New mitochondrial genome resource for resolving the grey langur species complex

Vipin Hiremath

Dr Babasaheb Ambedkar Marathwada University

Gulab Khedkar (✉ gdkhedkar@gmail.com)

Paul Hebert Centre for DNA Barcoding and Biodiversity Studies <https://orcid.org/0000-0002-2060-785X>

Chandrakant Chandrakant Jadhav

Dr Babasaheb Ambedkar Marathwada University

Raveendranathanpillai Sanil

Government Arts College

Chandraprakash Khedkar

Maharashtra Animal and Fishery Sciences University

Bharathi Prakash

Mangalore University

Research Article

Keywords: Semnopithecus hypoleucos, Colobinae, Mitochondrial genome, Phylogenetic analysis, comparative analysis

Posted Date: August 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-282571/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Among Indian primate species, the gray langur, is widely distributed and shows extraordinary morphological plasticity throughout its range of distribution. Various taxonomic assignments have been suggested and developed over the past century in resolving the systematic position of the gray langur. There was a substantial disagreements on number of species. To bring about some clarity, it has been proposed to use six morphotypes along with ecological niche modelling to form the basis of a taxonomic classification scheme for the grey langurs found in peninsular India. Nevertheless, the validation for these ambiguities in species numbers is necessary by employing independent tools and molecular data informatics. However, previous attempts at species authentication using molecular data usually rely on only one or a small number of genes. To develop a more reliable method for the reconstruction of phylogenetic relationships of these taxa, and to create an important genomic resource for future studies, we sequenced whole mitochondrial genome of the gray langur *Semnopithecus hypoleucos*. Our results substantiate the value of sequencing whole mitochondrial genomes for this purpose, and show that developing similar information for new members belonging to the species complex associated to the Hanuman langur clade is essential. Our study proposed a species boundary using different genomic locus to differentiate a species. This may aid to resolve taxonomy and systematic relationships of entire hanuman langur clade in the Indian subcontinent. Further data resource based validity is necessary in this regards.

Introduction

In taxonomy, the species designation underpins several areas of biology including behavioral, ecological, systematic, evolutionary, and conservation studies (Thomson et al. 2018; Lehman et al. 2017; de Queiroz 2007). In particular, initial studies of organisms may rely most heavily on accurate and definitive species identifications (Kaiser et al. 2013; Rhodin et al. 2009; Hey 2009). However, for many taxa, the basis for species identification and classification remain either unclear or unresolved due to limits of many common practices used in taxonomy (Thiele and Yeates 2002). This issue may further be compounded by the existence of complexes of closely related species where individual taxa are not readily distinguishable, especially when relying extensively on morphological characters for identification purposes.

Species of primates found in India are an example of one group of organisms exhibiting many such unresolved taxonomic issues. The taxonomy of this group has been revised several times based on morphological, reproductive, behavioral, adaptive, bio-geographical characters and through the fossil record, but the validity of many of these schemes continue to be challenged (Perelman et al. 2011; Zalmout et al. 2010).

In Indian subcontinent, the gray langur, often popularly known as the “Hanuman langur” is an important primate species. As stated above, various taxonomic assignments have been suggested and developed over the past century in resolving the systematic position of the gray langurs in relation to other primates (Hill, 1939; Napier and Napier 1967; Ellerman and Morrison-Scott 1966; Roonwal and Mohnot 1977; Groves 2001; Roonwal 1984; Brandon-jones 2004), but again, many questions continue to be raised about the classification of this species (Roonwal and Mohnot 1977; Nag et al. 2014).

In the Indian subcontinent, the Hanuman langur is known for its extensive distribution, but it also shows plentiful plasticity in its morphology (Bishop 1979; Newton 1988). Therefore, the number of “species” encompassed by this taxon still remains uncertain, and this confusion may have been compounded by dubious estimates of diversity (de Queiroz 2007). Furthermore, this lack of clarity has major implications for this as well as any other species considered to be endangered. Similar disagreements over species status have been highlighted recently in primates due to the frequent disparities seen between classifications based on phylogeny vs. taxonomy (Hey 2009; Sites and Marchall 2004) and the importance of identifying species for conservation purposes (Issac et al. 2004).

Until recently, the gray langurs (Genus: *Semnopithecus*, Subfamily: Colobinae) were considered to be classified either *S. entellus* as a sole species, or into several subspecies of *S. entellus* (Bennett and Davies 1994; Groves 2001). For example, from South India, Hill (1939) described three species of Hanuman langurs which include, *S. hypoleucos*; *S. entellus* and *S. priam* along with *S. johnii* as a sister taxa popularly known as Nilgiri langur. However, Groves (2001) recognized seven species (*S. ajax*, *S. dussumieri*, *S. entellus*, *S. hector*, *S. hypoleucos*, *S. priam* and *S. schistaceus*) belonging to the taxonomic group *Semnopithecus*. Recently Nag et al. (2014) used the MaxEnt modeling tool and argued that *S. hypoleucos* and *S. priam* each have at least three subspecies.

