

Interrogating Phylogenetic Discordance Resolves Deep Splits in the Rapid Radiation of Old World Fruit Bats (Chiroptera: Pteropodidae)

Nicolas Nesi, Georgia Tsagkogeorga, Susan M Tsang, Violaine Nicolas, Aude Lalis, Annette T Scanlon, Silke Riesle-Sbarbaro, Sigit Wiantoro, Alan T Hitch, Javier Juste, et al.

▶ **To cite this version:**

Nicolas Nesi, Georgia Tsagkogeorga, Susan M Tsang, Violaine Nicolas, Aude Lalis, et al.. Interrogating Phylogenetic Discordance Resolves Deep Splits in the Rapid Radiation of Old World Fruit Bats (Chiroptera: Pteropodidae). Systematic Biology, 2021, 10.1093/sysbio/syab013 . hal-03167896

HAL Id: hal-03167896

https://hal.sorbonne-universite.fr/hal-03167896

Submitted on 12 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Interrogating Phylogenetic Discordance Resolves Deep Splits in the Rapid Radiation of Old World Fruit Bats (Chiroptera: Pteropodidae)

NICOLAS NESI[1], GEORGIA TSAGKOGEORGA[1], SUSAN M. TSANG[2,3], VIOLAINE NICOLAS[4], AUDE LALIS[4], ANNETTE T. SCANLON[5], SILKE A. RIESLE-SBARBARO[6,7†], SIGIT WIANTORO[8], ALAN T. HITCH[9], JAVIER JUSTE[10], CORINNA A. PINZARI[11], FRANK J. BONACCORSO[12], CHRISTOPHER M. TODD[13], BURTON K. LIM[14], NANCY B. SIMMONS[2], MICHAEL R. McGOWEN[15] AND STEPHEN J. ROSSITER[1*]

[1]*School of Biological and Chemical Sciences, Queen Mary University of London, London, UK*

[2] *Department of Mammalogy, Division of Vertebrate Zoology, American Museum of Natural History, New York, USA*

[3] *Zoology Section, National Museum of Natural History, Manila, Philippines*

[4] *Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France*

[5] *School of Natural and Built Environments, University of South Australia, Mawson Lakes, SA, Australia*

[6] *Department of Veterinary Medicine, University of Cambridge, Cambridge, UK.*

[7] *Institute of Zoology, Zoological Society of London, London, UK.*

[8] *Museum Zoologicum Bogoriense, Research Center for Biology, Indonesian Institute of Sciences, Cibinong, Indonesia*

[9]*Department of Wildlife, Fish, and Conservation Biology, University of California Davis, California, USA*

[10] *Estación Biológica de Doñana (CSIC), Avda. Américo Vespucio, Sevilla, Spain*

[11] *University of Hawaii at Hilo, Hawaii, USA*

[12] *PO Box 417, Volcano, Hawaii, USA*

[13] *The Hawkesbury institute for the Environment, Western Sydney University, Australia*

[14] *Royal Ontario Museum, Toronto, ON M5S 2C6, Canada*

© The Author(s) 2021. Published by Oxford University Press, on behalf of the Society of Systematic Biologists.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

http://mc.manuscriptcentral.com/systbiol

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

[15] *Department of Vertebrate Zoology, Smithsonian National Museum of Natural History, Washington, DC, USA*

† Current address: Centre for Biological Threats and Special Pathogens, Robert Koch Institute, Berlin, Germany.

**Correspondence to be sent to: School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom; E-mail: s.j.rossiter@qmul.ac.uk*

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

ABSTRACT

The family Pteropodidae (Old World fruit bats) comprises >200 species distributed across the Old World tropics and subtropics. Most pteropodids feed on fruit, suggesting an early origin of frugivory, although several lineages have shifted to nectar-based diets. Pteropodids are of exceptional conservation concern with >50% of species considered threatened, yet the systematics of this group has long been debated, with uncertainty surrounding early splits attributed to an ancient rapid diversification. Resolving the relationships among the main pteropodid lineages is essential if we are to fully understand their evolutionary distinctiveness, and the extent to which these bats have transitioned to nectar-feeding. Here we generated orthologous sequences for >1400 nuclear protein-coding genes (2.8 million base-pairs) across 114 species from 43 genera of Old World fruit bats (57% and 96% of extant species- and genus-level diversity, respectively), and combined phylogenomic inference with filtering by information content to resolve systematic relationships among the major lineages. Concatenation and coalescent-based methods recovered three distinct backbone topologies that were not able to be reconciled by filtering via phylogenetic information content. Concordance analysis and gene genealogy interrogation shows that one topology is consistently the best supported, and that observed phylogenetic conflicts arise from both gene tree error and deep incomplete lineage sorting. In addition to resolving long-standing inconsistencies in the reported relationships among major lineages, we show that Old World fruit bats have likely undergone at least seven independent dietary transitions from frugivory to nectarivory. Finally, we use this phylogeny to identify and describe one new genus.

Running Head: PHYLOGENOMICS OF FRUIT BATS

Keywords: Chiroptera; coalescence; concordance; incomplete lineage sorting; nectar feeder; species tree; target enrichment

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

INTRODUCTION

The family Pteropodidae (Old World fruit bats) contains 45 genera with >200 extant species distributed across the Old World tropics and subtropics (Simmons 2005; Simmons and Cirranello 2019). Despite performing key roles in forest ecosystems as agents of seed dispersal and pollination (e.g., Hodgkinson et al. 2003), these bats face widespread threats that include habitat loss, hunting, and persecution (e.g., Fujita and Tuttle 1991). Over half of the recognised species are of conservation concern, with 13 (6.5%) listed as near threatened, 37 (18.5%) as vulnerable, 16 (8%) as endangered, eight (4%) as critically endangered, and four as recently extinct (IUCN Red List 2019). Past population declines have undoubtedly contributed to a scarcity of data for some species, several of which are known from just a few specimens.

Molecular phylogenies of pteropodids have revealed an early rapid radiation, with modern lineages arising in the late Oligocene *ca* 26-25 million years ago (Ma) (Teeling et al. 2005; Meredith et al. 2011; Amador et al. 2018). While phylogenetic analyses have clarified many relationships, and also overturned earlier morphology-based systematics (Bergmans 1997) in favour of newly proposed higher taxa (e.g., Giannini and Simmons 2005; Almeida et al. 2009, 2020; Nesi et al. 2013), many basal nodes remain unresolved. As a result, the relationships among the subfamilies and their relative timing of divergence remain poorly understood (Giannini and Simmons 2003, 2005; Almeida et al. 2011, 2020).

Unlike their closest relatives, all pteropodids are herbivorous, with most species feeding on fruit. Their early transition to a carbohydrate-rich plant-based diet is expected to have posed physiological and metabolic challenges, and several studies have uncovered molecular adaptations linked to diet (Liu et al. 2012; Sharma et al. 2018; Jiao et al. 2019; Wang et al. 2020). In this context, of particular interest is the evolution of nectarivory in several pteropodid lineages (see Freeman 1995), a dietary specialization that in mammals is also seen in two lineages of Neotropical leaf-nosed bats (family Phyllostomidae) (e.g., Potter et al. 2021) and the honey possum (Bradshaw and Bradshaw 2012). To date, few studies of pteropodid bats have focused on nectarivorous lineages, which are promising models for studying metabolic adaptations due to their ability to subsist on simple sugars without developing glucose toxicity (e.g., Gutiérrez-Guerrero et al. 2020; Potter et al. 2021). Resolving the relationships among pteropodid bat

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

lineages is essential for both inferring the number of independent origins of nectar-feeding, and for uncovering any lineage-specific molecular adaptations. Indeed, while molecular phylogenetic studies have overturned earlier proposals that nectar-feeding pteropodids constitute a single monophyletic subfamily (Hollar and Springer 1997, Kirsh and Lapointe 1997), these studies have been based on few genes, and have also omitted key taxa, including the putative nectarivore *Plerotes*.

The difficulties of resolving ancient rapid radiations are well known and can be attributed to two main factors (Whitfield and Lockhart 2007). First, small numbers of fixed substitutions among early lineages results in little historical signal, such that gene genealogies often differ due to ancestral polymorphisms and incomplete lineage sorting (ILS) (Maddison 1997; Oliver 2013). Second, over long periods of time, historical signal in sequence data may be overwritten by multiple substitutions, contributing to 'nonhistorical signal' (Ho and Jermiin 2004). Failure of phylogenetic models to capture true evolutionary processes - which are typically assumed to be stationary, reversible, and homogeneous ('SRH') - can lead to incorrect inference of topologies and branch lengths (Yang and Rannala 2012; Jermiin et al. 2017).

With advances in sequencing technologies, a common strategy to resolve long-standing problematic nodes has been to analyse larger datasets (e.g., Jarvis et al. 2014; Prum et al. 2015; McGowen et al. 2020), typically using concatenation (Gatesy and Springer 2014) and/or multispecies coalescent (MSC) based approaches (Edwards 2009; Liu et al. 2015; Jiang et al. 2020). The relative merits and disadvantages of both approaches have been strongly debated (Bryant and Hahn 2020). Generally, concatenation appears to perform poorly when incomplete lineage sorting is high (Degnan and Rosenberg 2006; Mendes and Hahn 2018), whereas summary tree inferences based on the MSC are prone to errors and higher levels of uncertainty in the presence of uninformative loci (Gatesy and Springer 2014; Xi et al. 2014; Springer and Gatesy 2016; Brown and Thomson 2017). Large multigene datasets can be further enhanced by filtering steps; for example, to reduce the effects of uninformative loci and other sources of error. Removing loci based on proportion of parsimony informative (PI) sites has been shown to improve the efficiency of summary species coalescence methods as well as increase topological concordance between inference methods (Hosner et al. 2016; Manthey et al. 2016; Blom et al.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

2017). Filtering can also be used to reduce systematic biases due to model misspecification, by removing model partitions that violate SRH assumptions (Jermiin et al. 2004; Duchene et al. 2017; Naser-Khdour et al. 2019; Jermiin et al. 2020). Such violations may lead to systematic error that, unlike stochastic error, cannot be remedied simply by increasing the size of a dataset (Ho and Jermiin 2004; Jermiin et al. 2004; Philippe et al. 2005; Kumar et al. 2012; Duchene et al. 2017). Where filtering steps fail to resolve recalcitrant nodes, the Gene Genealogy Interrogation (GGI) method (Arcila et al. 2017; Betancur-R et al. 2019) ranks competing topologies based on likelihood scores using the approximately unbiased (AU) test (Shimodaira 2002). Hypothetical topologies selected by a majority of genes are identified as the preferred resolution of the species tree for that node. This method is being recognised as a powerful tool for disentangling the relative contributions of information content, ILS, and data error in driving discordance.

