Research Article

Four species in one: multigene analyses reveal phylogenetic patterns within Hardwicke’s woolly bat, *Kerivoula hardwickii*-complex (Chiroptera, Vespertilionidae) in Asia

Vuong Tan Tu1,2,3,4*, Alexandre Hassanin1,2,*, Neil M. Furry5, Nguyen Truong Son3,4, Gábor Csorba6

1Institut de Systematique, Evolution, Biodiversité (ISTEB), UMR 7205 MNHN CNRS UPMC, Museum national d’Histoire naturelle, Case postale N°51–55, rue Buffon, 75005 Paris, France
2Service de Systematique Moleculaire, UMS 2700, Museum national d’Histoire naturelle, Case postale N°26–43, rue Cuvier, 75005 Paris, France
3Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet road, Cau Giay district, Hanoi, Vietnam
4Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet road, Cau Giay districts, Hanoi, Vietnam
5Fauna & Flora International, Cambodia Programme, 19 Street 360, Boeng Keng Kang 1, Chamkarmorn, Phnom Penh, Cambodia
6Department of Zoology, Hungarian Natural History Museum, Baross u. 13, H-1088, Budapest, Hungary

Keywords:
Kerivoulaeae
Asia
phylogeography
taxonomy
cryptic species

Abstract

We undertook a comparative phylogeographic study using molecular, morphological and morphometric approaches to address systematic issues in bats of the *Kerivoula hardwickii* complex in Asia. Our phylogenetic reconstructions using DNA sequences of two mitochondrial and seven nuclear genes reveal a distinct clade containing four small-sized species (*K. hardwickii* sensu stricto, *K. depressa*, *K. furva* and *Kerivoula sp. nov.) previously assigned to *K. hardwickii* sensu lato, as well as *K. kuchinensis*, a distinctly larger taxon. Among the four small species, morphological analyses showed that *K. hardwickii* s.s.s. differs significantly from the other three species in skull shape, whereas *K. depressa*, *K. furva* and the new species appear to be morphologically cryptic species. Molecular dating estimates suggest that species within the *hardwickii*-complex diversified during the Late Pliocene/Early Pleistocene period, most probably in different glacial refugia in Asia. Available evidence indicates that allopatric speciation within the complex included morphological, acoustic, and cytogenetic divergence, although these were not always mutually exclusive. Such adaptive changes would explain how different taxa with overlapping morphological features can share ecological niches and maintain their gene flow in sympathy. Accordingly, we also suggest that two subspecies of *K. hardwickii* sensu lato (*K. h. crypta* and *K. h. malpasi*) originally described as distinct species (from Southern India and Sri Lanka, respectively) can be re-elevated to species rank. Should these be found to be conspecific however, the name *crypta* would have priority and *malpasi* should be treated as its subspecies.

Introduction

Hardwicke’s woolly bat, *Kerivoula hardwickii* (Horsfield, 1824) (sensu Corbet and Hill, 1992) is a moderate-sized species within the *Kerivuinae* with a forearm length of 30.0–34.0 mm. The species is widely distributed throughout the mainland and on most islands in Southeast Asia (Sumatra, Java, Southwestern Europe (Borneo, and the Philippines). Isolated populations also exist on Hainan Island, Taiwan, and in Chongqing Province (central China), Fujian Province (southeast China), northern India (Jammu and Kashmir), northeastern Pakistan, southern India (Karnataka), and Sri Lanka. Across its range, the species has been found from sea level up to 2,100 m in elevation (Hill, 1965; Simmons, 2005; Rosell-Ambal et al., 2008; Yoshiyuki et al., 2010) (Fig. 1).

Several taxa are traditionally listed as subspecies or synonyms of *K. hardwickii*. These include *K. crypta*, *K. depressa*, *K. engana*, *K. fusca*, and *K. malpasi* (Corbet and Hill, 1992; Simmons, 2005; Rosell-Ambal et al., 2008) and can be divided into two morphological types, one comprising taxa with domed skulls (the nominal subspecies plus *crypta, engana, fusca, and malpasi*) and the other including a single taxon with a characteristically flattened skull (*depressa*). Individuals with domed skulls have been found in South India and Sri Lanka, and following a geographic hiatus of thousands of kilometres, from southern Indochina and southern Thailand, through Peninsular Malaysia to the islands of Southeast Asia (Borneo, Sumatra, Java, Kangean Is, Engano, South Sulawesi, and the Philippines). Individuals with flattened skulls have been recorded from Nepal through northeast India, southern China, and Taiwan to southern Thailand, and the range of the two morphological types overlaps (Fig. 1) (Hill, 1965; Corbet and Hill, 1992; Bates and Harrison, 1997).

*Corresponding author
Email addresses: vuvtu1984@ac.vn (Vuong Tan Tu), alexandre.hassanin@mnhn.fr (Alexandre Hassanin)
The systematics of *K. hardwickii* s.l. in Asia has been subjected to several studies in recent years. Bates et al. (2007) suggested that the correct names for a slightly larger taxon with a braincase height (BH) of over 5.1 mm and for a smaller species with a flattened skull (BH<5.1 mm) are *K. hardwickii* and *K. depressa*, respectively. Using mtDNA barcode sequences (COI), Francis et al. (2007, 2010) found five divergent mitochondrial haplogroups in bats tentatively identified as *K. hardwickii*. Khan et al. (2010) also detected high levels of divergence (5.7%) in Cytb sequences between populations of *K. hardwickii* s.l. in Borneo. In Thailand, *K. hardwickii* specimens were divided into two morphotypes (domed-skulls vs. flattened skulls) and COI analyses indicated that specimens in the domed-skull group formed a monophyletic clade, whereas those with flattened skulls resided in two divergent monophyletic clades (Doughangubhpa et al., 2015). Likewise, in Vietnam, Son et al. (2015) noted that populations of *K. hardwickii* s.l. formed three morphometrically separable groups.

