Short Communication

Additional molecular evidence strongly supports the distinction between the recently described African primate *Rungwecebus kipunji* (Cercopithecidae, Papionini) and *Lophocebus*

Link E. Olson, Eric J. Sargis, William T. Stanley, Kyndall B.P. Hildebrandt, Tim R.B. Davenport

1. Introduction

The kipunji (*Rungwecebus kipunji*), known only from the Southern Highlands and Udzungwa Mountains of southern Tanzania, was originally described from photographs (Jones et al., 2005). The authors placed the new species in the genus *Lophocebus* based on its arboreality and non-contrasting black eyelids, two characteristics it shares with *Lophocebus* species. Soon thereafter, the collection of a voucher specimen from Mt. Rungwe allowed molecular phylogenetic and comparative morphological analyses, which resulted in the placement of the kipunji in a new genus, *Rungwecebus* (Davenport et al., 2006).

By not acknowledging the kipunji in his recent taxonomic revision of the *Lophocebus albigena* group and description of the Uganda mangabey (*L. ugandae*), Groves (2007, p. 69) implicitly embraced the recognition of *Rungwecebus* and noted that “[d]espite the fact that genera (and families, and orders) are currently recognized in a fashion that is still quite arbitrary...they must be monophyletic.” This is consistent with the conclusions of Davenport et al. (2006), who described *Rungwecebus* due in part to its well-supported sister relationship to baboons (*Papio* spp.) rather than arboreal mangabeys (*Lophocebus* spp.). This grouping was consistently recovered in analyses of multiple mitochondrial genes and a single autosomal gene (although analyses of a y-linked marker were equivocal). However, in a review of previously published data on the kipunji, Ehardt and Butynski (2006, p. 81) stated that they “and several [unnamed] experts in primate taxonomy and molecular biology” were not in agreement with the decision of Davenport et al. (2006) to recognize *Rungwecebus*, citing an as-yet unpublished study (“Disotell et al., in prep.”) but offering no rationale for their disagreement. Given this, we conducted phylogenetic analyses on an expanded molecular dataset that now includes representative DNA sequence data from all four mammalian inheritance pathways (autosomal, x-linked, y-linked, and mitochondrial) and relevant papionin taxa.

Clarifying the evolutionary relationship of *Rungwecebus* to other primates will not only elucidate the origin of this unique taxon, but may illuminate the history of the endemic biota of the Southern Highlands and the Udzungwa Mountains. The enigmatic nature of this biota is illustrated not only by *Rungwecebus*, but also the recent discoveries of birds and shrews with ancient and geographically distant affinities (Dinesen et al., 1994; Stanley et al., 2005).

2. Materials and methods

2.1. Taxonomic and gene sampling

Because our objective was to further test the phylogenetic position of *Rungwecebus* within papionin primates, we took advantage of existing DNA sequence data from all four mammalian inheritance pathways available for relevant taxa on GenBank (see Table 1). Minimally, we sought genes for which sequences were available for *Papio*, *Lophocebus*, and at least one additional papionin. For this study, we sequenced and analyzed 578 base pairs (bp) of the nuclear autosomal lipoprotein, Lp(a) gene (LPA hereafter; Boffelli et al., 2003), 580 bp of the nuclear autosomal gene encoding the CD4
molecule used in Harris and Disotell (1998), and 1270 bp of the X chromosome intergenic region used in Tosi et al. (2005). In addition, we expanded the coverage of the y-linked testis-specific protein (TSPY) sequence for Rungwecebus reported by Davenport et al. (2006) from 573 bp to 1592 bp. These new or expanded datasets were analyzed separately and in two combined analyses with the data for three mitochondrial genes (12S, CO1, CO2) and an additional autosomal gene (X 1.3-GT) from Davenport et al. (2006). The combined analyses included one analysis maximizing taxon sampling (including Theropithecus but not including LPA, which has not been sequenced for that taxon) and another maximizing character sampling (excluding Theropithecus but including LPA).