Here we use target enrichment to capture >1400 nuclear protein-coding genes and the mitogenome (see Lemmon et al. 2012; Li et al. 2013) to perform the first phylogenomic reconstruction of Old World fruit bats. By applying node interrogation, we aim to resolve early splits in the context of a rapid radiation, and, in doing so, determine the higher-level systematic groups. As well as addressing long-standing inconsistencies in the reported relationships among major lineages, we aim to infer the number of dietary transitions between frugivory and nectarivory in this distinctly herbivorous family.

MATERIALS AND METHODS

Taxon Sampling, Marker Selection, DNA Extraction, and Sequencing

For generating new sequence data, we sampled 114 pteropodid species (135 individuals) from 43 genera, representing 57% and 96% of the respective total extant species- and genus-level diversity (IUCN Red List 2019). Two rare genera were not included due to unavailability of suitable material (*Aproteles* and *Mirimiri*). All tissue samples came from museum and other zoological collections (Supplementary Table S1 available on Dryad at https://doi.org/10.5061/dryad.7m0cfxpq5). We designed a new bait set comprising exons of 1545 putative protein-coding nuclear genes, selected to encompass a range of Gene Ontology (GO)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

categories (Supplementary Table S2 available on Dryad) using QuickGO (Binns et al. 2009). Orthologous coding sequences (CDSs) were downloaded from two high quality reference genomes of Pteropodidae, *Rousettus aegyptiacus* (Raegyp2.0, GCA_001466805.2, 169.2x) and *Pteropus alecto* (ASM32557v1, GCA_000325575.1, 110x). CDSs from *Rousettus aegyptiacus* were submitted to Arbor Biosciences (Formerly MYcroarray; Ann Arbor, Michigan, USA) for the development of 90 bp candidate RNA baits for each targeted gene. DNA extractions were conducted using DNeasy Blood and Tissue kits (Qiagen, USA) and Norgen's FFPE DNA Purification kit (Norgen) for one formalin-fixed sample (*Cynopterus horsfieldii*). Genomic library preparation, target enrichment, and sequencing preparation were performed by Arbor Biosciences. The libraries were prepared using the MYbaits-3 custom Kit (Arbor Biosciences, USA) and sequencing conducted using 150 bp paired-end reads on Illumina HiSeq 4000 platform.

Data Assembly and Sequence Alignment

For each nuclear protein-coding gene, Illumina reads were processed to construct multiple sequence alignments (MSA) using HybPhyloMaker 1.6.4 (Fer and Schmickl 2018). Off-target intronic sequences were trimmed to reduce alignment errors; however, off-target mitochondrial (mt) reads were recovered (see Picardi and Pesole 2012) using the pipeline HybPiper 1.3.1 (Johnson et al. 2016). We added published sequences for four mt loci (CO1, CYTB, 12S, 16S) from five additional species (*Aproteles bulmerae*, *Casinycteris campomaanensis*, *Epomophorus anselli*, *Epomops buettikoferi* and *Mirimiri acrodonta*) and two species (*Epomophorus minimus* and *Hypsignathus monstrosus*) for which we were unable to recover mitochondrial data using sequence capture (Supplementary Table S3 available on Dryad). Each nuclear and mt locus was individually aligned using MAFFT (Katoh and Standley 2013). We excluded nuclear loci with < 50% species completeness and > 60% missing data. We integrated orthologous sequences (NCBI Genbank) of two outgroup species, *Hipposideros armiger* and *Rhinolophus sinicus*, representatives of the closely-related families Hipposideridae and Rhinolophidae respectively. We visually inspected and corrected alignments for indels, incorrect reading frames, and premature stop codons in BioEdit 7.0.5 (Hall 1999), and retained 1455 CDSs as our baseline nuclear data matrix (referred to as 'all loci') for downstream analyses.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Phylogenetic Analyses and Tests for Model Misspecification

We inferred phylogenetic trees using maximum likelihood (ML) with IQ-TREE (Nguyen et al. 2015, Chernomor et al. 2016, Hoang et al. 2018) on our concatenated nuclear 'all loci' matrix as well as our complete mitochondrial dataset. We then reduced the 'all loci' dataset using a filtering strategy based on the parsimony informativeness (PI) content within each locus (Meiklejohn et al. 2016). This resulted in three additional modified CDS datasets including loci with ≥ 18% (PI18, n=1088 loci), 22% (PI22, n=728 loci) and 27% (PI27, n=341 loci) of PI sites, reflecting the respective first quartile value (17.3%), median (21.3%) and third quartile (25.8%) of PI variation in our baseline dataset. Thus, four nuclear CDS datasets and one concatenated mtDNA data set were used in subsequent analyses. Details of loci assembly, alignment and filtering strategy are available in the Supplementary Materials and Methods.

We also reduced our 'all loci' dataset by removing partitions that violated SRH assumptions. First, best-fit partitioning schemes and models of molecular evolution were determined with PartitionFinder 2.1.1 (Lanfear et al. 2016) using a partition-by-locus scheme for nuclear data sets and partition by locus and codon position for the mt dataset. Then, for each partition in each nuclear data set ('all loci', PI18, PI22, PI27), we tested for violation of the SRH assumptions. While new non-SRH models have been recently developed to account for data that have evolved under non-SRH conditions (e.g., Zou et al. 2012; Woodhams et al. 2015), they are not easy to use and require large computation power (e.g., Betancur-R et al. 2013). An alternative method is to identify data that evolved under non-SRH conditions before performing phylogenetic reconstruction. We used matched-pairs tests of homogeneity (Jermiin et al. 2017) implemented in IQ-TREE 2.0-rc1 (Naser-Khdour et al. 2019). Partitions rejecting the SRH assumptions ($P < 0.05$) were discarded before performing ML phylogenetic inferences.

In addition to concatenation analysis, coalescent-based species tree summary methods were performed with ASTRAL-III (Mirarab et al. 2014; Zhang et al. 2018) and MP-EST (Liu et al. 2010) on all gene trees, as well as the reduced PI datasets. Individual nuclear ML gene trees, based on the full coding sequence, inferred with IQ-TREE were used as inputs. Branch support as local posterior probabilities (LPP; Sayyari and Mirarab 2016), and quartet support for the main topology and alternative topologies were calculated with ASTRAL. For MP-EST analysis, 10

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

independent searches for the maximum pseudo-likelihood tree were performed. We estimated topological concordances among our concatenated and multispecies coalescent trees by performing weighted Shimodaira-Hasegawa (wSH) and Approximately Unbiased (AU) tests in IQ-TREE with 1,000 RELL replicates (Kishino et al. 1990). Topologies were imported in Dendroscope 3.5.10 (Huson and Scornavacca 2012) and compared to each other as a tanglegram.

Tree Discordance and Gene Genealogy Interrogation

To summarise discordance among ML gene trees, we estimated bipartitions based on topology using PhyParts (Smith et al. 2015) with the function phypartspiecharts.py (https://github.com/mossmatters). Internode Certainty All (ICA) values were calculated by comparing the individual ML gene trees to both the estimated concatenated 'all loci' tree and the ASTRAL species tree. We also quantified the gene concordance factor (gCF) and site concordance factor (sCF) in IQ-TREE 1.7-beta7 (Minh et al. 2020). The concatenated 'all loci' tree was used as the reference tree with 100 quartets computed to estimate sCF values. Individual genes in concordance with specific branches of the reference tree were identified using the --cf-verbose option. Conflict was visualised using a cloudogram generated via the *DensiTree* function in the R package *Phangron* v.2.5.5 (Schliep 2011), with CDS alignments reduced to 15 species from the main subfamilies of the Pteropodidae.

To further interrogate discordance among recalcitrant backbone relationships, we tested whether observed differences among gene tree topologies were likely due to gene estimation error (limited phylogenetic signal content), or to biological conflict arising from ILS. For this, we applied a Gene Genealogy Interrogation (GGI) approach to the key backbone nodes identified in our different analyses (nodes B-D; Figure 1). We tested 105 alternative rooted topologies that only differed with respect to the relationships among the five main lineages of Pteropodidae: (1) Rousettinae + Eidolinae, (2) Cynopterinae, (3) Harpyionycterinae, (4) Macroglossusinae, and (5) Nyctimeninae + Pteropodinae + Notopterisinae (Bergmans 1997; Almeida et al. 2020). We constrained the ancestral node of each of these five clades in all analyses. In addition, we repeated this analysis for a topology in which these lineages formed a polytomy. We ran ML searches on each of our 1455 loci for these 105 hypotheses (152,775 constrained ML searches) (Fig. 2). To assess topology support, we compared site wise likelihood scores across all 105 trees

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

with the AU test in IQ-TREE, and ranked trees by *P*-value. Finally, we conducted summary species-tree analyses with ASTRAL using as input all rank 1 constrained trees (1455 gene trees), as well as with only the set of 'rank 1' trees significantly better than the alternatives ($P < 0.05$; 54 gene trees).