Most recently, Kuo et al. (2017) found that bats of the *hardwickii-*complex belonged to four distinct clades based on COI sequences (A: central and south Laos, central and southern Vietnam and north Thailand; B: north and south Laos, south Vietnam, north and south Thailand; C: south Laos, south Vietnam, south Thailand, Malay and Borneo; and D: Taiwan, southeast China and north Myanmar). Specimens of the distinctly larger *K. kachinensis* were nested between these clades. In contrast, phylogenetic reconstructions based on RAG2 sequences supported the monophyly of *K. kachinensis* and of *K. hardwickii* s.l., with no signals indicating separation of the four mtDNA lineages comprising the latter. Morphometric analysis in the same study revealed three groups which included individuals of the A+B, C, and D haplogroups, respectively. Kuo et al. (2017) subsequently concluded that bats in clades A and B represented two mitochondrial lineages of a single taxon assigned to *K. depressa*, that bats in clade C represented *K. hardwickii* s.s., whereas bats in clade D were described as a new species: *K. furva*. However, the authors acknowledged that their morphometric analyses might be biased due to the limited number of specimens examined and that additional genetic makers should be included to test for discordance between their mtDNA and nuDNA phylogenies (Kuo et al., 2017).

Using the broadest geographic coverage and taxon sample to date, we employ molecular (mitochondrial and nuclear markers) and morphological analyses to further investigate phylogenetic relationships within the genus *Kerivoula* and address the following questions: (1) how many species occur within the *K. hardwickii-*complex in Southeast Asia; (2) what are their geographic distributions (3) what levels of intra- and interspecific variation occur in mtDNA and nuDNA sequences; (4) can gene flow (in particular mtDNA introgression) be characterised between the taxa; (5) how and when did the *hardwickii-*complex diversify; and (6) what are the phylogenetic relationships of these species and their affinities with other species within the genus *Kerivoula*?

Materials and methods

Taxonomic sampling

Sixty-seven tissue samples, including 64 *Kerivoula* spp. and three outgroup species were included in our analyses (Tab. S2). Most of these came from recent field expeditions undertaken by the authors. In the field, bats were captured with the use of mist nets (Ecotone, Poland) and four-bank harp-traps and handled in accordance with guidelines approved by the American Society of Mammalogists (Sikes et al., 2011). These were measured, photographed and initially identified using Francis (2008). Tissue samples were collected from muscles or patagia and preserved in 95% ethanol in 2 ml Eppendorf tubes. A few samples were taken from older specimens in museum collections. The study specimens are deposited in the following institutions: The Institute of Ecology and Biological Resources (Hanoi, Vietnam), the Hungarian Natural History Museum (Budapest, Hungary), the Centre for Biodiversity Conservation (Phnom Penh, Cambodia), the Museum national d’Histoire naturelle (Paris, France), and the Naturalis Biodiversity Center (Leiden, Netherlands) (Tab. S2). The three outgroup species were chosen on the basis of previous molecular studies (Hoofer et al., 2003; Khan et al., 2010; Ruedi et al., 2012) and represent three vespertilionid genera from two subfamilies: *Myotis muricola* belongs to Myotinidae, *Harpiocephalus harpaia*, and *Murina cyclotis* to Murinidae.

DNA extraction, amplification, sequencing

Total DNA was extracted using QIAGEN DNeasy Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol. Two mitochondrial genes were sequenced for this study: the COI barcode fragment and the complete Cytb gene. The primers used for PCR amplification of mitochondrial genes were published in Hassanin et al. (2012) and Hassanin (2014).

Seven nuclear genes (six introns and one exon) were sequenced: intron 7 of FGB (the nuclear β-fibrinogen gene), intron 6 of HDAC1 (histone deactetylase 1), intron 10 of HDAC2 (histone deactetylase 2), intron 6 of RIOK3 (RIO kinase 3), intron 9 of TUFM (elongation factor Tu, mitochondrial precursor), intron 6 of ZFYVE27 (zinc finger, FYVE domain containing 27), and RAG2, a recombination activating gene that encodes a protein involved in the V(D)J recombination. The primers used for PCR amplification of nuclear introns were detailed in Hassanin et al. (2013). The primer set used for amplifying RAG2 were published in Hassanin et al. (2018).

Amplifications were done in 20 µl using 3 µl of Buffer 10X with MgCl2, 2 µl of dNTP (6.6 mM), 0.12 µl of Taq DNA polymerase (2.5 U, Qiagen, Hilden, Germany) and 0.5–1 µl of the two primers at 10 µM. The standard PCR conditions were as follows: 4 min at 95 °C; 5 cycles of denaturation/annealing/extension with 45 s at 95 °C, 1 min at 60 °C and 1 min at 72 °C, followed by 30 cycles of 30 s at 95 °C, 45 s at 55 °C, and 1 min at 72 °C, followed by 10 min at 72 °C. PCR products were resolved by electrophoresis on a 1.5% agarose gel stained with ethidium bromide and visualized under UV light.

Both strands of PCR products were sequenced using Sanger sequencing on an ABI 3730 automatic sequencer at the Centre National de Séquençage (Genoscope) in Evry (France). The sequences were edited and assembled using CodonCode Alignment Version 3.7.1 (CodonCode Corporation) and Sequencher 5.1 (Gene Codes Corporation). Heterozygous positions (double peaks) were scored using the IUPAC ambiguity codes. Sequences generated for this study were deposited in the EMBL/DBJ/GenBank database (accession numbers MH137299-MH137612).