2.2. DNA extraction, PCR and sequencing

To date, Rungwecebus is represented by a single voucher specimen housed in the Field Museum of Natural History, and we used the same DNA extraction from that specimen (FMNH 187122, a subadult male) as in our previous study (Davenport et al., 2006). A 1273 bp fragment of the noncoding X chromosome intergenic region spanning positions 6193–7465 of the Papio hamadryas sequence on GenBank (AY899234) was amplified and sequenced in three broadly overlapping sections using primers LOF3/LOR2, KHInF1/KHInR2, and LOR2/LOR1 (a list of all primers used in this study is provided in Table 2). A 578 bp fragment of the LPA gene encompassing positions 68–647 of Papio sp. AF057388, which is primarily intronic but includes a short region of exon 5, was amplified and sequenced with primers F4 and R4; this includes spanning positions 1421–1843 of Papio hamadryas (Davenport et al., 2006), whose GenBank accession has now been updated (Table 1).

Amplification and sequencing followed the methods outlined in Davenport et al. (2006); all new Rungwecebus sequences have been deposited to GenBank (Table 1).

2.3. Alignment and phylogenetic analyses

Sequences were aligned by eye in either Sequencher (v. 4.6, GeneCodes) or MacClade (v. 4.08; Maddison and Maddison, 2005) with reference to the translated amino acid sequence for coding encompassing positions 636–1693 of P. hamadryas AF284277 was amplified and sequenced in two overlapping fragments using primers F2/IR3 and InF/InR. The resulting fragment overlaps the 3' end of the original TSPY sequence reported for Rungwecebus (Davenport et al., 2006), whose GenBank accession has now been updated (Table 1).

Table 2

<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer (direction)</th>
<th>Sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPA</td>
<td>F4 (forward)</td>
<td>GTACCAGAAATGGAGCTA</td>
</tr>
<tr>
<td></td>
<td>R4 (reverse)</td>
<td>CAATCCATCGGATGGCA</td>
</tr>
<tr>
<td>CD4</td>
<td>LOF1 (f)</td>
<td>TACATCTGATAGCGGAGCA</td>
</tr>
<tr>
<td></td>
<td>LOR1 (r)</td>
<td>ATGTCGACAGTCTAGGCA</td>
</tr>
<tr>
<td></td>
<td>LOF3 (f)</td>
<td>CTTATATAAAGGGGCAAGGA</td>
</tr>
<tr>
<td></td>
<td>KHInF1 (f)</td>
<td>TGGATGACGCTGGGAGTG</td>
</tr>
<tr>
<td></td>
<td>LOF2 (r)</td>
<td>AGTGTGACAGTCCAGGATA</td>
</tr>
<tr>
<td></td>
<td>LOR2 (r)</td>
<td>TGGTACCCATACCGGAGTACGTA</td>
</tr>
<tr>
<td></td>
<td>KHInR2 (r)</td>
<td>CCAAGGCTGCTGCAGATAC</td>
</tr>
<tr>
<td></td>
<td>LOR1 (r)</td>
<td>GACCAAGGCAAGTGTGAAAGCA</td>
</tr>
<tr>
<td>TSPY</td>
<td>F2 (f)</td>
<td>AGGGAAAGGAGCTATCCATGGTG</td>
</tr>
<tr>
<td></td>
<td>IR3 (r)</td>
<td>GTGGTGCATCTCTGTCAGTA</td>
</tr>
<tr>
<td></td>
<td>InF (f)</td>
<td>TGGTGGAAGCAGGAGACTG</td>
</tr>
<tr>
<td></td>
<td>InR (r)</td>
<td>TGGACGCCATGCAAACCAAGTGG</td>
</tr>
</tbody>
</table>

a Length (in bp) of included characters (minus alignment–ambiguous positions and end regions missing in relevant taxa).
b Generated for this study.
c Includes additional sequence from that reported by Davenport et al. (2006).
d Used in combined analyses (Fig. 1D).
regions. For the X chromosome intergenic region, two insertion/deletion (indel) events were inferred—a 6 bp deletion in *Mandrillus sphinx* spanning alignment positions 153–158 and a separate indel requiring 1–3 gaps inserted at positions 456–458 in several taxa. The latter was considered to be alignment-ambiguous and was excluded from all analyses. For the LPA gene, two single bp indels were inferred at positions 424 and 567, neither of which falls within the exonic region. The addition of the newly generated TSPY sequence necessitated single bp indels at positions 1093, 1360, 1489, and 1639 in various taxa. The total number of aligned positions used in each data set is provided in Fig. 1 and all alignments used in this study are provided in Appendix 1. Details on the alignment of sequences used in combined analyses here but first reported in *Davenport et al. (2006)* can be found therein.