Some of the conflicting results of these studies may reflect the use of different behavioral, morphological and or ecological data sets in deciding the species boundaries (Nag et al. 2014; Davies et al. 1994; Groves 2001; Napier 1970; Szalay 1979). These studies also involve highly complex procedures and questions have been raised when using classic procedures to validate the taxonomic status of individual specimens or offal materials.

Some studies have focused on the use of molecular data for delineating species boundaries in these primates based either on genes from the nuclear or mitochondrial genomes (Steinberg et al. 2009; Karanth 2010; Roos et al. 2007; Nadler et al. 2005; Wang et al. 2012; Geissmann et al. 2004). Such insights may be very useful since they allow for comparative genomics studies and provide dataset useful in taxonomic delineation and for reconstructing phylogenetic relationships. In these studies,

however, multilocus phylogenies appear to have advantages over those involving single loci from either nuclear or mitochondrial genome (Yoshida et al. 2019). To further this types of analysis, we sequenced whole the mitochondrial genome of the gray langur (*S. hypoleucos*), one of the important hanuman langur species, to create a genomic reference for future studies. The genome sequence obtained here was also annotated and compared with other members of langur family where genome datasets were available and to delineate phylogenetic relationships. These results obtained here will provide more insight and data for future studies on the taxonomy, phylogenetic relationships and the patterns of adaptive evolution of Hanuman langur species.

Materials And Methods

Ethical statement

To obtain samples we used a noninvasive method of collecting fallen hairs of *S. hypoleucos* from enclosures (floor and walls of night shelters) at the Pilikula Biological Park, Mangalore, Karnataka. Samples were collected during morning hours when animals were out for feeding to ensure no physical disturbances to them. As our sampling did not include any physical handling of animals or performing of experiments directly on them, ethical permissions were not required for this study.

Sample collection, DNA extraction and amplification

Collected hairs were disinfected by washing with ethanol and preservation in 70% ethanol before being brought to the laboratory. DNA was extracted and purified using the QIAamp DNA micro kit (Qiagen, Germany). Extracted DNA was quantified and checked for quality for downscale applications. For amplifying the mitochondrial genome, several primers were synthesized using sequences from published literature and tested. The primers which successfully amplified complete mitochondrial genome are depicted in supplementary table S1. Details of the PCR reactions mixes and thermocycling conditions are also included in supplementary table S1.

Qualitative and quantitative analysis of PCR Products

The quality of the PCR amplicons were checked on agarose gels (1%) and quantified on the Qubit® 2.0 Fluorometer (Thermo Scientific) system. Amplified products were subjected to gel extraction and purification using the Zymo large fragment DNA recovery kit to obtain DNA representing the ~12kb and 7.5kb bands. The remaining products were pooled alongside the 12kb and 7.5kb bands prior to library preparation.

Cluster Generation and Sequencing

After a quality check, the libraries were generated and sequenced on the MiSeq. Illumina platform (Illumina Inc. USA) following recommended guidelines. The template fragments were sequenced bi-directional employing paired-end sequencing. Gap regions not covered by the NGS reads were resequenced on a Sanger sequencing platform (ABI 3730xl, Foster city, CA, USA). For this, 12 new primers pairs were designed using a reference genome of *Trachypithecus johnii* (NCBI accession no. **NC_019583.1**) (Supplementary Table S1).

Assembly and annotation of mitogenome

Sequences obtained were aligned and assembled using the Codon Code aligner v4.0.3 (CodonCode, Dedham, MA, USA), and alignments were checked using BLAST. Both the Sanger and NGS sequences were combined as a hybrid genome assembly (using overlapping sequences) to cover the complete mitochondrial genome. Genome annotations were done through a web server MITO (Bernt et al. 2013). Data obtained were submitted to SRA (Short read archive) database (www.ncbi.nlm.nih.gov/sra) and assigned the accession id **SRR6027938**. For comparative analysis 25 pre published primate mitochondrial genomes were downloaded from NCBI Genbank (Supplementary Table S2)

Sequence alignment, data partition and phylogenetic analysis

Obtained DNA sequences of all putative protein-coding genes were aligned using the default settings in Muscle as implemented in MEGA7 (Kumar et al. 2016). Maximum likelihood (ML) algorithm and Bayesian inference (BI) procedures in MEGA v7.0 were followed for building the phylogenetic trees (Kumar et al. 2016). The GTRGAMMA model (1000X bootstrap) was selected for the concatenated datasets. For statistically balanced phylogenetic tree construction, MrBayes version 3.2.2 program (Ronquist et al. 2012) were also used and the Bayesian inference (BI) procedures were performed. Some 25% of trees were castoff, and the final resultant trees were used to build a majority consensus tree following high values of posterior probabilities. The reliability of the maximum likelihood phylogenetic tree was supported with 1000 bootstrap replications, and the best-fitting models were selected using Partition Finder V2.1.1.

