the first 25% were discarded, and the remaining were combined using TreeAnnotator v1.8.0 [Drummond & Rambaut, 2007].

RESULTS

In this study, one complete mitochondrial genome sequence and two *Numt* sequences greater than 16 kb in length were generated from *T. pileatus* and *T. shortridgei*. The new sequences are available under GenBank accession numbers KP834333–KP834335.

A 16,155–bp DNA fragment was amplified using primer pair (LL/LR) from the total genomic DNA of *T. pileatus* and termed *Numt*-Tpi. This sequence was composed of a 15,278 bp *Numt*, which was homologous to approximately 92.4% of the mitochondrial genome, a 753 bp 5' flanking region and a 124 bp 3' flanking region. The 13 protein-coding DNA sequences were translated into protein sequences using vertebrate mitochondrial genetic codes and seven genes contained premature stop codons. Of these, we found that four (*Numt*-nd1, *Numt*-atp8, *Numt*-atp6, and *Numt*-cytb) resulted from a 1 bp deletion, respectively, two (*Numt*-nd2 and *Numt*-nd5) from nonsense mutations, and one (*Numt*-cox1) from a 579 bp deletion. Also, we found a large deletion of an 843–bp fragment in the D-loop region and a reverse rearrangement block (394 bp) containing the entire tRNA-Phe gene, part of the gene 12S rRNA, and part of the D-loop region. We detected the existence or nonexistence of this *Numt* in 16 colobine taxa, including four species of the genus *Trachypithecus*, six species of the genus *Semnopithecus*, and all of the representative species of the six other genera belonging to subfamily Colobinae. However, this *Numt* sequence was not detected in any of the other colobine species except for *T. shortridgei*, which has the closest phylogenetic relationship with *T. pileatus*.

In the analyses of evolutionary divergence between two *Numts* and 17 organelle mtDNA sequences, the lowest sequence divergence was obtained between *Numt*-Tpi and *Numt*-Tsh with 0.034% p-distances, compared to the 0.874% p-distances of the mitochondrial genome between *T. pileatus* and *T. shortridgei*. The low divergence suggests that the mutation rate dropped radically after the *Numt* inserted into the nuclear genome. However, among all the organelle mtDNA, the highest similarity to two *Numt* sequences was found in all six species of the genus *Semnopithecus* (5.768–6.489%) rather than among all species of the genus *Trachypithecus* (14.433–15.07%) (Table II), indicating that the *Numt* might have been derived fromthe genus *Semnopithecus* and then invaded the genome of the genus *Trachypithecus* by interspecific hybridization.

The identical topology was obtained from both the ML and BI analyses (Fig. 2). The results strongly support a sister-group relationship between *Numt*-Tpi and *Numt*-Tsh, as well as between the five *Trachypithecus* species and three *Semnopithecus* species, each of which also form a monophyletic clade. Also, we found that the *Numts* clade clusters with all species of *Semnopithecus* forming a monophyletic group (Fig. 2). This is supported by a 100% bootstrap value and a 100% posterior probability in ML and BI analyses, respectively.

To determine the date of the proposed hybridization event between *Semnopithecus* and *Trachypithecus*, we estimated divergence times among the five *Trachypithecus* species. Three calibration points based on well-supported fossil dates were chosen to provide nodes temporally constrained in the phylogenetic tree. Most relationships and branching orders are strongly supported and congruent with higher level analyses of colobine phylogeny, indicating a high reliability of this dataset. The time of the most recent common ancestor of genus *Trachypithecus* dates to approximately 3.47 ± 0.55 million years ago (Mya), and the *T. cristatus* species group split from this basal

TABLE II. Pairwise Differences Between Sequences for Each Lineage Within Subfamily *Colobinae*