Each of the three new markers sequenced for this study, as well as the expanded TSPY data set, was analyzed separately with relevant taxa for which sequences were available on GenBank (Table 1). For the combined analyses, previously reported *Rungwecebus* sequences from *Davenport et al. (2006)* were also included. Phylogenetic analysis was performed under the maximum likelihood criterion using PAUP* (v. 4.0b10; *Swofford, 2003*) under models of nucleotide substitution estimated separately for each gene (for individual gene

![Fig. 1. Optimal trees recovered from maximum likelihood analysis of the (A) nuclear autosomal LPA gene, (B) nuclear autosomal CD4 gene, (C) y-linked TSPY gene (relationships among outgroup taxa not shown), (D) X-chromosome intergenic region, (E) combined dataset including all mitochondrial, nuclear autosomal, x-linked, and y-linked markers, and (F) combined dataset including all genes available for Lophocebus, Theropithecus, Papio, and Rungwecebus. Numbers next to nodes represent likelihood bootstrap proportions (top) and Bayesian posterior probabilities (bottom). Scale bars represent branch lengths in the indicated number of inferred substitutions per site. Length of each dataset (in base pairs) represents the number of sites included in the analysis (i.e., not including ambiguously aligned regions or end regions with missing data). See text and Table 1 for concatenation details and GenBank accession numbers.](image)
analyses; a separate model was estimated for the combined matrix) under the Akaike Information Criterion as implemented in ModelTest (v. 3.7; Posada and Crandall, 1998; optimal models and parameter values are available in Appendix 1). Likelihood searches employed 20 heuristic replicates using the tree bisection–reconnection (TBR) branch-swapping algorithm in PAUP*, with starting trees obtained via stepwise addition. Bootstrap support was estimated based on 1000 pseudoreplicates using the subtree pruning–regrafting (SPR) branch-swapping algorithm on starting trees obtained by stepwise addition. The Shimodaira–Hasegawa test (Shimodaira and Hasegawa, 1999) as implemented in PAUP* was used to evaluate alternative topologies. Because it incorporates individual site-likelihoods for all characters and is known to be overly conservative due to the minimization of type I error (Goldman et al., 2000; Shimodaira, 2002), we limited its use to the combined analyses only. Trees resulting from otherwise identical heuristic searches constrained to recover topologies in which Rungwecebus and Papio were not sister taxa were compared to trees resulting from unconstrained searches. The full optimization option in PAUP* was used to generate the test distribution, with 1000 bootstrap replicates.

Bayesian posterior probabilities were estimated using MrBayes (v. 3.1.2; Ronquist and Huelsenbeck, 2003). For the TSPY dataset, four partitions were specified (first, second, and third codon positions and introns; nonadjacent introns and exons were not considered separate partitions). For LPA and CD4, no partitions were specified due to the relatively small number of coding positions (94). Likewise, no functional domains or features have been recognized in the X chromosome intergenic region (Tosi et al., 2005), which was therefore not partitioned. In the combined analysis favoring character sampling, which included these three genes, as well as the mitochondrial 12S, CO1, and CO2 genes and the autosomal α1,3-GT gene analyzed in Davenport et al. (2006), 16 partitions were specified. These included (1–2) paired stem and nonpairing loop or bulge positions in the 12S gene according to the 12S rRNA secondary structure model of Springer and Douzery (1996); (3–5) first, second, and third codon positions in CO2; (6–8) first, second, and third codon positions in CO1; (9–12) the four aforementioned TSPY partitions; and (13–16) separate partitions for α, 1,3-GT, LPA, CD4, and the X chromosome intergenic region. In the separate combined analysis favoring taxon sampling, LPA (which has not been sequenced for Theropithecus) was not included and only 15 partitions were specified.