Comparative mitogenomics

We compared the annotated mitochondrial genome of *S. hypoleucos* with mitochondrial genomes of 25 species (Supporting Table S2). Mitochondrial genomes were downloaded from NCBI refseq database (Pruitt et al. 2007). The base composition of the mitogenomes in all taxa under study was calculated

using MEGA 7.0. (Kumar et al. 2016) AT and GC asymmetries were calculated by $AT\ skew = (A - T\%)/(A + T\%)$ and $GC\ skew = (G - C\%)/(G + C\%)$ formula (Hassanin et al. 2005). Rates of synonymous substitution (Ks), nonsynonymous substitutions (Ka) and ka/ks ratio were obtained using a computer program Dnasp ver. 4.0 (Librado and Rozas 2009).

Estimation of species boundary

Sequence divergence values within and among species were employed the Kimura two parameter (K2P) model (Kimura 1980) using analytical functions on MEGA 7.0. (Kumar et al. 2016) for deciding species boundary as suggested by Hebert et al. (2003). A neighbor joining (NJ) tree based on K2P distance, nearest neighbor analysis (NN), and nucleotide composition values were also obtained using MEGA 7.0. The analysis of genetic distances was complemented by downloading of related sequences from GenBank (Wiemers and Fiedler 2007).

Results

Mitochondrial Genome characteristics

Overview of the mitochondrial genome

This study reports on first annotated sequence of a complete mitochondrial genome of Hanuman langur, *S. hypoleucos*. The sequence dataset consisting of 16563 bp has been deposited in GenBank under accession **MG471386**. The annotated genome (Fig. 1) contains 37 genes which includes 13 genes that encode proteins (atp6 and 8, cytb, cox1, 2 and 3, and nad1 to nad6), 22 tRNAs, two rRNA genes (rrnL, rrnS), and one D loop (Table 1). Of these genes, 28 are located on positive strand and the other nine are found on the negative strand (Fig. 1).

tRNA Structure

The twenty two tRNAs show typical clover leaf shaped secondary structures having a length of 64 to 73 bp in size. Among these, three tRNAs have a common antisense codon (UGC), while the remaining 19 each use a unique antisense codon (Fig. 2).

Nucleotide composition

The nucleotide composition of *S. hypoleucos* mitochondrial genome showed skewed values for AT and GC reflecting strand asymmetries. The value for AT skew was found to be 0.056 while the GC skew is -0.338. The nucleotide compositions and GC percentage values for 26 primate species are given in Table 2.

Codons (Initiation and termination)

The initiation codon and termination codon usage bias for protein coding genes (PCG) was analyzed for 25 primate species and compared with *S. hypoleucos* (Table 3).

Codon usage

The analysis of relative synonymous codon usage (RSCU) for 26 species shown, *S. hypoleucos* uses more NNA and NNT codon compared to others. The codon AUG was used as start codon among PCGs in all 26 primates. The UAA and UAG stop codons, were used, and incomplete stop codons were not observed (Table 3).

Evolutionary rates

To examine the evolutionary modifications among the PCGs in primate species, the values of non-synonymous substitutions (K_a), synonymous substitutions (K_s), and the ratio of K_a/K_s , (ω) were analyzed. The evolutionary rate values (ω) of Cox1 was found to be lowest (0.02), while it was highest for ATP8 (0.22) (Fig. 3). From these observations, genes having lower evolutionary rate can be used in molecular differentiation of primate species.

Gene order in primates

Gene order is predicted for *S. hypoleucos* and compared it with other 25 primate species. The order of genes has no significant changes but there are slight differences in their gene length (Table 1).

Comparative accounts

To infer phylogenetic relationships of *S. hypoleucos* in the primate clade, a comparative study was conducted using previously published mitogenomes of 25 primate species (Supplementary Table S2). Based on this comparison, genetic distance was deduced between *S. hypoleucos* and other primate species based on different genes, a concatenated assembly of 13 PCGs and for complete mitochondrial genome (Table 4). From the genetic distance values it can be observed that most of the mitochondrial genes are useful in molecular differentiation of primate species except ATP6, 8 and ND 6 genes, whereas, whole mitochondrial genome based comparisons are also reliable for species differentiations (Table 4).