*NUMT*_Tpi	*NUMT*_Tsh	*S_vetulus*	*S_johnii*	*S_entellus*	*T_cristatus*	*T_pileatus*	*T_shortridgei*	*T_obscurus*	*T_francoisi*	*N_larvatus*	*S_concolor*	*R_Roxellana*	*P_melalophos*	*P_nemaeus*	*P_verus*	*P_badius*	*C_gureza*
*NUMT*_Tsh	0.00034																
S_vetulus	0.05799	0.05768															
S_johnii	0.06239	0.06208	0.07913														
S_entellus	0.06489	0.06472	0.08200	0.06501													
T_cristatus	0.14445	0.14433	0.15644	0.14592	0.15635												
T_pileatus	0.14691	0.14671	0.15677	0.14910	0.15833	0.08463											
T_shortridgei	0.14630	0.14609	0.15575	0.14819	0.15773	0.08466	0.00874										
T_obscurus	0.14697	0.14685	0.15913	0.14807	0.15783	0.06794	0.08326	0.08332									
T_francoisi	0.15070	0.15040	0.16126	0.15621	0.16196	0.08062	0.08897	0.08808	0.06890								
N_larvatus	0.14582	0.14562	0.15571	0.15187	0.15779	0.15395	0.15221	0.15251	0.15237	0.15477							
S_concolor	0.14589	0.14569	0.15932	0.15245	0.15867	0.15327	0.15228	0.15293	0.15423	0.15621	0.04951						
R_Roxellana	0.14855	0.14835	0.15997	0.15318	0.15967	0.15278	0.15027	0.15011	0.15293	0.16056	0.13414	0.13556					
P_melalophos	0.15279	0.15267	0.16304	0.15562	0.16505	0.15116	0.15003	0.14966	0.14989	0.15523	0.15658	0.15475	0.15677				
P_nemaeus	0.15504	0.15484	0.16282	0.15603	0.16737	0.16030	0.15768	0.15785	0.15866	0.16434	0.14018	0.13839	0.13835	0.16253			
P_verus	0.16895	0.16873	0.17573	0.17047	0.17843	0.16964	0.16743	0.16680	0.16740	0.16969	0.16401	0.16618	0.16671	0.17534	0.17458		
P_badius	0.16940	0.16910	0.17788	0.17510	0.18165	0.17427	0.16937	0.17010	0.17425	0.17489	0.17533	0.17029	0.17263	0.17979	0.17972	0.13276	
C_gureza	0.17692	0.17669	0.18231	0.17636	0.18574	0.17390	0.17289	0.17430	0.17573	0.17966	0.17535	0.17483	0.17846	0.18654	0.18019	0.15317	0.15939

lineage approximately at 2.76 ± 0.46 Mya. The most recent split among the four *Trachypithecus* species occurred between *T. obscurus* and *T. francoisi* (2.36 ± 0.42 Mya). Within the *T. pileatus* species group, the divergence time of *T. pileatus* and *T. shortridgei* occurred about 0.26 ± 0.08 Mya (Table III). These results suggest that hybridization between the *T. pileatus* species group and *Semnopithecus* is most likely to have occurred between 3.47 and 0.26 Mya.

Discussion

Origin of the *Numt*

In this study, we described an almost complete mtDNA in the *T. pileatus* nuclear genome representing approximately 92% of the entire mitochondrial genome. This *Numt* is a transposition of a bigger proportion of the mtDNA genome than previously reported, and it also represents the second longest *Numt*, next only to the one in *Arabidopsis thaliana* [Stupar et al., 2001]. The analyses of *Numts* from 17 of colobine species indicated that the *Numt*-Tpi, which is present in the *T. pileatus* nuclear genome, originated from the mitochondrial genome of the genus *Semnopithecus*. There are two plausible hypotheses to explain these results. One hypothesis is that an intergenomic transfer event initially occurred in the common ancestor of *Semnopithecus* and *Trachypithecus*, and this *Numt* was then eliminated from the nuclear genome in all extant species except *T. pileatus* and *T. shortridgei* after the two genera diverged. However, the sister-group relationship between the *Numt* and *Semnopithecus* implies that the formation of *Numt* is likely to have occurred posterior to the splitting of the two genera (Fig. 2). Therefore, this hypothesis should be rejected.