Phylogenetic analyses

Our sequences were compared to 129 COI, 67 Cytb, and 25 RAG2 sequences of the subfamily Kerivoulineae extracted from the nucleotide databases. The origins of all new samples and downloaded sequences are detailed in Tab. S3.
the mtDNA dataset was run using a GTR+I+G model for each of the seven nuclear genes; 6214 bp and 67 indels), mtDNA (8059 bp and 67 indels), nuDNA (combining all the seven nuclear genes; 9917 bp and 148 indels) were analysed separately: supermatrix (combining all the nine genes; 19,131 bp and 216 indels). Three codon positions for the two combined genes; and the supermatrix was run using selected models for each partition. The posterior probabilities (PP) were calculated using four independent Markov chains run for 10,000,000 Metropolis-coupled MCMC generations, with tree sampling every 1000 generations, and a burn-in of 25%.

The results obtained from the separate Bayesian analyses of the eight independent molecular markers (mtDNA and the seven nuclear genes) were analysed using the SuperTRI method (Ropiquet et al., 2009). The lists of bipartitions obtained from the eight Bayesian analyses were transformed into a weighted binary matrix for supertree construction using SuperTRI v.5.7 (available at http://www.normalesup.org/bli/Programs/programs.html). Each binary character corresponds to a node, which was weighted according to its frequency of occurrence in one of the eight lists of bipartitions. In that way, the SuperTRI method takes into account both principal and secondary signals, because all phylogenetic hypotheses found during the eight separate analyses are represented in the weighted binary matrix used for supertree construction. The reliability of the nodes was assessed using three different measures. The first value is the SuperTree Bootstrap Percentage (SBP), which was calculated under PAUP* v.4.10 (Swofford, 2003) after 1000 BP replicates of the weighted binary matrix reconstructed with SuperTRI (1014 characters; heuristic search). The second value is the “Mean Posterior Probability” (MPP) calculated using the lists of bipartitions obtained from Bayesian analyses of the eight data-sets. The third value is the index of reproducibility (Rep), which is the ratio of the number of datasets supporting the node of interest to the total number of data-sets. The MPP and Rep values were directly calculated on SuperTRI v.5.7.

We also applied Bayesian multi-species coalescent analysis in BEAST v.2.1.3 (Bouckaert et al., 2014) to coestimate the shared species tree from seven independent nuclear loci of 31 taxa (Heled and Drummond, 2010). The mtDNA dataset was excluded because the maternal history can be inconsistent with the species tree (e.g. Degnan and Rosenberg, 2002). Substitution models for each nuclear locus were those selected under jModelTest (see above). We ran a MCMC chain of 10 million generations, sampled every 1000 generations, with Yule speciation, a strict clock and a burn-in of 10%. Adequacy of chain mixing and MCMC chain convergence were assessed using the ESS values in Tracer v.1.6 (available in the BEAST package). The consensus topology was generated with TreeAnnotator v.1.7.5 (available in the BEAST package) and visualized with FigTree v.1.4.1 (Rambaut, 2009) and DensiTree v.2.0 (available in the BEAST package).

**Molecular dating**

Divergence times were estimated with the Bayesian approach implemented in BEAST v.2.1.3 (Bouckaert et al., 2014) using a Cytb alignment of 58 taxa and 1140 bp. As no calibration point (fossil record or biogeographic event) was accurate for Kerivoula, we used a mutation rate of 0.02 per site per lineage per Myr with a lower boundary of 0.01 and an upper boundary of 0.025, which is in agreement with previous studies on bats (e.g., HuIva et al., 2004; Mao et al., 2010). We applied a HKY+I+G model of evolution (based on jModelTest) and a relaxed-clock model with uncorrelated lognormal distribution for substitution rate. Node ages were estimated using a Yule speciation prior and 108 generations, with tree sampling every 1000 generations, and a burn-in of 25%. The outputs of BEAST analyses were checked with ESS values of >200 in Tracer v.1.6. The chronograms were generated and visualized by using TreeAnnotator v.1.7.5 and FigTree v.1.4.1 (Rambaut, 2009), respectively.

**Morphological and morphometric analyses**

A total of 89 adult specimens of K. hardwickii s.l. (n=53), K. titania (n=19) and K. kachinenensis (n=17), with intact skulls were examined using one external (forearm length, FA) and 17 craniodental measurements (Tab. S2). Craniodental measurements were taken to the nearest 0.01 mm using digital callipers under stereomicroscope and included: greatest length of skull (GLS), from the anterior of the 1st upper incisor to the most posteriorly projecting point of the occipital region; condylo-canine length (CCL), from the exoccipital condyle to the most anterior part of the canine; greatest width across the upper
canines from their buccal borders (CC); greatest width across the upper first premolars from their buccal borders (P1P2); greatest width across the upper second premolars from their buccal borders (P1P3); greatest width across the upper third premolars from their buccal borders (P1P4); greatest width across the crowns of the last upper molars from their buccal borders (M1M3); greatest width of the skull across the zygomatic arches (ZB); greatest distance across the mastoid region (MB); greatest width of the braincase (BC); braincase height (BH), from the basisphenoid at the level of the hamular processes to the most dorsal part of the skull, including the sagittal crest (if present); maxillary toothrow length (CM3), from the anterior of the upper canine to the posterior of the crown of the 3rd molar; distance from the anterior of the upper canine to the posterior of the crown of the last premolar (CP3); mandible length (ML), from the anterior rim of the alveolus of the 1st lower incisor to the most posterior part of the condyle; mandibular toothrow length (CM4), from the anterior of the lower canine to the posterior of the crown of the 3rd lower molar; distance from the anterior of the lower canine to the posterior of the crown of the last premolar (CP4); and height of the coronoid process (PCH) from the tip of the coronoid process to the apex of the indentation on the inferior surface of the ramus adjacent to the angular process.