Ancestral State Reconstruction of Diet

To reconstruct the evolutionary history of nectar feeding, we performed ancestral state reconstructions (ASR) of diet (frugivory vs. nectarivory) based on Topology 1 in which long-standing problematic nodes were resolved (see Result Fig. 1). For classifying nectar-feeding lineages, we followed Freeman (1995) supplemented with Bergmans (1989), Sheherazade et al. (2019) and Almeida et al. (2020). We used a Bayesian stochastic character mapping approach in the R package *phytools* version 0.7-47 (Revell 2012) with the function *make.simmap* (Bollback 2006). Akaike information criteria (AIC) were used to assess which of the three available hierarchical models (ER, SYM, and ARD) was the most appropriate for our data. Analyses were then run using the best model for 10,000 simulation replicates, with all other options set as default.

RESULTS AND DISCUSSION

Using sequence capture, we successfully recovered an average of 1536.9 loci (99.4%) per sample (n = 135). Following manual examination of our alignments, we discarded 90 problematic loci, yielding a final dataset of 1455 CDSs spanning 2,791,426 base-pairs ('all loci' dataset). Filtering the 'all loci' dataset on parsimony informativeness reduced the numbers of CDSs to 1088, 728 and 341 based on PI thresholds of ≥18%, 22% and 26%, respectively. Using PartitionFinder 2 we also divided the 1455 loci into 1028 partitions, of which 473 partitions met SRH assumptions (555 failed) based on a matched-pairs test of symmetry (MPTS) (Table 1) and we included in our 'SRH loci' dataset. Finally, we recovered sequences from up to 15 mitochondrial loci, totalling 13,950 bp, with a mean of 106 taxa (min = 65 and max = 131) (Table 1; Supplementary Tables S1, S4-S6 available on Dryad).

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Resolution of the Basal Pteropodid Relationships

All phylogenetic trees inferred using concatenation recovered the majority of recognised subfamilies and tribes with high bootstrap support (>90), with the exception of the tribe Rousettini (see *Implications for Systematics*). At the same time, we found that the topology reconstructed from the 'SRH loci' dataset (Topology 1) differed from that of the 'all loci' dataset (Topology 2) with respect to the position of Macroglossusinae, suggesting that SRH violation may affect phylogenetic inference (Figs. 1-2; Supplementary Fig. S1 available on Dryad). However, one alternative explanation for this difference is that removing loci that failed the matched-pairs test of symmetry may lead to improved model adequacy and accuracy. PI filtering schemes applied to both datasets also supported Topologies 1 and 2 (Fig. 2; Supplementary Figs. S2 to S4 available on Dryad). In contrast to concatenation, coalescent-based summary analyses of the complete dataset using both ASTRAL and MP-EST recovered another distinct topology (Topology 3), although Topology 2 was recovered with ASTRAL under the PI18 and PI27 filtering schemes (Fig. 2; Table 1; Supplementary Figs. S5-S12 available on Dryad). Finally, phylogenetic reconstruction based on concatenated mitochondrial genes produced a fourth topology with most major subfamilies retained (Fig. 2 Topology 4; Supplementary Fig. S13 available on Dryad). To assess the relative statistical support for these four alternative topologies, we performed an AU test using the complete concatenated dataset and found that Topology 1 was significantly most likely ($P < 0.05$, AU test) (Supplementary Table S7 available on Dryad). A potential driver of the observed topological discordance at the base of the tree is the presence of ILS. Thus, our finding that concatenation outperformed coalescent summary methods was somewhat unexpected given that the latter are often considered more robust than concatenation in the presence of high rates of ILS. However, coalescent methods have also been shown to be susceptible to the effects of gene tree estimation error, and simulations indicate that filtering to reduce this error is only effective in cases of low to moderate ILS (Molloy and Warnow 2018). In severe cases of ILS, short branches can produce gene trees that conflict with the species tree more frequently than they agree with it, leading to a so-called "anomaly zone" (Degnan and Rosenberg 2006).

To quantify the degree of topological discordance with respect to the backbone nodes A-F (Figure 2), we applied two analyses of concordance based on Topology 1 and found comparable

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

patterns of weak support. Specifically, concordance analysis using PhyParts showed low percentages of gene trees recovering nodes A-F (ICA: 0.09-0.10; Fig. 1; Supplementary Fig. S14 and Table S8 available on Dryad). We repeated the PhyParts concordance analysis using Topology 2 and Topology 3, and found that a similar number of genes supported these backbone relationships (Figure 2; Supplementary Figs. S15-S16 available on Dryad). Concordance analyses performed on Topology 1 using IQ-TREE also indicated that nodes A-F were supported by few gene trees (gCF=10.55-25.57; mean gCF=15.95) and a low proportion of parsimony-informative sites (sCF=31.14-40.41; mean sCF=35.68), despite very high bootstrap support (mean BS=100) (Figure 1). Furthermore, no alternative to Topology 1 emerged from the inferred gene trees, likely reflecting limited phylogenetic signal in individual loci (i.e. low gDF1 and gDF2 values; Supplementary Table S9 available on Dryad).

To evaluate the alternative topologies in greater depth, we reordered the branching pattern of the conflicting nodes B-D relative to the other early splits, and identified and compared support for 105 possible topologies using Gene Genealogy Interrogation (GGI) (Supplementary Fig. S17 available on Dryad). We found that each of the 105 hypotheses was supported by at least one gene, and a large number of alternative topologies (n=36) received reasonable support (>15 loci). While no topology showed strong support above all others, Topology 1 was consistently the best supported based on rank of p-values (46 of 1455 loci; 3.16%). In contrast, Topologies 2 and 3 were supported by 39 (2.68%) and 31 (2.13%) loci respectively, although both of these were outperformed by alternatives. For each of these three topologies, only a subset of loci was significantly better than the second ranked locus (5, 4, and 1 loci, respectively). We also compared each of these topologies to a null hypothesis in which the five major lineages were placed as a polytomy at the crown node of the Pteropodidae (Figure 2b; also see Alda et al. 2019), and found that none of the loci supported a resolved topology over the polytomy. On the other hand, more than a third of loci supported the polytomy over one of the alternatives (36%; 524 out of 1455 loci), with 74 showing significant support (Supplementary Table S10 available on Dryad). Finally, we performed coalescent analysis in ASTRAL using both the set of all gene trees of rank 1 (n = 1455), and a subset of these that showed statistical support above rank 2 ($P < 0.05$; n = 54). In both cases, the inferred tree was consistent with Topology 1 with the exception

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

of the relationship of *Notopteris*, which in the latter analysis was recovered as a sister taxon to the Nyctimeninae (Supplementary Figs. S18 and S19 available on Dryad).

By revealing that multiple evolutionary scenarios receive some degree of support, these results highlight the long-standing difficulties that researchers have faced in resolving the early splits in this group of mammals. Our analyses of concordance and node interrogation suggest that such uncertainty likely stems from two main causes. First, many single genes (around a third of those examined) appear to contain too little information for resolving ancestral nodes separated by short branches, a pattern that characterises the pteropodid tree. Second, for a larger proportion of genes sampled, we find that incongruence among loci with respect to the positions of backbone nodes falls into several discrete branching patterns, consistent with extensive ILS (Suh et al. 2015; Arcila et al. 2017; Alda et al. 2019). If ILS is indeed a major cause of this discordance, we might expect Topology 3 recovered with ASTRAL to be the best supported tree. Yet when coalescence-based analysis was repeated using only either the top ranked topologies in GGI, or the set of gene trees that performed statistically better than alternatives, we recovered the relationships among nodes A-F seen in Topology 1. Thus, our findings differ from those of Alda et al. (2019), who found that coalescence-based reanalysis using statistically-supported gene trees recovered the main coalescent-based species tree. We suggest that in the case of basal pteropodid lineages, while ILS is a major cause of discordance, gene tree error may have had a greater effect in misleading coalescence-based analyses. In most cases, Topology 1 remains the most highly-supported bifurcating tree.

Diversification and Diet

Our analyses of a large molecular dataset confirm the early rapid divergence of the major fruit bat lineages. The early diversification of this group has previously been attributed to a shift in feeding habits coupled with the potential loss of laryngeal echolocation (Amador et al. 2018; Almeida et al. 2020), although such scenarios are speculation given the highly depauperate fossil record (e.g., Gunnell and Manthi 2018). The ultimate drivers of this early diversification remain obscure, particularly since it appears to have occurred ~25 to 20 million years ago, long after the radiation of angiosperms. One scenario is that the phenotypic changes associated with a transition to frugivory, combined with their ability to fly, means that early pteropodids would have been

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

able to realise the ecological opportunities presented by an unevenly distributed and ephemeral resource such as fruiting trees. Indeed, frugivory was previously linked to high rates of diversification in primates (Gómez and Verdú 2012) and phyllostomid bats (Rojas et al. 2012), and it is noteworthy that these groups all exhibit visual and/or olfactory adaptations for detecting fruit (Zhao et al. 2009; Hayden et al. 2014; Wang et al. 2020).

Within the Pteropodidae, a best-fitting model (ER) of diet evolution supported seven independent transitions from frugivory to nectarivory (Fig. 3; Supplementary Figure S21 available on Dryad). Nectar-feeding is relatively rare in vertebrates, and our understanding of its evolution in pteropodids has changed markedly with improved resolution of systematic relationships (e.g., Hollar and Springer 1997; Kirsh and Lapointe 1997; Alvarez et al. 1999; Juste et al. 1999). Historically, Andersen (1912) grouped most nectarivorous Old World fruit bats into the distinct subfamily Macroglossinae, comprising five genera (*Eonycteris*, *Macroglossus*, *Melonycteris*, *Notopteris* and *Syconycteris*) from Indo-Australasia plus a sixth genus (*Megaloglossus*) endemic to Africa. Bergmans (1989) subsequently proposed that the little-known African genus *Plerotes* feeds on nectar based on evaluation of its palate and dentition, and Almeida et al. (2020) split *Nesonycteris* from *Melonycteris.* Finally, the Australian species *Pteropus scapulatus* was proposed to be nectarivorous based on behavioral data, and other *Pteropus* species have also been observed to either feed on nectar occasionally (e.g., Aziz et al. 2017; Sheherazade et al. 2019) or to possess morphological features that are consistent with flower-visiting and nectarivory (Almeida et al. 2014).