The other remaining hypothesis is a hybridization event resulting in the unidirectional introgression of genetic material from genus *Semnopithecus* to ancestor of *T. pileatus* and *T. shortridgei.* We estimate that this would have occurred at approximately between 3.47 and 0.26 Mya. This hypothesis also is supported by data reported in other studies [Roos et al., 2011; Sterner et al., 2006; Ting et al., 2008; Wang et al., 2012]. Based on mitochondrial genome analysis, Roos et al. [2011] argued that *Semnopithecus* occupies a basal position among Asian colobines. These authors also state that analyses of mobile elements and nuclear genes indicate that *Semnopithecus* has a sister relationship to *Trachypithecus* and proposed the possibility of a hybridization event to explain the topological incongruence [Roos et al., 2011]. This is consistent with biogeographical information. The genus *Semnopithecus* is widely distributed throughout the Indian subcontinent and is found to overlap with *T. pileatus* species group in a sandwich-like located in the area of northeast India

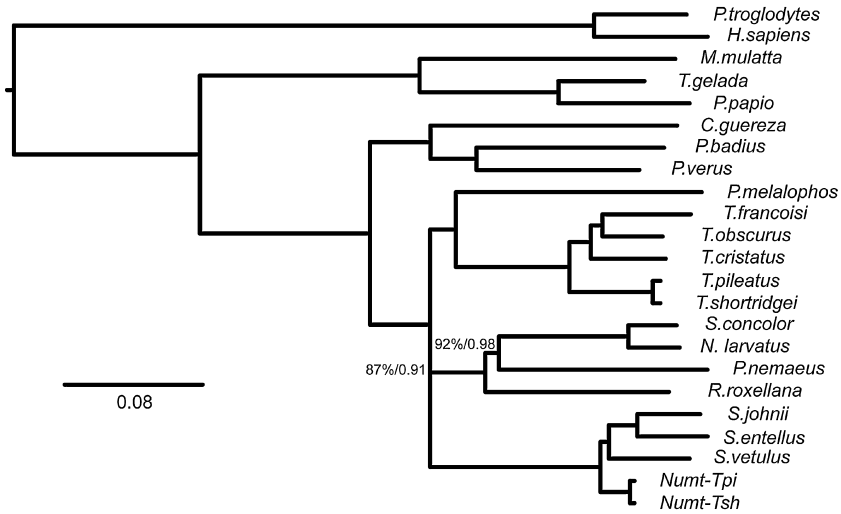

Figure 2. Phylogenetic tree inferred from the mt genome sequences. Note: Phylogenetic tree inferred by ML from the mt genome sequences. The identical topology resulted from Bayesian analysis of the same data. The nodal supports (BI/ML) are shown above the nodes. Support values are only shown with bootstrap/posterior probability values less than 100%/1.0.

[Kay & Davies, 1994; Mittermeier et al., 2013] (Fig. 1), making opportunities for hybridization possible. In addition, *T. pileatus* species group has a significant difference in body mass among all *Trachypithecus* species (Wilcoxon test: $P < 0.01$) [Delson et al., 2000]. The unusually heavy body of *T. pileatus* species group is much similar to *Semnopithecus* than *Trachypithecus*, possibly due to its hybrid origin, as seen in baboons [Brett et al., 1982; Phillips-Conroy & Jolly, 1981]. Finally, there is evidence that *S. entellus* can hybridize with *T. pileatus* in zoos [Finn, 2002], indicating that the divergences between two genera have not accumulated enough karyotypic differences leading to reproductive isolation. Accordingly, we suggested that the hybridization hypothesis resulting in the transfer of mitochondrial DNA between *Semnopithecus* and *Trachypithecus* offers a viable explanation for the incongruent phylogenetic position of genus *Semnopithecus* among Asian colobines.