In our study, bats of K. hardwickii s. 1. were initially assigned to different groups based on molecular data. Specimens without genetic information were classified into molecular groups on the basis of morphological similarity and geographic origin. Because sexual dimorphism in size between bats in each group was not significant (t-test), the morphological affinities of identified taxa were investigated using Principal Component Analyses (PCA) in PAST (Hammer et al., 2001) combining data for both sexes: (1) 10 log-transformed raw measurements and (2) 10 log-transformed standardized raw measurements that removed size variation by regression of raw score on geometric mean (Jungers et al., 1995; Lindenfors et al., 2007; Son et al., 2015). Statistically significant differences in PC scores between different groups were then tested using Kruskal-Wallis one-way analysis of variance (ANOVA, p<0.05) (Zar, 1999).

Results
Phylogenetic relationships of the genus Kerivoula revealed by mtDNA sequences
Bats were reconstructed from the three mtDNA datasets, i.e., COI and Cytb (Fig. 2) and the concatenation of COI and Cytb sequences, named mtDNA (Fig. S4), highly supported the monophyly of the genus Kerivoula (PP=1).

Within Kerivoula, K. picta and K. pellucida occurred outside of the clade uniting all other studied species (PP=1 in COI/Cytb/mtDNA). In this large clade, two major groups were recovered in all analyses, namely: one group named “papillosa” uniting K. papillosa, K. lenis, and a potentially undescribed species of Kerivoula from Borneo (see Khan et al., 2010 for details) (PP=1), and another group named “hardwickii” including K. intermedia, K. kachinensis, K. krauensis, K. minuta, K. titania, and K. hardwickii s.s. (PP=0.9/0.8/0.98).

With the major exception of K. hardwickii s.s. all species of the hardwickii-group were recovered as monophyletic with maximal support values (Fig. 2). Kerivoula hardwickii s.l. was found to be paraphyletic due to the inclusive position of K. kachinensis (Fig. 2 and S4). Four lineages of K. hardwickii s.l., named Kh1 to Kh4, were supported by maximal PP values in COI, Cytb and mtDNA analyses and all three datasets supported a basal divergence of K. hardwickii Kh1 (PP=0.9/0.8/0.91). In contrast, all other relationships were conflicting between COI, Cytb and mtDNA datasets. For instance, K. kachinensis was the sister-group of Kh4 in the COI tree (PP=0.6), but was grouped with Kh2 in the Cytb and mtDNA trees (PP=0.9/0.85).

Phylogenetic relationships of the genus Kerivoula inferred from RAG2 sequences
The topology of the Bayesian tree inferred from RAG2 sequences (Fig. S5) is similar to that of the mtDNA trees: the genus Kerivoula was recovered as monophyletic with maximal support (PP=1); two species K. picta and K. pellucida were divergent from the clade uniting the papillosa-group and the hardwickii-group (PP=1); and, within the hardwickii-group, three species, K. intermedia, K. minuta and K. titania were monophyletic. Although basal relationships within the hardwickii-group were unstable according to mtDNA data, the RAG2 tree supported the existence of two major clades: one uniting K. intermedia and K. minuta (PP=1), the other including K. titania and an assemblage of K. hardwickii s.l. and K. kachinensis sequences (PP=0.9).

Within the latter assemblage, K. kachinensis and the Kh3 lineage of K. hardwickii s.l. were monophyletic (PP=1). In contrast, the other three lineages of K. hardwickii s.l. were paraphyletic: Kh1 (PP=0.88), Kh2 (PP=1), and Kh4 (PP=0.35).

Multi-locus phylogeny of the genus Kerivoula
Multi-locus phylogenetic analyses were performed using 31 samples, including three outgroup species and several specimens characterized by divergent mtDNA haplotypes for the following species or lineages of Kerivoula: K. hardwickii Kh1 (4 specimens), K. hardwickii Kh2 (3 specimens), K. hardwickii Kh3 (4 specimens), K. hardwickii Kh4 (6 specimens), K. kachinensis (2 specimens), K. cf. papillosa A (1 specimen), K. cf. papillosa B (1 specimen), K. picta (2 specimens), and K. titania (5 specimens) (see details in Tab. S2).

With the exception of the interrelationships between K. kachinensis and the four lineages of K. hardwickii s.l., all of the following phylogenetic relationships were considered as reliable, because they were found in all analyses (mtDNA, nuDNA, BEAST, supermatrix and SuperTRI) and in most separate analyses of the eight independent markers.
The mtDNA distances can be ranked to four categories only. The three first correspond to nuDNA categories: (1) intraspecific (excluding comparisons between the four lineages of *K. hardwickii*): <0.54%; (2) interspecific, between closely related species of the *hardwickii* or *papillosa* groups: 8.13–13.4%; (3) between *hardwickii*-group and *papillosa*-group: 13.65–15.13%. The fourth category includes distances between *K. picta* and “*hardwickii*”/“*papillosa*” groups (15.63–18.70%), and the subfamilies Myotinae, Murininae and Kerivoulinae (17.24–20.10%).