For the 12S pairing partition, the doublet model of nucleotide substitution, which is appropriate for modeling stem regions of ribosomal genes (Ronquist and Huelsenbeck, 2003), was specified in MrBayes. For all remaining partitions, a model with six categories in MrBayes analyses is included at the end of each matrix in Appendix 1. Resulting trees were imported into PAUP* and, after discarding the first 10% as burn-in, combined in a majority-rule consensus tree to obtain posterior probabilities. All trees were rooted with Macaca (Table 1).

3. Results and discussion

3.1. Phylogenetic position of Rungwecebus with respect to other papionin genera

Results of the maximum likelihood and Bayesian analyses of the newly sequenced or expanded genes from Rungwecebus, as well as a new combined analysis of all sequences available for Rungwecebus and relevant ingroup taxa, are shown in Fig. 1. Rungwecebus was strongly supported as the sister taxon to baboons (Papio) by all markers except the autosomal CD4 and γ-linked TSPY genes, which failed to resolve relationships among the genera of interest (but did not favor a Rungwecebus + Lophocebus clade). The TSPY results (Fig. 1C) are unsurprising given the similar findings of Tosi et al. (2003) with respect to Lophocebus, Theropithecus, and Papio based on 2.2 kb of TSPY sequence (which includes the region analyzed in this study). Although relationships among these genera were fully resolved by both parsimony and likelihood analyses (their Figs. 2 and 3, respectively), parsimony bootstrap values and likelihood branch lengths supporting their Lophocebus + Theropithecus clade were as low as, or lower than, those associated with other nodes on their trees specifically highlighted as being weakly supported. Although our analyses of the TSPY data set failed to identify the kipunji’s closest relative, a Rungwecebus + Lophocebus grouping was only recovered in 12.4% of the bipartitions in Bayesian analyses (not shown).

As with the TSPY data, CD4 failed to resolve relationships among the aforementioned genera, although a Lophocebus + Theropithecus clade was supported with moderately high bootstrap support and a posterior probability of 0.93. Variation in this region of CD4 is very low, with no substitutions or indels between different species in the genera Mandrillus and Lophocebus and only a single variable position between Cercocebus galeritus and C. torquatus. This may account for the generally low bootstrap and posterior probability values for generic monophyly. More importantly, the sequence from Rungwecebus was identical to that published for Papio anubis. Although the inferred branch length between the ‘node’ connecting these two taxa and the next basal node is zero, their sequence identity constitutes obvious evidence of a close affinity between the two. With the exception of the inclusion of Rungwecebus, the topology in Fig. 1B is identical to that recovered by Harris and Disotell (1998, Fig. 2A).

The 578 bp region of the autosomal apolipoprotein [a] gene (LPA) in Rungwecebus was identical to that published for P. hamadryas. Not surprisingly, these two taxa were recovered as sister taxa (Fig. 1A) with large support values (the <100% bootstrap value can be attributed to the low number of variable positions [17]). This region of the LPA gene was targeted for sequencing in this study because it contained a relatively large number of positions representing inferred apomorphies in either Papio (three total) or Lophocebus (five total) with respect to the remaining non-Rungwecebus taxa shown in Fig. 1A. Our results therefore suggest that the three apomorphies we originally inferred in Papio represent synapomorphies shared with Rungwecebus. The alternative explanation, that these synapomorphies were present in the common ancestor of Papio, Rungwecebus, and Lophocebus but were subsequently obscured by apomorphies at these same three positions in Lophocebus, seems decidedly unlikely in light of the small total number of variable positions. We note that this gene, which is orthologically restricted to Old-World monkeys and hominoids (Lawn et al., 1995), has not been previously employed

| Macaca | . . . TAAATAA-AAAAAAAGATGT . . . |
| Cercocebus | . . . TAAATAA-AAAAAAAGATGT . . . |
| Mandrillus | . . . TAAATAA-AAAAAAAGATGT . . . |
| Lophocebus | . . . TAAATAA-AAAAAAAGATGT . . . |
| Theropithecus | . . . TAAATAA-AAAAAAAGATGT . . . |
| Papio | . . . TAAATAA-AAAAAAAGATGT . . . |
| Rungwecebus | . . . TAAATAA-AAAAAAAGATGT . . . |