Phylogenetic analysis

Data analysis resulted in four partitions best suited with an evolutionary model. The first partition GTR+I+G included five genes, specifically CytB, COIII, ND4L, COII and ND1. The second partition GTR+I+G included six genes: ATP6, ATP8, ND2, ND3, ND4 and ND5. The third partition GTR+I+G included one gene (COI) whereas the fourth partition HKY+G included a single gene ND6 (Table 5). Phylogenetic relationships among 26 primate species were inferred from the phylogeny constructed using concatenated datasets of protein coding gene sequences from mitochondrial genomes (Fig. 4). For further clarity on hanuman langur species complex, a pairwise between species genetic distance comparison was done (Table 4).

Discussion

Taxonomic classifications systems always have inherent strengths and weaknesses for studies of systematics (Mayden 1997; Wheeler and Meier 2000). In general, better taxonomic assignments are possible when combination of such methods are employed (Yoshida et al. 2016). Among primates, these issues have impacted taxonomic and systematic studies of Hanuman langurs for many years. Debates on taxonomic classification in the literature have shown disparities on the status of species vs. sub species of several populations of these primates (Nag et al. 2011; Nag et al. 2014). For example, Hanuman langurs was considered as a single species (*S. entellus*) by several authors, but these include classification schemes that encompass 14 to 16 subspecies. Other split these taxa into several individual species (Nag et al. 2011).

As in any taxonomic study, much depends on the characters used to decide on making classifications. Recently Nag et al. (2014), using an array of five morphological diagnostic characters, observed that Hanuman langurs comprises of at least six morphotypes in the peninsular India. This study also employed ecological niche modelling to validate the status of these six taxa, and recommended using Hill's (1939) classification scheme for the taxonomy of these langurs in the future. These findings of need to be appraised by employing independent tools and data informatics. However, the use of molecular data for species authentication often relies on only one, or a small number of genes.

Expanded molecular datasets developed from multiple loci or extended regions of the genome may help in facilitating the development of more robust analyses of phylogenetic relationships of taxa. For this purpose, here we have sought to develop the complete mitochondrial genome data resource in the hope that it will aid future studies attempting to resolve the taxonomic status of taxa in the Hanuman langur clade.

The mitochondrial genome sequenced studied here was obtained using material from a registered species (*S. hypoleucos*) of Hanuman langur (Pilicula Biological Part, Mangalore, Karnataka, India). The genome sequence and gene arrangement of this newly sequenced Hanuman langur mitochondrial genome is similar to other primates in terms of genomic features (Wang et al. 2012) (Table 1, 2 and 3). For example, most of the tRNA identified here appear to adopt classical secondary structures except that the anticodon stems of nine base pairs appear to possess a bulge that is a unique character reported here (Fig. 2). A highly conserved DHU stem, aminoacyl and anticodon arm and aminoacyl acceptor stems were noticed in contrasts to other Colobinae compared here. However, most of the sequence difference is confined to the pseudouridine (T ψ C), DHU loops. Whereas, the variable arms are having indel polymorphisms (Fig. 2). Some of this variation may be representative of a response to thermal adaptations by these organism as suggested by (Lorenz et al. 2017). Results also suggest that primate mitochondrial genomes are diverse in terms of nucleotide composition, codon usage and percent GC content (Table 2). The mitochondrial genome sequenced here is A+T rich overall as can be seen in the comparative accounts given in Table 2. Interestingly, the values for T (%) were exceptionally high in case of the ND6 gene in comparison with others. This may in part be correlated with the location of this gene on the L-strand here while others are found on the H-strand (Fig. 1) (Urbina et al. 2006).

The evolutionary rate (ω) estimated here for primate species revealed that ATP8 had the highest value (0.377) whereas COI, COII, COIII and CytB had the lowest evolutionary rates (<0.07). This suggests that they may be potentially useful markers for evolutionary studies and for molecular identification of organisms (Fig. 3) (Shi et al. 2016). Overall, however, the evolutionary rates for all the PCGs are substantially lower than 1.0 (<0.20), suggesting that these genes may be evolving under purifying selection (Shi et al. 2016; Jacobsen et al. 2016). Therefore, when drawing phylogenetic relationships within Colobinae, all mitochondrial PCGs should be used. Further, the comparison of RSCU (Relative Synonymous Codon Usages) for 26 species showed the newly sequenced mitochondrial genome had more UAU and AUU codons than others, it means desired codons are more frequent at sites having the conserved amino acid so as to retain it in the course of evolution (Stoletzki and Eyre-Walker 2007). From this observation it may be hypothesized that the species within the genus *Semnopithecus* may have more consistency in codon usage, and this can give clues for other species to be classified in the future (Table 3). However, based on best suited evolutionary model, mitochondrial PCGs were partitions into four according to which data was analyzed in this study (Table 5). Accordingly, use of different genetic locus data like nuclear genome and mitochondrial genome combine may endorsed to the co-evolutionary relic of most species, because of the evolutionary events such as meiotic sexual recombination in eukaryotes. In such cases phylogenetic trees based on different locus (either different gene or different genome