In light of molecular phylogenies, the shared characters that led nectar-feeding bats to be considered a single clade - including reduced cheek teeth, elongated rostra, and elongated brushy tongues (also see Marshall 1983, Courts 1998) - are now recognised as convergent traits (e.g., Hollar and Springer 1997; Almeida et al. 2011). Yet while nectarivory is often considered to be a highly-derived specialisation, we find that several lineages of nectar-feeding pteropodids appear to have long evolutionary histories, branching off early in their respective subfamilies. Moreover, nectar-feeding bats are also seen in four of the eight major pteropodid subfamilies in our phylogeny, with species affiliations implying at least two origins in Africa, and up to five in Asia. Currently it is not known whether such independent origins have involved convergent changes at

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

the molecular level, however, the well-supported topology presented here will provide the requisite phylogenetic framework for determining the likely evolutionary position of molecular adaptations that might have contributed to the transitions to fruit- and nectar-based diets, and subsequent diversification of the Pteropodidae.

Implications for Systematics

The inferred relationships among the main pteropodid lineages differ substantially from published phylogenetic trees (e.g., Giannini and Simmons 2005; Almeida et al. 2011; 2020; Fig. 3). Given that these earlier analyses were based on eight loci (~8000 bp), representing <1% of the dataset generated here (~2,791,500 bp), we attribute the greater resolution and support for basal nodes in our trees to the much larger volume of data.

We found that the tribes Pteralopini and Melonycterini grouped together (BS = 100, Fig. 1; Topologies 1-3) and not with other genera of the Pteropodini with which they have been traditionally allied (Bergmans 1997; Almeida et al. 2020). Of particular interest was the phylogenetic position of the genus *Eidolon*, which has long been problematic and unstable, even across molecular studies (e.g., Giannini and Simmons 2005, Almeida et al. 2011, Almeida et al. 2016, Hassanin et al. 2016). *Eidolon* is superficially similar to the other large members of the Pteropodini, yet a sister relationship between *Eidolon* and Pteropodini has never been strongly supported by either morphological characters (e.g., Giannini and Simmons 2005) or mtDNA (e.g., Hassanin et al. 2016). Our results indicate that *Eidolon* is in fact a sister taxon to the subfamily Rousettinae, with strong bootstrap support in all ML concatenated analyses based on nuclear DNA data (node A, Fig. 1), albeit with low concordance support across individual gene trees (n = 201, ICA = 0.09; Fig. 1). The exception was the mitochondrial analysis, which placed *Eidolon* at the base of the pteropodid tree, likely reflecting the lack of power with this marker (Supplementary Fig. S20 available on Dryad). *Eidolon* was previously included in Rousettinae by Bergmans (1997) based on morphological similarity, but Almeida et al. (2016) subsequently reclassified *Eidolon* in a new subfamily Eidolinae, which we also retain due to this taxon's distinctiveness. By moving *Eidolon* to a grouping with Rousettinae, we place all continental African pteropodids in one well-supported clade (Eidolinae + Rousettinae), although several

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

species within the Rousettinae (*Eonycteris*, species within *Rousettus*) are currently restricted to the Asian tropics.

Unexpectedly, and in disagreement with previous systematic assessments, the tribe Rousettini (*sensu* Almeida et al. 2016) was found to be paraphyletic. Specifically, we found that *Rousettus celebensis* - included here in a molecular analysis for the first time - did not group with other *Rousettus* species, but instead appeared as a sister taxon to a clade containing the tribes Stenonycterini, Myonycterini, Plerotini and Epomophorini. This arrangement, which was strongly supported and recovered in the majority of individual gene trees (BS = 100, ICA = 0.38; Fig. 1), is intriguing because it suggests that *R. celebensis* represents a distinct lineage from Wallacea that has strong affiliations with African lineages. This is the second time that an endemic Sulawesi species has been removed from *Rousettus*, with *Boneia* previously elevated to a separate genus. Although Wallacea is a hotspot of endemism for fruit bats, with at least three other genera known only from Sulawesi and its offshore islands (*Boneia*, *Neopteryx*, and *Thoopterus*), these genera are not sister taxa (also see Almeida et al. 2020), implying that pteropodids have likely crossed Wallace's line several times in their evolutionary history.

The systematics of *Rousettus* has undergone several revisions, largely following genetic analyses. *Boneia*, *Lissonycteris* (now synonymized to *Myonycteris*; Nesi et al. 2013), and *Stenonycteris* were all previously classified as subgenera of *Rousettus* on the basis of specific morphological characters (Andersen 1912; Bergmans and Rozendaal 1988; Koopman 1994), but were removed subsequently (Bergmans 1997; Nesi et al. 2013). To our knowledge, specimens of *R. celebensis* described to date do not show any morphological characters that distinguish this taxon from the other *Rousettus* species (Andersen 1912; Bergmans and Rozendaal 1988), and thus in-depth morphological re-examination is warranted in light of our results. In this respect, the future inclusion of *R. linduensis* - another Sulawesi endemic that was recently described (Maryanto and Yani 2003) - and the only *Rousettus* missing from our study - will be important.

Taxonomy

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

The results of this study have implications for taxonomy, with *R. celebensis* specimen MSB 93154 (Supplementary Table S1) found to represent a new monotypic genus of bat, described below for the first time.

Pilonycteris gen. nov. Nesi, Tsang, Simmons, McGowen & Rossiter

Rousettus Gray 1821

Etymology: Named from the combination of the Latin *pilo*, hairy, from its distinctly longer fur, and the Greek *nycteris*, bat. The gender of the genus is masculine.

Type species: *Rousettus celebensis* Andersen, 1907 (holotype BMNH 1897.1.2.8; type locality: Mt Masarang, 3500 ft, Celebes).

Diagnosis: Andersen (1912) noted that *Pilonycteris celebensis* is distinguishable from sympatric species *Rousettus amplexicaudatus*, *R. brachyotis* (now *R. amplexicaudatus brachyotis*), and *R. leschenaultii* by having a bony palate that is narrower posteriorly, last upper and lower premolar and molars that are unusually narrow, fur that is brighter, longer, and richer in colour, a notopatagium (= uropatagium) that is partly or wholly hairy, and a generally smaller size but with digits proportionally longer than other sympatric *Rousettus* species. Rookmaaker and Bergmans (1981) further indicated that *Pilonycteris celebensis* differs in hairiness, especially the absence of reduction of hair in the neck region, presence of a hairy notopatagium, and the greater length of its fur, which differentiates it from both *Rousettus amplexicaudatus* and *R. leschenaultii*. *Pilonycteris celebensis* also has a relatively longer rostrum, longer lower and upper tooth rows, and a smaller average distance between the two M^2. Particularly distinctive characters involve dimensions of the cheek teeth; *P. celebensis* has a longer P^4, narrower M^1, shorter and narrower M^2, longer and narrower M_1, narrower M_2, and an M_3 that is sometimes shorter and narrower than seen in *Rousettus amplexicaudatus* spp. (see Table 6 in Rookmaaker and Bergmans 1981).

Contents: Monotypic, comprising only *P. celebensis* (Andersen, 1907). May also contain *Rousettus linduensis*, but further analyses will be needed to confirm this.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

AUTHOR'S CONTRIBUTIONS

N.N. and S.J.R. conceived of the study; V.N., A.L., A.T.S., S.A.R., S.W., A.T.H., J.J., C.A.P., F.J.B., C.M.T., B.K.L., and N.B.S supplied tissues; N.N. and G.T. compiled genes and designed baits with the assistance of MYcroarray (now Arbor Biosciences); N.N. performed laboratory work; N.N. performed all bioinformatics and phylogenetic analyses; N.N., M.R.M. and S.J.R. wrote the paper with contributions from the other authors.

SUPPLEMENTARY MATERIAL

Detailed methods, supplementary tables and figures reported in this study are available in the Supplementary Material. Data sets, alignment files, and tree files are provided in the Dryad Digital Repository (https://doi.org/10.5061/dryad.7m0cfxpq5).

CONFLICT OF INTEREST

All authors declare no conflict of interest.

FUNDING

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 753367 awarded to NN, with additional support from a European Research Council Starting Grant (310482) awarded SJR, and Smithsonian NMNH Start-up funds to MRM.

ACKNOWLEDGMENTS

We thank the many museum curators and collection managers for their service and assistance with specimen samples: Scott Ginn and Sandy Ingleby at the Australian Museum (Sidney, Australia); Julie Feinstein at the Ambrose Monell Cryo Collection (New York, USA); Neil Duncan at the American Museum of Natural History (New York, USA); Alex Drew and Leo Joseph at the Australian National Wildlife Collection CSIRO (Australia); Suzanne McLaren at

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

the Carnegie Museum of Natural History (Pittsburgh, USA); Leigh Richards at the Durban Museum (Durban, South Africa); Adam Ferguson, Lawrence Heaney and Bruce Patterson at the Field Museum of Natural History (Chicago, USA); Malcolm Pearch at the Harrison Institute, Centre for Biodiversity research (Sevenoaks, United Kingdom); Maria Eifler at the Kansas University, Biodiversity Institute and Natural History Museum (Lawrence, USA); Jean-Marc Pons at the Museum Nationale d'Histoire Naturelle (Paris, France); Mariel Campbell and Joseph Cook at the Museum of Southwestern Biology, University of New Mexico (Albuquerque, USA); Andrew Engilis and Irene Engilis at the Museum of Wildlife and Fish Biology (Davis, USA); Louise Tomsett at the Natural History Museum (London, United Kingdom); Maria Josefa Veluz at the National Museum of the Philippines (Manila, Philippines); Heike Kappes and Christian Printzen at the Senckenberg Museum (Frankfurt, Germany); Heath Garner and Caleb Phillips at the Museum of Texas Tech University (Lubbock, USA); Link Olson at the University of Alaska, Museum of the North (Fairbanks, USA); Darrin Lunde at the Smithsonian National Museum of Natural History (Washington, USA); Paula Holahan and Laura Monahan at the University of Wisconsin Zoological Museum (Madison, USA); Rebecca Bray and Kevin Travouillon at the Western Australian Museum (Perth, Australia); Jan Decher at the Zoological Research Museum Alexander Koenig (Bonn, Germany); and Pepjin Kamminga at the Naturalis Biodiversity Center (Leiden, Netherlands). We also thank Holger Meinig and Steve Goodman for sending us tissue samples from specimens collected on the field and Sophie Petit for assistance with obtaining samples. This research utilized Queen Mary's Apocrita HPC Facility, supported by QMUL Research-IT (http://doi.org/10.5281/zenodo.438045) and the CIPRES Science Gateway (Miller et al. 2010).