Expanding on the hybridization hypothesis, we suggest that in areas of their range where both taxa are sympatric, male *Semnopithecus* may have successfully copulated with *Trachypithecus* females. Assuming that ancestral *T. pileatus* males had a body size of 7.4 kg, consistent with other *Trachypithecus* today, mean male body size (18.2 kg) for three *Semnopithecus* species, which have suture or overlap zone with *T. pileatus*, is nearly three times larger, providing *Semnopithecus* males with a tremendous advantage over resident *Trachypithecus* males in access to reproductive partners (Table IV). Furthermore, both *Semnopithecus* and *Trachypithecus* commonly exhibit a unimale–multifemale form of social organization [Mittermeier et al., 2013]. This means that larger *Semnopithecus* males would have an advantage over relatively diminutive ancestral *T. pileatus* males in fighting over mates. To sum up the above discussion, we rejected the first hypothesis that an intergenomic transfer event occurred in the common ancestor of *Semnopithecus* and *Trachypithecus* because the sister-group relationship between the *Numt* and *Semnopithecus* shows that the formation of *Numt* is posterior to the splitting of the two genera. Then we argued the second hypothesis from different aspects, including habitat distribution, morphological characteristics, mating system, and social organization, which provided compelling evidence for the unidirectional introgression of genetic material from genus *Semnopithecus* to *T. pileatus* due to a hybridization event.

TABLE III. Divergence Times of the Last Common Ancestors as Estimated from the mt Genome Sequences

Last common ancestor	Divergence time (mya)
Genus *Trachypithecus*	3.47 ± 0.55
T.cristatus–T.francoisi–T.obscurus	2.76 ± 0.46
T.francoisi–T.obscurus	2.36 ± 0.42
T.pileatus–T.shortridgei	0.26 ± 0.08
S.johnii–S.entellus	2.26 ± 0.39
S.vetulus–S.johnii–S.entellus	2.76 ± 0.48

Evaluation of the Hybridization Time

The hybridization hypothesis between *Semnopithecus* and *Trachypithecus* argues for a unidirectional exchange of genetic material from followed by extensive backcrossing over a period of many generations. This hypothesis predicts that a greater amount of genetic differences exists between *Semnopithecus* and

TABLE IV. Body Weight of *Semnopithecus* and *Trachypithecus*

Genus	Species	Mean body weight of males	Mean body weight of females
Semnopithecus	*S.entellus*	18.2	12.8
	S.ajax	20.0	12.7
	S.hector	17.2	13.6
	S.schistaceus	19.2	14.8
	S.hypoleucos	14.8	11.5
	S.priam	16.8	8.8
	S.johnii	11.7	11.1
Trachypithecus	*T.cristata*	6.7	5.8
	T.francoisi	8.0	7.8
	T.phayrei	7.6	6.2
	T.pileatus	12.1	10.0

Data from [Delson et al., 2000; Mittermeier et al., 2013; Phillips-Conroy & Jolly, 1981].