**Molecular dating**

Our molecular dating estimations based on a *Cyb* alignment (Fig. S11) suggest that *K. kachinensis* and the four lineages of *K. hardwickii* diverged from each other during the Late Pliocene/Early Pleistocene period, between 4.03 and 2.72 Mya.

**Discussion**

**How many species exist in the *hardwickii*-complex?**

There are considerable discrepancies between studies identifying numbers of cryptic species in the *hardwickii*-complex and their phylogenetic relationships with other taxa, particularly *K. kachinensis*. Using DNA barcodes, Francis et al. (2007) showed that specimens currently referred to as *K. hardwickii* fell into three unrelated clusters with nucleotide divergences comparable to those calculated among species. Two were related to *K. kachinensis* and the third was linked to *K. papillosa*. More recently, Francis et al. (2010) identified five divergent *COI* haplogroups within *K. hardwickii*. Four were related to *K. kachinensis*, whereas the position of the fifth, containing only one specimen, was uncertain. Likewise, the phylogenetic tree reconstructed from *COI* sequences by Douangboubha et al. (2015) showed that bats of *K. hardwickii* fall into three distinct haplogroups, two characterized by “flattened skulls” and one by “domed skulls”, whereas *K. kachinensis* appeared as sister to *K. papillosa*. Most recently, based on combined data from genetic (*COI* and *RAG2* genes) and morphological analyses, Kuo et al. (2017) concluded that the *hardwickii*-group includes *K. hardwickii* s.s., *K. krauensis*, and *K. kachinensis*, and divided *K. hardwickii* s.l. into *K. hardwickii* s.str., *K. furva* and *K. depressa*, although the latter species contained two divergent clades in a *COI* gene tree.

The results of our phylogenetic analyses of mtDNA genes (*COI* and *Cyb*) and *RAG2* marker are comparable with those reported by Kuo et al. (2017) (Fig. 2, Tab. S4 and S5). However, our analyses of additional nuclear markers confirm the monophyly of bats of *K. kachinensis* and four lineages of *K. hardwickii* s.l. (Kh1-Kh4) (PP=1; Fig. 3 and S6). All these lineages are supported by high *SuperTRI* values (SBP=100; 0.57<MP<0.88; 0.5<Rep=1) and by separate analyses of several independent nuclear genes i.e., *FBG*, *TUFM*, and *ZFYVE27* (Fig. S8). Nucleotide distances estimated with either mtDNA or nuDNA alignments from those of *K. hardwickii* s.l. into *Hardwickiinae*: 5.86–7.62%.

The results of our phylogenetic analyses of mtDNA genes (*COI* and *Cyb*) and *RAG2* marker are comparable with those reported by Kuo et al. (2017) (Fig. 2, Tab. S4 and S5). However, our analyses of additional nuclear markers confirm the monophyly of bats of *K. kachinensis* and four lineages of *K. hardwickii* s.l. (Kh1-Kh4) (PP=1; Fig. 3 and S6). All these lineages are supported by high *SuperTRI* values (SBP=100; 0.57<MP<0.88; 0.5<Rep=1) and by separate analyses of several independent nuclear genes i.e., *FBG*, *TUFM*, and *ZFYVE27* (Fig. S8). Nucleotide distances estimated with either mtDNA or nuDNA alignments from those of *K. hardwickii* s.l. into *Hardwickiinae*: 5.86–7.62%.
have demonstrated that the speciation of closely related species, including Asian bats, is not always accompanied by morphological change (Bickford et al., 2007; Tu et al., 2015, 2017). We therefore conclude that the K. hardwickii-complex comprises K. kachinensis and four distinct species (Kh1-Kh4), the latter being previously referred to collectively as K. hardwickii s.l., three of which are morphologically cryptic.

Current geographic distribution and habitat preferences of the five species within the hardwickii-complex

*Kerivoula hardwickii* s.l. was previously assumed to have a very wide geographic range but because our results indicate that the taxon contains four distinct species, it is necessary to determine their respective geographical ranges and environmental characteristics. According to our data, Kh1 and Kh4 are allopatric and separated around the 16th parallel north, with Kh4 encompassing a large area from Nepal to Taiwan, through northern Vietnam and southern China (Fig. 1). Kh2 and Kh3 occur in sympatry over most of their distribution along the Annamite Range, but the former has a wider range in Indochina. Within the Annamite Range, both are found in sympatry with Kh4 in the northern parts, whereas they co-occur with Kh1 in the southern areas (Fig. 1).

With the exception of Kh4 (= K. furva, see below), which is quite well studied in Taiwan (Chiang et al., 2006; Chang et al., 2010; Liao, 2013), almost nothing is known about the ecology of species within the hardwickii-complex, except that they are forest dwellers and usually occur in sympatry with other *Kerivoula* spp. (Kingston et al., 1999; Bates et al., 2007; Hassanin and Tu, 2010; Son et al., 2015). Francis (2008) indicated that *K. hardwickii* (in a broad sense) roosts in hollow trees or dead leaves in understory of various forest types ranging from lowland evergreen, deciduous or agricultural habitats to montane evergreen forests. However, given that *K. hardwickii* s.l. is a species complex, constituent taxa would be expected to show specialized ecological requirements or behaviours, as typically found between co-occurring sibling bat species (Kingston et al., 2001; Zhang et al., 2005; Soisook et al., 2008; Sun et al., 2008). Previous studies have shown that differences in habitat use, echolocation calls and diet partitioning of insectivorous bats can be explained by skull morphology (Freeman et al., 1981; Barlow et al., 1997; Bogdanowicz et al., 1999). Knowing this, interspecific variation in craniodental characters between Kh1 and Kh2-Kh4 in the complex suggests these may select different habitats and prey, although further investigation is needed to understand how the latter bats can share habitat and food sources by reducing interspecific competition, particularly in areas where several species occur in sympathy (Fig. 1).