resources like mitochondrial or nuclear DNA) do not show precisely coordinated topologies and chances to mislead species estimates cannot be overruled.

To test the robustness of using mitochondrial genome based on multilocus phylogenetics in the primate clade, a concatenated assembly of 13 PCGs was used. The maximum likelihood and Bayesian inferences resulted in consistent tree topologies supporting the appropriateness of evolutionary model selection for building the phylogeny. The node support specifies high values for posterior probability (> 0.95) (Fig. 3A) and high bootstrap support value for ML (>95) (Fig. 3B). Based on the taxon support values, species under NCBI accession NC_019583.1 and NC_019582.1 appear to belong to the genus *Simnopithecus*, as has been suggested by several others (Malviya et al. 2011; Nag et al. 2014; Groves 2017). Such annotations are also comparable with the observations made by Fabre et al. (2009) and Chatterjee et al. (2009) where authors noted that the published sequences culled from prior studies were characteristically inconsistent in terms of missing data in a great proportion relating to a taxon. However, our study has created a sound background to use multi locus genetic data for species comparison as shown in table 4 and 6. This comparison reflected that species belong to genus *Simnopithecus* (= *Trachypithecus*) are well resolved using molecular data following species boundary criteria of 3% (Hebert et al. 2003) at all four locus compared here. This reflects, utility and robustness of molecular data over other data sets such as behavioral, morphological and or ecological data sets in deciding the species boundaries (Nag et al. 2014; Davies et al. 1994; Groves 2001).

Results shown in Figure 3 also suggest that there is no sister clad relationship between the *Trachypithecus* and *Semnopithecus* genera (Fig. 3A & B and Table 6). This result refutes the earlier finding by Roos et al. (2011). Our phylogenetic analysis does, however, confirm the sister relationship of the Nilgiri langur (*S. hohnii*) with the hanuman langur (*S. hypoleucos*), consistent with the observations by Nag et al. (2014). Further, we can hypothesize that *S. enterllus* is a most recent common ancestor to *S. hypoleucos* based on the phylogenetic analysis as well as pairwise genetic comparison at various molecular locus (Table 4). According to the data presented in table 6, four species belonging to hanuman langur complex are well resolved and suggest a recent speciation event based on narrow genetic distances as compare to other species compared (Tables 4 and 6).

In overall, based on genetic distance values compared for primate clade at different locus, it can be proposed that species boundary values for COI would be $> 3\%$, for Cyt B would be $>5\%$, for concatenated assembly of 13 PCGs would be $>6\%$ and for complete mitochondrial genome data would be $>6\%$. Based on this assumptions, use of either single or multiple gene from mitochondrial genome may be advantageous based on their genetic resolution values (Table 6).

Prior to our study, there was only one complete mitogenome of a *Semnopithecus* species was available (NCBI accession DQ355297.1), and this was considered to be a limiting factor in resolving phylogenetic relationships of these taxa. In our this study we have created a new record of a complete primate mitochondrial genome, and this has aided in resolving the ambiguity for genus assignments of *T. johnii* as *Semnopithecus johnii* and *T. vetulus* as *Semnopithecus vetulus* (Fig. 4) in a manner consistent with

the opinions of various other authors (Wang et al. 2012; Osterholz et al. 2008; Brandon-Jones et al. 2004). In part this may be due to the fact that the earlier scientific efforts to resolve taxonomic relationships to these two genera relied on only the three genes, COI, CytB and ND2 for reconstructing phylogenies, an approach that has been found to be substantially insufficient in relation to statistical support to phylogenetic clades and percent identities (Karantha et al. 2008; Moore 1995). In addition, the pairwise genetic distance computations revealed that comparisons of *S. hypoleucos* with all 25 primates produced sufficiently high difference values (>3%) as required for species identification (Table 5) (Hebert et al. 2003). Based on inferences about sub species or morphotype species made by Nag et al. (2014), our study further substantiates the need for sequencing whole mitochondrial genome for other members referred as Hanuman langur to aid in rectifying their taxonomic status, testing the species boundary and resolving phylogeny of the entire hanuman langur clade in the Indian subcontinent.