REFERENCES

Alda F., Tagliacollo V.A., Bernt M.J., Waltz B.T., Ludt W.B., Faircloth B.C., Alfaro M.E., Albert J.S., Chakrabarty P. 2019. Resolving deep nodes in an ancient radiation of neotropical fishes in the presence of conflicting signals from incomplete lineage sorting. Syst. Biol. 68:573-593.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Almeida F.C., Giannini N.P., DeSalle R., Simmons, N.B. 2009. The phylogenetic relationships of cynopterine fruit bats (Chiroptera: Pteropodidae: Cynopterinae). Mol. Phylogenet. Evol. 53: 772-783.

Almeida, F.C., Giannini, N.P., DeSalle, R., Simmons N.B. 2011. Evolutionary relationships of the old world fruit bats (Chiroptera, Pteropodidae): another star phylogeny? BMC Evol. Biol. 11: 281.

Almeida F.C., Giannini N.P., Simmons N.B., Helgen K.M. 2014. Each flying fox on its own branch: a phylogenetic tree for *Pteropus* and related genera (Chiroptera: Pteropodidae). Mol. Phylogenet. Evol. 77:83-95.

Almeida F.C., Giannini N.P., Simmons N.B. 2016. The evolutionary history of the African fruit bats (Chiroptera: Pteropodidae). Acta Chiropterol. 18:73-108.

Almeida F.C., Simmons N.B., Giannini N.P. 2020. A species-level phylogeny of Old World fruit bats with a new higher-level classification of the family Pteropodidae. Am. Mus. Novit. 2020:1-24.

Alvarez Y., Juste J., Tabares E., Garrido-Pertierra A., Ibánez C., Bautista J.M. 1999. Molecular phylogeny and morphological homoplasy in fruitbats. Mol. Biol. Evol. 16:1061-1067.

Amador L.I., Arévalo L.R., Almeida F.C., Catalano S.A., Giannini N.P. 2018. Bat Systematics in the light of unconstrained analyses of a comprehensive molecular supermatrix. J. Mamm. Evol. 25:37-70.

Andersen K. 1912. Catalogue of the Chiroptera in the collection of the British Museum. Megachiroptera 1:1-854.

Arcila D., Ortí G., Vari R., Armbruster J.W., Stiassny M.L., Ko K.D., Sabaj M.H., Lundberg J., Revell L.J., Betancur-R. R. 2017. Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. Nat. Ecol. Evol. 1:0020.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Aziz S.A., Clements G.R., McConkey K.R., Sritongchuay T., Pathil S., Abu Yazid M.N.H., Campos-Arceiz A., Forget P-M., Bumrungsri, S. 2017. Pollination by the locally endangered island flying fox (*Pteropus hypomelanus*) enhances fruit production of the economically important durian (*Durio zibethinus*). Ecol. Evol. 7:8670-8684.

Bergmans W. 1989. Taxonomy and biogeography of African fruit bats (Mammalia, Megachiroptera). 2. The genera *Micropteropus* Matschie, 1899, *Epomops* Gray, 1870, *Hypsignathus* H. Allen, 1861, *Nanonycteris* Matschie, 1899, and *Plerotes* Andersen, 1910. Beaufortia. 39:89-153.

Bergmans W. 1997. Taxonomy and biogeography of African fruit bats (Mammalia, Megachiroptera). 5. The genera *Lissonycteris* Andersen, 1912, *Myonycteris* Matschie, 1899 and *Megaloglossus* Pagenstecher, 1885; general remarks and conclusions; annex: key to all species. Beaufortia. 47:11-90.

Bergmans W., Rozendaal F.G. 1988. Notes on collections of fruit bats from Sulawesi and some off-lying islands (Mammalia, Megachiroptera). Zoologische Verhandelingen. 248:1-74.

Betancur-R. R., Li C., Munroe T.A., Ballesteros, J.A., Ortí, G. 2013. Addressing gene tree discordance and non-stationarity to resolve a multi-locus phylogeny of the flatfishes (Teleostei: Pleuronectiformes). Syst. Biol. 62:763-785.

Betancur-R. R., Arcila D., Vari R.P., Hughes L.C., Oliveira C., Sabaj M.H., Ortí G. 2019. Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: The monophyly of characiform fishes. Evolution. 73:329-345.

Binns D., Dimmer E., Huntley R., Barrell D., O'Donovan C., Apweiler R. 2009. QuickGO: a web-based tool for Gene Ontology searching. Bioinformatics. 25:3045-3046.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Blom M.P., Bragg J.G., Potter S., Moritz C. 2017. Accounting for uncertainty in gene tree estimation: summary-coalescent species tree inference in a challenging radiation of Australian lizards. Syst. Biol. 66:352-366.

Bollback J.P. 2006. SIMMAP: stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics. 7:88.

Bradshaw D., Bradshaw F. 2012. The physiology of the honey possum, *Tarsipes rostratus*, a small marsupial with a suite of highly specialised characters: a review. J. Comp. Physiology B. 182:469-489.

Brown J. M., Thomson R.C. 2017. Bayes factors unmask highly variable information content, bias, and extreme influence in phylogenomic analyses. Syst. Biol. 66:517-530.

Bryant D., Hahn M.W. 2020. The concatenation question. In: Scornavacca, C., Delsuc, F., Galtier, N., editors. Phylogenetics in the genomic era. p. 3.4:1–3.4:23. No commercial publisher, Authors' open access book.

Chernomor O., Von Haeseler A., Minh B.Q. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Syst. Biol. 65:997-1008.

Courts S.E. 1998. Dietary strategies of Old World fruit bats (Megachiroptera, Pteropodidae): how do they obtain sufficient protein? Mammal Rev. 28:185-194.

Degnan J. H., Rosenberg N.A. 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2:e68.

Duchene D.A., Duchene S., Ho S. 2017. New statistical criteria detect phylogenetic bias caused by compositional heterogeneity. Mol. Biol. Evol. 34:1529-1534.

Edwards S.V., Liu L., Pearl D.K. 2007. High-resolution species trees without concatenation. Proc. R. Soc. B. 104:5936-5941.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1-19.

Fér T., Schmickl R.E. 2018. HybPhyloMaker: Target Enrichment Data Analysis From Raw Reads to Species Trees. Evol. Bioinform. 14:1176934317742613.

Freeman, P.W. 1995. Nectarivorous feeding mechanisms in bats. Biol. J. Linn. Soc. 56:439-463.

Fujita M.S., Tuttle M.D. 1991. Flying foxes (Chiroptera: Pteropodidae): threatened animals of key ecological and economic importance. Conserv. Biol. 5:455-463.

Gatesy J., Springer M.S. 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80:231-266.

Giannini N.P., Simmons N.B. 2003. A phylogeny of megachiropteran bats (Mammalia: Chiroptera: Pteropodidae) based on direct optimization analysis of one nuclear and four mitochondrial genes. Cladistics. 19:496-511.

Giannini N.P., Simmons N.B. 2005. Conflict and congruence in a combined DNA–morphology analysis of megachiropteran bat relationships (Mammalia: Chiroptera: Pteropodidae). Cladistics. 21:411-437.

Giannini N.P., Almeida F.C., Simmons N.B., Helgen K.M. 2008. The systematic position of *Pteropus leucopterus* and its bearing on the monophyly and relationships of *Pteropus* (Chiroptera: Pteropodidae). Acta Chiropterol. 10:11-20.

Gómez J.M., Verdú M. 2012. Mutualism with plants drives primate diversification. Syst. Biol. 61:567-577.

Gunnell G.F., Manthi F.K. 2018. Pliocene bats (Chiroptera) from Kanapoi, Turkana Basin, Kenya. J. Hum. Evol. 140:102440.

Gutiérrez-Guerrero Y.T., Ibarra-Laclette E., Martínez del Río C., Barrera-Redondo J., Rebollar

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

E.A., Ortega J., León-Paniagua L., Urrutia A., Aguirre-Planter E., Eguiarte L.E. 2020. Genomic consequences of dietary diversification and parallel evolution due to nectarivory in leaf-nosed bats. Gigascience 9.

Hassanin A., Nesi N., Marin J., Kadjo B., Pourrut X., Leroy É., Gembu G.-C., Akawa P., Ngoagouni C., Nakouné E., Ruedi M., Tshikung D., Shongo C., Bonillo C. 2016. Comparative phylogeography of African fruit bats (Chiroptera, Pteropodidae) provide new insights into the outbreak of Ebola virus disease in West Africa, 2014–2016. C. R. Biol. 339:517-528.

Hayden S., Bekaert M., Goodbla A., Murphy W.J., Dávalos L.M., Teeling E.C. 2014. A cluster of olfactory receptor genes linked to frugivory in bats. Mol. Biol. Evol. 31:917-927.

Ho S.Y., Jermiin L.S. 2004. Tracing the decay of the historical signal in biological sequence data. Syst. Biol. 53:623-637.

Hoang D.T., Chernomor O., von Haeseler A., Minh B.Q., Vinh L.S. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35:518-522.

Hodgkison R., Balding S.T., Akbar Z., Kunz T.H. 2003. Roosting ecology and social organization of the spotted-winged fruit bat, *Balionycteris maculata* (Chiroptera: Pteropodidae), in a Malaysian lowland dipterocarp forest. J. Trop. Ecol. 19:667-676.

Hollar L.J., Springer M.S. 1997. Old World fruitbat phylogeny: evidence for convergent evolution and an endemic African clade. Proc. Nat. Acad. Sci. USA 94:5716-5721.