Fig. 2. Positions 450–469 in the X chromosome intergenic region alignment. Shaded positions (456–458) were considered alignment-ambiguous but nonetheless likely represent a synapomorphetic insertion shared by Papio and Rungwecebus.
in phylogenetic studies of primates and future studies should include sequences from Theropithecus and the remaining Lophocebus species. Nonetheless, we consider the equality of our Rungwecebus sequence to that previously published for Papio to be additional strong evidence of their exclusive shared ancestry. The 1273 bp of X chromosome intergenic region sequence obtained from Rungwecebus differed at only two positions (one substitution and one single bp indel) from the corresponding sequence available for P. hamadryas. This is in striking contrast to the 14 substitutions and two indels observed between Papio and Rungwecebus. Indeed, the minimal variation between Papio and Rungwecebus is less than that observed between other congeneric cercopithecine species (e.g., Chlorocebus lhoesti AY899217 and C. solatus AY899218; see Tosi et al., 2003, on inclusion of lhoestiform species in Chlorocebus). In addition to overall sequence similarity and concomitantly strong support for a sister relationship (Fig. 1D), Rungwecebus and Papio are invariant at positions 456–458, which span a 1–3 bp indel (Fig. 2). We considered this short region to be too alignment–ambiguous to confidently infer positional homology and therefore excluded it from our analyses. However, the fact that Papio and Rungwecebus would be coded as identical to each other and different from all other taxa in Fig. 2 under any existing scheme developed for accommodating ambiguously aligned regions in phylogenetic analysis (e.g., Lutzoni et al., 2000) is unambiguous. Given the position of these two taxa in the topology shown in Fig. 1D, the most parsimonious explanation for their monomorphy is an insertion event, and we agree with Tosi et al. (2005) that indels in slowly evolving genes such as this should be considered “powerful indicators of shared ancestry” (p. 62).

Results of the combined analyses are shown in Fig. 1E–F. Although both data sets suffer more than any individual gene matrix in terms of sparse taxon sampling, all three papionin genera (Lophocebus, Papio, and Rungwecebus) relevant to our objective are represented, and both likelihood and Bayesian analyses of the resulting 5.9 and 6.5 kb of mitochondrial, autosomal, x-linked, and y-linked sequence recover a Papio + Rungwecebus clade with the highest possible support. Shimodaira–Hasegawa tests comparing these topologies to those resulting from searches constrained to recover nonmonophyly of the Papio + Rungwecebus clade strongly suggest a significantly worse explanation of the data (p = 0.005–0.007). Despite the fully resolved tree recovered in likelihood analyses and shown in Fig. 1F, the position of Theropithecus is poorly supported and the node uniting Lophocebus, Theropithecus, and Papio + Rungwecebus should therefore be considered an unresolved trichotomy. This is unsurprising in light of previous studies that have similarly failed to recover a consistent sister relationship between Theropithecus, Lophocebus, and Papio using multiple unlinked genes (e.g., Harris and Disotell, 1998). It is possible that Lophocebus is more closely related to the Papio + Rungwecebus clade than is Theropithecus, as there is some morphological evidence that Theropithecus is the most basally divergent taxon in this group (Gilbert and Rossie, 2007).