Declarations

Acknowledgement

VH was supported by University Grants Commission, New Delhi under JRF scheme, India. GDK was supported by Dr. Babasaheb Ambedkar Marathwada University Aurangabad, India (Grant Ref. No. STAT/IV/RC/Dept./2019-20/295-96 dated 13.06.3019) and Rashtriya Uchchatar Shiksha Abhiyan (RUSA), Maharashtra, India, (Grant no. PD/RUSA/Order/2018/127 dated February 14th, 2018). We sincerely thank the Pilikula Biological Park, Mangalore providing sample of hairs with follicle. We sincerely thank all staff member and students at Paul Hebert Centre for DNA Barcoding and Biodiversity Studies, Aurangabad for their assistance in completing this work.

Authors Contribution VH, GDK, CDK, RS and BP conceiving research idea, VH, BP collected samples and conducted laboratory experiments, VH analysed data and GDK, RS, CDK, BP wrote manuscript.

Competing interests: The authors do not have conflict of interest to declare.

References

Bennett EL, Davies AG (1994) The ecology of Asian colobines. In: Colobine Monkeys: Their Ecology, Behaviour and Evolution, AG Davies and JF Oates (eds.), pp.129–172. Cambridge University Press, Cambridge, UK.

Bernt MA, Donath FJ, Externbrink F, Florentz C, Fritzsche, Pütz GJ, Middendorf M, Stadler PF (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol Physiol Evol* 69:313-319. <https://doi.org/10.1016/j.ympev.2012.08.023>

Bishop NH (1979) Himalayan langurs: temperate colobines. *J Human Evol* 8:251–81.

Brandon-jones D (2004) A taxonomic revision of the Langurs and Leaf Monkeys (Primates: Colobinae) of South Asia. *Zoos' Print J* 19:1552–1594.

Chatterjee HJ, Ho SY, Barnes Groves C (2009) Estimating the phylogeny and divergence times of primates using a supermatrix approach. *BMC Evol Biol* 9:1-19. <https://doi.org/10.1186/1471-2148-9-259>

de Queiroz K (2007) Species concepts and species delimitation. *Sys Bio* 56:879-886.

Ellerman JR, Morrison-Scott TCS (1966) Checklist of palaearctic and Indian 455 mammals, 1758–1946 (2nd ed.). London: Trustees of the British Museum (Natural 456 History).

Fabre V, Condemi S, Degioanni A (2009) Genetic evidence of geographical groups among Neanderthals. *PLoS One*: 4. <https://doi.org/10.1371/journal.pone.0005151>

Geissmann T, Groves CP, Roos C (2004) The Tenasserim lutung, *Trachypithecus barbei* (Blyth, 1847) (Primates: Cercopithecidae): Description of a live specimen, and a reassessment of phylogenetic affinities, taxonomic history, and distribution. *Contributions Zoo* 73(4): 271-282.

Groves CP (2001) *Primate taxonomy*. Smithsonian Institution Press, Washington D.C.

Groves C, Napier John (2017) In *The International Encyclopedia of Primatology* (eds M Bezanson, KC MacKinnon, E Riley, CJ Campbell, K Nekaris, A Estrada, AF Di Fiore, S Ross, LE Jones-Engel, B Thierry, RW.

Sussman, C Sanz, J Loudon, S Elton, A Fuentes). doi:[10.1002/9781119179313.wbprim0042](https://doi.org/10.1002/9781119179313.wbprim0042)

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proc Natl Acad Sci USA* 100(12):12884-12889. doi:10.1098/rspb.2002.2218

Hey J (2009) Why Should We Care about Species? *Nature Education*, 2(5), 2 India. *International Journal of Primatology* 32:830–848.

Hill WC (1939) An annotated systematic list of the leaf-monkeys. *Ceylon J Sci* 21:277–305.

Isaac NJB, Agapow PM, Harvey PH, Purvis A (2003) Phylogenetically nested comparisons for testing correlates of species-richness: a simulation study of continuous variables. *Evol* 57: 18-26.

Jacobsen MW, Rute R, Bernatchez L, Hansen MM (2016) Molecular Phylogenetics and Evolution Comparative analysis of complete mitochondrial genomes suggests that relaxed purifying selection is driving high nonsynonymous evolutionary rate of the NADH2 gene in whitefish (*Coregonus ssp.*). *Mol Phy Evol* 95:161-170. <https://doi.org/10.1016/j.ympev.2015.11.008>

Kaiser H, Crother BI, Kelly CMR, Luiselli L, O'Shea M, Ota H et al. (2013) Best practices: in the 21st century, taxonomic decisions in herpetology are acceptable only when supported by a body of evidence and published via peer-review. *Herpetology Rev* 44(1):8-23.