Hosner P.A., Faircloth B.C., Glenn T.C., Braun E.L., Kimball R.T. 2016. Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). Mol. Biol. Evol. 33:1110-1125.

Huson D.H., Scornavacca C. 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61:1061-1067.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

IUCN 2019. The IUCN Red List of Threatened Species. Version 2019-1. http://www.iucnredlist.org. Downloaded on 21 March 2019.

Jarvis E.D., Mirarab S., Aberer A.J., Li B., Houde P., Li C., Ho S.Y., Faircloth B.C., Nabholz B., Howard J.T., Suh A., Weber C.C., da Fonseca R.R., Li J., Zhang F., Li H., Zhou L., Narula N., Liu L., Ganapathy G., Boussau B., Bayzid M., Zavidovych V., Subramanian S., Gabaldón T., Capella-Gutiérrez S., Huerta-Cepas J., Rekepalli B., Munch K., Schierup M., Lindow B., Warren W.C., Ray D., Green R.E., Bruford M.W., Zhan X., Dixon A., Li S., Li N., Huang Y., Derryberry E.P., Bertelsen M., Sheldon F.H., Brumfield R.T., Mello C.V., Lovell P.V., Wirthlin M., Schneider M., Prosdocimi F., Samaniego J., Velazquez A., Alfaro-Núñez A., Campos P.F., Petersen B., Sicheritz-Ponten T., Pas A., Bailey T., Scofield P., Bunce M., Lambert D.M., Zhou Q., Perelman P., Driskell A.C., Shapiro B., Xiong Z., Zeng Y., Liu S., Li Z., Liu B., Wu K., Xiao J., Yinqi X., Zheng Q., Zhang Y., Yang H., Wang J., Smeds L., Rheindt F.E., Braun M., Fjeldsa J., Orlando L., Barker K.F., Jønsson K., Johnson W., Koepfli K.-P., O'Brien S., Haussler D., Ryder O.A., Rahbek C., Willerslev E., Graves G.R., Glenn T.C., McCormack J., Burt D., Ellegren H., Alström P., Edwards S.V., Stamatakis A., Mindell D.P., Cracraft J., Braun E.L., Warnow T., Jun W., Gilbert T.M., Zhang G. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 346:1320-1331.

Jermiin L.S., Ho S.Y., Ababneh F., Robinson J., Larkum A.W. 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst. Biol. 53:638-643.

Jermiin L.S., Jayaswal V., Ababneh F.M., Robinson J. 2017. Identifying optimal models of evolution. In Bioinformatics (pp. 379-420). Humana Press, New York, NY.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Jermiin L.S., Lovell D.R., Misof B., Foster P.G., Robinson J. 2020. Detecting and visualising the impact of heterogeneous evolutionary processes on phylogenetic estimates. bioRxiv. 828996.

Jiang X., Edwards S.V., Liu L. 2020. The multispecies coalescent model outperforms concatenation across diverse phylogenomic datasets. Syst. Biol. 69:795-812.

Jiao H., Zhang L., Xie H.W., Simmons N.B., Liu, H., Zhao H. 2019. Trehalase gene as a molecular signature of dietary diversification in mammals. Mol. Biol. Evol. 36:2171-2183.

Johnson M.G., Gardner E.M., Liu Y., Medina R., Goffinet B., Shaw J.A., Zerega N.J., Wickett N.J. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. Appl. Plant. Sci. 5:1600016.

Jones M.R., Good J.M. 2016. Targeted capture in evolutionary and ecological genomics. Mol. Ecol. 25:185-202.

Juste J., Álvarez Y., Tabarés E., Garrido-Pertierra A., Ibáñez C., Bautista J.M. 1999. Phylogeography of African fruitbats (Megachiroptera). Mol. Phylogenet. Evol. 13:596-604.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772-780.

Kirsch J.A., Lapointe, F.J. 1997. You aren't (always) what you eat: Evolution of nectar-feeding among Old World fruitbats (Megachiroptera: Pteropodidae). In: Molecular Evolution and Adaptive Radiation, T. J. Givnish, K. J. Sytsma, eds., pp. 313-330, Cambridge University Press, New York.

Kishino H., Miyata T., Hasegawa M. 1990. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. J. Mol. Evol. 31:151-160.

Koopman K.F. 1994. Handbuch der Zoologie (Handbook of Zoology), volume VIII, part 60. Mammalia. Walter de Gruyter, Berlin, Germany.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Kumar S., Filipski A.J., Battistuzzi F.U., Kosakovsky Pond S.L., Tamura K. 2012. Statistics and truth in phylogenomics. Mol. Biol. Evol. 29:457-472.

Lanfear R., Calcott B., Kainer D., Mayer C., Stamatakis A. 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. 14:82.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34:772-773.

Lemmon A.R., Emme S.A., Lemmon E.M. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst. Biol. 61:727-744.

Li C., Hofreiter M., Straube N., Corrigan S., Naylor G.J. 2013. Capturing protein-coding genes across highly divergent species. Biotechniques. 54:321-326.

Liu L., Yu L., Edwards S.V. 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evol. Biol. 10:302.

Liu Y., Xu H., Yuan X., Rossiter S.J., Zhang S. 2012. Multiple adaptive losses of alanine-glyoxylate aminotransferase mitochondrial targeting in fruit-eating bats. Mol. Biol. Evol. 29:1507-1511.

Liu L., Wu S., Yu L. 2015. Coalescent methods for estimating species trees from phylogenomic data. J. Syst. Evol. 53:380-390.

Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46:523-536.

Manthey J.D., Campillo L.C., Burns K.J., Moyle R.G. 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (Aves, Genus: Piranga). Syst. Biol. 65:640-650.

Marshall A.G. 1983. Bats, flowers and fruit: evolutionary relationships in the Old World. Biol. J. Linn. Soc. 20:115-135.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Maryanto I., Yani M. 2003. A new species of *Rousettus* (Chiroptera: Pteropodidae) from Lore Lindu, Central Sulawesi. Mamm. Study. 28:111-120.

McGowen M.R., Tsagkogeorga G., Alvarez-Carretero S., dos Reis M., Struebig M., Deaville R., Jepson P.D., Jarman S., Polanowski A., Morin P.A., Rossiter S.J. 2020. Phylogenomic resolution of the cetacean tree of life using target sequence capture. Syst. Biol. 69:479-501.

Meiklejohn K.A., Faircloth B.C., Glenn T.C., Kimball R.T., Braun E.L. 2016. Analysis of a rapid evolutionary radiation using ultraconserved elements: evidence for a bias in some multispecies coalescent methods. Syst. Biol. 65:612-627.

Mendes F. K., Hahn M.W. 2018. Why concatenation fails near the anomaly zone. Syst. Biol. 67:158-169.

Meredith R.W., Janečka J.E., Gatesy J., Ryder O.A., Fisher C.A., Teeling E.C., Goodbla A., Eizirik E., Simão T.L., Stadler T., Rabosky D.L., Honeycutt R.L., Flynn J.J., Ingram C.M., Steiner C., Williams T.L., Robinson T.J., Burk-Herrick A., Westerman M., Ayoub N.A., Springer M.S., Murphy W.J. 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science. 334:521-524.

Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 gateway computing environments workshop:1-8.

Minh B.Q., Hahn M., Lanfear R. 2020. New methods to calculate concordance factors for phylogenomic datasets. Mol. Biol. Evol. 37:2727-2733.

Mirarab S., Reaz R., Bayzid M.S., Zimmermann T., Swenson M.S., Warnow T. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics. 30:i541-i548.

Mirarab S., Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics. 31:i44-i52.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Molloy E.K., Warnow T. 2018. To include or not to include: the impact of gene filtering on species tree estimation methods. Syst. Biol. 67:285-303.

Naser-Khdour S., Minh B.Q., Zhang W., Stone E.A., Lanfear R. 2019. The prevalence and impact of model violations in phylogenetic analysis. Genome Biol. Evol. 11:3341-3352.

Nesi N., Kadjo B., Pourrut X., Leroy E., Shongo C.P., Cruaud C., Hassanin A. 2013. Molecular systematics and phylogeography of the tribe Myonycterini (Mammalia, Pteropodidae) inferred from mitochondrial and nuclear markers. Mol. Phylogenet. Evol. 66:126-137.

Nguyen L.T., Schmidt H.A., von Haeseler A., Minh B.Q. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32:268-274.

Oliver J.C. 2013. Microevolutionary processes generate phylogenomic discordance at ancient divergences. Evolution. 67:1823-1830.

Philippe H., Delsuc F., Brinkmann H., Lartillot N. 2005. Phylogenomics. Annu. Rev. Ecol. Evol. Syst. 36:541-562.

Picardi E., Pesole G. 2012. Mitochondrial genomes gleaned from human whole-exome sequencing. Nat. Methods. 9:523-524.

Potter J.H.T., Davies K.T.J., Yohe L.R., Sanchez, M.K.R., Rengifo, E.M., Struebig M., Warren K., Tsagkogeorga G., Lim B.K., Mario dos Reis M., Dávalos L.M., Rossiter S.J. 2021. Dietary diversification and specialisation in New World bats facilitated by early molecular evolution. Mol. Biol. Evol. Advance Access https://doi.org/10.1093/molbev/msab028.

Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature. 526:569-573.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217-223.

Rojas D., Vale A., Ferrero V., Navarro L. 2012. The role of frugivory in the diversification of bats in the Neotropics. J. Biogeogr. 39:1948-1960.

Rookmaaker L.C., Bergmans W. 1981. Taxonomy and geography of *Rousettus amplexicaudatus* (Geoffroy, 1810) with comparative notes on sympatric congeners (Mammalia, Megachiroptera). Beaufortia. 31:1-29.

Sayyari E., Mirarab S. 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33:1654-1668.

Schliep K.P. 2011. Phangorn: phylogenetic analysis in R. Bioinformatics. 27:592–593.

Sharma V., Hecker N., Roscito J.G., Foerster L., Langer B.E., Hiller M. 2018. A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. Nat. Comm. 9:1215.