Trachypithecus in their mitochondrial genome than in their nuclear genome (nuclear swamping). However, when the hybrid zone first occurred remains controversial. Roos et al. [2011] suggested that the hybridization was an ancestral event that occurred prior to the time their mitochondrial lineages diverged (~8.47 Mya) because of monophyly in both mitochondrial and nuclear phylogenetic tree among all *Trachypithecus* species. In contrast, our results indicated that the *Numt* sequence only exits in *T. pileatus* species group, implying the introgression hybridization, which led to the interspecific transfer of mitochondrial DNA was a relatively recent event, posterior to the time that *T. pileatus* species group first diverged from other species of *Trachypithecus* (approximately 3.47 Mya). Because no signal of nuclear swamping was detected in the nuclear genome of *T. pileatus*, we suggest that the introgression hybridization period was a short-term event. Generally, the timing of a hybridization event is estimated based on the presence of orthologous sequences, such as the two *Numt* sequences across the studied taxa. However, although these sequences are present in *T. pileatus* species group, we did not find this *Numt* in any of the six species of *Semnopithecus* studied, indicating that if an ancestral *Semnopithecus* species contained these sequences, it either became extinct or is *S. schistaceus*, which has an overlap habitat with *T. pileatus* species group but was not sampled. As a consequence, we adopted an indirect method to estimate the time horizon for hybridization. According to the *Numt* present in *T. pileatus* species group and absent in other *Trachypithecus*, the earliest time point for hybridization would have occurred during the time of the most recent common ancestor of genus *Trachypithecus*, approximately 3.47 Mya. The identical *Numt* size and insert position between *T. pileatus* and *T. shortridgei* indicate the hybridization predates their divergence, which occurred approximately 0.26 Mya; in other words, the most recent time point for hybridization should be no later than 0.26 Mya.

MtDNA is considered to be a very useful molecular marker for evolutionary studies due to its lack of recombination, high copy numbers, haploid character, and maternal inheritance [Harrison, 1989; Moritz et al., 1987; Zhang & Hewitt, 1996a]. However, the presence of *Numt* sequences in the nuclear genome could contaminate the authentic mtDNA during PCR amplification and lead to erroneous results in phylogenetic and population genetic analyses [Collura & Stewart, 1995; van der Kuyl et al., 1995]. Also, *Numt* sequences are usually more conserved relative to their mitochondrial counterparts due to the lower mutation rate in the nuclear genome than in mtDNA, and might have been preferentially amplified using universal primers for cross-species amplification [Zhang and Hewitt, 1996a]. This might help explain why the evolutionary history of *T. pileatus* was misunderstood in previous study [Osterholz et al., 2008]. In the present study, we described the full sequence of what we feel is best considered a *Semnopithecus*-derived *Numt* in the *Trachypithecus* nuclear genome. This provides evidence in support of the existence of hybridization between *Semnopithecus* and *Trachypithecus*. The full sequence of the *Numt* also facilitated the confirmation of authentic *T. pileatus* mtDNA. Furthermore, a clearer understanding of hybridization zone that existed or possibly continues to exist between *T. pileatus* and *Semnopithecus* will provide a valuable contribution for the conservation and the management of this threatened species.

ACKNOWLEDGMENT

We thank Chrissie McKenney for the English revision.

REFERENCES

Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. 2004. Parallel metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics 20:407–415.

Arnold ML, Meyer A. 2006. Natural hybridization in primates: one evolutionary mechanism. Zoology 109:261–276.