Taxonomy of the hardwickii-complex

Within the hardwickii-complex, *K. kachinensis* is easily distinguishable from other species by its distinct morphology e.g. FA (mm): 41.85±1.09 vs ≤34.3; GLS (mm): 16.79±0.44 vs ≤14.28 (Fig. 5 and Tab. S12). However, the lack of genetic information for the holotype of *K. hardwickii* (which has a badly damaged braincase) makes evaluation of affinities among the Kh1-Kh4 clades challenging.

Based on our combined molecular and morphological evidence and distributional data, *K. h. engana* can be identified as the same taxon as Kh1 in this study. These bats are characterized by the “domed skull” typical of *K. hardwickii* (sensu Bates et al., 2007) which separates this taxon from its “flattened skull” congeners Kh2, Kh3 and Kh4 (or *K. depressa* Miller, 1906b sensu Bates et al., 2007) that occur in mainland Southeast Asia (Fig. 1, 5 and 6). Thus we suggest that, in line with the results of (Kuo et al., 2017), our Kh1 haplotype bats represent the genuine *K. hardwickii* with *K. engana* as its synonym.

Previous studies (Hill, 1965; Corbet and Hill, 1992; Bates and Harrison, 1997; Kuo et al., 2017) regarded two additional taxa (*K. crypta* and *K. malpasi*) as synonyms of *K. hardwickii* due to similarities in certain external and craniodental measurements but did not consider variation in other morphological features (i.e. pelage colour, ear structure) (see Wroughton and Ryley, 1913; Phillips, 1932). Both *K. crypta* and *K. malpasi*, described from southern India and Sri Lanka, respectively, are characterised by high BH values and are isolated from other populations of *K. hardwickii* s.l. in northern India and Southeast Asia by long distances (Fig. 1). Similar to Douangboubpha et al. (2015), who found that the range of “domed skull” bats within the hardwickii-complex in Thailand is restricted to the southern part of the country, our study shows that the occurrence of “domed skull” bats in northern Indochina is unlikely. Further studies including specimens from south-
Phylogeny and taxonomy of the Kerivoula hardwickii complex

Figure 5 – Lateral view of skulls and plot of greatest length of skull against braincase height for study Kerivoula spp. a – Kerivoula h. engana (holotype, ♂); b – K. depressa (holotype, ♀); c – K. titania (♀); d – K. kachinensis (♀); e – Kh1 (≡K. hardwickii s. str., ♀); f – Kh2 (≡K. depressa, ♀); g – Kh3 (≡K. dongduongana sp. nov., holotype, ♀); and h – Kh4 (≡K. furva, ♀). Scale bar=10 mm.

ern Myanmar and other regions of the Indian Subcontinent are evidently required to test for recent gene flow among K. hardwickii s.str. and K. crypta/K. malpasi. On biogeographical grounds, the latter taxa could be regarded as separate species but should these be found to be conspecific, the name crypta would have priority and malpasi should be treated as its subspecies.

Regarding the taxonomy and nomenclature of the three cryptic “flattened skull” species within the hardwickii-complex, our Kh2 and Kh3 clades correspond to clades A and B of Kuo et al. (2017), whereas those of Kh4 represent K. furva. Although Kuo et al. (2017) discriminated K. furva from K. depressa morphologically, our morphological comparisons showed significant overlap between Vietnamese specimens of K. furva, Kh2 and Kh3 bats, and the type specimens of K. depressa. The discrepancy between our study and Kuo et al. (2017) in distinguishing taxa within the complex can be explained by the different datasets used (e.g., Kuo et al., 2017 did not include Vietnamese specimens of K. furva) or by the inherent difficulty in separating morphologically cryptic taxa (e.g., Tu et al., 2017).

Although Kuo et al. (2017) found K. furva from Northeast India to Taiwan with the southern limit in Guangxi (China) and Kachin (Myanmar), our data indicate that the southern and western distributional limits for the taxon can be drawn around the 18°N parallel in northern Vietnam and in eastern Nepal, respectively. However, there are no records of K. furva or Kh3 bats between the Karin Hills and northern Vietnam, and this region is occupied exclusively by bats within the Kh2 clade (Fig. 1). Although further investigations are needed to confirm this pattern, our results and biogeographic considerations corroborate the validity of the recently described K. furva and suggest that bats within the Kh2 clade can be assigned to K. depressa. As a consequence, bats within the Kh3 clade belong to a previously unnamed species which is formally described below.

Systematic description

Kerivoula dongduongana sp.n. (as Kh3 in Fig. 5, 6 and 7)

Description of holotype: Mass: 4.5 g. Measurements (in mm) are as follows: FA: 33.0; Tail: 38.0; Tibia: 18; GLS: 13.79; CCL: 12.76; CC: 3.33; M3M3: 5.10; ZB: 8.25; MB: 7.56; BW: 7.2; BH: 4.76; CM3: 5.45; ML: 9.78; CM3: 5.76, and PCH: 2.92. Its DNA sequences have been deposited in the EMBL/GenBank/DDBJ nucleotide databases with accession numbers shown in Tab. S2.

Type locality: Ngoc Linh Nature Reserve, Kon Tum province, Vietnam (15°4.766′N, 107°49.822′E; 1117 m a.s.l.).