Sheherazade., Ober H.K., Tsang S.M. 2019. Contributions of bats to the local economy through durian pollination in Sulawesi, Indonesia. Biotropica. 51:913-922.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51:492-508.

Simmons M.P., Sloan D.B., Gatesy J. 2016. The effects of subsampling gene trees on coalescent methods applied to ancient divergences. Mol. Phylogenet. Evol. 97:76-89.

Simmons N.B. 2005. Order Chiroptera. In: Wilson D. E., Reeder D. M., editors. Mammal species of the world: a taxonomic and geographic reference. 3rd ed. Baltimore: Johns Hopkins University Press. p. 312-529.

Simmons N.B., Cirranello A.L. 2019. Bat Species of the World: A taxonomic and geographic database. (batnames.org). Accessed on 06/25/2019

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Smith S.A., Moore M.J., Brown J.W., Yang Y. 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15:150.

Springer M.S., Gatesy J. 2016. The gene tree delusion. Mol. Phylogenet. Evol. 94:1-33.

Teeling E.C., Springer M.S., Madsen O., Bates P., O'Brien S.J., Murphy W.J. 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. Science. 307:580-584.

Suh A., Smeds L., Ellegren H. 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. PLoS Biol. 13:e1002224.

Wang K., Tian S., Galindo-González J., Dávalos L.M., Zhang Y., Zhao H. 2020. Molecular adaptation and convergent evolution of frugivory in Old World and neotropical fruit bats. Mol. Ecol. 29:4366-4381.

Wilson D. E., Mittermeier R. A. 2019. Handbook of the Mammals of the World. Vol. 9. Bats. Lynx Edicions, Barcelona.

Whitfield J.B., Lockhart P.J. 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22:258-265.

Woodhams M.D., Fernández-Sánchez J., Sumner J.G. 2015. A new hierarchy of phylogenetic models consistent with heterogeneous substitution rates. Syst. Biol. 64:638-650.

Xi Z., Liu L., Rest J.S., Davis C.C. 2014. Coalescent versus concatenation methods and the placement of Amborella as sister to water lilies. Syst. Biol. 63:919-932.

Yang Z., Rannala B. 2012. Molecular phylogenetics: principles and practice. Nat. Rev. Genet. 13: 303-314.

Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 19:153

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Zhao H., Rossiter S.J., Teeling E.C., Li C., Cotton J.A., Zhang S. 2009. The evolution of color vision in nocturnal mammals. Proc. National Acad. Sci. 106:8980-8985.

Zou L., Susko E., Field C., Roger A. J. 2012. Fitting nonstationary general-time-reversible models to obtain edge-lengths and frequencies for the Barry–Hartigan model. Syst. Biol. 61:927-940.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

TABLES

Table 1. Summary of datasets and methods of analyses used in the study. Number of taxa, loci, size for each dataset and number of partitions following the SRH assumptions are indicated, as well as the recovered topology by each method. Topology code follows Figure 2.

					Concatenation approach	Species tree coalescent-based	
Dataset	Nb of taxa	Nb of loci	Size alignment (bp)	Nb partitions pass SRH $p \geq 0.05$	ML	ASTRAL	MP-EST
Nuclear CDSs							
Complete	137	1455	2,791,426	473	Topology 1	Topology 2	Topology 3
Parsimony Informative filtering							
PI ≥ 18%	137	1088	2,100,753	314	Topology 1	Topology 2	Topology 3
PI ≥ 22%	137	728	1,405,743	201	Topology 1	Topology 3	Topology 3
PI ≥ 27%	137	341	631,544	103	Topology 2	Topology 2	Topology 3
Mitochondrial markers							
Complete 15 loci	142	15	13,950	NA	Topology 4	NA	NA

NA, not applicable; Nb, number; ML, Maximum Likelihood.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

FIGURES

Figure 1. Phylogram based on Maximum Likelihood inference of the complete CDS dataset (1455 loci). Backbone nodes are indicated by the letters A to F. Node bootstrap support values are ≥ 90% unless indicated otherwise. Pie charts on backbone nodes show the proportion of genes trees that support that clade (blue), the proportion that support the main alternative for that clade (green), the proportion that support the remaining alternatives (red), and the proportion of uninformative gene trees for that node (grey). Numbers next to the pie charts correspond to ICA, gCF and sCF values for backbone nodes. For two species, *Aproteles bulmerae* and *Mirimiri acrodonta*, the putative placement here using dashed lines is based on our mitochondrial tree (Supplementary Fig. S13 available on Dryad). The tribes, following Bergmans (1997) and Almeida et al. (2020), are listed to the right, and are depicted by a solid line if recovered as monophyletic in our analysis, or by a dashed line if not monophyletic. Green branches correspond to nectar-feeding pteropodid lineages. Representatives species from subfamilies and tribes across the family from top to bottom are: *Epomophorus gambianus* (Gambian epauletted fruit bat), *Plerotes anchietae* (Anchieta's broad-faced fruit bat), *Myonycteris brachycephala* (São Tomé collared fruit bat), *Stenonycteris lanosus* (long-haired rousette), *Rousettus aegyptiacus* (Egyptian rousette), *Eonycteris spelaea* (lesser dawn bat), *Scotonycteris occidentalis* (Hayman's tear-drop fruit bat), *Eidolon dupreanum* (Malagasy straw-colored fruit bat), *Dyacopterus rickarti* (Philippine large-headed fruit bat), *Cynopterus sphinx* (greater short-nosed fruit bat), *Aproteles bulmerae* (Bulmer's fruit bat), *Harpyionycteris celebensis* (Sulawesi harpy fruit bat), *Pteropus samoensis* (Samoan flying fox), *Pteralopex pulchra* (montane monkey-faced fruit bat), *Melonycteris melanops* (black-bellied fruit bat), *Notopteris macdonaldii* (Fijian long-tailed fruit bat), *Nyctimene robinsoni* (Queensland tube-nosed fruit bat), and *Syconycteris carolinae*

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

(Halmaheran blossom bat). Illustrations by I. Velikov from Wilson and Mittermeier (2019), reproduced with permission.

Figure 2. a) Phylogenetic relationships of the family Pteropodidae inferred using concatenation and species-tree coalescent-based methods on the complete CDS, the parsimony informative (PI) filtered CDS and the concatenated mtDNA datasets. Backbone nodes are named as in Figure 1. Coloured clades correspond to: Rousettinae, turquoise; Cynopterinae, orange; Harpiyonycterinae, green; Pteropodinae, yellow; Nyctimeninae, blue; and Macroglossusinae, red. b) On the left, the DensiTree cloudogram inferred from nuclear genes for the reduced 16-taxon data set. For contrast, Topology 1 is shown in black. On the right, Gene Genealogy Interrogation (GGI) results testing alternative hypotheses of the backbone relationships within the Pteropodidae. Plot: lines correspond to the cumulative number of genes (*x*-axis) supporting topology hypothesis with highest probability and their *P*-values (*y*-axis) from the approximately unbiased (AU) topology tests. Values above the dashed line indicate topology hypotheses that are significantly better than the alternatives ($P < 0.05$).

Figure 3. Tanglegram of the Topology 1 recovered in this study (see Table 1) on the left and the pteropodids topology presented in Almeida et al. 2011 (based on Maximum Likelihood tree, Figure 2) on the right. Branches are coloured by major clades: Rousettinae, turquoise; Eidolinae, dark purple; Cynopterinae, orange; Harpiyonycterinae, green; Pteropodinae, yellow; Nyctimeninae, blue; Notopterisinae, purple; and Macroglossusinae, red. Grey lines connect taxa between trees. Genera only present in our study are indicated in bold and nectar feeder genera are coloured in green. Bootstrap support is reported for nodes of the Almeida et al. 2011 topology. Dashed branches correspond to relationships with node support value < 50. * indicates *Micropteropus* synonymized to *Epomophorus*.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