Bensasson D, Zhang DX, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecology & Evolution 16:314–321.
Bigoni F, Stanyon R, Wimmer R, Schempp W. 2003. Chromosome painting shows that the proboscis monkey (*Nasalis larvatus*) has a derived karyotype and is phylogenetically nested within Asian Colobines. American Journal of Primatology 60:85–93.
Brandon-Jones D, Eudey AA, Geissmann T, et al. 2004. Asian primate classification. International Journal of Primatology 25:97–164.
Brett FL, Turner TR, Jolly CJ, Cauble RG. 1982. Trapping baboons and vervet monkeys from wild, free-ranging populations. The Journal of Wildlife Management 46:164–174.
Brunet M, Guy F, Pilbeam D, et al. 2005. New material of the earliest hominid from the Upper Miocene of Chad. Nature 434:752–755.
Burrell AS, Jolly CJ, Tosi AJ, et al. 2009. Mitochondrial evidence for the hybrid origin of the kipunji, *Rungwecebus kipunji* (Primates: Papionini). Molecular Phylogenetics and Evolution 51:340–348.
Bynum EL, Bynum DZ, Supriatna J. 1997. Confirmation and location of the hybrid zone between wild populations of *Macaca tonkeana* and *Macaca hecki* in Central Sulawesi, Indonesia. American Journal of Primatology 43:181–209.
Collura RV, Stewart CB. 1995. Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. Nature 378:485–489.
Delson E, Terranova CJ, Jungers WL, et al. 2000. Body mass in Cercopithecidae (Primates, Mammalia): estimation and scaling in extinct and extant taxa. Anthropological Papers of the American Museum of Natural History 83:1–159.
Detwiler KM, Burrell AS, Jolly CJ. 2005. Conservation implications of hybridization in African cercopithecine monkeys. International Journal of Primatology 26: 661–684.
Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4:e88.
Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.
Dunbar RIM, Dunbar EP. 1974. On hybridization between *Theropithecus gelada* and *Papio anubis* in the wild. Journal of Human Evolution 3:187–192.
Finn F. 2002. Sterndale's mammalia of India. New Delhi: Cosmo. p 347.
Forest F. 2009. Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. Annals of Botany 104:789.
Groves CP. 2001. Primate taxonomy. Washington, DC: Smithsonian Institution Press. p 300–350.
Hall BG. 2004. Comparison of the Accuracies of Several Phylogenetic Methods Using Protein and DNA Sequences. Molecular Biology and Evolution 22:792–802.
Harrison RG. 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends in Ecology & Evolution 4:6–11.
Ho SY, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Systematic Biology 58:367.
Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310–2314.
Karanth KP, Singh L, Collura RV, Stewart CB. 2008. Molecular phylogeny and biogeography of langurs and leaf monkeys of South Asia (Primates: Colobinae). Molecular Phylogenetics and Evolution 46:683–694.
Kay RN, Davies AG. 1994. In: Davies G, Oates A, editors. Digestive physiology. Cambridge: Cambridge University Press. p 229–249.
Leakey MG. 1993. In: Jablonski NG, editor. Evolution of *Theropithecus* in the Turkana Basin. Cambridge: Cambridge University Press. p 85.
Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. 1994. N*umt*, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. Journal of Molecular Evolution 39:174–190.
Messier W, Stewart CB. 1997. Episodic adaptive evolution of primate lysozymes. Nature 385:151–154.
Mittermeier RA, Wilson DE, Rylands AB. 2013. Handbook of the mammals of the world: primates. Barcelona: Lynx Edicions. p 733–758.
Moritz C, Dowling TE, Brown WM. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annual Review of Ecology and Systematics 18:269–292.
Nag KS, Pramod P, Karanth KP. 2011. Taxonomic implications of a field study of morphotypes of Hanuman Langurs (*Semnopithecus entellus*) in peninsular India. International Journal of Primatology 32:830–848.
Nagel U. 1973. A comparison of anubis baboons, hamadryas baboons and their hybrids at a species border in Ethiopia. Folia Primatologica 19:104–165.
Napier JR, Napier PH. 1967. A handbook of living primate. London: Academic Press.
Nuin P. 2008. MrMTgui: cross-platform interface for Model Test and MrModeltest http://www.genedrift.org/mtgui.php. [Accessed May 26, 2014].
Nylander JA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey J. 2004. Bayesian phylogenetic analysis of combined data. Systematic Biology 53:47–67.
Osterholz M, Walter L, Roos C. 2008. Phylogenetic position of the langur genera *Semnopithecus* and *Trachypithecus* among Asian colobines, and genus affiliations of their species groups. BMC Evolutionary Biology 8:58.
Packer L, Gibbs J, Sheffield C, Hanner R. 2009. DNA barcoding and the mediocrity of morphology. Molecular Ecology Resources 9:42–50.
Peng Z, Elango N, Wildman DE, Soojin VY. 2009. Primate phylogenomics: developing numerous nuclear non-coding, non-repetitive markers for ecological and phylogenetic applications and analysis of evolutionary rate variation. BMC Genomics 10:247.
Perelman P, Johnson WE, Roos C, et al. 2011. A molecular phylogeny of living primates. PLoS Genetics 7:e1001342.
Phillips-Conroy JE, Jolly CJ. 1981. Sexual dimorphism in two subspecies of Ethiopian baboons (*Papio hamadryas*) and their hybrids. American Journal of Physical Anthropology 56:115–129.
Pocock RI. 1928. The langurs or leaf monkeys of British India. Journal of the Bombay Natural History Society 32:472–504.
Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817.
Rambaut A, Drummond AJ, Suchard M. 2013. Tracer v1. 6 http://beast.bio.ed.ac.uk/Tracer [Accessed September 14, 2014].
Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
Roonwal ML, Mohnot SM. 1977. Primates of South Asia: ecology, sociobiology, and behavior. Cambridge, MA: Harvard University Press
Roos C, Zinner D, Kubatko LS, et al. 2011. Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. BMC Evolutionary Biology 11:77.
Rowe N. 1996. The pictorial guide to the living primates. East Hampton, NY: Pogonias Press. p 263.
Rutschmann F, Eriksson T, Salim KA, Conti E. 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. Systematic Biology 56:591–608.