Karanth KP, Singh L, Collura, RV, Stewart CB (2008) Molecular phylogeny and biogeography of langurs and leaf monkeys of South Asia (Primates: Colobinae). *Mol Phy Evol* 46:683-694.

<https://doi.org/10.1016/j.ympev.2007.11.026>

Karanth KP (2010) Molecular systematics and conservation of the langurs and leaf monkeys of South Asia. *J Gene* 89:393-399.

Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J Mol Evol 29(2):170-179.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870-1874. <https://doi.org/10.1093/molbev/msw054>

Lehmann AW, Devriese H, Tumbrinck T, Skejo J, Lehmann GUC, Axel Hochkirch A (2017) The importance of validated alpha taxonomy for phylogenetic and DNA barcoding studies: a comment on species identification of pygmy grasshoppers (Orthoptera, Tetrigidae), Zoo Keys: 679:139-144. <https://doi.org/10.3897/zookeys.679.12507>

Lorenz C, Lunse CE, Morl M (2017) tRNA Modifications: Impact on Structure and Thermal Adaptation. Biomolecules 7(2):35. <https://doi.org/10.3390/biom7020035>

Osterholz M, L Walter, C Roos (2008) Phylogenetic position of the langur genera Semnopithecus and Trachypithecus among Asian colobines, and genus affiliations of their species groups. BMC Evol Biol 8:1-12.

Malviya M, Srivastav A, Nigam P, Tyagi PC (2011) Indian National Studbook of Nilgiri Langur (*Trachypithecus johnii*). Wildlife Institute of India, Dehradun and Central Zoo Authority, New Delhi.

Mayden RL (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In Claridge MF, Dawah HA, Wilson MR (eds.) Species: The units of diversity, Chapman & Hall. pp. 381-423.

Moore SW (1995) Inferring Phylogenies from Mt DNA variation: Mitochondrial gene trees versus nuclear gene trees, Evolution 49(4):718-726.

Nag KSC, Pramod P, Karanth KP (2011) Taxonomic Implications of a Field Study of Morphotypes of Hanuman Langurs (*Semnopithecus entellus*) in Peninsular India. *Int J Primatology* 32: 830-848.
<https://doi.org/10.1007/s10764-011-9504-0>

Nag C, Karanth KP, Gururaja KV (2014) Delineating ecological boundaries of Hanuman Langur species complex in peninsular India using MaxEnt modeling approach. *Plosone* 9(2):e87804.
<https://doi.org/10.1371/journal.pone.0087804>

Napier JR, Napier PH (1967) A handbook of living primates: morphology, ecology and behaviour of nonhuman primates. Academic Press.

Napier JR, Napier PH (1970) Old World monkeys. London: Academic Press.

Newton PN (1988) The variable social organization of hanuman langurs (*Presbytis entellus*), infanticide and the monopolization of females. *Int J Primatol* 9(1):59-77.

Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290:470-474.

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 Evol Bio* 8:1-12.
<https://doi.org/10.1186/1471-2148-8-58>

Panwar B, Raghava G (2014) Prediction of uridine modifications in tRNA sequences. *BMC Bioinfo* 15(1):326.

Perelman P, Johnson WE, Roos C, Seuánez HN, Horvath JE, Moreira MAM, Kessing B, Pontius J, Roelke M, Rumpler Y, Schneider MPC, Silva A, O'Brien SJ, Pecon-Slattery J (2011) A molecular phylogeny of living primates, *PLoS Gen* 7:1-17. <https://doi.org/10.1371/journal.pgen.1001342>

Pocock RI (1928) The langurs or leaf monkeys of British India. J Bom Nat His Soc 32:472-504.

Rhodin AGJ, Kaiser H, van Dijk PP, Wuster W, O'Shea M, Archer M, et al. (2015) Comment on *Spracklandus* Hoser, 2009 (Reptilia, Serpentes, ELAPIDAE): request for confirmation of availability of the generic name and for the nomenclatural validation of the journal in which it was published. Bull Zool Nom 72(1):65-78.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Sys Bio 61:539-542.

Roonwal ML, Mohnot SM (1977) Primates of South Asia: Ecology, Sociobiology, and Behavior. Harvard University Press: Cambridge, MA.