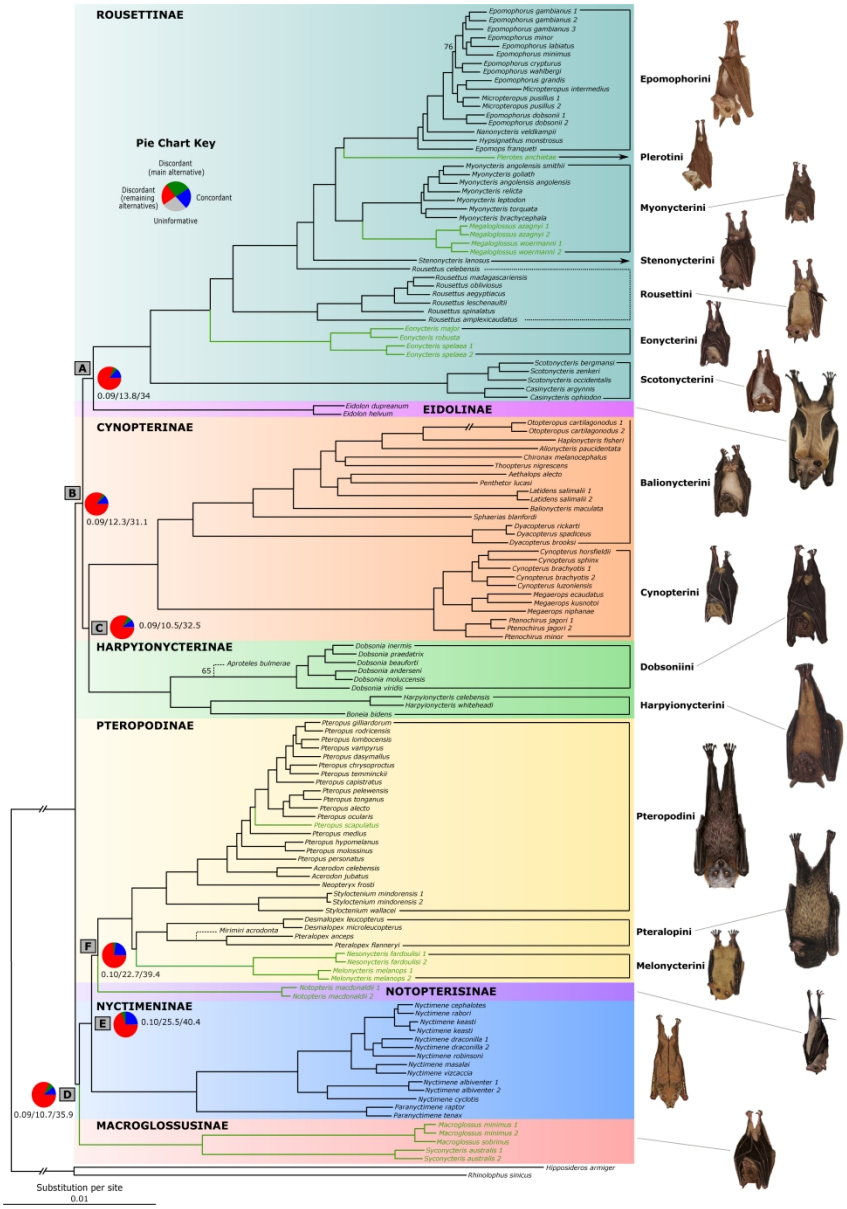


Figure 1. Phylogram based on Maximum Likelihood inference of the complete CDS dataset (1455 loci). Backbone nodes are indicated by the letters A to F. Node bootstrap support values are ≥ 90% unless indicated otherwise. Pie charts on backbone nodes show the proportion of genes trees that support that clade (blue), the proportion that support the main alternative for that clade (green), the proportion that support the remaining alternatives (red), and the proportion of uninformative gene trees for that node (grey). Numbers next to the pie charts correspond to ICA, gCF and sCF values for backbone nodes. For two species, Aproteles bulmerae and Mirimiri acrodonta, the putative placement here using dashed lines is based on our mitochondrial tree (Supplementary Fig. S13 available on Dryad). The tribes, following Bergmans (1997) and Almeida et al. (2020), are listed to the right, and are depicted by a solid line if recovered as monophyletic in our analysis, or by a dashed line if not monophyletic. Green branches correspond to nectar-feeding pteropodid lineages. Representatives species from subfamilies and tribes across the family from top to bottom are: Epomophorus gambianus (Gambian epauletted fruit bat), Plerotes anchietae (Anchieta's broad-faced fruit bat), Myonycteris brachycephala (São Tomé collared fruit bat), Stenonycteris lanosus (long-haired rousette), Rousettus aegyptiacus (Egyptian rousette), Eonycteris spelaea (lesser dawn bat),

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

Scotonycteris occidentalis (Hayman's tear-drop fruit bat), Eidolon dupreanum (Malagasy straw-colored fruit bat), Dyacopterus rickarti (Philippine large-headed fruit bat), Cynopterus sphinx (greater short-nosed fruit bat), Aproteles bulmerae (Bulmer's fruit bat), Harpyionycteris celebensis (Sulawesi harpy fruit bat), Pteropus samoensis (Samoan flying fox), Pteralopex pulchra (montane monkey-faced fruit bat), Melonycteris melanops (black-bellied fruit bat), Notopteris macdonaldii (Fijian long-tailed fruit bat), Nyctimene robinsoni (Queensland tube-nosed fruit bat), and Syconycteris carolinae (Halmaheran blossom bat). Illustrations by I. Velikov from Wilson and Mittermeier (2019), reproduced with permission.

303x430mm (300 x 300 DPI)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

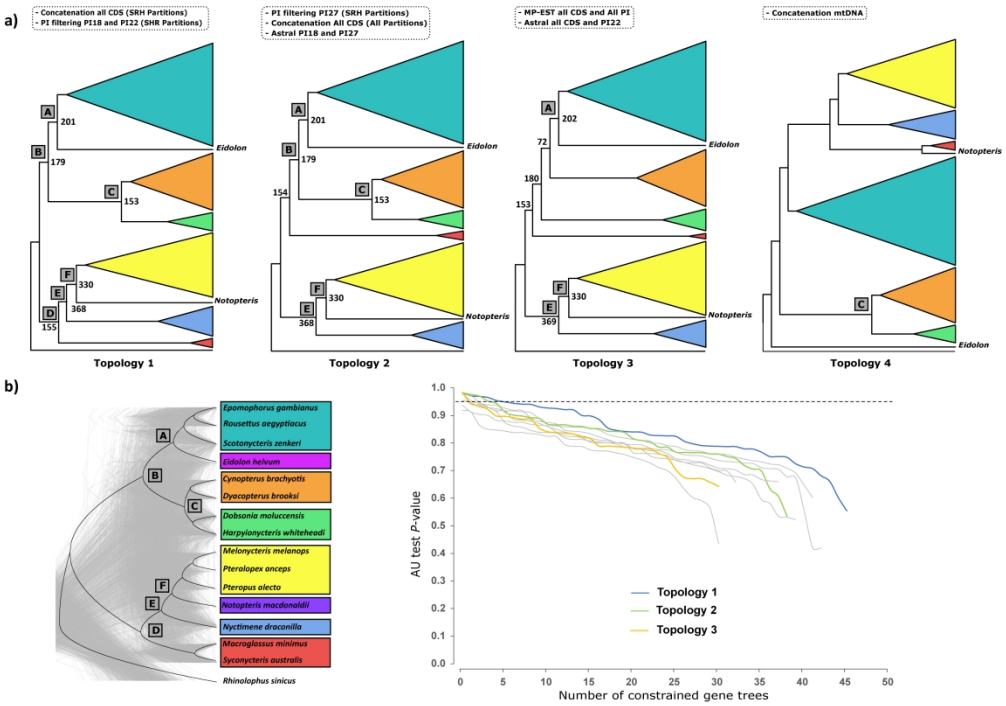


Figure 2. a) Phylogenetic relationships of the family Pteropodidae inferred using concatenation and species-tree coalescent-based methods on the complete CDS, the parsimony informative (PI) filtered CDS and the concatenated mtDNA datasets. Backbone nodes are named as in Figure 1. Coloured clades correspond to: Rousettinae, turquoise; Cynopterinae, orange; Harpiyonycterinae, green; Pteropodinae, yellow; Nyctimeninae, blue; and Macroglossusinae, red. b) On the left, the DensiTree cloudogram inferred from nuclear genes for the reduced 16-taxon data set. For contrast, Topology 1 is shown in black. On the right, Gene Genealogy Interrogation (GGI) results testing alternative hypotheses of the backbone relationships within the Pteropodidae. Plot: lines correspond to the cumulative number of genes (x-axis) supporting topology hypothesis with highest probability and their P-values (y-axis) from the approximately unbiased (AU) topology tests. Values above the dashed line indicate topology hypotheses that are significantly better than the alternatives ($P < 0.05$).

286x199mm (300 x 300 DPI)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

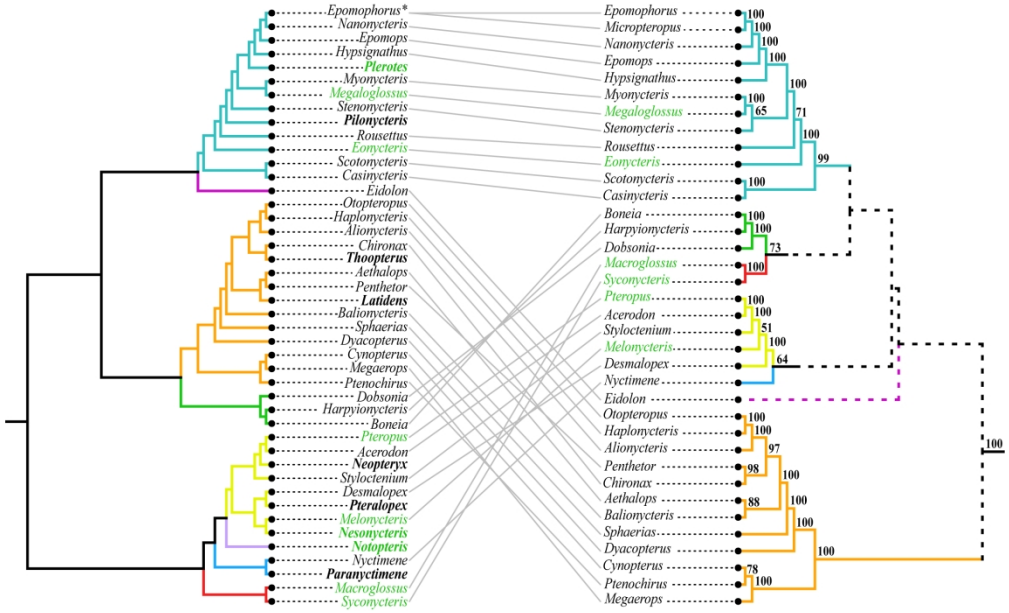


Figure 3. Tanglegram of the Topology 1 recovered in this study (see Table 1) on the left and the pteropodids topology presented in Almeida et al. 2011 (based on Maximum Likelihood tree, Figure 2) on the right. Branches are coloured by major clades: Rousettinae, turquoise; Eidolinae, dark purple; Cynopterinae, orange; Harpiyonycterinae, green; Pteropodinae, yellow; Nyctimeninae, blue; Notopterisinae, purple; and Macroglossusinae, red. Grey lines connect taxa between trees. Genera only present in our study are indicated in bold and nectar feeder genera are coloured in green. Bootstrap support is reported for nodes of the Almeida et al. 2011 topology. Dashed branches correspond to relationships with node support value < 50. * indicates Micropteropus synonymized to Epomophorus.

199x121mm (300 x 300 DPI)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021

The data associated with this paper are available for review via Dryad. The following is a temporary direct download link. Please copy and paste it directly into a web browser to download the data files to your computer (unfortunately this may not work as a link to click on)

https://datadryad.org/stash/share/lV9DLg4fDAX8SVJXp0s7BlTY37eD32wFnGE-ZHrqmOI

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab013/6161783 by Sorbonne Université user on 12 March 2021