Paratypes: IEBR-VN11-1178 (=MNHN (CG) 2017-2724) (Field n: VN11-1501; adult ♀), body in alcohol, skulls extracted; collected in Ngoc Linh Nature Reserve (15°4.149′N, 107°49.822′E; 1533 m a.s.l.); MNHN (CG) 2017-2725 (Field n: CPV10-292, adult, ♂), MNHN (CG) 2017-2726 (Field n: CPV10-295, adult, ♂), MNHN (CG) 2017-2727 (Field n: CPV10-297, adult, ♂), body in ethanol, skull removed, collected in Vinh Chai National Park, Ratanakiri, Cambodia; NHNM 2012.30.17 (adult, ♂), Tissue code: 23028), and NHNM 2012.30.23 (adult, ♂, tissue code: 23036), body in ethanol, skull removed, collected in Pu Huong Nature Reserve, Nghe An, Vietnam.

Referred specimens: All specimens identified as belonging to clade Kh3 from Vietnam and Laos (Tab. S3) are referred to the new species.

Etymology: Named to denote its restricted distribution range in the Annamite Mountains of Cambodia, Laos and Vietnam. These three countries were formerly known as Indochina (=“Đông Dương” [pronounced as ‘đông ūng’] in Vietnamese). We propose “Indochina’s woolly bat” as the English name, “chuae-souris lainée d’Indochine” as the French name, and “Đoél lòng xỉ Đồ Dòng” as the Vietnamese name.

Description: This is a moderate-sized Kerivoula species with a FA of 32.00±1.73 mm. Ears are small and rounded and the posterior margin of the ear has a deep, smoothly concave emargination just below the apex. The pelage is characterized by long and woolly hairs which are buff brown to dark brown. Dorsal hairs are dark gray. Ventral hairs are paler and grayish. Wing membranes are dark brown and translucent (Fig. 7).

The skull is small with a GLS of 13.51±0.38 mm, lightly built and relatively domed. The lateral profile of the skull is flattened from the rostrum to the forehead compared with K. hardwickii s.str. (Kh1 or K. h. engana in Fig. 5 and 6). The rostrum is short. The second upper incisor (I2) is situated posterior to the first (I1). I2 is one third to half of I2 in height. I2 is one half the height of the upper canine. The maxillary toothrows are convergent anteriorly. The width and crown area of anterior and posterior premolars (P2P2 and P4P4) exceed those of P3P3. The width and crown area of upper molars are relatively similar in size (Fig. 7).

Comparisons with other taxa: As a moderate-sized Kerivoula, K. dongduongana sp. n. differs morphologically from both its larger (K. flora, K. papillosa, K. lenis, K. titania and K. kachinensis) and smaller (K. intermedia, K. minuta, and K. whiteheadi) congeners. The new species differs from K. pellucida in various characters i.e. pelage, degree of wing transparency and craniodental characters (Hill, 1965; Hill and Rozedal, 1989; Francis, 2008) and from K. kraensis by both pelage and craniodental characters (Francis et al., 2007). In relation to similar-sized species within the hardwickii-complex, the new species can be differentiated from K. hardwickii s.str. by its skull shape, although it overlaps morphologically with K. depressa and K. furva (Fig. 5 and 6, Tab. S12). Genetically, K. dongduongana sp. n. differs from other species in the Kerivoulinae by more than 8% in COI and 8.4% in Cytb sequences, and in particular, by two specific deletions in nuDNA genes (deletion of a T at position 534 of HADAC2; deletion of “RW” at position 103 of TUFM).

Distribution: Currently, K. dongduongana sp. n. is known only from the Annamite Mountains of former Indochina (Fig. 1).

How has speciation occurred in the hardwickii-complex? Relationships among the five species of the hardwickii-complex remain problematic, because we found robust conflicting signals (PP>0.98) between independent markers. For instance, Kh2 (=K. depressa) is related to either Kh1 (=K. hardwickii s.str.) (node supported by RJIO3 and ZFYVE27) or K. kachinensis (node supported by HDAC1), whereas Kh4 (=K. furva) is related to Kh3 (=K. dongduongana sp. n.) (node supported by mtDNA and RAG2) or Kh2 (=K. depressa) and K. kachinensis (node supported by HDAC1). Topological conflicts between gene trees can usually be explained by the retention of ancestral polymorphism due to incomplete lineage sorting or genetic introgression (Pamilo and Nei, 1988; Funk and Omland, 2003). All five species within the hardwickii-complex are monophyletic with most markers (Fig. S8), suggesting that lineage sorting may be complete between taxa. This inference agrees with our estimation of their species divergence that took place during late Pliocene and early Pleistocene epochs (4.03–2.72 Mya) (Fig. S11). Regarding the genetic introgression hypothesis, several lines of evidence that support ancient event(s) of introgressive hybridization contributed to the pattern we obtained, whereas recent gene flow between extant taxa in the complex can be ruled out. For instance, as a consequence of their old divergent event(s), species within the complex may have evolved reproductive barriers that prevent recent gene flow between these. Support for this comes from the fact that significant differences in morphology (i.e. body size) between K. kachinensis and the four smaller sized species (K. hardwickii s.str., K. depressa, K. furva, and K. dongduongana sp. n.) might represent boundaries preventing copulation and causing mortality in pregnant females of smaller species carrying an overly large hybrid foetus (Nesi et al., 2011). Recent hybridization between K. hardwickii and K. furva might also be precluded by their cyto-genetic distinctness (see below) and current geographic separation (Fig. 1). The available evidence thus suggests that ancient introgression event(s) between incipient species of the hardwicki complex might have occurred alongside geographical range shifts driven by climate and vegetational change during the Pliocene-Pleistocene transition.