3.2. The generic status of Rungwecebus

The results obtained both here and in Davenport et al. (2006) strongly support shared common ancestry between Papio and Rungwecebus relative to all other living primates. A number of phenomena might account for any incorrect phylogenetic results obtained from a single gene or linkage partition, including incomplete lineage sorting, paralogy, homoplasy, methodological errors, or incorrect model choice. However, it is decidedly unlikely that one or more of these would consistently yield the same incorrect topology when multiple unlinked genes, sequenced by several investigators in different labs, are analyzed using a variety of phylogenetic inference methods. Furthermore, despite the skepticism raised by the description of Rungwecebus (Ehardt and Butynski, 2006; von Buol 2006), no evidence that would contradict the findings of Davenport et al. (2006) has yet been adduced. The original description of the kipunji relied on photographs, and its placement in the genus Lophocebus was based on a single morphological feature (non–contrasting black eyelids) and substrate preference (arboreality) (Jones et al., 2005). To our knowledge, the phylogenetic utility of eyelid color in primates has never been rigorously or explicitly investigated. In a cladistic study of papionin phylogeny based on morphology, only one of 36 characters (substrate preference) was found to be variable (and therefore potentially informative) among Lophocebus, Theropithecus, and Papio (Strasser and Delson, 1987). However, superficial morphology and substrate preference, as well as craniodental data (Collard and Wood, 2000; but see Gilbert and Rossie, 2007), have been notoriously misleading in attempts to reconstruct papionin relationships, particularly with respect to these genera (see Disotell et al., 1992). Moreover, since the publication of Jones et al. (2005), research in the field has shown that the kipunji, while still arboreal, does spend a larger proportion of its time on the forest floor than had previously been thought (Davenport et al., 2006, in press). We therefore conclude that the morphological evidence for the kipunji's inclusion in Lophocebus is moot. Rungwecebus and Lophocebus may superficially resemble one another because of their relatively small body size; i.e., their similarities are due to an allometric effect (see Gilbert, 2007; Gilbert and Rossie, 2007) and similar arboreal niche, rather than the result of a close phylogenetic relationship. Our original conclusion (Davenport et al., 2006) that Rungwecebus is the sister taxon to Papio is further supported by x-linked data and continues to be supported by mitochondrial and nuclear autosomal DNA, whether analyzed separately or together. Despite a threefold increase in the amount of y-linked data, relationships among Lophocebus, Theropithecus, Papio, and Rungwecebus are not resolved by this marker, although we reiterate that the majority of splits in the Bayesian analysis do not support a Lophocebus–Rungwecebus sister relationship.

The minimal (X-chromosome intergenic region, TSPY) to nonexistent (LPA, CD4) variation observed between Rungwecebus and Papio in the markers reported here, together with the consistent recovery of a sister relationship between the two, suggest that an argument might be made for placing the kipunji in the genus Papio. Significant morphological differences aside, this is not necessary (at least given the evidence adduced to date) to maintain monophyly of Papio and would require a thorough reevaluation of its diagnosis, with potentially vexing implications for naming and describing constituent fossil taxa. Furthermore, we can think of no compelling reason to do so. We continue to regard the kipunji's unique morphology relative to baboons (smaller body size, low level of sexual dimorphism, lack of an elongated rostrum, absence of deep fossae on the mandible and rostrum, and presence of deep suborbital fossae) as legitimate justification for not including it in the genus Papio (see Szalay and Delson, 1979, on Papio). The kipunji's unique vocalization, its predominantly arboreal nature, and strict preference for evergreen forest habitat, all further distance it from Papio (Davenport et al., 2006, in press). One possibility that may explain the strikingly low levels of genetic variation between Papio and Rungwecebus is past hybridization, which has been documented among other papionin genera (e.g., Jolly et al., 1997). Finally, several of the analyses and interpretations presented both here and in Davenport et al. (2006) have implicitly assumed monophyly of the genus Papio. Although baboon monophyly has been recovered in two recent studies (Newman et al., 2004; Wildman et al., 2004), both were conducted prior to the discovery and description of Rungwecebus. In light of growing evidence of close kinship between these genera, a reexamination of baboon interrelationships may be warranted.
Although debate on the supraspecific taxonomy of *Rungwecebus kipunji* may continue, we hope that it will not detract from the very pressing conservation issues facing this unquestionably threatened and unique taxon (Davenport et al., in press).

**Acknowledgments**

This research was supported by the National Science Foundation Grant DEB-0542725/0542532 to L.E.O. and E.J.S. E.J.S. was also supported by a Yale University Social Science Faculty Research Fund. TRBD was supported by the Wildlife Conservation Society. The Tanzania Wildlife Research Institute and Tanzania Commission for Science and Technology granted research permits to WTS and TRBD. DNA sequencing was conducted in the University of Alaska Fairbanks Institute of Arctic Biology Core Facility for Nucleic Acid Analysis with support from NSF EPSCoR Grant EPS-0346770. Analyses were performed on a parallel cluster administered by the UAF Biotechnology Computing Research Group, a core research resource supported by Grant RR016466 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). We thank Trina Roberts, Jack Fooden, Al Gardner, Karen and Dan Pritzker, an anonymous donor, Sofie Lamantia, Derek Wildman, Chris Gilbert, and an anonymous reviewer for their insightful discussion, helpful comments, and support.

**Appendix A. Supplementary data**


**References**