Smith MF, Thomas WK, Patton JL. 1992. Mitochondrial DNA-like sequence in the nuclear genome of an akodontine rodent. Molecular Biology and Evolution 9:204–215.

Stebbins GL. 1959. The role of hybridization in evolution. Proceedings of the American Philosophical Society 103:231–251.

Sterner KN, Raaum RL, Zhang YP, Stewart CB, Disotell TR. 2006. Mitochondrial data support an odd-nosed colobine clade. Molecular Phylogenetics and Evolution 40:1–7.

Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S. 2001. Complex mtDNA constitutes an approximate 620-kb insertion on Arabidopsis thaliana chromosome 2: implication of potential sequencing errors caused by large-unit repeats. Proceedings of the National Academy of Sciences of the United States of America 98:5099–5103.

Swofford DL. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). version 4 Sunderland: Sinauer Associates.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEG A6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30:2725–2729.

Thalmann O, Hebler J, Poinar HN, Pääbo S, Vigilant L. 2004. Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. Molecular Ecology 13:321–335.

Ting N, Tosi AJ, Li Y, Zhang YP, Disotell TR. 2008. Phylogenetic incongruence between nuclear and mitochondrial markers in the Asian colobines and the evolution of the langurs and leaf monkeys. Molecular Phylogenetics and Evolution 46:466–474.

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 17:133–144.

van der Kuyl AC, Kuiken CL, Dekker JT, Perizonius WR, Goudsmit J. 1995. Nuclear counterparts of the cytoplasmic mitochondrial 12S rRNA gene: a problem of ancient DNA and molecular phylogenies. Journal of Molecular Evolution 40:652–657.

Vignaud P, Duringer P, Mackaye HT, et al. 2002. Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad. Nature 418:152–155.

Wang XP, Yu L, Roos C, et al. 2012. Phylogenetic relationships among the colobine monkeys revisited: new insights from analyses of complete mt genomes and 44 nuclear non-coding markers. PLoS One 7:e36274.

Won YJ, Hey J. 2005. Divergence population genetics of chimpanzees. Molecular Biology and Evolution 22:297–307.

Zalmout IS, Sanders WJ, MacLatchy LM, et al. 2010. New Oligocene primate from Saudi Arabia and the divergence of apes and Old World monkeys. Nature 466:360–364.

Zhang DX, Hewitt GM. 1996a. Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: some implications for population studies. Molecular Ecology 5:295–300.

Zhang DX, Hewitt GM. 1996b. Nuclear integrations: challenges for mitochondrial DNA markers. Trends in Ecology & Evolution 11:247–251.

Zinner D, Arnold ML, Roos C. 2011. The strange blood: natural hybridization in primates. Evolutionary Anthropology: Issues, News, and Reviews 20:96–103.