The lack of resolution for interspecies relationships within the hardwickii-complex suggests that their most recent common ancestor rapidly diversified into several species. Our dating estimates based on
Cytb sequences indicate that species diversification within the complex took place in the Late Pliocene-Early Pleistocene (4.03–2.72 Mya) (Fig. S11). During the Miocene and until the early Pliocene, Southeast Asia was a single rainforest block as a result of warm and humid climatic conditions (Morley, 2000; Meijaard and Groves, 2006). The period of diversification therefore appears to coincide with the onset of extensive glaciations during the Late Pliocene, which led to the development of more open vegetation and isolation of forest dependent species including common ancestors of taxa within the hardwickii-complex into several refugia across Asia. Since Kerivoula spp. are characterised by low vagility (Rossiter et al., 2012), vicariance of forest habitats into distinct ancestral refugia might act as casual factors preventing gene flow among isolated ancestral populations of the complex and facilitating allopatric speciation events. In contrast, during interglacial periods of the Pliocene-Pleistocene epochs, warmer and humid conditions promoted the expansion of rain forests (Morley, 2000; Meijaard and Groves, 2006) and subsequently also the distribution range of allopatric incipient taxa from isolated glacial refugia (Khan et al., 2010; Tu et al., 2017). As a consequence, secondary contact(s) between these taxa (via the overlapping geographic range of some extant taxa in Fig. 1) might have facilitated the exchange of genetic material through hybridization (Berthier et al., 2006; Mao et al., 2010). However, based on the inferred ancient rapid radiation, introgression event(s) between incipient taxa in the hardwickii-complex might have ceased since the early phases of glacial-interglacial cycles of the Pliocene-Pleistocene boundary.

Based on our results and available literature, it can be hypothesized that the interspecific diversification involved mechanisms that were not always mutually exclusive. For instance, different isolated ancestral populations adaptively evolved under divergent selection pressures imposed by special paleo-environmental conditions of allopatric ancestral refugia. The larger body size of K. kachinensis suggest that its ancestors might have evolved in a given refugium that selected for gigantism (Taylor et al., 2012), whereas the significant overlap in examined morphological characters among the small sized species (K. hardwickii s.str., K. depressa, K. furva, and K. dongduongana sp.n.) suggest that their ancestors evolved under similar natural selections on these traits (Tu et al., 2015, 2017). However, K. hardwickii s.str. specimens from Borneo have 2n=26 chromosomes (Khan et al., 2010), whereas K. furva specimens from Hainan and Taiwan have 2n=32 chromosomes (Wu et al., 2012; originally identified as K. titania). These cytogenetic differences suggest that chromosomal divergence may have been associated with allopatric speciation within the complex. Thus, further cytogenetic analyses of bats of K. depressa, and K. dongduongana sp.n. are needed to test this hypothesis. If confirmed, it might explain the absence of recent gene flow between morphologically and similarly-sized species living in sympathy.

It should be noted that small taxa within the hardwickii-complex may have divergent echolocation calls, as an indirect result of allopatric speciation. For example, Douangboubpha et al. (2015) found in Thailand that echolocation calls of K. hardwickii A (with “flattened” skulls) differ significantly from those of K. hardwickii C (with domed skull) in parameters such as MaxF (maximum frequency, kHz): 222.0–234.0 vs. 194.0–245.0; MinF (minimum frequency, kHz): 74.0–87.0 vs. 62.0–86.0; and MaxEF (maximum energy frequency, kHz): 145.9–180.7 vs. 129.5–186.4. In addition, based on Kuo et al. (2017), echolocation calls emitted by bats of K. furva found in Taiwan can be differentiated from those of K. hardwickii A (with “flattened” skulls) recorded in Thailand by Douangboubpha et al. (2015) such as MaxF: 182.4–222.1 vs. 222.0–234.0; MinF: 105.7–118.3 vs. 74.0–87.0; and MaxEF: 141.8–163.3 vs. 145.9–180.7. Although divergence in echolocation calls among sister bat taxa could explain how they can share their ecological niches, particularly in sympathy (Kingston et al., 2001), it also suggests that allopatrically ancestral populations might have evolved their own Specific-Mate-Recognition Systems (SMRS) during speciation (Cotterill, 2002; Taylor et al., 2012). These SMRS might have subsequently served as boundaries to gene flow among sister taxa within the hardwickii-complex where their geographical ranges expanded into sympathy. 6%

References


Supplemental information

Additional Supplemental Information may be found in the online version of this article:

**Table S1** Numbers assigned for collecting localities of specimens.

**Table S2** Summary of genetic markers used and specimens that were morphologically examined.

**Table S3** Genbank accession numbers included in phylogenetic reconstructions.

**Figure S4** Bayesian trees reconstructed from mtDNA genes for *Kerivoula* spp. and associated outgroups.

**Figure S5** Bayesian tree of *RAG2* for *Kerivoula* spp. and associated outgroups.

**Figure S6** Bayesian tree of Supermatrix and SuperTRI for *Kerivoula* spp. and associated outgroups.

**Figure S7** SuperTRI analyses.

**Figure S8** Bayesian analyses of the eight independent genes.

**Figure S9** Consensus (A) and Densitree (B) species tree of the genus *Kerivoula*.

**Table S10** Uncorrected pairwise p-distances.

**Figure S11** Chronogram reconstructed from the cytochrome b dataset.

**Table S12** External and craniodental measurements (in mm) of studied *Kerivoula* spp.

**Table S13** Character factor loadings for PCs, obtained from PCAs on log-transformed raw and standardized craniodental measurements of studied *Kerivoula hardwickii* s.l.