Roos C, Zinner D, Kubatko LS et al. (2011) Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. BMC Evol Biol 11: 77. <https://doi.org/10.1186/1471-2148-11-77>

Shi Y, Chu Q, Wei DD, Qiu YJ, Shang F, Dou W, Wang JJ (2016) The mitochondrial genome of booklouse, *Liposcelis sculptilis* (Psocoptera: Liposcelididae) and the evolutionary timescale of *Liposcelis*. Sci Rep 6:1-11. <https://doi.org/10.1038/srep30660>

Sites JW Jr, Marshall JC (2004) Operational criteria for delimiting species. Annual Reviews in Ecology. Evo Syst 35:199–227.

Steinberg ER, Nieves M, Ascunce MS, Palermo AM, Mudry MD (2009) Morphological and genetic characterization of *Saimiri boliviensis* (Primates: Cebidae): Implications for management and conservation. Cons Gen 30:29-41.

Sterner KN, Raaum RL, Zhang YP, Stewart CB, Disotell TR (2006) Mitochondrial data support an odd-nosed colobine clade. *Mol Phylogenet Evol* 40:1-7. <https://doi.org/10.1016/j.ympev.2006.01.017>

Stoletzki N, Eyre-Walker A (2007) Synonymous codon usage in *Escherichia coli*: Selection for translational accuracy. *Mol Biol Evol* 24:374-381.

Szalay FS, Rosenberger AL, Dagosto M (1987) Diagnosis and differentiation of the order primates. *Yearbook of Physical Anthropology* 30:75-105.

Thiele K, Yeates D (2002) Tension arises from duality at the heart of taxonomy. *Nature* 419(6905):337. [pmid:12353005](https://pubmed.ncbi.nlm.nih.gov/12353005/).

Thomson SA, Pyle RL, Ahyong ST, Alonso-Zarazaga M, Ammirati J, Araya JF, et al. (2018) Taxonomy based on science is necessary for global conservation. *PLoS Biol* 16(3): e2005075. <https://doi.org/10.1371/journal.pbio.2005075>

Urbina D, Tang B, Higgs PG (2006) The response of amino acid frequencies to directional mutational pressure in mitochondrial genome sequences is related to the physical properties of the amino acids and to the structure of the genetic code. *J Mol Evol* 62:340-361.

Wang XP, L Yu, C Roos, N Ting, 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(4):1-12.

Wheeler QD, Meier R (2000) Species concepts and phylogenetic theory: a debate. Columbia University Press, New York

Yoshida R, Fukumizu K, Vogiatzis C (2019) Multilocus phylogenetic analysis with gene tree clustering. *Ann Oper Res* 276:293-313. <https://doi.org/10.1007/s10479-017-2456-9>

Zalmout IS, Sanders WJ, MacLatchy LM, Gunnell GF, Al-Mufarreh YA, Ali MA, Nasser AAH, Al-Masari A M, Al-Sobhi, SA, Nadhra AO, Matari AH, Wilson JA, Gingerich PD (2010) New oligocene primate from Saudi Arabia and the divergence of apes and old world monkeys. *Nature* 466:360-364.
<https://doi.org/10.1038/nature09094>

Tables

Please see the supplementary files section to view the tables.

Figures

Figure 1

Mitochondrial genome of *S. hypoleucos* (Circular map) (Abbreviations of gene names are: atp6 and atp8 for ATP synthase subunits 6 and 8, cox1-3 for cytochrome oxidase subunits 1-3, cob for cytochrome b, nad 1-6 for NADH dehydrogenase subunits 1-6, rrnL and rrnS for large and small rRNA subunits) a. First inner circle: light strand b. Second circle from inner side: graphic directed towards centre- GC skew values on light strand and graphic directed towards outside- GC skew values on heavy strand c. Third circle from inner side: graphic directed towards centre- GC content on light strand and graphic directed towards outside- GC content on heavy strand d. Outermost circle: Heavy strand

Figure 2

Graphical representation of structural details of tRNA structures of mitochondrial genome of *S. hypoleucos*

Figure 3

Box plot showing evolutionary rates of mitochondrial protein coding genes. Dotted lines show the mean values.

Figure 4

Phylogenetic relationship of primate species using a concatenated assembly of 13 protein coding genes from mitochondrial genome. (Node support is indicated by black arrow and blue circle. Black arrow indicates high Bayesian posterior probability support (> 0.95) as well as a high maximum likelihood bootstrap support (>95). Blue circle indicates only a high Bayesian posterior probability support (< 0.75))

A. Bayesian phylogeny of primates B. Maximum Likelihood phylogenetic tree

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.docx](#)
- [Table2.docx](#)
- [Table3.docx](#)
- [Table4.docx](#)
- [Table5.docx](#)
- [Table6.docx](#)
- [SupplementaryTableS1.docx](#)
- [SupplementaryTableS2.docx](#)